

Universität der Bundeswehr München  
Fakultät für Luft- und Raumfahrttechnik  
Institut für Systemdynamik und Flugmechanik

## Zur visuellen Umfeldwahrnehmung autonomer Fahrzeuge

Dipl.-Ing. Ulrich Hofmann

Vollständiger Abdruck der bei  
der Fakultät für Luft- und Raumfahrttechnik  
der Universität der Bundeswehr München  
zur Erlangung des akademischen Grades eines

Doktor-Ingenieurs (Dr.-Ing.)

eingereichten Dissertation.

Vorsitzender: Prof. Dr. rer. nat. Bernd Häusler  
1. Berichterstatter: Prof. Dr.-Ing. Ernst-Dieter Dickmanns  
2. Berichterstatter: Prof. Dr.-Ing. Georg Färber (TU München)

Die Dissertation wurde am 23.06.2004 bei der Universität  
der Bundeswehr München, 85577 Neubiberg, Werner-Heisenberg-Weg 39,  
eingereicht und durch die Fakultät für Luft- und Raumfahrttechnik  
am 30.06.2004 angenommen.

Tag der Promotion: 19.11.2004



# Vorwort

Die vorliegende Arbeit entstand im Wesentlichen im Rahmen meiner Tätigkeit als wissenschaftlicher Mitarbeiter am Institut für Systemdynamik und Flugmechanik (ISF) an der Fakultät für Luft- und Raumfahrttechnik der Universität der Bundeswehr München.

Ich möchte mich an dieser Stelle bei meinem Doktorvater Prof. Dr.-Ing. Ernst-Dieter Dickmanns für die sehr abwechslungsreiche und interessante Arbeit am ISF und für seine Neugier und Ausdauer in den oft sehr engagierten fachlichen Diskussionen bedanken. Er sorgte stets für ein hervorragendes Arbeitsumfeld, in dem wissenschaftliches Arbeiten und Dienstreisen zu internationalen Konferenzen in guter Relation standen.

Ferner gebührt mein Dank allen Kollegen und ehemaligen Kollegen, namentlich Stefan Baten, Simon Fürst, Dr.-Ing. Rudolf Gregor, Dr.-Ing. Michael Lützeler, Dr.-Ing. Markus Maurer, Dr.-Ing. Martin Pellkofer, Dr.-Ing. André Rieder, Dr.-Ing. Karl-Heinz Siedersberger und Dr.-Ing. Volker von Holt für die kollegiale und produktive Zusammenarbeit im Rahmen der verschiedenen Projekte.

Mein besonderer Dank gilt Dr.-Ing. André Rieder, mit dem das Arbeiten gerade in Phasen höchster Anspannung aufgrund zeitlicher Rahmenbedingungen riesig Spaß gemacht hat. Diese extrem intensiven Arbeitsphasen werden mir in bester Erinnerung bleiben.

Ferner möchte ich die sehr angenehme und produktive Zusammenarbeit mit Dr.-Ing. Karl-Heinz Siedersberger hervorheben, der als Antreiber mit zuweilen deutlichen Worten, Organisator in Verwaltungsfragen, Motivator mit seinem niederbayrisch direkten Humor und als Verantwortlicher für die Versuchsfahrzeuge und deren Fahrzeugführung die gute Seele so manches Projekts war.

Der grauen Eminenz Dr.-Ing. Klaus-Dieter Otto danke ich sehr für die abwechslungsreichen, zumeist mathematische Problemstellungen und MATLAB betreffenden Diskussionen, aber auch für die über das Alltagsgeschäft hinausgehenden politisch-gesellschaftlichen Gespräche, welche mir sehr viel Freude bereitet haben.

Außerdem möchte ich mich sehr herzlich bei den Professoren R. Onken, F. Svaricek, A. Schulte und H.-J. Wünsche dafür bedanken, dass sie mir nach der Emeritierung von Prof. Dickmanns für die Fertigstellung meiner Arbeit ein funktionierendes Arbeitsumfeld ermöglicht haben. Hermann Lex gilt mein Dank für die zuverlässige Unterstützung in allen Rechnerfragen. Die Arbeit am Institut und die Zusammenarbeit mit allen Kollegen hat mir sehr viel Spaß gemacht. Es war eine schöne Zeit - vielen Dank dafür.

Für die Übernahme des Koreferats möchte ich mich recht herzlich bei Prof. Dr.-Ing. Georg Färber bedanken und für die Übernahme des Vorsitzes der Prüfungskommission danke ich Prof. Dr. rer. nat. Bernd Häusler.

Mein besonderer Dank gebührt den Verantwortlichen von INI.TUM (Ingolstadt Institute der TU München) und Dr.-Ing. Markus Maurer, welche es mir ermöglichten, meine Erfahrungen und Erkenntnisse in einem von der Themenstellung ähnlichen Teilbereich der Umfeldwahrnehmung im Rahmen von INI.TUM in Zusammenarbeit mit Audi (Entwicklung Fahrerassistenzsysteme) einzubringen und damit auch die zügige Fertigstellung meiner Dissertation förderten.

Zu guter Letzt danke ich meinen Eltern Georg und Klara, meinem Bruder Georg und meinen Erlangerer Studentenwohnheim- und Studienfreunden dafür, dass es für sie stets außer Frage stand, „dass das schon klappt mit der Promotion.“

Mit dieser Zuversicht im Rücken, lässt es sich unbeschwert arbeiten - Danke.

# Kurzfassung

Für die Realisierbarkeit intelligenter autonomer Fahrfunktionen ist die Wahrnehmung des Fahrzeugumfeldes zur Detektion von Objekten und zur Bestimmung der eigenen Position relativ zu anderen Objekten grundlegend.

Hierfür werden in dieser Arbeit im Wesentlichen modellbasierte Bildverarbeitungsansätze verwendet. Ferner werden neue Verfahren zur Extraktion von Bildmerkmalen entwickelt, welche die Zuordnung gemessener Bildmerkmale bezüglich einer aktuellen Modellvorstellung erleichtern sollen. Unter anderem werden zwei neue Verfahren vorgestellt, mit deren Hilfe streifenförmige Bildausschnitte in homogene Segmente mit linearem Helligkeitsverlauf unterteilt werden können, um bei Bedarf mit benachbarten Segmenten zu Kantenverläufen bzw. Bereichen gruppiert zu werden.

Die entwickelten Bildmessverfahren werden zur Detektion und unter Verwendung räumlich zeitlicher Modelle zur Verfolgung von Objekten eingesetzt. In verschiedenen Anwendungen werden die erarbeiteten Ansätze zur Objekterkennung exemplarisch in realen Verkehrsszenen mit Versuchsfahrzeugen für autonome Mobilität und Rechnersehen validiert:

- Es werden die Synergien für ein hybrides ACC System (Hybrid Adaptive Cruise Control (HACC)) aufgezeigt, welche sich aus einer Kombination eines radarbasierten ACC mit visueller Fahrspur- und Objekterkennung ergeben.
- Es werden Strategien und Bildverarbeitungsoperatoren zur rein visuellen Detektion und Verfolgung von Pkws im Nahbereich für ein Stauassistenzsystem entwickelt und auf Videosequenzen verifiziert. Die Detektion von Pkws bei seitlicher Ansicht wird mittels aspektabhängiger generischer Muster zur Detektion von Rädern verwirklicht.
- Ferner wird als ein Teilaspekt einer größeren vollautonomen Mission mit querfeldein GPS-Wegpunktnavigation die Erkennung eines negativen Hindernisses mit anschließendem Ausweichmanöver unter Verwendung der entwickelten Bildverarbeitungsverfahren und aktiver Blickrichtungssteuerung realisiert.

Als Rahmen für die Entwicklung der Fahrerassistanzanwendungen dient das EMS-Vision System (Erwartungsbasiertes Multifokales Sakkadisches - Sehen) des Instituts für Systemdynamik und Flugmechanik der UniBwM.



# Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>1</b>
1.1	Gedanken zur Wahrnehmung . . . . .	1
1.2	Was ist ein Merkmal? . . . . .	3
1.3	Erkennen, Wiedererkennen und Verfolgen . . . . .	4
1.4	Ziele und Beiträge der Arbeit . . . . .	5
1.5	Struktur der Arbeit . . . . .	6
<b>2</b>	<b>Stand der Technik</b>	<b>9</b>
2.1	Systeme mit Sensorkombinationen aus Radar und Vision . . . . .	10
2.1.1	Carnegie Mellon University . . . . .	10
2.1.2	Daimler-Benz Forschung . . . . .	12
2.1.3	Fraunhofer Institut für Informations- und Datenverarbeitung . . . . .	13
2.2	Systeme zur Detektion negativer Hindernisse . . . . .	14
2.2.1	Jet Propulsion Laboratory . . . . .	14
2.2.2	UniBwM und Sarnoff Corporation . . . . .	15
2.2.3	Sarnoff Corporation . . . . .	16
2.2.4	National Institute of Standards and Technology . . . . .	16
<b>3</b>	<b>Sensorsysteme zur maschinellen Umgebungserfassung</b>	<b>17</b>
3.1	Der Szenenbaum . . . . .	18
3.2	Ultraschall, Lidar und Radar . . . . .	19
3.3	Bildgebende Sensoren . . . . .	20
3.3.1	Anforderungen an die Kamerakalibrierung . . . . .	20
3.3.2	Kameramodelle . . . . .	21
3.3.2.1	Lochkameramodell . . . . .	21

3.3.2.2	Kameramodell nach [TSAI 87] . . . . .	23
3.3.3	Vorgehensweise bei der Kalibrierung komplexer Kameraanordnungen im EMS-Vision System . . . . .	25
3.3.4	Kamerakonfigurationen . . . . .	26
3.3.4.1	Monokulares visuelles Sensorsystem . . . . .	26
3.3.4.2	Bifokale Kameraanordnung . . . . .	27
3.3.4.3	Parallele Stereoanordnung . . . . .	27
3.3.4.4	Sensorkonzept MarVEye . . . . .	28
3.4	Kameraplattformen der UniBwM für aktive Blickrichtungssteuerung . . . . .	28
3.4.1	Ein-Achsen-Kameraplattform im Versuchsfahrzeug VaMP . . . . .	28
3.4.2	Zwei-Achsen-Kameraplattform im Versuchsfahrzeug VaMoRs . . . . .	28
<b>4</b>	<b>Extraktion von Bildmerkmalen aus Grauwertbildern</b>	<b>29</b>
4.1	Richtungsselektive Kantensuche mittels ternärer Masken . . . . .	33
4.2	Extraktion von Segmenten homogener Helligkeit durch hierarchische Segmentierung . . . . .	35
4.2.1	Funktionsweise des hierarchischen Segmentierungsverfahrens . . . . .	35
4.2.2	Anwendung des Verfahrens zur Bestimmung von Klothoidensegmenten aus dem Krümmungsverlauf von Feldwegen . . . . .	41
4.2.3	Vor- und Nachteile des hierarchischen Segmentierungsverfahrens . . . . .	42
4.3	Segmentierung homogener Bereiche durch Extraktion gerichteter Kantenpunkte . . . . .	43
4.3.1	Bestimmung der Richtung eines Kantenpunktes bei ungeradzahli- ger nxm-Korrelationsmaske . . . . .	44
4.3.2	Effiziente Summation durch Zerlegung der Maske in Elementarblöcke	44
4.3.3	Bestimmung der Richtung eines Kantenpunktes für eine geradzah- lige nxm-Korrelationsmaske . . . . .	45
4.3.4	Bestimmung der Helligkeitswerte zur Linken und zur Rechten einer Kante in Kantenrichtung . . . . .	47
4.3.5	Suche nach den Segmentgrenzen mit gleichzeitiger Nachbildung des Helligkeitsverlaufs . . . . .	48
4.3.6	Parametrisierung und Ergebnisdatenstruktur . . . . .	49
4.3.7	Effiziente Auswertung eines größeren Bildbereichs . . . . .	51
4.3.8	Beispiel: Auswertung eines Ausschnitts eines Schwarz-Weiß-Bildes . . . . .	52

4.3.9	Rechenzeitbedarf des HOFMANN-Operators . . . . .	58
4.3.10	Statistik der photometrischen Eigenschaften eines Bildbereichs aus der Beschreibung homogener Segmente . . . . .	58
4.3.11	Beispiel: Ermittlung einer näherungsweise Abbildungsvorschrift für Helligkeiten . . . . .	61
4.3.12	Zuordnung korrespondierender Merkmale . . . . .	63
4.3.13	Gruppierung von gerichteten Kantenpunkten zu Kantenverläufen . .	64
4.3.13.1	Lineare Verkettung von gerichteten 2D-Kantenpunkten . .	65
4.3.13.2	Gruppierung von Kantenpunkten zu Punktemengen durch lokale Verkettung . . . . .	67
4.3.13.3	Gruppierung von homogenen Segmenten zu Bereichen . .	67
4.4	Explizite Suche nach Subobjekten mittels generisch erzeugter Muster . . . .	68
4.5	Suche nach Häufungen unspezifischer Helligkeitsübergänge . . . . .	70
<b>5</b>	<b>Umfeldwahrnehmung mittels integraler räumlich-zeitlicher Modelle</b>	<b>73</b>
5.1	4-D Ansatz . . . . .	73
5.1.1	Rekursive Zustandsschätzung . . . . .	74
5.1.2	Messmodell . . . . .	74
5.1.3	Generisches Formmodell und Dynamikmodell . . . . .	75
5.1.4	Numerische Bestimmung der partiellen Ableitung des Messmodells .	76
5.1.5	Festlegung der Filterparameter . . . . .	76
5.2	Repräsentation von a priori Expertenwissen zur Wahrnehmung . . . . .	77
5.2.1	Generisches Formmodell eines Fremdfahrzeugs . . . . .	77
5.2.2	Klassifizierung der Aspektbedingung . . . . .	79
5.2.3	Sichtbarkeit von 3-D Messpunkten im Bild . . . . .	80
5.2.4	Verdeckungen . . . . .	81
5.2.5	Aspektabhängige Beobachtbarkeit von Zustandsgrößen . . . . .	82
5.3	Berücksichtigung unspezifischer Merkmale bei der Zustandsschätzung . . .	82
5.3.1	Erwartungsbasierte Bestimmung des Merkmalflusses . . . . .	83
5.3.2	Erwartungsbasierte Stereoauswertung bei divergenten optischen Achsen . . . . .	86
5.4	Fahrspurmodellierung und Modellierung der relativen Eigenposition . . . .	89
5.5	Koordinatensysteme zur visuellen Fahrspur- und Objekterkennung . . . . .	90

<b>6</b>	<b>Hybrider adaptiver Tempomat</b>	<b>93</b>
6.1	Funktionsbeschreibung . . . . .	93
6.2	Systemdesign mit skalierbarer Systemleistung . . . . .	94
6.3	Gegenüberstellung der Eigenschaften eines Radars und eines visuellen Sensors	96
6.4	Verwendete Hardware . . . . .	98
6.5	Datenbasis des HACC Systems . . . . .	99
6.6	Verfahren zur Erkennung und Verfolgung der Fahrspur . . . . .	100
6.7	Verfahren zur Fahrzeugdetektion, Hypothesengenerierung und deren Ver- folgung . . . . .	102
6.8	Auswahl des relevanten Fahrzeugs für die Fahrzeuglängsregelung . . . . .	103
6.9	Experimentelle Ergebnisse . . . . .	107
6.10	Zusammenfassung und Ausblick . . . . .	108
<b>7</b>	<b>Erkennung von Fremdfahrzeugen im Nahbereich</b>	<b>111</b>
7.1	Anforderungen an eine Umfeldwahrnehmung im Nahbereich . . . . .	112
7.2	Verfahren zur monokularen Detektion von Längsverkehr im Nahbereich . .	114
7.2.1	Detektion von Fremdfahrzeugen durch explizite Suche nach dem Schatten unter dem Fahrzeugaufbau und nach Rädern . . . . .	115
7.2.2	Generierung einer Objekthypothese aufgrund des optischen Merk- malflusses zur Überholerdetektion . . . . .	117
7.3	Verfolgung einer Objekthypothese . . . . .	118
7.4	Zusammenfassung . . . . .	120
<b>8</b>	<b>Erkennung negativer Hindernisse</b>	<b>121</b>
8.1	Hardware . . . . .	123
8.1.1	Kamerakonfiguration . . . . .	123
8.1.2	Aktive Zwei-Achsen-Kameraplattform . . . . .	124
8.1.3	Acadia-Stereobildverarbeitungshardware . . . . .	125
8.2	Prinzipielle Vorgehensweise zur Erkennung negativer Hindernisse . . . . .	126
8.2.1	Festlegung eines interessanten Bereichs für die Bildauswertung . . .	126
8.2.2	Adaptive Festlegung der Maskengröße für Bildoperatoren . . . . .	128
8.2.3	Kantensuche im Grauwertbild . . . . .	128
8.2.4	Validierung der Kantenpunkte im Disparitätenbild . . . . .	129
8.2.5	Verkettung der Kantenpunkte zu Kanten . . . . .	131

8.2.6	Verknüpfung der Kanten zu Linien . . . . .	131
8.2.7	Erzeugung einer Grabenhypothese . . . . .	131
8.3	Hypothesenverwaltung . . . . .	134
8.3.1	Sichtbarkeitsanalyse . . . . .	135
8.3.2	Verkürzen oder Löschen einer bestehenden Hypothese . . . . .	136
8.3.3	Steuerung der Innovation . . . . .	137
8.4	Zusammenspiel von Wahrnehmung, Blickrichtungssteuerung und Fahrzeugführung . . . . .	141
8.4.1	Wahrnehmungsmodi zur Steuerung der Aufmerksamkeitsbereiche . . . . .	142
8.4.2	Parametrisierung der Wahrnehmung . . . . .	147
8.4.3	Anknüpfung an das Fähigkeitsnetz . . . . .	147
8.5	Ergebnisse einer realen Versuchsfahrt . . . . .	147
8.5.1	Erläuterungen zum Ausweichen an einem Graben anhand einer Bildfolge . . . . .	149
8.5.2	Verlauf der Parametrisierung der Wahrnehmung . . . . .	153
8.5.3	Vogelperspektive des Ausweichmanövers an einem Graben . . . . .	154
8.5.4	Diskussion des Verlaufs der Modellparameter bei der Grabenerkennung . . . . .	154
8.5.5	Anmerkungen zur Stereohardware . . . . .	156
8.6	Zusammenfassung der Ergebnisse . . . . .	157
<b>9</b>	<b>Zusammenfassung</b>	<b>159</b>
<b>A</b>	<b>Mathematische Grundlagen</b>	<b>161</b>
A.1	Ausgleichsrechnung . . . . .	161
A.2	Bestimmung einer Ausgleichsgeraden durch eine Punktemenge . . . . .	162
A.3	Varianz der Abweichung der Werte einer Zahlenfolge zur Ausgleichsgeraden . . . . .	165
A.4	Verknüpfung zweier Ausgleichsgeraden zu einer . . . . .	165
A.4.1	Verknüpfung zweier Zahlenfolgen beliebiger Länge . . . . .	166
A.4.2	Verknüpfen von Zahlenfolgen gleicher Länge . . . . .	168
A.4.3	Rekursive Bestimmung einer Ausgleichsgeraden . . . . .	169
A.5	Einbringung einer Zwangsbedingung bei der Regression . . . . .	171
A.6	Subpixelgenauigkeit . . . . .	173
A.7	Bestimmung einer Ausgleichsebene . . . . .	175
<b>B</b>	<b>Zustandsschätzung</b>	<b>179</b>
B.1	Das Kalmanfilter . . . . .	179
B.2	Anmerkungen zur verwendeten Implementierung des EKF . . . . .	182
<b>Literatur</b>		<b>183</b>

# Kapitel 1

## Einleitung

### 1.1 Gedanken zur Wahrnehmung

Was ist eigentlich „Wahrnehmung“?

Einen ersten Eindruck dessen, was alles mit diesem Begriff im Zusammenhang steht, ermöglicht die Betrachtung gebräuchlicher Redewendungen. Man spricht z.B. von *etwas für wahr nehmen (etwas als Tatsache auffassen)* oder *sich etwas gewahr werden (etwas ins Bewusstsein bringen)* oder *jemanden in Gewahrsam nehmen (jemanden unter Beobachtung stellen)*.

Bereits Aristoteles unterschied zwischen sinnlicher Wahrnehmung und geistiger Wahrnehmung. Die sinnliche Wahrnehmung stellt dabei die Grundlage der geistigen Wahrnehmung dar. In [MIETZEL 02] wird die sinnliche Wahrnehmung des Menschen und der Begriff der Aufmerksamkeit folgendermaßen umschrieben:

Damit eine Wahrnehmung entstehen kann, muss ein Reiz auf ein Sinnesorgan treffen. Der Reiz muss eine bestimmte Stärke aufweisen, um vom Menschen wahrgenommen werden zu können. Würden alle wahrnehmbaren Reize wahrgenommen werden, wäre es dem Menschen nicht mehr möglich, wichtige Informationen von unwichtigen zu trennen.

Der Prozess der Aufmerksamkeit hat deshalb die Aufgabe, eine solche Unterscheidung zu treffen. Bei der „Aufmerksamkeit“ handelt es sich um einen Prozess, der mitbestimmt, welche Informationen zur Verarbeitung ausgewählt werden. Wenn sich Ereignisse vom Umfeld abheben (z.B. durch Bewegung, Größe, Wiederholung etc.), geraten diese in den Brennpunkt der Aufmerksamkeit, werden beachtet, während das Umfeld in den Hintergrund gerückt wird.

Unter Wahrnehmung verstehen wir jene Prozesse, die eintreffende Informationen von den Sinnesorganen auswählen, ordnen und interpretieren. Wenn wir das Wort „Wahrnehmen“ genau nehmen, bedeutet es, dass das, was wir wahrnehmen, tatsächlich **wahr** ist.

Um etwas auswählen und interpretieren zu können, muss sich ein Individuum oder eine Maschine über den Gehalt einer Information bewusst werden. Dieses „bewusst werden“ stellt den Übergang von der sinnlichen zur geistigen Wahrnehmung dar.

Nach [METZINGER 01]

benötigt jede Maschine, die Bewusstsein haben soll, ein integriertes dynamisches Weltmodell (Anm.: erstes Kriterium). Sie muss eine *einheitliche* innere Darstellung der Welt als Ganzes besitzen und die in dieser Darstellung integrierte Information muss global verfügbar sein. Bewusste Information ist nämlich genau die Information im System, die gerade global - also für eine Vielzahl von Verarbeitungsmechanismen *gleichzeitig* - verfügbar ist...

Je mehr Information das System als *bewusste* Information verarbeitet, desto höher ist der Grad an Flexibilität und Kontextsensitivität, mit dem es auf Herausforderungen aus der Umwelt reagieren kann...

*Bewusste Systeme* sind alle Systeme, die mit global verfügbarer Information operieren und die sich selbst deshalb als *in einer einzigen Welt lebend* erfahren...

Die Globalität des bewussten Erlebens besteht also darin, dass alle Bewusstseinsinhalte immer in ein einheitliches *Realitätsmodell* (Anm.: entspricht dem situationalen Kontext) integriert sind. Es gibt ein einziges, ganzheitliches Bild der Wirklichkeit. Aus der Introspektive ist diese höchststufige phänomenale Ganzheit ganz einfach die Welt, in der ich mein Leben lebe - und die Grenzen dieser Welt sind die Grenzen meiner Wirklichkeit...

Ausnahmslos alle unsere Bewusstseinszustände sind dadurch gekennzeichnet, dass alles, was wir erleben - unabhängig von dem konkreten Inhalt, *den* wir erleben - immer als *jetzt* erlebt wird. Dass eine Maschine oder ein Mensch Bewusstsein hat, wird immer bedeuten, dass es für sie eine Gegenwart gibt: Gegenwärtigkeit bedeutet, dass einem System ein bestimmter geistiger Inhalt als aktuell gegeben erscheint...

Diese spezielle Form eines globalisierten Kurzzeitgedächtnisses - eine „Jetzt-Erinnerung“ - ist das, was jede auch bewusste Maschine bräuchte: Sie bräuchte eine repräsentationale Ressource, in der verschiedene repräsentationale Inhalte zusammengeführt und in ihrer scheinbar direkten Gegebenheit als *gleichzeitig* gegeben dargestellt werden. Ich nenne dieses zweite Kriterium den Besitz eines „virtuellen Gegenwartsfensters“...

Wenn das globale Weltmodell - oder ein Teil davon - in das virtuelle Gegenwartsfenster des Systems eingebettet wird, dann ist der so erzeugte repräsentationale Inhalt *die Gegenwart einer Welt ...*

Bewusstes Erleben ist die Gegenwart einer Wirklichkeit. Jetzt kann man sich auch gut vorstellen, wie ein System zusätzlich ein umfassendes *unbewusstes* Modell der Realität haben könnte, nämlich den Teil, der gerade nicht global verfügbar und in sein bewusstes Gegenwartsfenster eingebettet ist ...

Dies bedeutet nicht, dass ein bewusstes System bereits ein Bewusstsein besitzt. Hierfür sind nach Metzinger weitere Kriterien zu erfüllen, wie z.B. die Existenz eines Selbstmodells oder die Transparenz der Modelle. Transparenz bedeutet, dass die internen Modelle nicht mehr als solche erkannt werden, sondern selbst zur erlebten Wirklichkeit werden. Für den interessierten Leser sei hier die Lektüre des vollständigen Artikels [METZINGER 01] empfohlen. Die Auszüge aus dem Artikel geben aber einen guten Eindruck über die Komplexität dessen, was Wahrnehmung und eine bewusste Interpretation von Information alles beinhaltet.

Die Interpretation von Information kann stets nur so gut sein, wie der Kontext bekannt ist, in dem die Information steht, und es das Wissen und die Erfahrung des Individuums bzw. der Maschine zulässt. Aus der Sicht des einzelnen Individuums, das aufgrund seiner persönlichen Erfahrungen und Kenntnisse die eintreffenden Informationen interpretiert, stellt das Wahrgenommene die momentane „Wahrheit“ oder Wirklichkeit dar. Die Grenzen dieser Welt bzw. Wahrnehmung sind die Grenzen der individuellen Wirklichkeit.

Damit ein maschinelles Wahrnehmungssystem zumindest in den derzeit eingeschränkten Szenarien aus der Sicht des Menschen akzeptable Leistungen erbringen kann, kommt der Informationsgewinnung als Grundlage für eine spätere Interpretation auf Objektebene große Bedeutung zu.

Maschinelle Wahrnehmung beinhaltet neben der reinen Extraktion von Merkmalen aus einem Sensorsignal auch eine Vorstellung über die Gestalt und das aspektabhängige Erscheinungsbild im Situationskontext und Vorstellungen über das zeitliche Verhalten von Objekten und Subjekten. Der Mensch verfügt je nach Erfahrung über mehr oder weniger detaillierte Dynamikmodelle über das Verhalten von Subjekten und Objekten, die je nach Situationskontext parametrisiert und aktiviert werden.

## 1.2 Was ist ein Merkmal?

Nach Brockhaus [BROCKHAUS 98] ist ein Merkmal

allgemein ein charakteristisches Zeichen - ein Kennzeichen. In der Logik bezeichnet ein Merkmal den Inhalt eines Begriffs, der eine bestimmte Eigenschaft jeden Gegenstands, der unter den Begriff fällt, ausdrückt. Es werden dabei wesentliche oder notwendige von unwesentlichen oder zufälligen Merkmalen unterschieden. Der Begriff Merkmal bezeichnet auch eine Bestimmung, durch die sich Dinge der gleichen Art oder Arten der gleichen Gattung unterscheiden. In der Statistik wird zwischen qualitativen Merkmalen, wie z.B. das Geschlecht der Einheit Person, und quantitative Merkmale, wie z.B. der Durchmesser einer Schraube, unterschieden. Die möglichen Werte, die Merkmale annehmen können, heißen Merkmal-Ausprägungen (-Werte oder -Modalitäten).

Ein Merkmal ist eine eindeutig vermessbare Eigenschaft eines Objekts oder Subjekts. Ein Merkmal wird demnach erst dadurch zu einem Merkmal, dass es einen potentiellen Sensor für es gibt. Dies gilt auch, wenn die Bedingungen für eine Messung nicht gegeben sind, z.B. wenn im Wahrnehmungssystem ein geeigneter Sensor zu einem bestimmten Zeitpunkt nicht verfügbar oder gar nicht vorhanden ist.

Der Begriff der Wahrnehmung auf der Ebene der Merkmale bedeutet die Extraktion der Parameter einer Merkmalsbeschreibung mittels der dafür vorgesehenen Verfahren aus dem Ausgangssignalverlauf eines Sensors. Die notwendige Sensorik und Algorithmik zur Vermessung hängt dabei von der physikalischen Beschaffenheit einer Eigenschaft ab. Die Parameter einer Merkmalsbeschreibung könnten z.B. die Helligkeit der Projektion eines Teilobjekts in den Bildbereich eines visuellen Sensors oder seine Farbe, aber auch die Art eines akustischen Signals oder seine Temperatur uvm. sein.

In der Regel ist nicht das Vorhandensein *aller* erwarteten Merkmale das Kriterium dafür, ob ein Objekt zu einer bestimmten Klasse an Objekten gehört oder nicht, sondern der Anteil und die Signifikanz der gefundenen Merkmale ist wichtig. Signifikanz soll hierbei ausdrücken, dass nicht alle Merkmale die gleiche Wertigkeit für eine Klassifizierung haben. Für einen Pkw ist z.B. das Vorhandensein von Rädern eines der wichtigsten Merkmale.

### 1.3 Erkennen, Wiedererkennen und Verfolgen

Das Erkennen eines Merkmals setzt eine Vorstellung voraus und die Möglichkeit, eine Ähnlichkeit mit einer Vorstellung zu quantifizieren, um ein Objekt oder Merkmal zu klassifizieren. Das Wiedererkennen erfordert demgegenüber eine noch detailliertere Vorstellung und die Möglichkeit, bezüglich einer Menge gleichartiger Objekte eine eindeutige Abgrenzung und Unterscheidung zu treffen. Für das Verfolgen eines Merkmals wird dabei die Kontinuität der Bewegung in Raum und Zeit angenommen.

Das Verfolgen eines ganzen Objektes erfordert eine gewisse Anzahl verfolgter Merkmale, welche über der Zeit variieren können, allerdings stets eine eindeutige Abgrenzung gegenüber der Umgebung zulassen. Zu dieser variablen Menge kommen immer wieder neue Merkmale hinzu, andere fallen nach einer gewissen Zeit weg und tauchen eventuell später wieder auf. Diese Merkmale zeichnen sich durch eine ähnliche Dynamik und Konstellation aus. Die Beschreibung der idealerweise starren räumlichen Konstellation erfolgt durch ein generisches Formmodell in 3-D und durch die Anpassung seiner Formparameter anhand der Merkmalzuordnung. Die gleichförmige Bewegung wird durch das Dynamikmodell jedes Objekts beschrieben.

Sobald die Abgrenzung zur Umgebung nicht mehr möglich ist, kann zwar die Erwartung bezüglich des Vorhandenseins eines Objekts aufrecht erhalten werden und, sofern ein Dynamikmodell vorliegt, über die Zeit dynamisch fortgeschrieben werden, das Verfolgen eines Objektes ist für diesen Zeitraum aber unterbrochen, da keine Aktualisierung der beschreibenden Zustandsgrößen möglich ist, außer man fasst den Begriff des Verfolgens so

auf, dass bereits die Absicht zur Verfolgung und das Festhalten an einer Objekthypothese für das Aufrechterhalten einer Verfolgung ausreichend ist.

Bei der Erkennung (Detektion) und beim Verfolgen (Tracking) werden z.T. unterschiedliche Anforderungen an die Wahrnehmung gestellt. Bei der Detektion liegt in der Regel kein a priori Wissen über die Position eines Objekts oder Subjekts vor. Der Suchraum kann deshalb lediglich durch Anfangsannahmen eingeschränkt werden. Beispielsweise wird meist angenommen, dass sich das Eigenfahrzeug zum Zeitpunkt der Aktivierung einer Fahrspurerkennung innerhalb der aktuellen Fahrspur befindet und jede Fahrspur z.B. auf Autobahnen durch Spurmarkierungen zur Linken und zur Rechten der Fahrspur gekennzeichnet ist. Innerhalb eines gewissen Toleranzbereiches kann damit eine zuverlässige Erkennung erreicht werden.

Die Vorgehensweisen zur Detektion und zur Verfolgung legen es nahe, unterschiedliche Ausprägungen generischer Modelle bzw. Modelle, die fließend ineinander übergehen, zu verwenden. Ein Grundgedanke bei der Verfolgung besteht darin, dass meist nur eine Untermenge aller erwarteten Merkmale gefunden werden kann und in Abhängigkeit von den Ergebnissen der Messung festgelegt wird, welche Zustandsgrößen oder Formparameter daraus beobachtbar sind und deshalb eine Innovation erfahren. In einem Detektionsmodell müsste eine hierarchisch aufgebaute Logik repräsentiert sein, wie durch eine Kombination oder einen Wechsel von Verifizieren und Falsifizieren von Merkmalen eine zuverlässige erste Hypothese generiert werden kann, die dann an das Trackingmodell weitergereicht wird bzw. in diesem mündet.

Beispiel: Als ersten Hinweis für ein Fremdfahrzeug auf der Straße kann bei Tag ein dunkler Bereich dienen, welcher der Schatten eines Fahrzeugs sein könnte. Wenn dem so ist und sofern keine Verdeckungen vorliegen, dann müssten an den Rändern des Schattens irgendwo Räder zu finden sein. Wenn es sich bei einem gefundenen Rad um das Hinterrad handelt, dann müsste bei einer angenommenen Ausrichtung entlang der Fahrbahn innerhalb eines begrenzten Bereichs im Bild das Vorderrad zu finden sein. Ist dies nicht der Fall, so war das gefundene Rad vielleicht nicht das Hinterrad, sondern das Vorderrad, dann müsste in entgegengesetzter Richtung im Raum in einem begrenzten Bereich im Bild das Hinterrad zu finden sein. Ist dies nicht der Fall, sollte diese Hypothese verworfen werden.

Beim Verfolgen steht nicht mehr das Erkennen, sondern nur noch das Wiedererkennen im Vordergrund, wobei der Suchraum in Erwartung einer räumlich-zeitlichen Kontinuität der Bewegung eingeschränkt wird.

## 1.4 Ziele und Beiträge der Arbeit

Eines der Ziele dieser Arbeit war die Entwicklung von Bildverarbeitungswerkzeugen zur Steigerung des Informationsgewinns aus der Bildfolgenanalyse. Diese Bildverarbeitungswerkzeuge (Operatoren) sollen sowohl für die Detektion von Objekten als auch fürs Verfolgen bestehender Objekthypothesen konzipiert sein. Häufig wird hierbei entlang horizontaler bzw. vertikaler Bildstreifen nach Helligkeitsübergängen gesucht, wobei diese in

der Regel die Grenzen homogener Helligkeitsbereiche bilden. Speziell hierfür wurden zur Segmentierung von Helligkeitsverläufen zwei neue Verfahren entwickelt. Die Beschreibung eines Bildstreifens soll dabei gegenüber den bisherigen Verfahren durch die Fülle der extrahierten Information deutlich an Aussagekraft gewinnen, wodurch auch verbesserte Möglichkeiten zur Selektion und Gruppierung von Merkmalen zur Verfügung stehen. Weiterhin sollen Konzepte und Verfahren zur effizienten Positionsbestimmung typischer, durchaus komplexer, objektspezifischer Merkmale entwickelt werden.

Ein weiteres Ziel bestand darin, Konzepte zur Ausnutzung vorhandener Information für eine effiziente Steuerung der Wahrnehmung zu entwickeln. Hierunter fallen statistische Auswertungen der Helligkeit interessanter Bildbereiche zur Festlegung von Helligkeitsschwellen für die Bildverarbeitung bzw. für die Beurteilung von Korrespondenzen. Es wurde in verschiedenen Anwendungen eine Sichtbarkeits- und Verdeckungsanalyse in Abhängigkeit von den Aspektbedingungen und der Positionen der Messpunkte im Raum umgesetzt. In Abhängigkeit der Sichtbarkeit wird wiederum entschieden, für welche Modellparameter eine Aktualisierung aufgrund der Messung sinnvoll und erlaubt ist. Es sind die Parameter der Bildverarbeitungsoperatoren in Abhängigkeit von der erwarteten Größe der Abbildung eines gesuchten Merkmals im Bild festzulegen. Bei Vorhandensein einer aktiven Blickrichtungssteuerung sind relevante Bereiche zu definieren, welche für die Erfüllung der Wahrnehmungsaufgabe und die Interaktion mit der Blickrichtungssteuerung notwendig sind. Das Zusammenspiel zwischen Wahrnehmung, Blickrichtungssteuerung und Fahrzeugführung wird dabei speziell am Beispiel der Erkennung eines negativen Hindernisses in Kapitel 8 erläutert.

Ferner sollten Wahrnehmungsmodule zur Durchführung autonomer und semiautonomer Missionen mittels des EMS-VISION Systems (Erwartungsbasiertes Multifokales Sakkadisches Sehen) der UniBwM entwickelt werden. In dieser Arbeit werden deshalb verschiedene Anwendungen zur Lösung ausgewählter Teilaspekte zur Umfeldwahrnehmung beschrieben.

## 1.5 Struktur der Arbeit

Nach den einleitenden Gedanken zur maschinellen Wahrnehmung folgt in Kapitel 2 eine Übersicht über den derzeitigen Stand der Technik. In Kapitel 3 schließt daran eine kurze Beschreibung der Wissensrepräsentation im EMS-VISION System an. Über diese wird der Bezug zu den in Fahrerassistenzsystemen häufig verwendeten Sensoren zur Umfelderkennung hergestellt, wobei hier speziell der Fokus auf bildgebende Sensoren (d.h. Kameras) und die Vorgehensweise zu deren Kalibrierung gelegt wird. In den Abschnitten 3.3.4 und 3.4 wird der Zusammenhang zwischen der Auflösbarkeit eines Merkmals einer bestimmten Größe, der Pixelanzahl des Sensorchips und dem horizontalen bzw. vertikalen Öffnungswinkel einer Sensorchip-Objektiv Kombination erläutert, der dann zur Gewährleistung einer gewünschten Wahrnehmungsleistung entsprechende Kamerakonfigurationen mit gegebenenfalls aktiver Blickrichtungssteuerung nach sich zieht.

In Kapitel 4 wird in Abschnitt 4.1 die erwartungsbasierte Bildauswertung mittels richtungsselektiver ternärer Masken (dem bisherigen Standard am Institut für Systemdynamik und Flugmechanik (ISF)) erläutert. Daran schließen sich in den Unterkapiteln 4.2 und 4.3 zwei neu entwickelte Verfahren zur Segmentierung von Helligkeitsverläufen an. Die dafür notwendigen, vergleichsweise einfachen mathematischen Grundlagen werden im Anhang A beschrieben. In Abschnitt 4.4 wird hiernach die explizite Suche nach Subobjekten entsprechend den aktuellen Aspektbedingungen mittels generisch erzeugter Muster am Beispiel der Raderkennung erläutert. Das Kapitel 4 endet mit Abschnitt 4.5, in dem der von [THOMANEK 96] entwickelte Betragsoperator kurz beschrieben wird, welcher für die Vermessung der Außenkanten von Fremdfahrzeugen zum Einsatz kommt.

Kapitel 5 behandelt die Umfeldwahrnehmung mittels integraler räumlich-zeitlicher Modelle. Hierfür wird in Abschnitt 5.1 zunächst das Prinzip des 4-D Ansatzes beschrieben, gefolgt von Abschnitten über die rekursive Zustandsschätzung, das Messmodell und die Modellierung der Form und Dynamik von Objekten. Im Anhang B werden die Grundzüge des Erweiterten Kalmanfilters kurz beschrieben. Die Abschnitte 5.1.4 und 5.1.5 befassen sich mit der numerischen Bestimmung der für eine Linearisierung bezüglich eines Arbeitspunktes notwendigen Jacobimatrizen des Messmodells und der Parametrisierung des Erweiterten Kalmanfilters.

Am Beispiel eines Pkw-Fremdfahrzeugs werden dann in Abschnitt 5.2.1 und den daran anschließenden Abschnitten die Formmodellierung und Parametrisierung der Bildverarbeitungsoperatoren zur Extraktion der erwarteten Merkmale erläutert. Es wird die Klassifizierung der Aspektbedingungen für eine Sichtbarkeits- und Verdeckungsanalyse und eine aspektabhängige Steuerung der Zustandsinnovation beschrieben. Daran anschließend werden in den Abschnitten 5.3.1 und 5.3.2 zwei Vorgehensweisen zur Berücksichtigung unspezifischer Merkmale bei der Zustandsschätzung vorgestellt. Dabei handelt es sich zum einen um eine Methode zur erwartungsbasierten Berücksichtigung des Merkmalflusses bei der Zustandsschätzung und zum anderen um die erwartungsbasierte Stereoauswertung bei divergenten optischen Achsen. Den Abschluß des Kapitels 5 bilden eine Kurzübersicht bezüglich der Fahrspurmodellierung und der Modellierung der relativen Eigenposition sowie ein Abschnitt bezüglich der Koordinatensysteme zur visuellen Fahrspur- und Objekterkennung.

Die Kapitel 6, 7 und 8 beschreiben die Konzeption, die Umsetzung und die Ergebnisse dreier Anwendungen zur maschinellen Umfelderkennung. Kapitel 6 behandelt dabei ein System, das zum Ziel hat, die Wahrnehmungsleistung eines adaptiven Tempomaten durch eine Kombination eines Radar-basierten ACC Systems (Adaptive Cruise Control) mit visueller Wahrnehmung zu verbessern. Das Kapitel 7 befasst sich zur Verwirklichung einer Stauassistentenfunktion speziell mit der Verbesserung der Erkennung von Pkw-Fremdfahrzeugen im Nahbereich des Eigenfahrzeugs. Die Detektion und das Verfolgen mehrerer Pkws sollte dabei rein visuell erfolgen, was speziell bei Seitenansicht eine besondere Herausforderung darstellt. Das Kapitel 8 verlässt thematisch die Domäne befestigter Fahrbahnen und wendet sich einem Teilaspekt zu, der insbesondere beim autonomen Querfeldeinfahren mittels GPS-Wegpunktnavigation von Bedeutung ist. Es wird ein Wahrnehmungsmodul

vorge stellt, das der Erkennung und Vermessung negativer Hindernisse unter Verwendung aktiver Blickrichtungssteuerung [PELLKOFER 03] dient, wodurch der Fahrzeugexperte in die Lage versetzt wird, eine Ausweichtrajektorie zu berechnen und ein Ausweichmanöver durchzuführen [SIEDERSBERGER 03].

Als Abschluß werden in Kapitel 9 zusammenfassend nochmals die wesentlichen Beiträge der Arbeit aufgeführt.

# Kapitel 2

## Stand der Technik

Die Anzahl der Forschergruppen, die auf dem Gebiet der Fahrerassistenzsysteme und autonomen Fahrzeugführung arbeiten, nimmt trotz der über zwanzig jährigen Entwicklungsgeschichte der Automatisierung von Fahrfunktionen durch maschinelles Sehen noch immer stetig zu. Neben der reizvollen Thematik liegt dies mit Sicherheit auch daran, dass durch den technologischen Fortschritt der letzten Jahre der Einstieg in die Thematik der Echtzeit-Bildverarbeitung auch ohne Spezialhardware deutlich einfacher und kostengünstiger geworden ist. Hier sind insbesondere die Entwicklungen auf dem Gebiet der Kamerahardware mit standardisierten Kameraschnittstellen, z.T. getrieben durch den Boom der digitalen Fotokameras, und die bereits in handelsüblichen PC-Systemen mit CPU-Taktfrequenzen über 3 GHz verfügbare Rechenleistung zu nennen. Deshalb ist es kein leichtes Unterfangen, einen Stand der Technik für dieses Gebiet, das interdisziplinär die Bereiche Fahrzeugtechnik, Regelungstechnik, Elektrotechnik und Informatik verbindet, zu verfassen.

Einen sehr guten Überblick über die Aktivitäten der verschiedenen Forschergruppen bieten die Konferenzbände der IEEE Intelligent Vehicles Symposien, die seit 1992 jährlich stattfinden und als Forum für den Austausch neuester Erkenntnisse auf dem Gebiet der Intelligenten Fahrzeuge und Fahrfunktionen dienen. Im Bereich der militärischen Forschung sind die Konferenzbände der „SPIE-AeroSense: Unmanned Ground Vehicle Technology“ Konferenzen in Orlando (Florida) zu empfehlen.

In [BISHOP 00] wird eine Übersicht über die weltweiten Aktivitäten auf dem Gebiet der intelligenten Fahrfunktionen gegeben. Der Konferenzbeitrag befasst sich mit Anwendungen im Personenkraftwagen, Anwendungen für den Schwerlastverkehr, Anwendungen im öffentlichen Personennahverkehr und besonderen Fahrzeuganwendungen.

In [KASTRINAKI 03] wird versucht, einen aktuellen Überblick über die etablierten Systeme und Bildverarbeitungsverfahren in Verkehrsanwendungen zu geben. Die Autoren unterteilen ihren Artikel nach der Einführung in drei größere Kapitel. Im ersten Teil beschreiben sie Systeme zur automatischen Fahrspurerkennung. Im zweiten Teil befassen sie sich mit Systemen zur Objekterkennung und im dritten Teil bringen sie eine tabellarische

Übersicht der nach ihrer Ansicht repräsentativen Systeme und zukünftiger Trends, wobei sie die Hauptmerkmale der Systeme stichpunktartig auflisten. Die ersten beiden Kapitel werden ferner jeweils in Systeme mit statischer Kamera, Systeme mit bewegter Kamera und Ansätze zur Detektion der Fahrspur bzw. von Objekten unterteilt. Die Autoren listen dabei elf Systeme mit statischer Kamera zur Verkehrsüberwachung und 25 Systeme für autonome Fahrfunktionen mit bewegter Kamera auf.

Der Fortschritt auf dem Gebiet der visuellen Objekterkennung vollzieht sich sehr schnell, darum wird und kann dieser Stand der Technik nur einen kleinen Teilbereich des gesamten Spektrums der Aktivitäten im zivilen und militärischen Bereich genauer beleuchten.

## 2.1 Systeme mit Sensorkombinationen aus Radar und Vision

### 2.1.1 Carnegie Mellon University

Am Robotik Institut der Carnegie Mellon University (CMU) in Pittsburgh wird seit vielen Jahren auf dem Gebiet der autonomen Fahrzeugführung geforscht. In [LANGER 96] wird beispielsweise ein System vorgestellt, das die Daten eines Radar-basierten Objekterkennungsmoduls mit der Spurinformaton eines visuellen Fahrspurerkennungssystems kombiniert. Das Radarsystem hat eine Sichtweite von bis zu 200m bei einem horizontalen Öffnungswinkel von 12°. Für die visuelle Spurerkennung wird das Modul RALPH (Rapidly Adapting Lateral Position Handler) [POMERLEAU 95], das ebenfalls vom Robotik Institut der CMU entwickelt wurde, eingesetzt. Die Fusion der beiden Sensorsysteme erfolgt dabei auf Objektebene, d.h. die Objektdaten des Radarsystems werden unter Verwendung der Fahrspurdaten des visuellen Systems interpretiert, die beiden Systeme arbeiten aber eigentlich unabhängig voneinander.

Die beiden Sensorsysteme sind über eine Modul „Map and Navigator (MaN)“ und ein Modul „Control Flow Coordinator (CFC)“ miteinander verbunden. Das CFC ist in Abwesenheit anderer Objekte dafür da, die von RALPH bestimmte Ablage des Eigenfahrzeugs gegenüber der Fahrspurmitte dem Fahrzeugregler zu übermitteln. Im Fall von detektierten Objekten wird im Modul MaN die Zuordnung der Objekte zu den Fahrspuren durchgeführt und entschieden, ob ein Spurwechsel notwendig wird. Ist dies der Fall, wird es dem Modul CFC mitgeteilt und dieses entscheidet je nach aktuell befahrener Fahrspur, ob ein Spurwechsel nach links oder rechts ausgeführt wird.

Mit diesem System wurde 1996 mediengerecht eine Langstreckenfahrt „No Hands Across America“ [POMERLEAU 96] durchgeführt, in der der Versuchsträger NAVLAB 5 der CMU unter Verwendung von RALPH 98, 1% der 2850 Meilen langen Strecke von Washington, D.C. nach San Diego automatisch quergeregelt (die Längsregelung wurde weiterhin vom Fahrer vorgenommen) zurückgelegt hat.

Die Erkennung der Fahrspur durch RALPH erfolgt in vier Schritten. Im ersten Schritt wird ein in 3-D ca.  $7m$  breiter und in der Vorausschau zwischen  $20m$  und  $70m$  gelegener Bereich, der sich ins Bild einer Videokamera als Trapezoid abbildet, durch Unterabtastung in ein niedrig aufgelöstes Bild mit  $30 \times 32$  Pixeln derart transformiert, dass die wesentlichen Merkmale, welche im Originaltrapezoid in Richtung des Horizonts zusammenlaufen, nun als parallele Strukturen erscheinen. Daraus wird in einem zweiten Schritt die Krümmung der Fahrspur dadurch bestimmt, dass ihr Einfluß durch inkrementelle Verschiebung der Zeilen des niedrig aufgelösten Bildes von oben nach unten ausgeglichen wird, bis sich aus den gekrümmten parallelen Strukturen nun gerade parallele Strukturen (kolinear mit den Spalten) ergeben haben. Hier ist anzumerken, dass es für diese Technik zur Bestimmung der Fahrspurkrümmung vollkommen unerheblich ist, welcher Art die im Bild existierenden Merkmale sind, solange die sichtbaren Merkmale parallel zur Fahrspur in 3-D verlaufen. In einem dritten Schritt wird daraus durch Summation der Spalten das aktuelle Intensitätsprofil der Fahrbahnoberfläche bestimmt. Dieses wird mit einem Referenzintensitätsverlauf korreliert, wodurch die Ablage des Eigenfahrzeugs relativ zu der Lage, in der der Referenzintensitätsverlauf aufgenommen wurde, bestimmt werden kann. In einem vierten Schritt wird der Referenzintensitätsverlauf den aktuellen Veränderungen der Fahrspuroberfläche nachgeführt. Hierfür stehen vier Mechanismen zur Verfügung:

- Bei der ersten Methode begibt sich der Fahrer mit dem Eigenfahrzeug in die Mitte der Fahrspur und drückt einen Knopf, worauf RALPH ein Referenzmuster erzeugt.
- Bei der zweiten Methode wählt RALPH ein Referenzmuster aus einer Bibliothek gespeicherter Referenzintensitätsverläufe aus, das am besten mit dem aktuellen Erscheinungsbild der Fahrspur korreliert.
- Durch Methode drei wird das aktuelle Referenzmuster durch Addition eines kleinen prozentualen Anteils des aktuellen Intensitätsprofils der Fahrspur den Veränderungen im Erscheinungsbild nachgeführt.
- Die vierte Methode betrachtet das Intensitätsprofil in einer größeren Vorausschau, wodurch sprungartige Änderungen im Erscheinungsbild, beispielsweise bei der Einfahrt in einen Tunnel, erkannt werden können und durch den Wechsel zu dem dann gültigen Referenzprofil die kontinuierliche Spurverfolgung gewährleistet wird.

Trotz der geringen Auflösung des Intensitätsprofils und der einfachen Grundidee scheint das Verfahren vergleichsweise robust gegenüber einer Variation der Beleuchtungsverhältnisse und sich ändernder Markierungsarten zu sein. Allerdings sind die ermittelten Größen der Fahrspurgeometrie und die Ablage vergleichsweise ungenau, was die Regelung des Fahrzeugs erschweren wird. Die Autoren berichten von Ablagefehlern im Bereich von ca.  $13cm$  bei einer Standardabweichung von ca.  $16cm$ . Ferner scheint eine Interpretation von Spurmarkierungen als Verkehrszeichen, wie es beispielsweise bei durchgezogenen bzw. gestrichelten Linien durchaus angebracht wäre, nicht vorgesehen oder möglich.

Weitere interessante Veröffentlichungen der CMU zu den Themen visuelle Fahrzeug- und Fahrspurerkennung sind beispielsweise [DELLAERT&THORPE 97] und [DELLAERT 98].

### 2.1.2 Daimler-Benz Forschung

In [FRANKE 97] wird eine Kombination einer visuellen Spurerkennung und eines radarbasierten AICC Systems (Autonomous Intelligent Cruise Control) zur Verbesserung der Spurerkennung vorgestellt. Ferner kann auch die Systemleistung des AICC Systems durch die Zuordnung der Fahrzeuge zu den Fahrspuren unter Verwendung der visuellen Fahrspurinformation gesteigert werden. Ein reines AICC System kann nämlich nicht zwischen einem Fahrzeug, das gerade eine Fahrspur wechselt, und einem Fahrzeug, das in eine Kurve einfährt, unterscheiden. Diese Mehrdeutigkeit kann durch die visuelle Fahrspurerkennung beseitigt werden. Die Autoren schlagen zur Verbesserung der Fahrspurerkennung vor, die Fahrtrajektorien der vorausfahrenden Fahrzeuge bei der Schätzung der Fahrspurgeometrie zu berücksichtigen. Dies geschieht unter der Annahme, dass die Querbewegungen der vorausfahrenden Fahrzeuge relativ zu ihren Fahrspuren klein sind, d.h. dass sie ihren aktuellen Fahrspuren folgen. Eine Verbesserung der Positionsbestimmung bzw. eine Bestimmung der Dimensionen der vorausfahrenden Fahrzeuge durch Vision wird nicht durchgeführt.

In Versuchsfahrten auf deutschen Autobahnen wurden Videodaten, Eigenfahrzeugdaten und die Objektlisten eines Radar-basierten AICC Systems aufgezeichnet und die rein visuelle Fahrspurerkennung anschließend im Labor mit der fusionierten Variante der Fahrspurerkennung bezüglich des ermittelten Krümmungswertes  $c_0$  verglichen. Die Autoren kommen zu dem Ergebnis, dass sich insbesondere bei guten Sichtverhältnissen (die mögliche Vorausschauentfernung wurde auf  $60m$  gesetzt) kaum nennenswerte Unterschiede ergeben, allerdings bei schlechten Sichtverhältnissen (die mögliche Vorausschauentfernung wurde auf  $20m$  gesetzt) die fusionierte Variante deutlich bessere Krümmungswerte liefert (wobei es bei den üblichen Krümmungen deutscher Autobahnen eigentlich nicht möglich erscheint, mit einer rein visuellen Fahrspurerkennung bei einer Vorausschauentfernung von  $\leq 20m$  überhaupt noch sinnvolle Krümmungswerte zu bestimmen). In der fusionierten Variante konnte gegenüber der Referenzkurve mit  $60m$  Vorausschau sogar gezeigt werden, dass sich durch die vorausfahrenden Fahrzeuge eine virtuell größere Vorausschau ergibt, wodurch in den gezeigten Krümmungsverläufen die fusionierte Variante Krümmungsänderungen früher registrieren konnte.

In [GERN 00] wird der Ansatz nach [FRANKE 97] zur Verbesserung der Fahrspurzustandsschätzung durch Fusion der mittels Radar verfolgten vorausfahrenden Fahrzeuge mit der visuellen Fahrspurerkennung nochmals vorgestellt. Als Radar-basiertes ACC System (Adaptive Cruise Control) für die Objekterkennung dient hier das in der Mercedes S-Klasse 1999 in Serie eingeführte DISTRONIC System. Entgegen [FRANKE 97] beschreiben die Autoren hier das Problem, dass die reflektierten Radarimpulse auf dem Heck eines Fahrzeuges stark in ihrer lateralen Position variieren, wodurch die fusionierte Krümmungsschätzung nur ungenau bestimmt werden kann. Die Autoren schreiben weiter, dass sie gerade dabei sind, ein System zu entwickeln, welches (ähnlich [HOFMANN ET AL. 00]) die im Radar detektierten Hypothesen mittels monokularer Bildauswertung verifiziert und ihre lateralen Positionen hier unter Verwendung eines Symme-

trieoperators nach [KÜHNLE 91] bestimmt. Der bisherige Ansatz zur visuellen Fahrspurerkennung beruhte auf der Annahme einer ebenen Fahrbahn. Unter Verwendung einer Stereoanordnung wird hier ein vertikales Krümmungsmodell entsprechend [MYSLIWETZ 90] und ein Nickwinkelschätzer ähnlich [BEHRINGER 96] integriert.

In Erweiterung von [GERN 00] wird in [GERN 02] zur Verbesserung der Fahrspurerkennung bei schlechten Witterungsbedingungen vorgeschlagen, neben den etablierten Merkmalen zur Fahrspurerkennung, wie z.B. Spurmarkierungen, den horizontalen optischen Fluß als Merkmal zur Berücksichtigung von zur Straße parallelen Strukturen zu verwenden. In aufeinanderfolgenden Bildern werden hierbei in definierten Bildzeilen die Helligkeitsverläufe korreliert und daraus der horizontale optische Fluß bestimmt, wobei dem Text nicht zu entnehmen ist, wie die Abweichung gegenüber dem erwarteten horizontalen optischen Fluss ins Kalmanfilter eingespeist wird.

### 2.1.3 Fraunhofer Institut für Informations- und Datenverarbeitung

In [ENKELMANN 02] wird ein System eines Spurwechselassistenten vorgestellt, das in Zusammenarbeit mit der Robert Bosch GmbH entwickelt wurde. Das System überwacht den rückwärtigen und seitlichen linken Verkehrsraum eines Pkw unter Verwendung eines Long Range Radar Systems (Reichweite bis zu 150m), eines Stereovision Systems (Reichweite bis 40m) und zweier seitlich angebrachter Short Range Radar Systeme (Reichweite ca. 8 – 10m). Das Spurwechselsystem ist als Warnsystem konzipiert, das den Fahrer im Falle eines gefährlichen Objekts in der benachbarten Spur vor der Durchführung eines Spurwechsels warnt. Gefährliche Objekte können dabei Fahrzeuge sein, die sich sehr schnell von hinten annähern oder sich bereits in unmittelbarer Nähe des Eigenfahrzeugs, z.B. im toten Winkel, befinden. Die Fusion der verschiedenen Objektlisten der einzelnen Sensoren erfolgt, indem zunächst alle Objekte in ein gemeinsames Fahrzeugkoordinatensystem transformiert und die Abstände zwischen allen Objekten aller Sensoren berechnet werden. Für zwei Messungen wird angenommen, dass sie genau dann vom gleichen Objekt herrühren, wenn ihr Abstand kleiner einer vorgegebenen Schwelle ist. Nachdem solche Korrespondenzen ermittelt wurden, werden die Messungen der verschiedenen Sensoren unter Verwendung fest vorgegebener Prioritäten (Gewichtungen, die das Vertrauen in den Wert einer Messung eines bestimmten Sensors widerspiegeln sollen) kombiniert, wobei nicht genau beschrieben ist, welcher Art diese Kombination ist. Die kombinierten (vermutlich gewichtet gemittelten) Werte werden dann Kalman gefiltert, um die Objekte über der Zeit zu tracken. Dabei werden die Position, Geschwindigkeit und Beschleunigung in Längs- und Querrichtung unabhängig voneinander geschätzt.

## 2.2 Systeme zur Detektion negativer Hindernisse

### 2.2.1 Jet Propulsion Laboratory

Eine Forschergruppe des JPL (Jet Propulsion Laboratory „*managed for NASA by the California Institute of Technology*“) gibt in [MATTHIES 03] zunächst einen guten Überblick, warum die Erkennung negativer Hindernisse im Vergleich zu positiven Hindernissen deutlich schwieriger ist. Sie sagen sogar „Detecting negative obstacles is one of the most difficult problems in perception for autonomous, off-road navigation.“

Dies liegt daran, dass sich der für die Wahrnehmung notwendige vertikale Raumwinkel bei positiven Hindernissen proportional  $\frac{1}{R}$  und bei negativen Hindernissen proportional  $\frac{1}{R^2}$  mit der Entfernung  $R$  verkleinert. Für die Erkennung eines negativen Hindernisses gehen sie davon aus, dass mindestens sechs Pixel Überdeckung im Bild notwendig sind, wodurch sie bei ihrer Sensorconfiguration für eine frühestmögliche zuverlässige Erkennung auf eine Entfernung von lediglich  $6m$  kommen. Nachdem dies bei einem ausgedehnten Objekt zu spät für ein Ausweichmanöver ist, kann bei geringen Geschwindigkeiten lediglich vor einem negativen Hindernis angehalten werden.

Zur früheren Erkennung negativer Hindernisse schlagen sie deshalb vor, die auf Stereoentfernungsmessungen basierende Objekterkennung durch alternative Merkmale zu ergänzen. Dabei legen sie ihren Fokus auf die Verbesserung der Erkennung bei Nacht. Sie zeigen anhand verschiedener systematischer Versuche, dass sich bei fehlender Sonneneinstrahlung und Trockenheit die Umgebung schneller abkühlt als ein negatives Hindernis, wodurch die Erkennung negativer Hindernisse bei Nacht durch Extraktion ihrer thermischen Signatur möglich wird. In Kombination mit den Detektierverfahren aus Stereoentfernungsdaten vergrößert sich hierdurch nach ihren Angaben die mögliche Vorausschau um 45%. Für die Wärmebildaufnahmen verwenden sie gekühlte Ferninfrarotkameras.

In ihrem Algorithmus suchen sie spaltenweise nach hellen Spots in den Temperaturbildern, indem sie zunächst den durchschnittlichen Helligkeitswert des Bodens ermitteln und dann alle Pixel, welche relativ zu diesem Wert eine Helligkeitsschwelle überschreiten, als potentielle Pixel eines negativen Hindernisses markieren. Die dazugehörigen Entfernungswerte werden dazu verwendet, um eine Hypothese zu stützen oder zu verwerfen. Nachdem alle Spalten derart abgearbeitet wurden, werden benachbarte Pixelbereiche zu Objektbereichen gruppiert. Als problematisch beschreiben die Autoren den Umstand, dass die potentiellen Grabenpixel im Temperaturbild häufig so hell sind, dass die Kameras dort in die Sättigung gehen, wodurch das Stereoverfahren bei der Berechnung der Entfernungswerte scheitert. In diesem Fall werden nur die Entfernungswerte der Berandungen für die Verifikation der Hypothese verwendet.

In [BELLUTTA 00] wird die Erkennung negativer und positiver Hindernisse bei Tag und Nacht mittels Stereoauswertung zweier FLIR-Kameras durchgeführt, wobei die einzelnen Spalten des daraus errechneten Entfernungsbildes nach Lücken und Diskontinuitäten durchsucht werden. Ferner versuchen sie, eine Gelände-Typ-Klassifikation in Abhängigkeit

von seiner Farbe durchzuführen, um beispielsweise bei einem Grasbüschel eine Entscheidung bezüglich der Überfahrbarkeit treffen zu können. Um den begrenzten Sichtbereich der Kameras optimal zu nutzen, wird hier eine aktive Kameraplattform eingesetzt, deren Blickrichtung stets so gesteuert wird, dass die Kameras auf den zukünftigen Pfad des Fahrzeugs blicken.

Die Autoren weisen aufgrund der extremen Geländeformationen, die von der Objekterkennung beherrscht werden sollen, explizit darauf hin, dass es nicht empfehlenswert wäre, die Stereobilder in Bezug auf eine horizontale *Horopterebene* (eine Ebene, die in die Bodenebene gelegt wird, relativ zu der die Disparitäten bestimmt werden) zu analysieren, da hohe Hindernisse (wie z.B. ein Ast) große Disparitäten in Bezug auf die Horopterebene hervorrufen würden, wodurch bei begrenzter Suchfenstergröße diese nicht detektiert werden könnten. Nachdem JPL in diesem Artikel ihre eigenes Stereoverfahren gegenüber anderen Forschergruppen etablieren will, kann diese Aussage durchaus auch als kleiner Seitenhieb gegenüber der Gruppe um Peter Burt (Sarnoff Corporation, Princeton, New Jersey, USA) verstanden werden, welche ebenfalls auf dem Gebiet tätig sind und zu dieser Zeit insbesondere mit ihrer Spezialhardware dem VFE200 (Vision Front End) zur Berechnung der Horopterebene und Stereoauswertung in Echtzeit große Erfolge hatten.

### 2.2.2 UniBwM und Sarnoff Corporation

Beispielsweise wird in [SIEDERSBERGER 01] die Kombination des von Sarnoff entwickelten VFE200 mit dem EMS-Vision System der UniBwM beschrieben. Das VFE200 ist eine speziell zur Echtzeit Bildverarbeitung konzipierte Spezialhardware, welche pro Sekunde 25 Frames einer 320x240 Pixel großen Disparitätenkarte berechnet und auswertet. Eine Schlüsseltechnik ist hierbei die Berechnung der bereits erwähnten Horopterebene in Echtzeit. Diese Spezialhardware wurde unter anderem zur Erkennung eines negativen Hindernisses verwendet, wobei die Objektposition und -ausdehnung in einem Gittermodell durch zeitlich akkumulierte belegte Gitterbereiche repräsentiert wurde, ohne dass die Form und Ausdehnung des Hindernisses explizit modelliert wurde. Die Entfernung zum Hindernis wurden dann in die Repräsentation und den Kontrollfluß des EMS-VISION Systems in den Versuchsträger VAMORS der UniBwM integriert, wodurch ein autonomes Anhalten vor einem negativen Hindernis verwirklicht werden konnte.

Auf der Abschlussdemonstration des Forschungsprojekts IFF3 (Intelligente Fahrzeugfunktionen Phase 3) zur autonomen Fahrzeugführung im Oktober 2001 konnte vor internationalem Publikum unter Verwendung des Nachfolgemodells des VFE200, einer PCI-PC-Einsteckkarte mit der Bezeichnung *Acadia*, wieder in Kombination mit dem EMS-VISION System das autonome Querfeldeinfahren mittels GPS-Wegpunktnavigation mit Anhalten vor einem negativen Hindernis gezeigt werden. Hierbei wurde erstmalig eine explizite Repräsentation für ein negatives Hindernis verwendet, wobei erst durch die Erweiterungen der Wahrnehmungsfähigkeiten entsprechend Kapitel 8 auch ein Ausweichen an einem Graben verwirklicht werden konnte.

### 2.2.3 Sarnoff Corporation

In [RIEDER 02] werden Versuchsträger für autonomes Querfeldeinfahren mit Erkennung positiver und negativer Hindernisse beschrieben. Als Versuchsträger dienen zwei Honda ATVs (all-terrain vehicles), die auf einer drehbaren Plattform, die nicht zur Bildstabilisierung gedacht ist, mit drei Stereokamera paaren und einer „EBK Scanning Ladar Range Camera“ ausgestattet wurden, wobei der Fokus des Papers auf der Beschreibung der Stereoanordnungen und der Objekterkennung aus Entfernungsdaten liegt. Nur ganz kurz: Die EBK Scanning Ladar Range Camera ist ein scannendes Lasersystem, das z.B. einen  $30^\circ \times 60^\circ$  großen Sichtbereich in  $0.5^\circ$  Schritten scannt und für jede Position den Entfernungswert ermittelt, wodurch ein Entfernungsbild entsteht. Die Bildwiederholrate liegt dabei im Bereich 1Hz - 4Hz.

Die drei Stereokamera paare sind so angeordnet, dass sie einen möglichst großen horizontalen Sichtbereich abdecken. Hierfür sind die Kameras eines monochromen Kamera paares horizontal mit Stereobasis  $55\text{cm}$  und parallelen optischen Achsen angeordnet und die beiden anderen Kamera paare (Farbkameras) jeweils vertikal übereinander mit Stereobasis  $23\text{cm}$  und ebenfalls parallelen optischen Achsen, jedoch gegenüber der ersten Kameraanordnung nach links und rechts um jeweils  $27^\circ$  gegiert. Die divergenten Stereokamera paare decken einen horizontalen Sichtbereich von ca.  $100^\circ$  ab. Die Kameras des zentralen horizontalen Stereopaars sind mit Objektiven größerer Brennweite ausgestattet, um eine verbesserte Tiefenaufösung zu erhalten. Alle Kameras haben eine Auflösung von  $720 \times 480$  Pixeln. Für die Echtzeit-Stereoauswertung wird die Acadia Spezialhardware eingesetzt.

Mittels dieser Sensoranordnung wurden systematische Versuche zur Bestimmung der Erkennungsleistung des Systems durchgeführt. Hierbei wurde der Abstand zu positiven und negativen Hindernissen von ca.  $3\text{m}$  bis  $16\text{m}$  variiert. Die ermittelten Entfernungen für eine frühestmögliche Erkennung lagen dabei für ein kleines positives Hindernis (Baumstumpf mit Durchmesser  $15\text{cm}$  und  $25\text{cm}$  Höhe) bei ca.  $10\text{m}$  und für ein negatives Hindernis (rechteckiger Graben mit  $0.6\text{m}$  Breite und  $1.2\text{m}$  Länge) bei ca.  $10.5\text{m}$ . Die Autoren führten auch reale Versuchsfahrten mit Ausweichen an Hindernissen durch. Die durchschnittliche Geschwindigkeit beim Durchfahren der Versuchsstrecke lag dabei bei ca.  $0.5 \frac{\text{m}}{\text{s}}$ .

### 2.2.4 National Institute of Standards and Technology

Eine Forschergruppe des NIST (National Institute of Standards and Technology, USA)) verwendet in [COOMBS 00] ebenfalls ein scannendes LADAR System zur Erkennung von positiven und negativen Hindernissen. Der Schwerpunkt ihres Papers liegt nicht im Bereich der Wahrnehmung, sondern unter der Annahme, dass die Wahrnehmung alle relevanten Hindernisse erkennt, mit möglichst hoher Geschwindigkeit ( $35 \frac{\text{km}}{\text{h}}$ ) definierte GPS-Wegpunkte anzufahren und dabei in ausreichendem Abstand an Hindernissen vorbei zu navigieren. Sie verweisen zum Vergleich auch auf das deutsche Projekt Primus-C [SCHWARTZ 99], in dem das Kettenfahrzeug „Wiesel2“ eine autonome Mission mit bis zu  $25 \frac{\text{km}}{\text{h}}$  absolvierte.

# Kapitel 3

## Sensorsysteme zur maschinellen Umgebungserfassung

Die maschinelle Umgebungserfassung dient der Bestimmung der Lage und Ausdehnung von Objekten im Raum mittels Sensoren. Wahrnehmung besteht dabei neben der Erfassung der Ausgangssignale der verwendeten Sensorik insbesondere aus der Interpretation der Sensordaten. Bei mehreren Sensoren ist ferner eine Zusammenschau zu einem konsistenten Umfeldmodell durchzuführen.

Für diese Zusammenschau müssen verschiedene Sensorsignale in Bezug zueinander gesetzt werden können. Hierfür ist zum einen die Möglichkeit zur Transformation in andere Koordinatensysteme notwendig und zum anderen muss die Abbildungsvorschrift der Objekte der realen Welt in den Sensordatenraum des jeweiligen Sensors bekannt sein. Diese Abbildungsvorschrift (das *Messmodell*) besteht aus dem Sensormodell und der Position und Ausrichtung des Sensors innerhalb der realen Welt. Für bildgebende Sensoren werden in Abschnitt 3.3.2 zwei Kameramodelle vorgestellt und in Abschnitt 3.3.3 die Vorgehensweise zu deren Kalibrierung erläutert.

Aufgrund unterschiedlicher physikalischer Funktionsprinzipien und Empfindlichkeitsbereiche besitzen verschiedene Sensorarten meist unterschiedliche charakteristische Eigenschaften. Die Möglichkeiten zur Wahrnehmung bestimmter Arten von Merkmalen können dabei sogar über die Fähigkeiten des Menschen hinausgehen. Ein Radarsystem ist beispielsweise in der Lage, die Entfernung und relative Geschwindigkeit gegenüber metallischen Objekten auch bei Dunkelheit oder durch Nebel hindurch wahrzunehmen, wobei allerdings die Ausmaße eines Objekts nicht bestimmt werden können. Eine Laserentfernungsmessung kann punktuell mit sehr hoher Genauigkeit die Entfernung gegenüber einem Reflexionspunkt auf einem Objekt bestimmen.

Für unterschiedliche Entfernungsbereiche, horizontale und vertikale Öffnungswinkel, geforderte Genauigkeiten, Totzeiten, Empfindlichkeiten gegenüber Störeinflüssen werden deshalb meist Kombinationen von Sensoren eingesetzt. Die Kunst bei der Auswahl einer Sensorkonfiguration besteht darin, die Schwächen des einen Sensors möglichst durch

Stärken eines anderen Sensors zu kompensieren, um im Gesamtsystem die notwendige Wahrnehmungsleistung für eine gewünschte Funktion gewährleisten zu können.

### 3.1 Der Szenenbaum

Zur Repräsentation der Relativlagen von Koordinatensystemen dient im EMS-VISION System ein gerichteter Graph, der sogenannte *Szenenbaum* [DICKMANN 97]. In Abbildung 3.1 ist ein Beispiel eines solchen Szenenbaums dargestellt. Die Knoten im Szenenbaum symbolisieren die Koordinatensysteme.

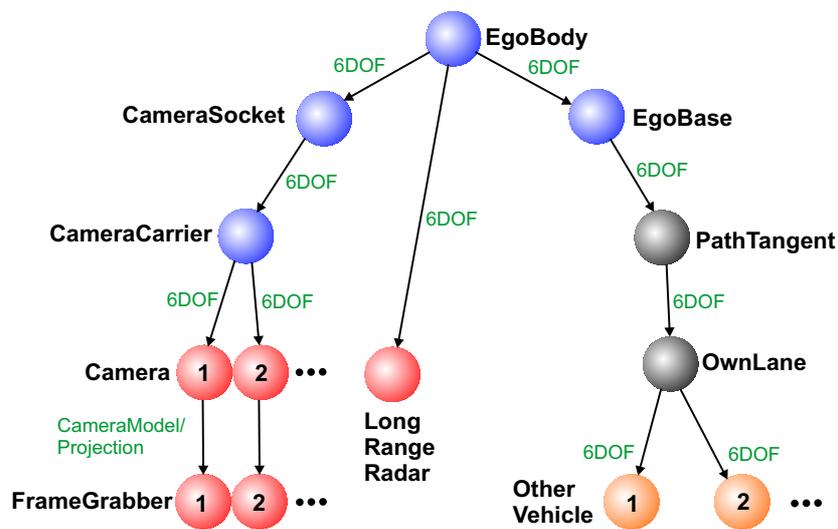


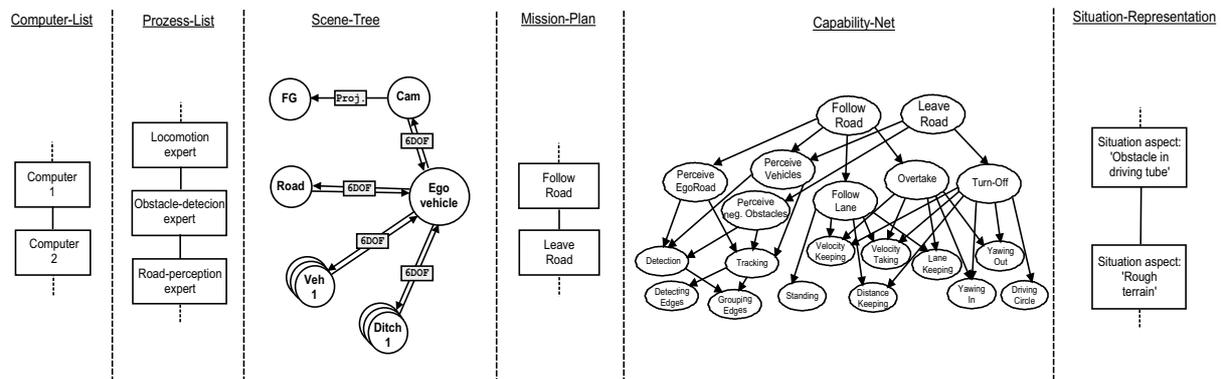
Abbildung 3.1: Ein Beispiel eines Szenenbaums des EMS-VISION Systems.

Die Relativlagen der Knoten, beispielsweise der Sensoren zueinander oder die Einbaulagen im Versuchsträger, werden hierbei mittels homogener Transformationsmatrizen (HTMs) mit sechs Freiheitsgraden (engl., 6 degrees of freedom = 6DOF) beschrieben. Dabei spielt es keine Rolle, ob ein Knoten ein physikalisches Objekt oder lediglich ein virtuelles Koordinatensystem repräsentiert. Details zu den Koordinatensystemen im EMS-VISION System können in [SIEDERSBERGER 01] nachgelesen werden. Ferner finden sich dort, in [RIEDER 00] und in [DICKMANN 97] Beschreibungen homogener Koordinaten und der homogenen Koordinatentransformationen.

Die Pfeile geben hier die jeweils positive Transformationsrichtung eines Punktes vom Vater- in ein Sohnkoordinatensystem an. Für die umgekehrte Transformationsrichtung sind die homogenen Transformationsmatrizen, die die Koordinatentransformation beschreiben, zu invertieren. Für die Translation und Rotation entsprechend den sechs Freiheitsgraden ist dies einfach zu realisieren. Dies gilt allerdings nicht für die Projektion eines

Raumpunktes mittels Kameramodell in die Bildebene bzw. den Bildspeicherbereich eines Framegrabbers, denn durch die Projektion ins Bild geht die Entfernungsinformation eines Raumpunktes verloren.

Der Szenenbaum ist Teil der Wissensrepräsentation des EMS-VISION Systems. Er ist eingebettet in die sogenannte *Dynamische Objektdatenbasis* (DOB) [RIEDER 00], welche neben dem Szenenbaum noch eine Repräsentation der im System verfügbaren Fähigkeiten und ihrer statischen Abhängigkeiten, das sogenannte *Fähigkeittennetz*, besitzt ([SIEDERSBERGER 03] und [PELLKOFER 03]). Weiterhin sind in ihr zur Verwaltung der Rechenressourcen die im System verfügbaren Computer und die dort jeweils gestarteten Prozesse eingetragen. Die Wissensrepräsentation verfügt ferner über einen Zweig, in dem die Missionselemente der Missionsplanung eingetragen werden [GREGOR 02] und einen Zweig, in dem die Situationsanalyse relevante Situationsaspekte mittels linguistischer Variablen beschreibt [PELLKOFER 03]. In Abbildung 3.2 sind die Zweige der Wissensrepräsentation nochmals schematisch dargestellt.



**Abbildung 3.2:** Wissensrepräsentation des EMS-VISION Systems. Diese Abbildung ist aus [HOFMANN&SIEDERSBERGER 03].

Die DOB stellt ferner Funktionalitäten zur Kommunikation zwischen Prozessen und zur Synchronisation der Wissensrepräsentationen in einem verteilten Rechnernetz zur Verfügung. Außerdem bietet sie die Möglichkeit zur automatisierten Berechnung der homogenen Transformationsmatrizen zwischen beliebigen Knoten im Szenenbaum. Die Qualität der Kalibrierung der Koordinatensysteme ist dabei mitentscheidend für die erreichbare Genauigkeit der internen Modellvorstellung über die Umwelt.

## 3.2 Ultraschall, Lidar und Radar

Für sehr kurze Distanzen zu Objekten ( $< 2m$ ) kommen z.B. für die Funktion einer Einparkhilfe häufig Ultraschallsensoren zum Einsatz. Für mittlere Entfernungen und bei

schnellerer Fahrt sind diese Sensoren für die Objekterkennung eher ungeeignet. Als Entfernungssensor im Nahbereich ( $< 20m$ ), für mittlere Entfernungen ( $20m...60m$ ) und bei guten Sichtbedingungen im Fernbereich ( $60m...150m$ ) können Laserentfernungsmesser in Form von Scannern oder als Multibeam Ausführung mit einer Entfernungsmessung in fest vorgegebenen Raumrichtungen verwendet werden. Diese werden in Anlehnung an die Bezeichnung des Radar (engl. „radio detection and ranging“) häufig auch als Lidar Systeme (engl. „light detection and ranging“) bezeichnet.

Für den nahen und mittleren Entfernungsbereich wird neuerdings der Einsatz von 24 GHz Short Range Radar Systemen untersucht, welche einen Öffnungswinkel von bis zu  $\pm 40^\circ$  besitzen. Für den Fernbereich sind bereits bei vielen Automobilherstellern 77 GHz Long Range Radar Systeme mit Reichweiten bis zu  $150m$  als Sensor einer intelligenten Tempomatfunktion käuflich zu erwerben.

In dieser Arbeit sollen hauptsächlich Videokameras als bildgebende Sensoren zum Einsatz kommen. In Kapitel 6 wird zwar ein Radarsystem mit einer bifokalen Kameraanordnung kombiniert, der Schwerpunkt liegt jedoch auch dort auf der Verbesserung der Erkennungsleistung durch visuelle Fahrspurerkennung und Objektverifikation. Deshalb wird in Abschnitt 6.3 lediglich eine Gegenüberstellung der wesentlichen Sensoreigenschaften eines Radarsystems und eines Kamerasystems durchgeführt und auf eine detaillierte Beschreibung der physikalischen Prinzipien bewusst verzichtet.

### 3.3 Bildgebende Sensoren

Der Wellenlängenbereich des sichtbaren Lichts befindet sich zwischen  $390nm$  (blau) und  $770nm$  (rot). Für die Bildauswertung werden hier nur monochrome Kameras eingesetzt, welche im Bereich des sichtbaren Lichts empfindlich sind. Bildfolgen von Kameras, welche über diesen sichtbaren Bereich hinaus noch im daran anschließenden nahen Infrarotbereich zwischen  $770nm$  und  $1400nm$  empfindlich sind, bzw. Bildfolgen spezieller (häufig gekühlter) Wärmebildkameras für den fernen Infrarotbereich wurden nicht untersucht.

#### 3.3.1 Anforderungen an die Kamerakalibrierung

Die Durchführung komplexer Fahrmanöver mit aktiver Blickrichtungssteuerung erfordert bei Verwendung mehrerer Kameras mit großen und kleinen Brennweiten eine genaue Kamerakalibrierung unter Berücksichtigung der Linsenverzerrungen und eine genaue Bestimmung der Relativlagen der verschiedenen Sensor-Koordinatensysteme (KS).

Größere Ungenauigkeiten bei der Bestimmung der Kameraparameter, ein ungenügendes Kameramodell oder zu ungenaue Einträge in den homogenen Transformationsmatrizen für die Transformationen zwischen den Koordinatensystemen würden dazu führen, dass sich die durch die visuelle Wahrnehmung aufgebaute interne Modellvorstellung über die

Ausmaße von Objekten und die ermittelten Entfernungen und Positionen zu Objekten durch die akkumulierten Fehler mehr oder weniger stark von der Realität unterscheidet.

Für das Kalibrieren einer Kamera ist zunächst ein geeignetes Kameramodell festzulegen. Die Parameter des Kameramodells werden als *interne* oder *intrinsische Parameter* bezeichnet im Gegensatz zu den *externen* oder *extrinsischen Parametern*, welche die Lage des Kamerakoordinatensystems relativ zum Kalibrierobjekt-KS mit drei translatorischen und drei rotatorischen Freiheitsgraden beschreiben. In Kapitel 3.3.2 wird zunächst das einfache Lochkameramodell erläutert und daran anschließend das Kameramodell nach [TSAI 87].

Weiterhin ist ein geeignetes Kalibrierobjekt zu definieren und die Lage der Kalibrierpunkte im Kalibrierobjekt-KS zu modellieren. Aus den dann bekannten Lagen der Kalibrierpunkte auf dem Kalibrierobjekt, beschrieben im objekteneigenen 3-D Koordinatensystem, und den diesen Punkten zugeordneten in die Bildebene projizierten Kalibrierpunkten kann die Bestimmung der internen und externen Parameter erfolgen.

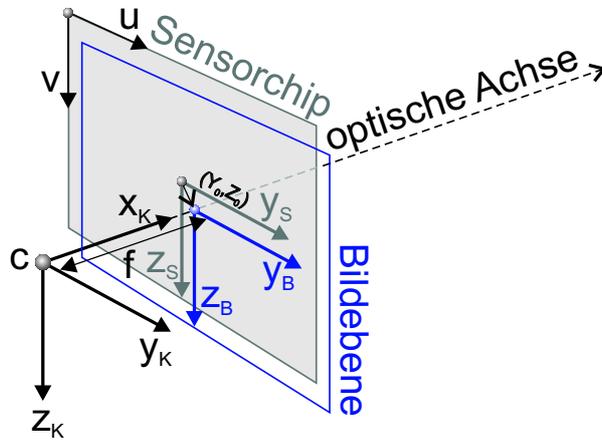
Die Kalibrierpunkte sollten in 3-D im Kalibrierobjektkoordinatensystem einfach zu modellieren sein, und die Bestimmung der in die Bildebene projizierten Kalibrierpunkte in 2-D sollte möglichst zu automatisieren sein. Es soll deshalb ein Kalibrierverfahren zum Einsatz kommen, welches sowohl die internen als auch externen Kameraparameter bezüglich des Kalibrierobjektes bestimmt. Die 2-D Kalibrierpunkte sollen durch möglichst wenige Vorgaben ohne 3-D Information aus dem 2-D Bild bestimmt werden. Die Vermessung soll auf Subpixel genau erfolgen. Anschliessend soll eine Datei erzeugt werden, welche unmittelbar als Eingangsdatum für die eigentliche Kalibrierroutine verwendet werden kann. In dieser Datei wird die Zuordnung der 2-D Bildpunktkoordinaten zu den im Objekt-KS beschriebenen 3-D Kalibrierpunktkoordinaten hergestellt.

## 3.3.2 Kameramodelle

### 3.3.2.1 Lochkameramodell

Das Lochkameramodell beschreibt die perspektivische Projektion mittels einer idealen Lochkamera. Linsenverzerrungen bleiben beim Lochkameramodell unberücksichtigt. Eine Lochkamera besteht aus einer Lochblende, die sich im Abstand  $f$ , auch *Brennweite* genannt, vor einer Bildebene befindet. Die Position der Lochblende wird als *optisches Zentrum*  $C$  oder auch als *Brennpunkt* bezeichnet. Ein Raumpunkt wird auf den Schnittpunkt des Projektionsstrahls, der durch diesen Raumpunkt und den Brennpunkt verläuft, mit der Bildebene abgebildet. Diese Abbildung wird *perspektivische Projektion* genannt. Der Projektionsstrahl, der normal zur Bildebene durch den Brennpunkt verläuft, wird als *optische Achse* bezeichnet. Der Schnittpunkt der optischen Achse mit der Bildebene wird *Hauptpunkt* genannt. Häufig wird die Bildebene als *virtuelle Bildebene* dargestellt, welche zwischen Brennpunkt und abzubildendem Raumpunkt im Abstand  $f$  vom Brennpunkt liegt und um  $180^\circ$  um die optische Achse gedreht wurde.

In der 2-D Bildebene sind drei Koordinatensysteme zu unterscheiden, welche sich durch Skalierung und Translation ineinander überführen lassen. Dies ist in Abbildung 3.3 dargestellt.



**Abbildung 3.3:** Schematische Darstellung der Kamerakoordinatensysteme.

Der Ursprung des  $y_B$ - $z_B$ -Bildkoordinatensystems liegt in der Bildebene im Durchstoßpunkt der optischen Achse, wobei seine  $y_B$ -Koordinatenachse parallel zu den Zeilen des Sensorchips von links nach rechts und seine  $z_B$ -Koordinatenachse parallel zu den Spalten von oben nach unten verläuft. Die Einheit der Koordinaten ist [Millimeter]. Durch Skalierung der  $Y_B$ -Koordinate mit dem Faktor  $\frac{\text{Pixelanzahl}_Y [\text{Pixel}]}{\text{Sensorchip}_Y [\text{mm}]}$  bzw. der  $Z_B$ -Koordinate mit dem Faktor  $\frac{\text{Pixelanzahl}_Z [\text{Pixel}]}{\text{Sensorchip}_Z [\text{mm}]}$  und Translation des Koordinatensystemursprungs um  $(-Y_0, -Z_0)$  gegenüber dem  $y_B$ - $z_B$ -Koordinatensystem ergibt sich das  $y_S$ - $z_S$ -Sensorkoordinatensystem, dessen Ursprung in der Mitte des Bildsensors liegt. Das heißt, im  $y_S$ - $z_S$ -Koordinatensystem besitzt der Hauptpunkt die Koordinaten  $(Y_0, Z_0)$ . Die Koordinaten werden in [Pixel] angegeben.

Ein drittes Koordinatensystem ist das  $u$ - $v$ -Bildkoordinatensystem, dessen Ursprung sich in der linken oberen Ecke des jeweiligen Bildsensors befindet, mit der  $u$ -Achse nach rechts und der  $v$ -Achse nach unten orientiert. Die Einheit seiner Koordinatenachsen ist [Pixel].

Der Ursprung des 3-D  $x_K$ - $y_K$ - $z_K$ -Kamerakoordinatensystems befindet sich im optischen Zentrum, d.h. im Brennpunkt. Die  $y_K$ - und  $z_K$ -Achsen des Kamera-KSs seien dabei kollinear zum  $y_B$ - $z_B$ -Bild-KS. Die  $x_K$ -Achse fällt demnach mit der optischen Achse zusammen. Da alle Projektionsstrahlen durch den Brennpunkt verlaufen, kann mit Hilfe des Strahlensatzes eine Beziehung zwischen den 3-D Raumkoordinaten und den 2-D Bildkoordinaten hergestellt werden.

$$\frac{y_B}{y_K} = \frac{z_B}{z_K} = \frac{f}{x_K} \quad (3.1)$$

Bei der perspektivischen Projektion werden also die Kamerakoordinaten beim Übergang zu den Bildkoordinaten mit dem Skalierfaktor  $\frac{f}{x_K}$  multipliziert. Die Bildkoordinaten  $y_B$  und  $z_B$  enthalten daher keine absolute Information über die Koordinaten eines Raumpunktes. Das heißt, sowohl die Entfernung als auch die wahre Größe eines Objektes sind daraus nicht bestimmbar.

### 3.3.2.2 Kameramodell nach [TSAI 87]

In [TSAI 87] wird ein Kameramodell zur Berücksichtigung radialer Linsenverzerrungen beschrieben. Die Transformation eines Punktes im 3-D Objektkoordinatensystem ins 2-D Bildkoordinatensystem erfolgt dabei in vier Schritten.

Es seien  $(x_w, y_w, z_w)$  die 3-D Koordinaten eines Punktes  $P$  im 3-D Objektkoordinatensystem. Entsprechend Abschnitt 3.3.2.1 seien  $(x, y, z)$  die 3-D Koordinaten dieses Punktes im 3-D  $x_K$ - $y_K$ - $z_K$ -Kamerakoordinatensystem.  $(Y_u, Z_u)$  seien die Bildkoordinaten bei Verwendung eines idealen Lochkameramodells.  $(Y_d, Z_d)$  seien die tatsächlichen Bildkoordinaten, welche sich von  $(Y_u, Z_u)$  durch die Linsenverzerrungen unterscheiden.

Im ersten Schritt erfolgt die Transformation eines Punktes im 3-D Objektkoordinatensystem ins 3-D Kamerakoordinatensystem. Unter Verwendung homogener Koordinaten und Transformationsmatrizen gilt:

$$\begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \mathbf{M} \cdot \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} = \mathbf{T}^{-1} \cdot \mathbf{R}_z^{-1} \cdot \mathbf{R}_y^{-1} \cdot \mathbf{R}_x^{-1} \cdot \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix}, \quad (3.2)$$

wobei

$$\mathbf{M} = \begin{bmatrix} \cos \theta \cos \psi & -\cos \phi \sin \psi + \sin \phi \sin \theta \cos \psi & \sin \phi \sin \psi + \cos \phi \sin \theta \cos \psi & x \\ \cos \theta \sin \psi & \cos \phi \cos \psi + \sin \phi \sin \theta \sin \psi & -\sin \phi \cos \psi + \cos \phi \sin \theta \sin \psi & y \\ -\sin \theta & \sin \phi \cos \theta & \cos \phi \cos \theta & z \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

mit

$$\mathbf{R}_x^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \phi & -\sin \phi & 0 \\ 0 & \sin \phi & \cos \phi & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \mathbf{R}_y^{-1} = \begin{bmatrix} \cos \theta & 0 & \sin \theta & 0 \\ 0 & 1 & 0 & 0 \\ -\sin \theta & 0 & \cos \theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\mathbf{R}_z^{-1} = \begin{bmatrix} \cos \psi & -\sin \psi & 0 & 0 \\ \sin \psi & \cos \psi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \mathbf{T}^{-1} = \begin{bmatrix} 1 & 0 & 0 & T_x \\ 0 & 1 & 0 & T_y \\ 0 & 0 & 1 & T_z \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

ist. Die zu kalibrierenden Parameter sind die Translationskomponenten  $T_x, T_y, T_z$  und die Rotationen um die Winkel  $\phi, \theta, \psi$  entsprechend der angegebenen Drehreihenfolge.

Im zweiten Schritt werden die  $(x, y, z)$ -Koordinaten im 3-D Kamera-KS unter Verwendung des idealen Lochkameramodells auf ideale (unverzerrte) Bildkoordinaten abgebildet:

$$\begin{bmatrix} Y_u \\ Z_u \end{bmatrix} = f \cdot \begin{bmatrix} y/x \\ z/x \end{bmatrix} \quad (3.3)$$

Der hier zu kalibrierende Parameter ist die effektive Brennweite  $f$ .

Im dritten Schritt werden aus den unverzerrten Koordinaten die verzerrten Koordinaten

$$\begin{bmatrix} Y_d \\ Z_d \end{bmatrix} = \begin{bmatrix} Y_u \\ Z_u \end{bmatrix} + \begin{bmatrix} D_{yradial} \\ D_{zradial} \end{bmatrix} + \begin{bmatrix} D_{ytan} \\ D_{ztan} \end{bmatrix} \quad (3.4)$$

berechnet, mit

$$\begin{bmatrix} D_{yradial} \\ D_{zradial} \end{bmatrix} = \begin{bmatrix} Y_u \\ Z_u \end{bmatrix} \cdot (\kappa_1 r^2 + \kappa_2 r^4 + \dots) \quad (3.5)$$

und

$$\begin{bmatrix} D_{ytan} \\ D_{ztan} \end{bmatrix} = \begin{bmatrix} 2 \cdot \tau_1 Y_u Z_u + \tau_2 \cdot (r^2 + 2 \cdot Y_u^2) \\ \tau_1 \cdot (r^2 + 2 \cdot Z_u^2) + 2 \cdot \tau_2 Y_u Z_u \end{bmatrix} \quad (3.6)$$

mit

$$r^2 = Y_u^2 + Z_u^2. \quad (3.7)$$

In der Kamerakalibrier-Toolbox der Open Computer Vision Library, die im Internet unter <http://sourceforge.net/projects/opencvlibrary/> heruntergeladen werden kann, bzw. in der Matlab Calibration Toolbox von Janne Heikkilä, University of Oulu, Finnland, im Internet unter <http://www.ee.oulu.fi/~jth/calibr/>, werden neben den radialen Linsenverzerrungen entsprechend Gleichung (3.5) noch zwei tangentiale Linsenverzerrungsparameter entsprechend Gleichung (3.6) bestimmt. Zu kalibrieren sind hier die radialen Verzerrungsparameter  $\kappa_i$  und tangentialen Verzerrungsparameter  $\tau_i$ . Bei geringen radialen Linsenverzerrungen reicht dabei häufig die Berücksichtigung eines  $\kappa_1$ . Bei stärkeren radialen Linsenverzerrungen macht es Sinn, auch einen  $\kappa_2$  Parameter zu bestimmen. Die tangentialen Verzerrungsparameter können häufig vernachlässigt werden. Die Kalibrieroutine der Open Computer Vision Library bietet ferner noch die Möglichkeit, einen Parameter  $s$  für einen Schrägversatz zu berücksichtigen.

Im vierten Schritt ergeben sich die resultierenden verzerrten Koordinaten im Sensorkoordinatensystem demnach mit einem Parameter  $s_{Ratio} \approx 1$  für die Unsicherheit im horizontalen Pixelskalierfaktor und nach Berücksichtigung der Hauptpunktverschiebung zu

$$Y_S = s_{Ratio} \frac{Pixelanzahl_Y [Pixel]}{Senorchip_Y [mm]} Y_d + s \cdot Z_d + Y_0 \quad (3.8)$$

und

$$Z_S = \frac{Pixelanzahl_Z [Pixel]}{Senorchip_Z [mm]} Z_d + Z_0. \quad (3.9)$$

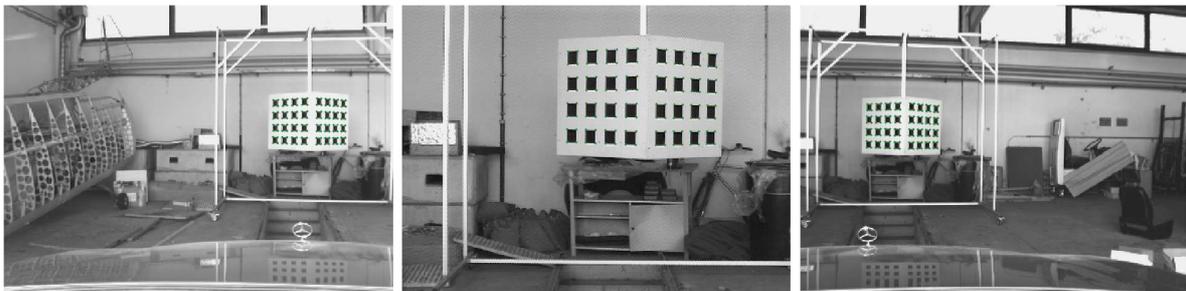
Die zu kalibrierenden Parameter sind  $s_{Ratio}$  und im Bedarfsfall  $s$  für  $s \neq 0$ .

### 3.3.3 Vorgehensweise bei der Kalibrierung komplexer Kameraanordnungen im EMS-Vision System

Als Basis für die Kalibrierung der internen Kameraparameter wurde das bereits erwähnte Kalibrierverfahren (Version 2.1b) von Janne Heikkilä, University of Oulu, Finland verwendet. Die Matlab-Toolbox nutzt als Kameramodell ein erweitertes Modell nach [TSAI 87] mit mehreren Parametern zur Modellierung von radialen und tangentialen Linienverzerrungen. Für nähere Details siehe auch [HEIKKILÄ 97].

Die Bestimmung der internen und der externen Kameraparameter relativ zum Kalibrierobjekt erfolgt dabei losgelöst von der Bestimmung der Einbaulage der Kameras im Fahrzeug. Das heißt, zunächst werden unter Verwendung eines 3-D Kalibrierobjektes entsprechend Abbildung 3.4 die internen Parameter der Kameras bestimmt. Die Kalibrierroutine optimiert die Parameter des Kameramodells dabei derart, dass die Abweichung zwischen den ins Bild projizierten modellierten Kalibrierpunkten (die Eckpunkte der schwarzen Quadrate) und den aufs Subpixel vermessenen Positionen minimiert wird.

Die Kalibrierroutine ermittelt dabei gleichzeitig die jeweilige Lage des Kalibrierobjektes (die externen Parameter) im Kamerakoordinatensystem der jeweiligen Kamera. Die Kalibrierpunkte sollten dabei die Bildebene möglichst vollständig überdecken. Alternativ können auch mehrere Ansichten des Kalibrierobjektes, die die Bildebene gemeinsam möglichst vollständig überdecken, gleichzeitig verwendet werden. Die daraus ermittelte Entfernung (bzw. bei mehreren Ansichten die ermittelten Entfernungen) wird(werden) zur Kontrolle mit der(den) mittels Laserentfernungsmesser bestimmten Entfernung(en) verglichen. Weiterhin werden entsprechend der jeweiligen Ansicht unter Verwendung der ermittelten internen und externen Kameraparameter die modellierten 3-D Kalibrierpunkte ins Bild der Kamera projiziert (grün). Diese sollten die im Bild vermessenen Positionen

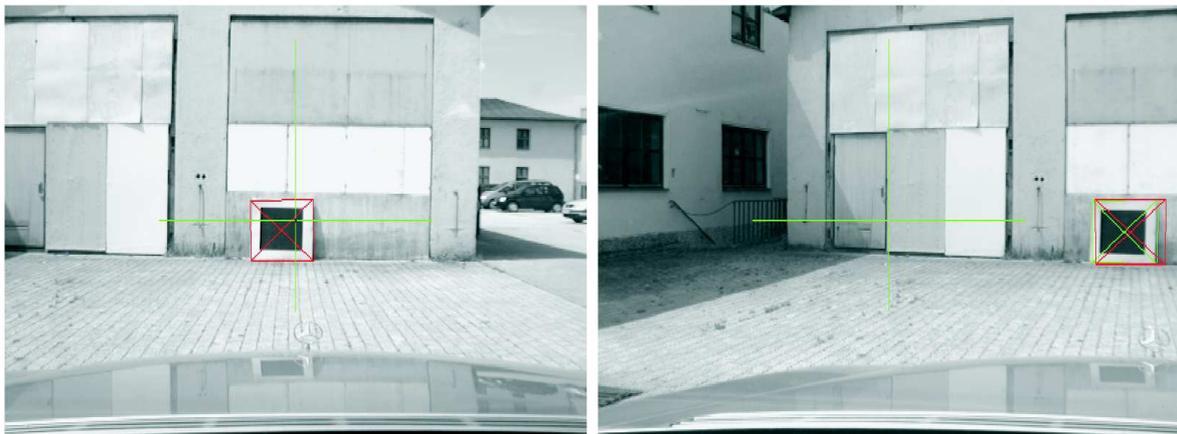


**Abbildung 3.4:** Visualisierung der ins Bild zurückprojizierten 3-D Kalibrierpunkte (grün) nach Bestimmung der internen und externen Kameraparameter: a) linke Nahbereichskamera (Öffnungswinkel  $\pm 22^\circ$ ), b) milde Telekamera (Öffnungswinkel  $\pm 11^\circ$ ), c) rechte Nahbereichskamera (Öffnungswinkel  $\pm 22^\circ$ ). Die Nahbereichskameras sind hier leicht divergent mit  $22^\circ$  horizontalem Überlappungsbereich angeordnet.

der Kalibrierpunkte (rot) im Idealfall vollständig verdecken. Siehe Abbildung 3.4. Die Ein-

baulage der Kameras im Fahrzeug und relativ zueinander ist zu diesem Zeitpunkt noch unbekannt.

Für die Ermittlung der Einbaulage einer Kamera wird zunächst mittels Meterstab die Lage des Ursprungs des Kamerakoordinatensystems relativ zum Kameraträger bestimmt. An dieser Stelle wird davon ausgegangen, dass die Einbaulage der Plattformbasis (CameraSocket) und des Kameraträgers (CameraCarrier) bereits mit hoher Genauigkeit in allen 6 Freiheitsgraden relativ zum Fahrzeugaufbaukoordinatensystem (EgoBody) vermessen wurde. Das heißt, es sind nun lediglich noch die Winkellagen der Kamera zu bestimmen. Zunächst wird der Rollwinkel gegenüber dem Horizont zu Null gedreht. Anschließend wird ein Testobjekt, dessen Lage relativ zum Fahrzeug mittels Laserentfernungsmesser exakt vermessen wird, an der vermessenen Position in den Szenenbaum aufgenommen und in das Kamerabild projiziert. Nun werden der Gier- und Nickwinkel derart angepasst, dass das visualisierte Testobjekt und das real abgebildete Testobjekt deckungsgleich werden. In Abbildung 3.5 ist dies beispielhaft für zwei Plattformstellungen dargestellt. Das grüne Quadrat stellt dabei die Projektion unter Verwendung des Kameramodells mit Berücksichtigung der Linsenverzerrungen und das rote Quadrat unter Verwendung des Lochkameramodells dar. Es ist zu erkennen, dass sich in den Randbereichen durchaus deutliche Abweichungen zwischen den Modellen zeigen.



**Abbildung 3.5:** Bestimmung der Einbaulage einer Kamera einer komplexen Kameraanordnung durch Anpassung des Gier- und Nickwinkels, indem das reale und das visualisierte Testobjekt zur Deckung gebracht werden: a) zentrale Ansicht, b) Randbereich

### 3.3.4 Kamerakonfigurationen

#### 3.3.4.1 Monokulares visuelles Sensorsystem

Das Ziel bei der Auslegung einer visuellen Sensorkonfiguration besteht darin, die Auflösbarkeit eines Merkmals einer bestimmten Größe in einem vorgegebenen Entfer-

nungsbereich und innerhalb eines gewünschten Sichtbereichs (Öffnungswinkel) zu gewährleisten. Ob ein Merkmal in einer bestimmten Entfernung zu detektieren (d.h. auflösen) ist, hängt von dem für die Detektion verwendeten Bildverarbeitungsverfahren ab. Dabei ist speziell die minimale Anzahl an Pixeln von Bedeutung, die ein bestimmter Bildverarbeitungsoperator für die Erzeugung einer Maskenantwort mit ausreichendem Signal-Rausch-Abstand benötigt. Je größer der Öffnungswinkel einer Kamera-Objektiv-Kombination, desto größer wird der Raumwinkelbereich, der auf ein einzelnes Pixel des Sensorchips abgebildet wird. Das heißt, der Entfernungsbereich, in dem ein bestimmtes Merkmal gerade noch aufgelöst werden kann, wird verkürzt, da das Merkmal auf eine geringere Anzahl an Pixeln abgebildet wird. Der Öffnungswinkel einer Kamera-Objektiv-Kombination hängt zum einen von der Chipgröße des lichtempfindlichen Sensors und zum anderen von der Brennweite des verwendeten Objektivs ab. Der horizontale bzw. vertikale Öffnungswinkel in [Grad] berechnet sich aus

$$\begin{aligned}\alpha_{hor} &= \frac{180^\circ}{\pi} \cdot \arctan\left(\frac{0.5 \cdot s_{hor}}{f}\right) \\ \alpha_{ver} &= \frac{180^\circ}{\pi} \cdot \arctan\left(\frac{0.5 \cdot s_{ver}}{f}\right)\end{aligned}\quad (3.10)$$

mit der Brennweite  $f$  und der horizontalen bzw. vertikalen Chipgröße  $s_{hor}$  bzw.  $s_{ver}$ . Möchte man bei vergrößertem Öffnungswinkel den Entfernungsbereich, in dem ein bestimmtes Merkmal aufgelöst werden kann, konstant halten, dann muss die Anzahl der Pixel auf dem Sensorchip entsprechend erhöht werden.

### 3.3.4.2 Bifokale Kameraanordnung

Alternativ können auch mehrere Kameras mit unterschiedlichen Objektiven und damit Öffnungswinkeln eingesetzt und in Abhängigkeit vom Entfernungsbereich die jeweilige Kamera ausgewertet werden. In Kapitel 6 wird hierfür eine bifokale Kameraanordnung ähnlich [THOMANEK 96] verwendet. Im Nahbereich wird eine Kamera mit einem horizontalen Öffnungswinkel von  $\pm 22^\circ$  und im Fernbereich werden die Bilder einer Kamera mit Öffnungswinkel  $\pm 5.5^\circ$  benutzt.

### 3.3.4.3 Parallele Stereoanordnung

Ist man daran interessiert, aus den Videobildern explizit Entfernungswerte zu berechnen, so werden häufig zwei identische Kameras mit parallelen optischen Achsen und kolinearen Bildzeilen im Abstand  $b$ , der *Stereobasis*, angeordnet. Durch Triangulation einander zugeordneter Bildbereiche bzw. Schnittpunktberechnung der korrespondierenden Sehstrahlen im linken und rechten Kamerabild kann daraus ein Entfernungswert bestimmt werden.

#### 3.3.4.4 Sensorkonzept MarVEye

Ein nach Entfernung- und Sichtbereichen noch stärker differenzierendes Sensorkonzept ist das MARVEYE-Sensorkonzept (Multi-focal active/reactive Vehicle Eye) nach [DICKMANN 95], welches zum einen für den Nahbereich ein divergent angeordnetes Kamerapaar mit einem großen Öffnungswinkel und Überlappungsbereich für eine Nahbereichsstereoauswertung besitzt, und zum anderen für mittlere und große Entfernungen jeweils Kameras mit den entsprechenden Objektivbrennweiten vorsieht. Ein Grundgedanke für den flexiblen Einsatz der Kameraanordnung besteht in der Kombination mit einer Kameraplattform für aktive Blickrichtungssteuerung, so dass ein relevanter 3-D Bereich jeweils mit derjenigen Kamera betrachtet wird, deren Auflösung für den entsprechenden Entfernungsbereich gerade am geeignetsten ist.

### 3.4 Kameraplattformen der UniBwM für aktive Blickrichtungssteuerung

#### 3.4.1 Ein-Achsen-Kameraplattform im Versuchsfahrzeug VaMP

Ein Beispiel einer Ein-Achsen-Kameraplattform ist in Kapitel 6 in Abbildung 6.3b dargestellt. Auf dieser Plattform sind vier Kameras zur Umfeldwahrnehmung befestigt. Die Plattform ist in der Lage, schnelle Blickzuwendungen - sogenannte *Sakkaden* - durchzuführen. Ausgehend von der Nullstellung, welche parallel zur Fahrzeuglängsachse verläuft, kann die Plattform dabei um bis zu  $\pm 40^\circ$  nach links und nach rechts gegiert werden.

#### 3.4.2 Zwei-Achsen-Kameraplattform im Versuchsfahrzeug VaMoRs

In Kapitel 8 ist in Abbildung 8.2 eine Zwei-Achsen-Kameraplattform dargestellt. Diese kann um mehr als  $\pm 90^\circ$  gieren und ist im Vergleich zur Ein-Achsen-Kameraplattform zusätzlich in der Lage, den Nickwinkel entsprechend einem vorgegebenen Sollwinkel einzustellen. Damit ist es möglich, beispielsweise ein bestimmtes Objekt im Raum zu fixieren oder zwischen verschiedenen Objekten durch schnelle Blickzuwendungen zu wechseln. Die Zwei-Achsen-Plattform wird in Kapitel 8 unter Verwendung aktiver Blickrichtungssteuerung zur Vermessung eines negativen Hindernisses mit anschließender Blickfixierung während eines Ausweichmanövers eingesetzt. Im Abschnitt 8.5.1 wird das Ausweichen an einem Graben anhand einer Bildfolge erläutert.

# Kapitel 4

## Extraktion von Bildmerkmalen aus Grauwertbildern

In diesem Kapitel werden die hier zur Extraktion von Bildmerkmalen verwendeten Verfahren vorgestellt. Aufgrund der Fülle der in der Literatur beschriebenen Verfahren wird auf eine detaillierte Darstellung anderer Verfahren zur Bildauswertung verzichtet - dies würde den Rahmen dieser Arbeit sprengen. Der interessierte Leser kann sich z.B. mittels [BÄSSMANN 89], [SCHALKOFF 89] oder [BALLARDBROWN 82] in großer Ausführlichkeit einen ersten Überblick verschaffen.

Für die Bildaufnahme wurden Schwarz-Weiß-Kameras verwendet, welche nach dem Halbbildverfahren arbeiten. Ein Vollbild setzt sich dabei aus zwei durch Unterabtastung mit dem Faktor zwei in Zeilenrichtung aus dem Vollbild entstandenen, um eine Zeile versetzten, zeitlich nacheinander (alle  $20ms$ ) aufgenommenen Halbbildern (engl. *frames*) zusammen. Es liegen demnach alle  $40ms$  - dem Videotakt - zwei Frames eines Vollbildes vor. Das Vollbild kann daraus durch Verschränkung der beiden Halbbilder gebildet werden, indem abwechselnd eine Zeile des ersten, dann des zweiten Halbbildes aneinander gefügt werden. Bei schnellen Kamera- oder Objektbewegungen führt das dazu, dass vertikale Kanten ausgefranst erscheinen, da die beiden Halbbilder ja zu unterschiedlichen Zeitpunkten aufgenommen wurden. Unter Berücksichtigung des Zeilenversatzes könnte alle  $20ms$  jeder Frame einzeln ausgewertet werden. Nachdem der zeitliche Aufwand für den Bildeinzug eines Halbbildes unabhängig vom Videotakt konstant ist, verbleibt nach dem Bildeinzug nur ein vergleichsweise kleines Zeitfenster, das für Berechnungen zur Verfügung steht. Aus diesem Grund sei hier das Eingangssignal für die Verfahren zur Bildauswertung das erste Halbbild oder das aus diesem Halbbild durch Unterabtastung mit dem Faktor zwei in Spaltenrichtung entstandene Viertelbild. Das jeweils zweite Halbbild wird nicht eingelesen. Die Visualisierung eines Halbbilds ist für den menschlichen Betrachter ungewohnt, da alle Objekte in horizontaler Richtung um den Faktor zwei gestreckt erscheinen.

Die Auswertung des Bildes soll zielgerichtet oder erwartungsbasiert erfolgen. Das heißt, ein Bildoperator wird für die Extraktion eines bestimmten Bildmerkmals einer bestimmten

Klasse und Ausdehnung innerhalb eines Suchbereichs angepasst. Das Ziel besteht darin, die genaue Position und je nach Merkmalart beschreibende Parameter zu extrahieren. Wichtige Klassen von Merkmalen eines Grauwertbildes sind z.B.

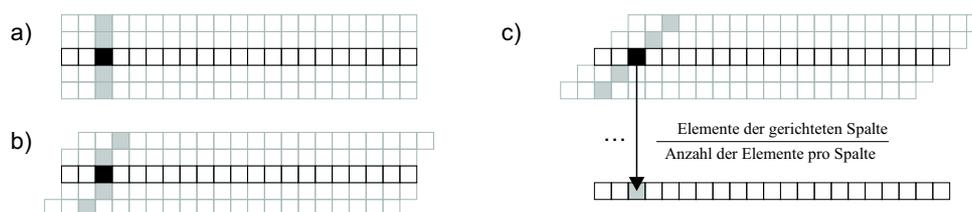
- starke Helligkeitsübergänge oder Diskontinuitäten, sogenannte *Kanten*,
- gewinkelte 2-D Helligkeitsübergänge, sogenannte *Ecken*,
- gleichmäßige kontinuierliche Helligkeitsverläufe, sogenannte *homogene Segmente*, aus denen durch Verknüpfung *homogene Bereiche* gebildet werden können,
- spezielle typische Helligkeitsverteilungen, sogenannte *Muster* (engl. *pattern*)
- und Bereiche mit einer sich wiederholenden Struktur, sogenannte *Texturen*.

Ein Bildverarbeitungsoperator ist in aller Regel sehr speziell für die Detektion bestimmter Merkmalklassen konzipiert. In dieser Arbeit werden keine Eckenoperatoren eingesetzt und es werden aufgrund der hohen Anforderungen an die Rechenleistung Texturen als Bildmerkmal nicht weiter betrachtet. Entweder sei die Struktur einer Textur so fein, dass sie als verrauschter homogener Bereich interpretiert werden kann, oder sie sei so grob, dass die Strukturen wiederum in homogene Segmente und begrenzende Kanten zerlegt werden können.

Die Parametrisierung eines Bildverarbeitungsoperators wird durch das gesuchte Merkmal im 3-D Raum festgelegt. Das Zentrum des Suchbereichs resultiert aus der Projektion in die Bildebene. Die Auswertung erfolgt dann entlang eines Suchpfads. Die Suchpfadrichtung darf im Prinzip beliebig gewählt werden. Sie wird häufig in 3-D definiert und durch Projektion ins Bild in 2-D festgelegt. Sinnvollerweise wird je nach Merkmalart auf die Besonderheiten des gesuchten Merkmals eingegangen. Die Suche nach einer Kantenposition erfolgt z.B. am besten orthogonal zu einer erwarteten Kantenrichtung. Aus Effizienzgründen werden häufig Suchpfadrichtungen parallel zu den Zeilen oder Spalten eines Bildes verwendet. Die Ausdehnung des ins Bild projizierten erwarteten Merkmals dient als Anhaltspunkt dafür, wie groß eine Suchmaske am besten gewählt werden sollte. In der Regel sind aus Verfahrens- oder Rechenzeitgründen sinnvolle Ober- und Untergrenzen definiert.

Ein Bildoperator wertet demnach das Bild entlang eines *Bildstreifens* aus, dessen Anfang und Ende bzw. Zentrum, Suchpfadrichtung und positive bzw. negative Toleranz in bzw. entgegen der Suchpfadrichtung durch den Suchpfad festgelegt wird und dessen Breite sich durch die Größe der Suchmaske ergibt. Bei der Extraktion von Kanten oder der Bestimmung homogener Segmente, deren Grenzen durch Kanten begründet sind, kann zwischen *richtungsselektiver* und *richtungssensitiver* Suche entlang des Bildstreifens unterschieden werden.

Bei der *richtungsselektiven* Suche wird der Bildstreifen im Prinzip zunächst einem Vorverarbeitungsschritt unterzogen. Durch gerichtete Mittelung entsprechend einer erwarteten Kantenrichtung wird wie in Abbildung 4.1 ein eindimensionaler gemittelter Helligkeitsverlauf erzeugt, der dann anschließend durch einen eindimensionalen Bildoperator bearbeitet wird.



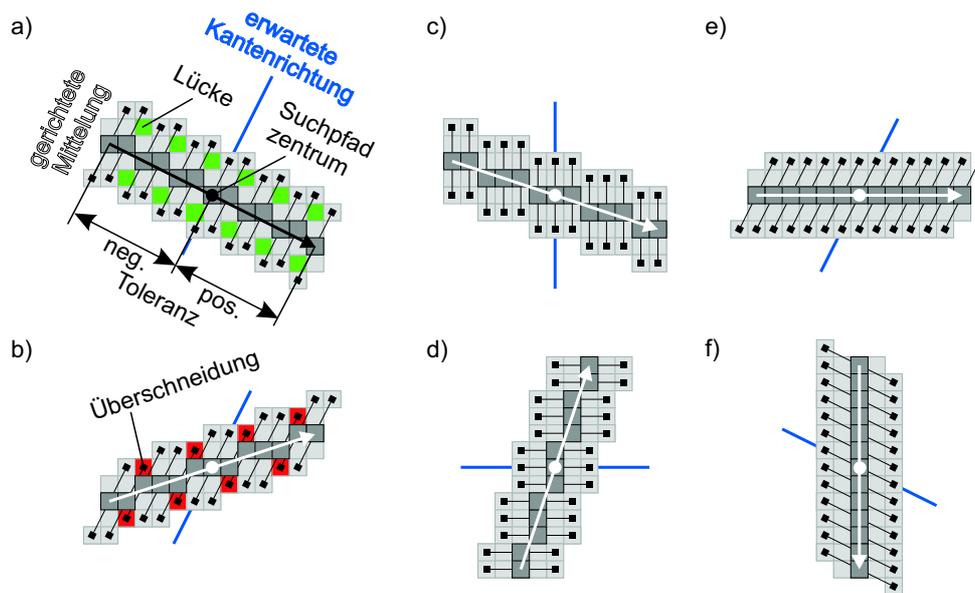
**Abbildung 4.1:** Erzeugung einer eindimensionalen Zahlenfolge aus einem Bildstreifen durch Mittelung über gerichtete Spalten z.B. a)  $90^\circ$  b)  $-67,5^\circ$  c)  $-45^\circ$

Helligkeitsübergänge bzw. Diskontinuitäten im Bildstreifen, deren Richtungen nicht der gerichteten Mittelung entsprechen, werden dabei abgeschwächt d.h. verwaschen, so dass diejenigen Übergänge, deren Kantenrichtung der Erwartung ähneln, im Verhältnis verstärkt werden. Diese gerichtete Tiefpassfilterung kann aber nicht verhindern, dass dennoch ein in Suchpfadrichtung sehr starker Helligkeitsübergang mit gegenüber der Erwartung abweichender Kantenrichtung im Vergleich zu einem schwachen Helligkeitsübergang mit erwarteter Kantenrichtung eine stärkere Maskenantwort bei der anschließenden Auswertung des eindimensionalen Ergebnisvektors hervorrufen und damit zu einer Fehlzuordnung führen kann. Ein Rückschluss auf die tatsächliche Kantenrichtung im Bildstreifen ist anhand des gemittelten Helligkeitsverlaufs nicht mehr möglich.

In Abbildung 4.2 sind alle möglichen Fälle zur gerichteten Mittelung dargestellt. Abbildung 4.2a zeigt die gerichtete Mittelung bei näherungsweise orthogonaler Suchpfadrichtung gegenüber einer erwarteten Kantenrichtung. In diesem Fall werden vereinzelte Pixel (*Lücken*) innerhalb des Bildstreifens bei der Auswertung nicht berücksichtigt. Dagegen kommt es in 4.2b zu *Überschneidungen* bei der Mittelung. Das heißt, es existieren Pixel, die in mehr als eine Mittelung eingehen und damit ein höheres Gewicht als andere Pixel besitzen. Um dies zu vermeiden, wird häufig entweder die gerichtete Mittelung parallel zu den Zeilen bzw. Spalten des Bildes durchgeführt (siehe Abb. 4.2c und 4.2d) oder die Suchpfade werden so gewählt, dass diese parallel zu den Zeilen bzw. Spalten des Bildes verlaufen (siehe Abb. 4.2e und 4.2f).

Unter Verwendung eines durch gerichtete Mittelung erzeugten eindimensionalen Helligkeitsverlaufs als Eingangssignal wird in Abschnitt 4.1 die Kantensuche mittels ternärer Masken erläutert. Daran anschließend wird in Abschnitt 4.2 ein neu entwickeltes hierarchisches Verfahren zur Segmentierung eines eindimensionalen Helligkeitsverlaufs in homogene Bereiche durch den Aufbau einer Segmentierungspyramide vorgestellt.

Bei der *richtungssensitiven* Auswertung eines Bildstreifens wird auf den richtungsselektiven Vorverarbeitungsschritt bewusst verzichtet. Mittels orthogonaler Differenzmasken wird an Helligkeitsübergängen die Kantenrichtung als kontinuierlicher Winkel im Intervall  $]-\pi; +\pi]$  bestimmt. Speziell bei der Ermittlung der Grenzen homogener Segmente kann nämlich nicht davon ausgegangen werden, dass alle Segmentgrenzen entlang eines Suchpfades kollinear zu *einer* erwarteten Kantenrichtung sind. Durch eine gerichtete Tiefpassfilterung würde das Eingangssignal im Extremfall so stark verfälscht, dass bestimmte



**Abbildung 4.2:** Gerichtetes Mittelung entsprechend erwarteter Kantenrichtung für richtungsselektive Kantensuche

Segmentgrenzen überhaupt nicht mehr gefunden werden könnten. Dies war der Grund dafür, neben dem richtungsselektiven hierarchischen Verfahren aus Abschnitt 4.2 ein weiteres Verfahren zur Bestimmung homogener Segmente zu entwickeln. Dieses wird in Abschnitt 4.3 eingehend erläutert und anhand von Beispielen veranschaulicht.

Homogenität sei hierbei so definiert, dass der Verlauf der Helligkeit entlang eines Suchpfads innerhalb eines Segments keine signifikante Diskontinuität aufweist. Die beiden Verfahren zur Segmentierung wurden so konzipiert, dass aus den Ergebnissen bereits durchgeführter Segmentierungen aufgrund eines hierarchischen Aufbaus bzw. durch Rekursion die beschreibenden Parameter neuer Segmentierungen gebildet werden können, ohne auf die Originaldaten ein zweites Mal zugreifen zu müssen. Dadurch wird die Bildauswertung sehr effizient.

Die zugrundeliegende relativ einfache Mathematik wird im Anhang A erläutert. Das Kapitel A.1 zeigt z.B. allgemein eine Möglichkeit, wie für ein überbestimmtes lineares Gleichungssystem durch Anwendung der Moore-Penrose-Pseudoinversen eine optimale Lösung bezüglich der kleinsten Summe der Fehlerquadrate gefunden werden kann. In den Abschnitten A.2 und A.3 wird die Theorie dann dazu verwendet, für eine Zahlenfolge eine lineare Näherung, d.h. eine Ausgleichsgerade, und als Maß für die Qualität der Näherung die Varianz der Abweichung der Originaldaten von einer Ausgleichsgeraden zu bestimmen. In den Abschnitten A.4.1, A.4.2 und A.4.3 wird beschrieben, wie die Ergebnisse zum Aufbau einer Segmentierungspyramide bzw. zur rekursiven Bestimmung einer Ausgleichsgeraden benutzt werden können. Abschnitt A.5 zeigt als Erweiterung zu Abschnitt A.1, wie zusätzliche Zwangsbedingungen bei der Berechnung einer Näherungslösung berück-

sichtigt werden können. Das beschriebene Beispiel wird später bei der Erkennung eines rechteckigen Grabens zur Trennung relevanter Merkmale von irrelevanten Bildmerkmalen zum Einsatz kommen. Anschließend wird in Abschnitt A.6 eine Polynominterpolation zur Bestimmung einer Merkmalsposition auf Subpixelgenauigkeit beschrieben, gefolgt von grundsätzlichen mathematischen Überlegungen zur Approximation des Helligkeitsgebirges eines rechteckigen Bildausschnitts durch eine Ausgleichsebene im Abschnitt A.7.

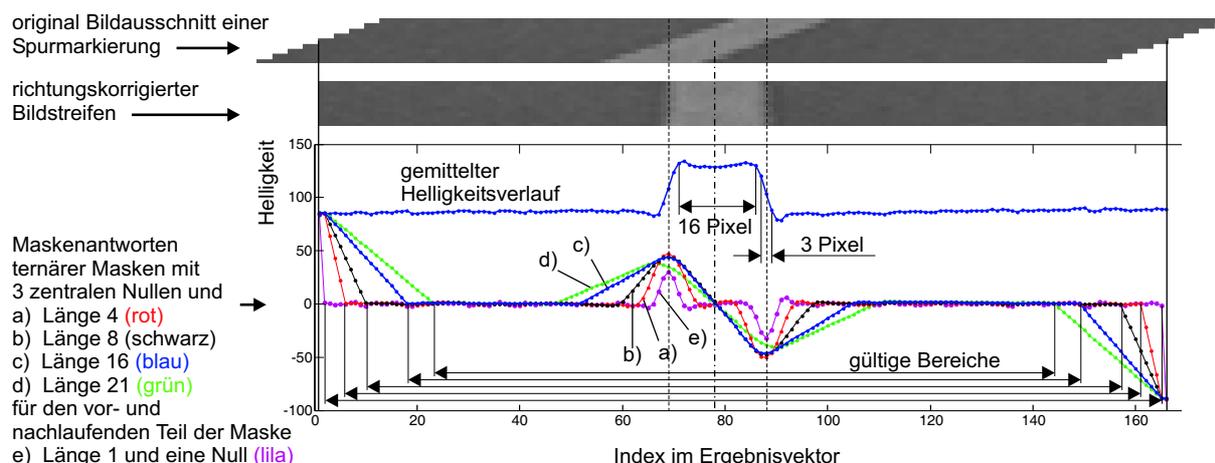
In Abschnitt 4.4 wird die Erzeugung generischer Merkmalmuster am Beispiel eines Rades und die Suche nach dem Muster entlang eines Suchpfads im Bild erläutert. Der Rad-detektor führt dabei eine Maximumsuche der Korrelation in (tendenziell) horizontaler Suchpfadrichtung und anschließend in vertikaler Richtung durch. Solch ein Operator wird hier als *Komplexoperator* bezeichnet, da er z.B. für das komplexe Merkmal „Rad“, d.h. für ein komplettes Subobjekt, die Positionsbestimmung im Suchbereich durchführt.

Zuletzt wird in Abschnitt 4.5 noch kurz ein Verfahren nach [THOMANEK 96], der sogenannte *Betragsoperator*, zur Extraktion vertikaler Häufungen von unspezifischen Kantenübergängen beschrieben.

## 4.1 Richtungsselektive Kantensuche mittels ternärer Masken

Als Eingangssignal dient ein mittels gerichteter Mittelung generierter eindimensionaler Helligkeitsverlauf, wie z.B. in Abbildung 4.3 dargestellt. Dieser Helligkeitsverlauf wird nun bezüglich starker Helligkeitsübergänge bzw. Diskontinuitäten im Helligkeitsverlauf untersucht. Hierfür wird die Korrelation zwischen dem Eingangssignal und einer speziellen Maske - einer sogenannten *ternären Maske* - berechnet. In der Regel wird eine ternäre Maske symmetrisch aufgebaut und besteht aus drei Abschnitten. Einem mit minus Einsen besetzten nachlaufenden Teil, einem vorlaufenden Teil mit plus Einsen und einem zentralen Teil, der mit Nullen belegt ist. Durch den symmetrischen Aufbau ergibt die Korrelation der Maske mit einem konstanten Helligkeitsverlauf eine Maskenantwort null. Die Längen des vorlaufenden, nachlaufenden und zentralen Teils der Maske werden im Idealfall an den Verlauf eines erwarteten Helligkeitsübergangs angepasst. Die Breite des zentralen Teils wird dabei am besten kleiner gleich der Anzahl der Pixel auf der erwarteten steilen Flanke (ohne Anfangs- und Endpunkt) des gesuchten Helligkeitsübergangs gewählt. Der vorlaufende Teil sollte nicht länger als die erwartete Ausdehnung des Merkmals in Suchpfadrichtung sein und die Länge des nachlaufenden Teils wird gleich der des vorlaufenden Teils gewählt.

Für das Beispiel aus Abbildung 4.3 bedeutet das, dass die für die Detektion der abgebildeten Spurmarkierung ideale ternäre Maske im zentralen Bereich drei Nullspalten besitzt und einen vor- und nachlaufenden Teil der Länge 16. Die auf die Länge des vor- bzw. nachlaufenden Teils normierte Maskenantwort der Korrelation dieser Maske mit dem



**Abbildung 4.3:** Maskenantwort durch Korrelation ternärer Masken mit gemitteltem Helligkeitsverlauf

Helligkeitsverlauf ist in der Abbildung 4.3 unter c) blau dargestellt. Der für eine Extremwertbestimmung gültige Bereich ist jeweils markiert. An der Position des Maximums hat der Helligkeitsverlauf seine größte ansteigende Flanke bzw. an der Position des Minimums seine größte fallende Flanke.

Die Maskenantworten a), b) und e) haben ihrer Extremwerte ebenfalls an den gesuchten Positionen. Anhand ihrer Verläufe können noch ein paar grundsätzliche Zusammenhänge erklärt werden. Je kürzer die Länge des vor- bzw. nachlaufenden Teils der Maske ist, desto unruhiger wird aufgrund der geringeren Mittelungstiefe der Verlauf in den eher homogenen Bereichen. In den Fällen a) und e) führt z.B. der Bereich des hellen Plateaus zu einer Maskenantwort nahe 0, wohingegen dies bei den Fällen b), c) und d) nicht beobachtet werden kann.

Im Fall d) ist die Länge des vor- bzw. nachlaufenden Teils größer als die Ausdehnung des Merkmals. Dies führt dazu, dass die Positionen der Flanken nicht mehr exakt bestimmt werden können, d.h. die Spurmarkierung nicht mehr korrekt aufgelöst werden kann. Die Positionen der Extremwerte wandern von der Symmetrieachse aus nach außen.

Im Fall e) ist zu erkennen, dass bei einer Gesamtlänge der ternären Maske (hier gleich 3 wegen  $[-1, 0, 1]$ ) kleiner der Länge der Flanke einschließlich Anfangs- und Endpunkt (hier 5 Pixel) die Amplitude der Maskenantwort sinkt und damit der Signal-Rausch-Abstand im Vergleich zu den anderen Fällen geringer wird.

Die Entwicklungsgeschichte der richtungselektiven Verfahren zur Kantenextraktion an der UniBwM bis hin zum Bildverarbeitungspaket *CRONOS* (dem bisherigen Standard am ISF) begann mit dem Konzept für eine effiziente Verwendung richtungselektiver ternärer Masken zur Kantenextraktion für konventionelle Mikroprozessoren [KUHNERT 88]. Eine echtzeitfähige Implementierung wurde durch [MYSLIWETZ 90] erstellt und auf natürlichen

Szenen zur Fahrspurerkennung eingesetzt. Im Speziellen wurden zusätzlich ternäre Masken mit alternativen Richtungen für die Kantensuche eingeführt, die um eine Winkelquantisierungsstufe bezüglich der erwarteten Kantenrichtung variierten und zur Nachführung der gerichteten Tiefpassfilterung entsprechend dem Kantenverlauf im Bild dienten. Die entwickelten Verfahren wurden in verallgemeinerter Form für transputerbasierte Systeme in das OCCAM-basierte Bildverarbeitungspaket *KRONOS* integriert [DICKMANN 94] und anschließend für PC-basierte Systeme nach C in das Bildverarbeitungspaket *CRONOS* portiert [FÜRST 98-01].

## 4.2 Extraktion von Segmenten homogener Helligkeit durch hierarchische Segmentierung

Das hierarchische Segmentierungsverfahren ist ein neues Verfahren zur Segmentierung einer eindimensionalen Zahlenfolge. Diese könnte z.B. durch eine gerichtete Mittelung eines Bildstreifens, wie in Abbildung 4.1 dargestellt, entstanden sein. Die Segmentierung erfolgt durch eine paarweise Verknüpfung abschnittsweise homogener benachbarter Daten. Hierdurch entsteht eine pyramidenartige Struktur. Über ein noch zu definierendes Kriterium der Homogenität eines Segments wird hier quasi vom Innern eines homogenen Bereichs ausgehend zu den Rändern hin durch Verknüpfung benachbarter Daten nach den Grenzen eines Segments gesucht.

### 4.2.1 Funktionsweise des hierarchischen Segmentierungsverfahrens

Das hierarchische Segmentierungsverfahren nutzt die Möglichkeit zur Bestimmung von Ausgleichsgeraden durch Verknüpfung von Zahlenfolgen entsprechend Abschnitt A.4.1 und A.4.2. Beim Aufbau einer Segmentierungspyramide stellen hierbei die Originaldaten der Eingangszahlenfolge die unterste Ebene (Hierarchieebene 1) dar. Einträge in einer darüberliegenden Hierarchieebene ergeben sich durch paarweises Zusammenfassen darunterliegender Einträge. Jeder Eintrag hat eine Datenstruktur entsprechend Tabelle 4.1. Einträge der Ebene 1 und Ebene 2 sind stets *homogen*, wobei der Begriff homogen hier derart definiert sei, dass die Varianz bzw. Standardabweichung der Differenz zwischen den Originaldaten und einer Ausgleichsgeraden kleiner als eine vorgegebene Schwelle ist, d.h.

$$\sqrt{Var} < Schwelle. \quad (4.1)$$

Diese Schwelle kann man sich anschaulich als eine Art Epsilon-Schlauch der erlaubten Abweichung um eine Ausgleichsgerade vorstellen. Die Varianzen der Einträge auf Ebene 1 und 2 sind stets null, da auf Ebene 1 die Größe der Segmente  $n = 1$  und auf Ebene 2 die Größe  $n = 2$  ist. Somit kann erst ab Ebene 2 eine Gerade gebildet werden. Der Wert der Steigung  $a$  eines Eintrags auf Ebene 1 sei null.

**Tabelle 4.1:** Datenstruktur eines Eintrags in der Segmentierungspyramide

<i>sy</i>	Summe der Grauwerte
<i>sy<sup>2</sup></i>	Summe der Quadrate der Grauwerte
Anfang	Anfangsindex in darunterliegender Ebene
Ende	Endindex in darunterliegender Ebene
Anfang1	Anfangsindex in unterster Ebene
Ende1	Endindex in unterster Ebene
min	Minimalwert
max	Maximalwert
<i>a</i>	Steigung <i>a</i> der Ausgleichsgeraden
<i>b</i>	Mittelwert <i>b</i> der Ausgleichsgeraden
<i>Var</i>	Varianz der Differenz zwischen den Daten und der Ausgleichsgeraden
<i>n</i>	Größe des Segments
flag	homogen=1, inhomogen=4

**Schritt 1:** Aufbau einer Segmentierungspyramide entsprechend Abschnitt A.4.2

Es werden jeweils zwei benachbarte Einträge einer Hierarchieebene zu einem Eintrag in der nächst höheren Ebene zusammengefasst. Jedem Eintrag der Hierarchieebene *i* werden demnach

$$n = 2^{(i-1)} \quad (4.2)$$

Originaldaten der Eingangszahlenfolge zugeordnet. Alle in der Hierarchie unterhalb eines homogenen Bereichs liegenden Einträge erfüllen die Bedingung (4.1). Neben der Steigung *a*, dem Mittelwert *b* und der Größe *n* einer Ausgleichsgeraden werden je Segment die Varianz *Var* der Abweichung bezüglich der Ausgleichsgeraden, der Minimalwert und der Maximalwert bestimmt. Sobald für eine Verknüpfung einer der beiden oder beide unterlagerten Bereiche die Bedingung (4.1) nicht mehr erfüllen, kann die Verknüpfung nur inhomogen sein. Das heißt, kein darüberliegender Bereich kann noch homogen werden. Darum werden diese zur Einsparung von Rechenzeit automatisch als inhomogen '4' gekennzeichnet. Die sukzessive Verknüpfung der Segmente wird solange durchgeführt, bis in der aktuell obersten Ebene kein homogenes Segment mehr gebildet werden kann. Tabelle 4.2 zeigt beispielhaft eine durch diese Vorgehensweise entstandene Struktur der Segmentierungspyramide nach Schritt 1. Die '.'-Punkte in der Tabelle stehen für homogene Teilsegmente, welche zu einem in der Hierarchie höher gelegenen, größeren Segment zusammengefasst werden konnten. In Abbildung 4.4 sind für ein Beispiel die Originaldaten mit dem Ergebnis der Segmentierung nach Schritt 1 überlagert. Es konnten acht Segmente gebildet werden.

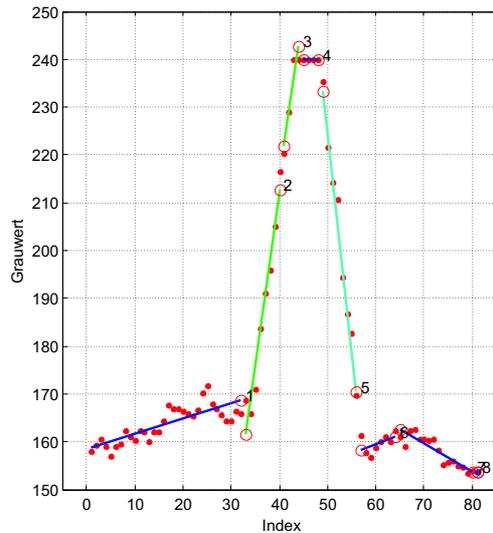
**Schritt 2:** Auslesen der homogenen Segmente der Segmentierungspyramide

Im Schritt 2 wird das Resultat der Pyramidenbildung ausgelesen, um dann in Schritt

**Tabelle 4.2:** Beispiel für die Struktur einer Segmentierungspyramide mit 1=homogen, 4=inhomogen und '.'=unterlagert homogen

```

.....1
. . . . . 1 1 . . . . .
. . . . . 1 4 1 1 . . . .
1 . . . . 4 4 4 1 . . . .
4 . . . . 4 . . . . 4
    
```



**Abbildung 4.4:** Ergebnis der Segmentierung nach Aufbau der Segmentierungspyramide

3 weiterverarbeitet zu werden. Das Auslesen erfolgt, indem beginnend mit der 1. Spalte der Pyramide die Hierarchieebene bestimmt wird, in der das größte homogene Segment, gekennzeichnet durch eine '1', zu finden ist. Der Index  $i$  der Hierarchieebene legt nach Gleichung (4.2) fest, in welcher Spalte relativ zur aktuellen Position nach dem nächsten Segment zu suchen ist. Dieser Vorgang ist in Abbildung 4.5 dargestellt. In dem Beispiel findet sich der Eintrag für das größte Segment der Spalte 1 auf Ebene 6, somit ist das nächste Segment in Spalte  $33 = 1 + 2^{(6-1)}$  und das übernächste Segment in Spalte  $41 = 33 + 2^{(4-1)}$  usw. zu finden.

**Schritt 3:** Verknüpfen benachbarter Segmente entsprechend Abschnitt A.4.1

Wird hier als Ziel die Bestimmung von Segmenten verfolgt, die in ihrer Gesamtheit linear homogen sind, dann werden die in Schritt 2 ausgelesenen benachbarten Segmente von links nach rechts beginnend sukzessive paarweise verknüpft und mittels Kriterium (4.1) auf Homogenität überprüft. Ist dieses erfüllt, dann wird das Segment verlängert und die Verknüpfung mit dem nächsten Segment durchgeführt usw..

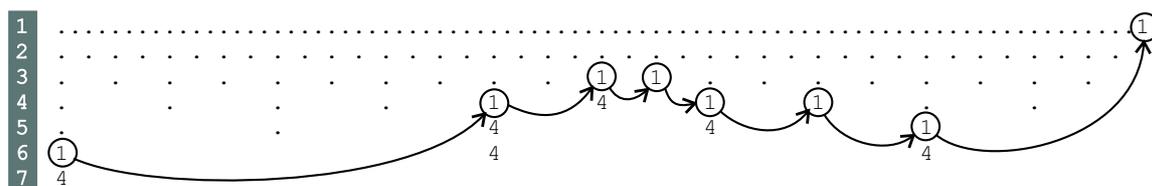


Abbildung 4.5: Auslesen der Segmentierungspyramide

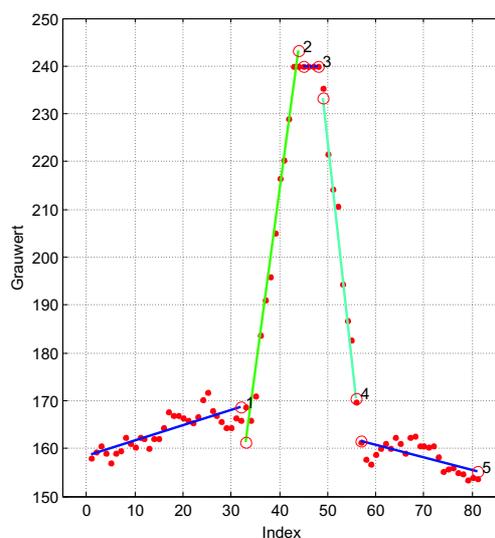
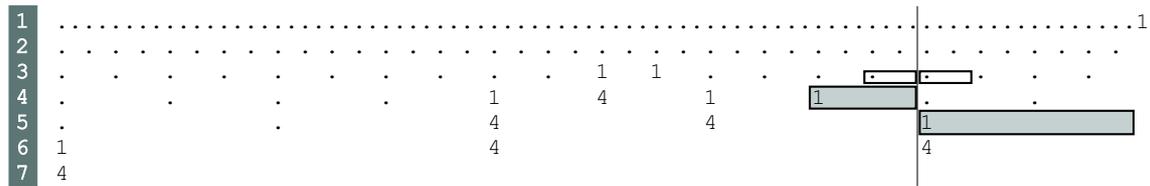


Abbildung 4.6: Ergebnis der Segmentierung nach Verknüpfung benachbarter Segmente

Sobald das Resultat die Homogenitätsschwelle überschreitet, wird das vorletzte Ergebnis der Verknüpfung gespeichert und mit dem aktuellen Segment die paarweise Verknüpfung neu begonnen. Abbildung 4.6 zeigt am Beispiel die aus Schritt 3 resultierenden fünf Segmente.

Reicht es dagegen aus, bei der Segmentierung einen kontinuierlichen Verlauf der Originaldaten sicherzustellen und wird diese Kontinuität als 'lokale' Homogenität definiert, dann kann z.B. die Homogenität einer Verknüpfung untergeordneter Teilstimente der Pyramide als Kriterium für eine sukzessive Verknüpfung der Segmente der jeweils obersten Ebenen verwendet werden. Dies wird in Abbildung 4.7 anhand der Segmente 6 und 7 aus Abbildung 4.4 und ihren unterlagerten homogenen Teilstimenten auf Ebene  $l = 3$  verdeutlicht. In diesem einfachen Beispiel ergibt sich als Resultat ebenfalls Abbildung 4.6.

Das Zusammenfügen benachbarter Bereiche auf jeweils oberster Ebene würde demnach solange fortgesetzt, wie die Verknüpfung benachbarter Segmente auf einer vor-



**Abbildung 4.7:** Verknüpfung bei lokaler Homogenität unterlagerter Teilsegmente

gegebenen Ebene  $l$  links und rechts der aktuellen Grenze lokal homogen ist. Wenn eines der oberen Segmente auf einer Hierarchieebene kleiner  $l$  liegt, dann wird der Test der lokalen Homogenität auf der niedrigeren Ebene durchgeführt. Alternativ dazu kann der Winkel zwischen zwei benachbarten untergeordneten Teilsegmenten der Pyramide als Kriterium für eine fortlaufende kontinuierliche Verknüpfung der jeweils obersten Segmente verwendet werden.

Da bisher die Endpunkte der Segmente nicht mit den Anfangspunkten des jeweils nächsten Segments zusammenfallen, haben die Schritte 4 bis 7 zum Ziel aus der Menge der Segmente einen zusammenhängenden Polygonzug zu erzeugen. Hierfür stehen vier verschiedene Methoden zur Verfügung. Diese werden nachfolgend erläutert.

#### Schritt 4: Schnittpunktberechnung zwischen aufeinanderfolgenden Segmenten

Es wird versucht, für zwei aneinandergrenzende Segmente mittels Schnittpunktberechnung einen gemeinsamen End- bzw. Anfangspunkt zu bestimmen. Dies ist in Abbildung 4.8 dargestellt. Liegt ein Schnittpunkt sowohl innerhalb eines Intervalls von  $\pm 0,25 \cdot len_L$  vom Endpunkt des linken Segments entfernt, wobei  $len_L$  die Länge des linken Segments ist, als auch innerhalb eines Intervalls von  $\pm 0,25 \cdot len_R$  vom Anfangspunkt des rechten Segments entfernt, dann sei das *Kriterium der Nähe* erfüllt. Somit wird der Schnittpunkt als gemeinsamer End- bzw. Anfangspunkt akzeptiert. Die Werte für die Mittelwerte  $b$  und die Größen  $n$  werden jeweils angepasst, die Steigungen  $a$  bleiben unverändert.

#### Schritt 5: Vermeidung kurzer Segmente

Findet sich kein geeigneter Schnittpunkt, wird der Abstand zwischen End- und Anfangspunkt aufeinanderfolgender Segmente bestimmt. Ist dieser Abstand kleiner als eine vorgegebene Schwelle, so wird als gemeinsamer End- bzw. Anfangspunkt die Mitte zwischen dem bisherigen End- bzw. Anfangspunkt gewählt. Die Werte für die Steigungen  $a$ , Mittelwerte  $b$  und Größen  $n$  werden jeweils angepasst. Es entsteht eine geringfügige Abweichung gegenüber der bzgl. des quadratischen Fehlers optimalen Lösung. In Abbildung 4.9 ist dies an einem Beispiel illustriert. Die schwarze Ellipse soll hierbei den Schwellenwert für den Abstand symbolisieren.

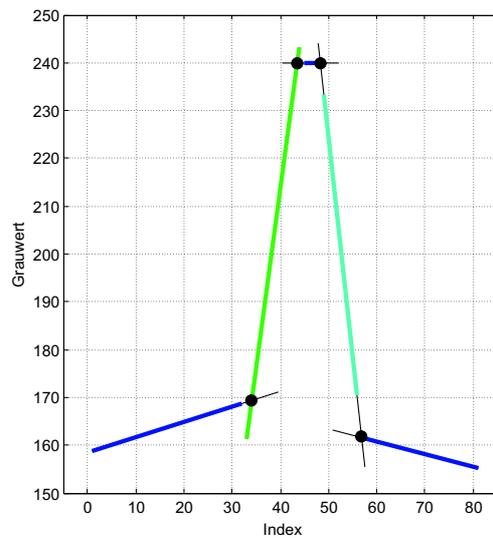


Abbildung 4.8: Schnittpunktberechnung zwischen aufeinanderfolgenden Segmenten

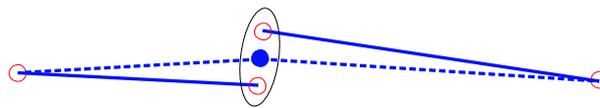


Abbildung 4.9: Vermeidung kurzer Segmente

#### Schritt 6: Einfügen von Verbindungssegmenten

Übersteigt der Abstand die vorgegebene Schwelle wird ein Verbindungssegment eingefügt. Seine Steigung  $a$  wird mittels Steigungsdreieck ermittelt und der Mittelwert  $b$  ergibt sich aus dem Mittelwert des Anfangs- und Endpunkts des Verbindungssegments. Die Größe des Segments ist  $n = 2$ . Siehe Abbildung 4.10.

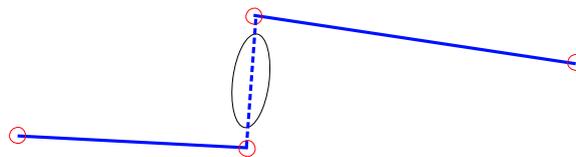
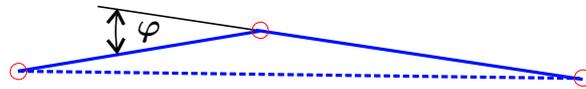


Abbildung 4.10: Einfügen eines Verbindungssegments

#### Schritt 7: Verknüpfung von Segmenten mit kleinem Zwischenwinkel

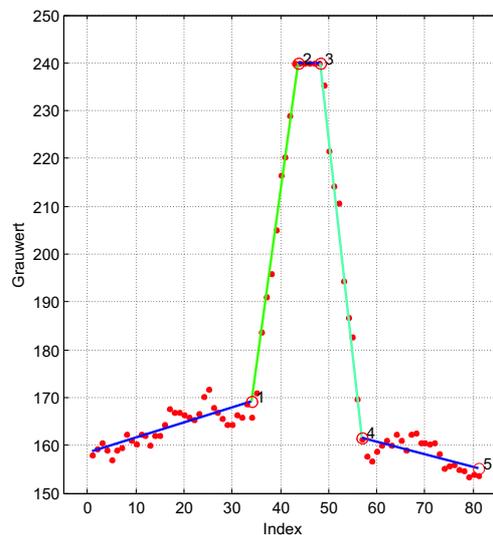
Als weiterer Parameter kann vom Benutzer eine Schwelle für den Betrag des Winkels  $\varphi$  zwischen zwei aufeinanderfolgenden Segmenten angegeben werden. Unterschreitet der Zwischenwinkel die Schwelle für den Winkel, so werden die beiden Segmente

zusammengefasst. Dies geschieht mit der Vorgabe, dass der Anfangspunkt des linken Segments und der Endpunkt des rechten Segments ihre Positionen beibehalten. Siehe Abbildung 4.11.



**Abbildung 4.11:** Verknüpfung von Segmenten mit kleinem Zwischenwinkel

Die Abbildung 4.12 zeigt das Ergebnis der Segmentierung nach dem Durchlaufen aller Verfahrensschritte. Als Segmentierungsparameter wurden eine Standardabweichung von 4, eine Winkelschwelle von  $10^\circ$  und ein minimaler Abstand zur Vermeidung kurzer Segmente von 6 Grauwertstufen vorgegeben.



**Abbildung 4.12:** Ergebnis der Segmentierung nach Durchlaufen aller Verfahrensschritte

#### 4.2.2 Anwendung des Verfahrens zur Bestimmung von Klothoidensegmenten aus dem Krümmungsverlauf von Feldwegen

Das hierarchische Segmentierungsverfahren wurde zwar vornehmlich als Bildverarbeitungsoperator zur Segmentierung der Helligkeitsverläufe von Bildstreifen entwickelt. Allerdings kann das Verfahren prinzipiell für die Segmentierung beliebiger Zahlenfolgen verwendet werden. Beispielsweise wurde es zur Verbesserung und Erweiterung der Wissensbasen zur Missionsplanung in der Dissertation "Fähigkeiten zur Missionsdurchführung

und Landmarkennavigation” [GREGOR 02] zur offline-Aufbereitung der Krümmungsdaten realer Versuchsfahrten auf Feldwegen verwendet. Ziel der Segmentierung war hierbei, den Krümmungsverlauf in Segmente mit konstanter Krümmungsänderung zu unterteilen, daraus den Straßenverlauf nachzubilden und anschließend mit den aus dem Kartenmaterial bekannten Krümmungsverläufen abzugleichen. Nach [GREGOR 02] (Seite 110) lieferte das hierarchische Segmentierungsverfahren sehr gute Ergebnisse und bekam deshalb den Vorzug gegenüber einem in [GREGOR 98] vorgestellten Verfahren.

### 4.2.3 Vor- und Nachteile des hierarchischen Segmentierungsverfahrens

Mit dem hierarchischen Segmentierungsverfahren wurde eine Möglichkeit vorgestellt, Zahlenfolgen wie z.B. Krümmungsverläufe oder Helligkeitsverläufe in homogene Segmente zu unterteilen, wobei als Hauptkriterium für die Homogenität eines Bereichs die Varianz der Abweichung der Originaldaten von der linearen Näherung dient.

Durch die hierarchische Struktur zur paarweisen Verknüpfung benachbarter Bereiche und der konsequenten sukzessiven Wiederverwendung der Ergebnisse unterlagerter Ebenen geht das Verfahren sehr effizient mit den Rechenressourcen um und ist dadurch sehr schnell.

Das Ergebnis der Segmentierung einer Zahlenfolge ist ein durchgehender Polygonzug homogener Segmente. Das Segmentierungsverfahren bildet dabei auch den Verlauf steiler Helligkeitsübergänge nach. Teilweise ist man allerdings nicht an dieser Nachbildung der steilen Übergangsegmente interessiert, sondern lediglich an einer Folge von Segmenten geringer Steigung und unterschiedlicher Helligkeitsniveaus. Dies kann dadurch erreicht werden, dass jedes Segment geringer Steigung um die halbe Größe anschließender steiler Segmente verlängert wird. Die Mitte des steilen Segments wird demnach als Grenze zwischen den unterschiedlichen Niveaus links und rechts des Übergangs verwendet.

Bei wiederholter Anwendung des Verfahrens auf eine Folge von Bildstreifen, die einen größeren Bildbereich überdecken, können sich durch den hierarchischen Aufbau bei kontinuierlich gekrümmten Helligkeitsverläufen teilweise Häufungen von Segmentgrenzen an Stellen entsprechend den Unterteilungen der Hierarchieebenen ergeben. Dies kann durch Verwendung der lokalen Homogenität unterlagerter Teilsegmente als Kriterium für die Verknüpfung benachbarter Segmente verbessert werden.

Der größte Nachteil des hierarchischen Verfahrens ist die fehlende Möglichkeit zur Bestimmung der Kantenrichtung an den Segmentgrenzen. Durch eine Verkettung mehrerer benachbarter Suchstreifen im Bild wäre dies zwar im Prinzip möglich, allerdings wird solch eine Verkettung durch die fehlende Richtungsinformation deutlich erschwert und funktioniert außerdem nur an Kantenverläufen, die sich mindestens über zwei besser drei benachbarte Bildstreifen erstrecken.

Weiterhin werden diejenigen Segmentgrenzen, deren Kantenrichtungen nicht mit der erwarteten bzw. bei der Detektion initialen Kantenrichtung übereinstimmen, durch den

meist durchgeführten Vorverarbeitungsschritt der gerichteten Spalten- bzw. Zeilensum-  
mation verfälscht. Im Extremfall kann dadurch ein Helligkeitsübergang derart stark ver-  
waschen werden, dass die entsprechende Segmentgrenze nicht mehr gefunden werden kann.  
Als Fazit kann hier festgehalten werden, dass ein Verfahren sehr wünschenswert wäre,  
welches gleichzeitig die Richtung der Kanten an Segmentgrenzen und eine Beschreibung  
der dazwischenliegenden homogenen Segmente liefert.

Dies war die Motivation für die Entwicklung eines weiteren Segmentierungsverfahrens,  
welches im Abschnitt 4.3 beschrieben wird. Dieser Operator soll sowohl bei der Detekti-  
on von Objekten, d.h. bei der Suche nach Objekten ohne Vorwissen, und beim Verfolgen  
bestehender Objekthypothesen Anwendung finden. Durch die extrahierten Kantenrichtun-  
gen sollte die Beschreibung eines Bildstreifens deutlich an Aussagekraft und Genauigkeit  
gewinnen. Weiterhin sind verbesserte Möglichkeiten zur Selektion von Merkmalen durch  
Verkettung zu erwarten.

### 4.3 Segmentierung homogener Bereiche durch Ex- traktion gerichteter Kantenpunkte

Das Ziel dieses Segmentierungsverfahrens besteht darin, die Richtungen von Segmentgren-  
zen zu bestimmen und gleichzeitig eine exakte Nachbildung des Helligkeitsverlaufs entlang  
eines Bildstreifens zu erstellen, um damit im Vergleich zum hierarchischen Verfahren aus  
Abschnitt 4.2 eine genauere und aussagekräftigere Beschreibung eines Bildstreifens zu  
erhalten. Die Kombination aus richtungssensitiver Kantenextraktion und gleichzeitiger li-  
nearer Approximation des Helligkeitsverlaufs homogener Segmente wird in dieser Arbeit  
selbstbewusst als HOFMANN-Operator bezeichnet, um in den nachfolgenden Abschnitten  
eine eindeutige Benennung zur Verfügung zu haben.

Dieses Verfahren führt im Gegensatz zum hierarchischen Segmentierungsverfahren kei-  
ne gerichtete Mittelung zur Vorverarbeitung der Daten durch, sondern bestimmt entlang  
eines Bildstreifens gerichtete Kantenpunkte, deren Richtungen mittels orthogonaler Kor-  
relationsmasken bzw. (je nach Sichtweise) über Ausgleichsebenen rechteckiger Bildaus-  
schnitte errechnet werden. Diese dienen gleichzeitig als Grenzen zwischen zwei homoge-  
nen Bereichen. Die Segmentierung erfolgt demnach über die Suche nach den Grenzen  
homogener Helligkeitsverläufe. Während des Suchvorgangs werden je Segment neben den  
Kantenrichtungen die Helligkeitswerte zur Linken und zur Rechten der gerichteten Seg-  
mentgrenzen, die Steigung und der Mittelwert des Helligkeitsverlaufs, die Positionen der  
Segmentgrenzen und die Segmentlängen ermittelt werden. Letztere werden sogar auf Sub-  
pixel genau bestimmt.

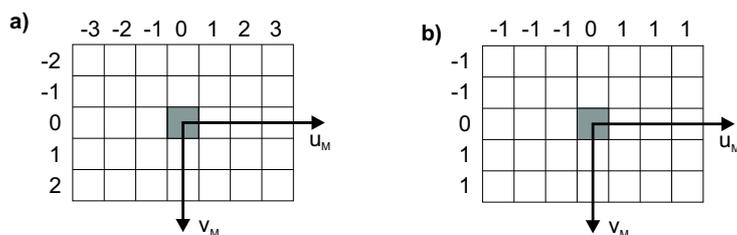
Neben der Genauigkeit und dem Informationsgehalt der extrahierten Daten stellt dabei  
natürlich wieder die Geschwindigkeit des Verfahrens ein wichtiges Entwurfskriterium dar.  
Durch eine möglichst effiziente Nutzung der Daten und eine gute Strukturierung der Er-  
gebnisse soll dabei gewährleistet werden, dass auch weiterführende Verarbeitungsschritte  
z.B. Verkettungsalgorithmen effizient implementiert werden können.

Zunächst stellt sich nun die Frage, wie kann die Richtung eines Kantenpunktes ermittelt werden?

### 4.3.1 Bestimmung der Richtung eines Kantenpunktes bei ungeradzahligem $n \times m$ -Korrelationsmaske

Nach Abschnitt A.7 lässt sich für das Helligkeitsgebirge eines rechteckigen Bildausschnitts eine Ausgleichsebene bestimmen, deren Normalenvektor anhand der Gleichungen (A.103), (A.104), (A.105) und (A.107) zu berechnen ist. Die Richtung einer gesuchten Kante ist dann orthogonal zur Projektion des Normalenvektors der Ausgleichsebene in die Bildebene. Der Richtungsvektor  $\vec{r}$  der Kante am betrachteten Bildpunkt kann nach Gleichung (A.109) durch das Kreuzprodukt des Helligkeitseinheitsvektors mit dem Normalenvektor  $\vec{n}$  der Ausgleichsebene berechnet werden.

In Abbildung 4.13a ist die symmetrische Indizierung nach Gleichung (A.106) und in Abbildung 4.13b die vereinfachte symmetrische Indizierung nach Gleichung (A.111) am Beispiel eines rechteckigen Bildausschnittes mit  $n=5$  und  $m=7$  veranschaulicht.

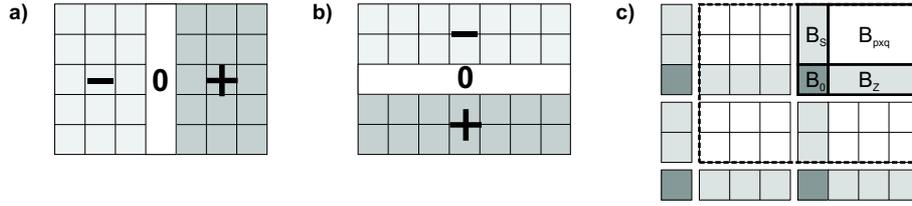


**Abbildung 4.13:** Beispiel für die Indizierung einer ungeradzahligem  $n \times m$ -Korrelationsmaske für die Berechnung einer Ausgleichsebene a) symmetrische Indizierung b) vereinfachte symmetrische Indizierung

Die vereinfachte symmetrische Indizierung hat gegenüber der ursprünglichen symmetrischen Indizierung den großen Vorteil, dass die in den Gleichungen (A.103) und (A.104) zu berechnenden Zeilen- und Spaltensummen in Blöcke gruppiert werden können, welche eine Wiederverwendung der Zwischenergebnisse ermöglichen. In Abbildung 4.14a ist die für Gleichung (A.103) notwendige Summation der Spalteneinträge und Gewichtung mittels der vereinfachten Indizierung dargestellt. Entsprechend symbolisiert Abbildung 4.14b die für die Gleichung (A.104) notwendige Zeilensummation und Gewichtung durch die vereinfachte Indizierung.

### 4.3.2 Effiziente Summation durch Zerlegung der Maske in Elementarblöcke

Durch eine geschickte Zerlegung in logische Einheiten kann die Summation so strukturiert werden, dass sie möglichst effizient durchgeführt werden kann. Speziell vor dem Hinter-



**Abbildung 4.14:** Differenzenbildung und Zerlegung einer ungeraden Maske in logische Einheiten für effiziente Summenbildung

grund, dass ja nicht nur an einer Stelle der Normalenvektor und daraus die Richtung einer potentiellen Segmentgrenze berechnet werden soll, sondern ein längerer Bildstreifen ausgewertet werden soll, kann durch eine sinnvolle Systematik deutlich Rechenzeit eingespart werden. In Abbildung 4.14c ist eine mögliche Zerlegung in vier Blöcke dargestellt, welche wiederum aus vier Elementarblöcken  $\mathbf{B}_0$ ,  $\mathbf{B}_S$ ,  $\mathbf{B}_Z$ ,  $\mathbf{B}_{pxq}$  bestehen.

Der Normalenvektor  $\vec{n}$  der Ausgleichsebene ergibt sich in diesem Beispiel zu

$$\vec{n} = \begin{bmatrix} -\frac{(\sum \mathbf{B}_{pxq}^{ol} + \sum \mathbf{B}_Z^{ol} + \sum \mathbf{B}_{pxq}^{ul}) + (\sum \mathbf{B}_{pxq}^{or} + \sum \mathbf{B}_Z^{or} + \sum \mathbf{B}_{pxq}^{ur})}{n \cdot (m-1) \cdot (m+1)/4} \\ -\frac{(\sum \mathbf{B}_{pxq}^{ol} + \sum \mathbf{B}_S^{or} + \sum \mathbf{B}_{pxq}^{or}) + (\sum \mathbf{B}_{pxq}^{ul} + \sum \mathbf{B}_S^{ur} + \sum \mathbf{B}_{pxq}^{ur})}{m \cdot (n-1) \cdot (n+1)/4} \\ 1 \end{bmatrix}, \quad (4.3)$$

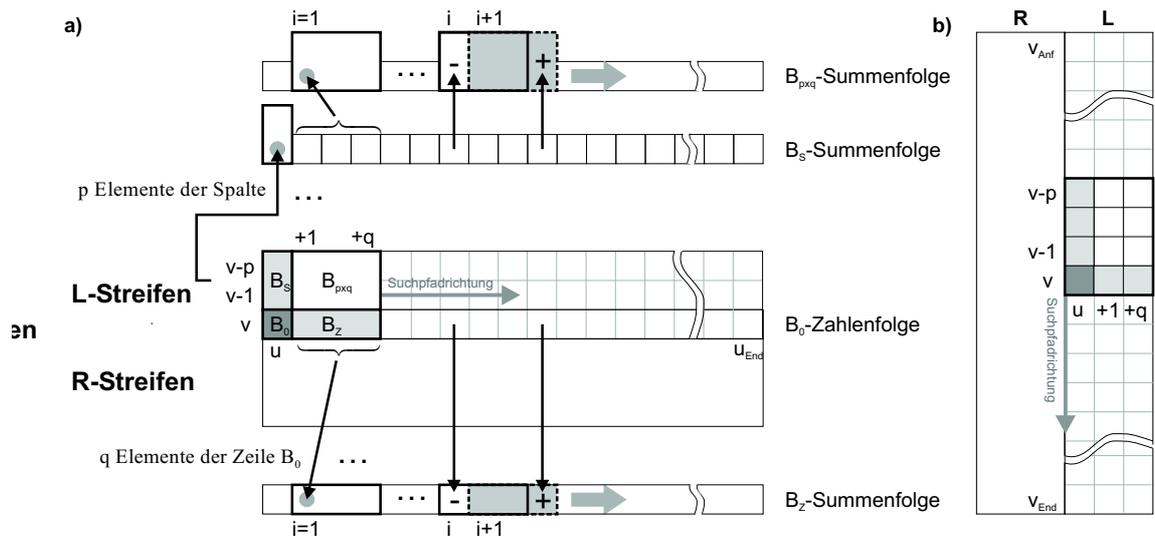
wobei die Hochindizes ol für oben links, or für oben rechts, ul für unten links und ur für unten rechts stehen.

Bei ungerader  $n \times m$ -Korrelationsmaske ist  $p = \frac{n-1}{2}$  und  $q = \frac{m-1}{2}$ . Ein Bildstreifen hat demnach bei horizontaler bzw. vertikaler Suchpfadrichtung eine Höhe von  $2 \cdot (p+1)$  bzw.  $2 \cdot (q+1)$  Pixeln. Siehe Abbildung 4.15a und 4.15b.

Entsprechend Abbildung 4.15a lässt sich die Berechnung der Elementarblöcke sehr effizient gestalten. Der ursprüngliche Bildstreifen wird dabei in einen in Suchpfadrichtung von der Mitte aus gesehen linken L- und rechten R-Streifen geteilt und dafür jeweils drei Summenfolgen für  $\mathbf{B}_S$ ,  $\mathbf{B}_Z$  und  $\mathbf{B}_{pxq}$  zum Teil rekursiv bestimmt. Für alle Positionen  $(u, v)$  mit  $u \in [u_{Anf} + q; u_{End} - q]$  innerhalb der  $\mathbf{B}_0$ -Zahlenfolge des L-Streifens kann nun der Normalenvektor und damit die Richtung einer potentiellen Kante entsprechend Gleichung (4.3) errechnet werden. Bei vertikalen Suchpfaden kann vergleichbar verfahren werden.

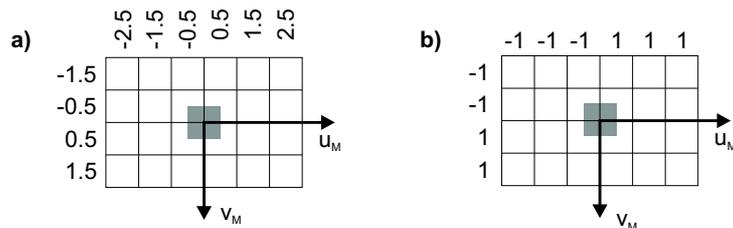
### 4.3.3 Bestimmung der Richtung eines Kantenpunktes für eine geradzahlige $n \times m$ -Korrelationsmaske

Dennoch ist bei Verwendung von Korrelationsmasken mit ungeraden Dimensionen der Verwaltungsaufwand noch recht erheblich. Wie der Vergleich mit den Abbildungen 4.16



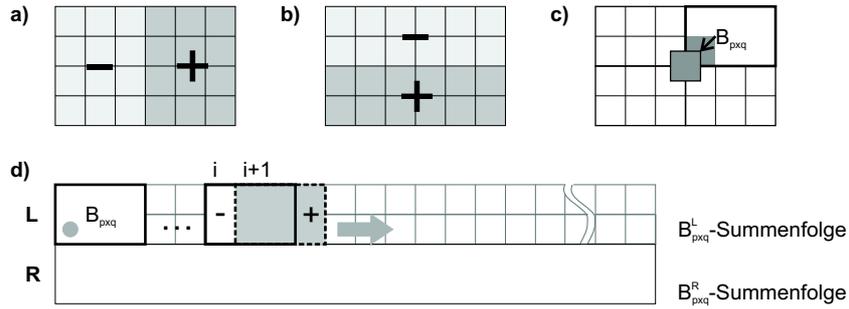
**Abbildung 4.15:** Effiziente Summenbildung a) bei horizontaler Suchpfadrichtung b) bei vertikaler Suchpfadrichtung

und 4.17 zeigt, kann durch Verwendung von Korrelationsmasken mit geraden Dimensionen und damit ohne Nullspalte und Nullzeile der Verwaltungsaufwand deutlich reduziert werden. Einziger Nachteil ist, dass sich nun der Ursprung des Maskenkoordinatensystems zwischen benachbarten Pixeln befindet. Nachdem die Positionen der Segmentgrenzen jedoch sowieso auf Subpixel genau bestimmt werden, hat diese Verschiebung keine negative Auswirkung.



**Abbildung 4.16:** Beispiel für die Indizierung einer geradzahigen  $n \times m$ -Korrelationsmaske für die Berechnung einer Ausgleichsebene a) symmetrische Indizierung b) vereinfachte symmetrische Indizierung

Der Richtungsvektor  $\vec{r}$  einer potentiellen Kante an der Position  $(u, v)$  ergibt sich bei horizontaler Suchpfadrichtung und vereinfachter symmetrischer Maske mit geradzahiger



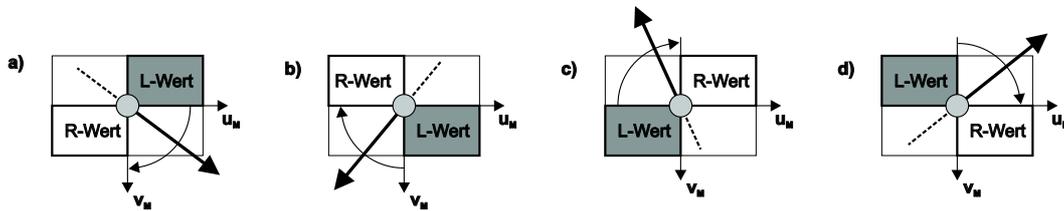
**Abbildung 4.17:** Differenzenbildung und Zerlegung einer geraden Maske in logische Einheiten für eine effiziente Summenbildung

Dimension  $n \times m$  mit  $p = \frac{n}{2}$  und  $q = \frac{m}{2}$  nach Gleichung (A.109) und (A.114) zu

$$\vec{r} = \begin{bmatrix} \frac{-\sum \mathbf{B}_{pxq}^L(u-q) - \sum \mathbf{B}_{pxq}^L(u) + \sum \mathbf{B}_{pxq}^R(u-q) + \sum \mathbf{B}_{pxq}^R(u)}{m \cdot n^2 / 4} \\ -\frac{-\sum \mathbf{B}_{pxq}^L(u-q) - \sum \mathbf{B}_{pxq}^R(u-q) + \sum \mathbf{B}_{pxq}^L(u) + \sum \mathbf{B}_{pxq}^R(u)}{n \cdot m^2 / 4} \\ 0 \end{bmatrix}. \quad (4.4)$$

#### 4.3.4 Bestimmung der Helligkeitswerte zur Linken und zur Rechten einer Kante in Kantenrichtung

Je nachdem in welchen Quadranten der Richtungsvektor  $\vec{r}$  zeigt, werden relativ zur Kantenrichtung ein Parameter *L-Wert* für den linken Helligkeitswert bzw. *R-Wert* für den rechten Helligkeitswert derart bestimmt, dass ihnen die auf die Blockgröße normierten Blocksummen  $\sum_{p,q} \mathbf{B}_{pxq}$  entsprechend Abbildung 4.18 zugewiesen werden. Wie in Abschnitt A.7 ausgeführt, wird dabei nach Definition der R-Wert stets größer als der L-Wert sein.



**Abbildung 4.18:** Bestimmung der Intensität zur Linken und zur Rechten eines Kantenpunktes in Kantenrichtung

### 4.3.5 Suche nach den Segmentgrenzen mit gleichzeitiger Nachbildung des Helligkeitsverlaufs

Die Suche nach gerichteten Segmentgrenzen geschieht nun wie folgt. Zunächst werden für einen Bildstreifen die zugehörigen linken und rechten Blocksummenfolgen berechnet. Dann wird in Suchpfadrichtung sukzessive der Suchpfad durchlaufen und für jede gültige Position der Richtungsvektor  $\vec{r}$  und der Normalenwinkel  $\alpha$  zwischen dem Normalenvektor und dem Dateneinheitsvektor mittels der Gleichungen (A.109) und (A.110) bestimmt.

Ist das aktuelle Segment das erste Segment im Suchpfad und seine Segmentlänge noch kleiner als die halbe Maskenlänge, dann wird je nach Kantenrichtung derjenige L-Wert bzw. R-Wert, der der aktuellen Position nachläuft, dem aktuellen Segment als Helligkeitswert für die rekursive Ausgleichsgeradenberechnung nach Gleichung (A.67) zugeordnet. Ist das aktuelle Segment nicht das erste Segment, aber die Länge des homogenen Segments noch kleiner als die halbe Maskenlänge, so wird je nach Kantenrichtung derjenige L-Wert bzw. R-Wert des vorhergehenden gerichteten Kantenpunkts dem aktuellen Segment als Helligkeitswert zugeordnet, der dem aktuellen Segment zugewandt ist. Übersteigt die Länge des aktuellen Segments die halbe Maskenlänge, dann wird jeweils der nachlaufende L- bzw. R-Helligkeitswert für die rekursive Bestimmung der Ausgleichsgerade verwendet.

Für die Detektion einer Segmentgrenze werden jeweils drei aufeinanderfolgende Positionen dahingehend überprüft, ob der Normalenwinkel der mittleren Position

1. einen Maximalwert im Vergleich zu den beiden anderen darstellt,
2. der Normalenwinkel einen Schwellenwert übersteigt und ob
3. der L-Wert kleiner einem prozentualen Anteil des R-Werts abzüglich einer minimalen Helligkeitsdifferenz zwischen L- und R-Wert ist.

$$\text{L-Wert} < \text{R-Wert} \cdot \text{RatioLR} - \text{DiffLR} \quad (4.5)$$

Optional kann mittels eines Parameters *Selectivity* der erlaubte Winkelbereich einer Kantenrichtung relativ zur Suchpfadrichtung eingeschränkt werden. Hierbei bedeutet ein Wert von  $0^\circ$ , dass keine Beschränkung vorliegt. Sind alle Kriterien erfüllt, dann wird mittels Gleichung (A.91) die Subpixelposition dieses gerichteten Kantenpunktes ermittelt, die Position im Suchpfad aufaddiert und das Ergebnis in der Ergebnisstruktur für das aktuelle Segment gespeichert. Weiterhin wird die Länge des Segments um den Subpixelanteil des vorhergehenden Segments verringert und um den aktuellen Subpixelanteil vergrößert. Damit sind sowohl die Position der gerichteten Segmentgrenze als auch die Segmentlänge auf Subpixel genau bestimmt.

**Tabelle 4.3:** Parameter für eine zielgerichtete Auswertung eines Bildstreifens

NormAngleThreshold	Schwellenwert für den Normalenwinkel
RatioLR	Helligkeitsschwelle als prozentualer Anteil des R-Werts
DiffLR	minimale Helligkeitsdifferenz zwischen L- und R-Wert
Selectivity	begrenzt den Winkelbereich der erlaubten Kantenrichtungen relativ zur Suchpfadrichtung ( $0^\circ$ =keine Begrenzung, z.B. $25^\circ$ = Betrag der Winkeldifferenz zur Suchpfadrichtung muss mindestens $25^\circ$ betragen)
SearchPathDirection	Suchpfadrichtung $0^\circ=0$ , $90^\circ=1$ , ( $\pm 180^\circ=2$ , $-90^\circ=3$ )
CenterY	Y-Position des Suchpfadzentrums im Bildkoordinatensystem
CenterZ	Z-Position des Suchpfadzentrums im Bildkoordinatensystem
TolPlus	Länge des Suchpfads in Suchpfadrichtung ausgehend vom Suchpfadzentrum
TolMinus	Länge des Suchpfads entgegen der Suchpfadrichtung ausgehend vom Suchpfadzentrum
HalfMaskSize	halbe Höhe des auszuwertenden Bildstreifens = Höhe des L- bzw. R-Streifens
HalfMaskLen	halbe Maskenlänge in Suchpfadrichtung bei gerader nxm-Korrelationsmaske

### 4.3.6 Parametrisierung und Ergebnisdatenstruktur

Eine vollständige Liste der Parameter zur zielgerichteten Anpassung des Verfahrens für die Auswertung eines Bildstreifens zeigt Tabelle 4.3.

Nachfolgend sind die Kerngedanken beschrieben, die dem Entwurf der Ergebnisdatenstruktur zugrunde liegen.

Jeder gerichtete Kantenpunkt entlang eines Bildstreifens stellt einen Endpunkt eines homogenen Segments dar. In Suchpfadrichtung werden die Kantenpunkte entsprechend ihrer Reihenfolge im Bildstreifen in einem eindimensionalen Ergebnisarray abgelegt. Am Ende jedes Suchstreifens wird ein unechter Kantenpunkt erzeugt und als solcher markiert. Dieser ist dazu da, die Parameter des letzten homogenen Segments eines jeden Suchpfads aufzunehmen.

Im Ergebnisarray besitzt jedes homogene Segment einen Eintrag. Die Datenstruktur dieses Eintrags kann Tabelle 4.4 entnommen werden.

Im Bedarfsfall könnte als eine weitere charakteristische Größe eines Segments auch noch die Varianz der Abweichung von der Ausgleichsgerade in die Struktur mit aufgenommen werden. Hierfür müsste lediglich neben der Summe der Helligkeitswerte entlang des Helligkeitsverlaufs die Summe der Quadrate der Helligkeitswerte mitberechnet werden. Die Varianz ergäbe sich dann unter Verwendung der Gleichung (A.30).

**Tabelle 4.4:** Ergebnisdatenstruktur eines homogenen Segments

YPos, ZPos	Subpixelgenaue Position des Abschlusskantenpunkts eines Segments in Bildkoordinaten.
dY, dZ	Richtungsvektor des Kantenpunkts, wobei jeweils derjenige Anteil auf Länge 1 normiert wird, der orthogonal zur Suchpfadrichtung zeigt. Dies erleichtert die Verkettung.
AngleNorm	Für Kontrollzwecke wird hier der Wert des Normalenwinkels gespeichert.
AngleEdge	Kantenwinkel in Grad im Bildkoordinatensystem
LeftIntensity	Intensitätswert zur Linken der Kante in Kantenrichtung = L-Wert
RightIntensity	Intensitätswert zur Rechten der Kante in Kantenrichtung = R-Wert
Used	Used-Flag für Verkettungsalgorithmen
SegLen	Subpixelgenaue Länge $l$ des homogenen Segments
SegMean	mittlere Helligkeit $b$ des homogenen Segments
SegMeanDiff	Steigung $a$ des Helligkeitsverlaufs im Segment
SegMin	Minimalwert innerhalb des Segments
SegMax	Maximalwert innerhalb des Segments
MeanSegPosY	y-Position der Segmentmitte
MeanSegPosZ	z-Position der Segmentmitte
MeandY	mittlere Richtung des Segments dY für Bereichsverkettung
MeandZ	mittlere Richtung des Segments dZ für Bereichsverkettung
TTStart	Klassifizierung nach Transitionstyp am Anfang des Segments (0=n.d. nicht definiert, 1=HD hell-dunkel, 2=DH dunkel-hell)
TTEnd	Klassifizierung nach Transitionstyp am Ende des Segments (0=n.d. nicht definiert, 1=HD hell-dunkel, 2=DH dunkel-hell)

**Tabelle 4.5:** Verwaltungsstruktur für die schnelle Navigation im Ergebnisarray

LineNumStrip	Zeile bzw. Spalte im Bildkoordinatensystem je nach Suchpfadrichtung
FeatureIndexBegin	Index des ersten Segments des betrachteten Suchstreifens in der Ergebnisdatenstruktur
NumOfFeatures	Anzahl der Segmente des betrachteten Suchstreifens

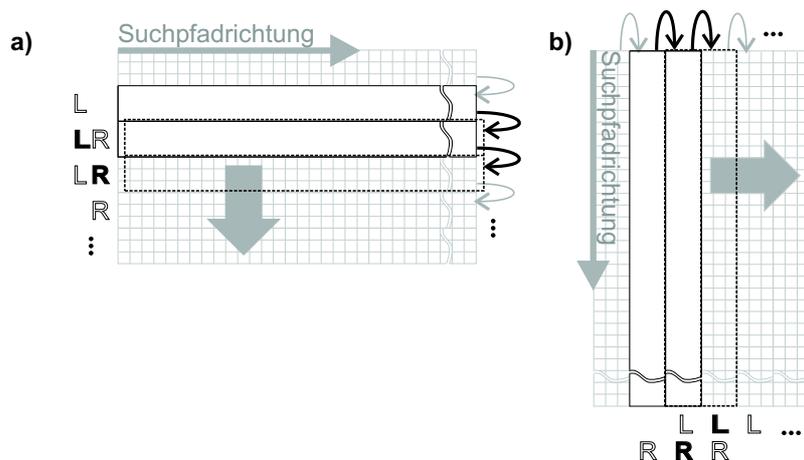
### 4.3.7 Effiziente Auswertung eines größeren Bildbereichs

Für den Fall, dass ein größerer rechteckiger Bereich eines Bildes ausgewertet werden soll, kann dies durch Zerlegung in mehrere parallele Bildstreifen erfolgen. Besonders ressourcenschonend bei möglichst hoher Auflösung ist dabei die Variante, bei der der Abstand der parallelen Suchpfade genau der Breite eines L- bzw. R-Streifens entspricht. Bereits berechnete Summenfolgen der L- bzw. R-Streifen können dabei wiederverwendet werden. Wie in Abbildung 4.19a dargestellt, kann bei horizontaler Suche durch Zuweisung des R-Streifens des letzten Suchpfades auf den L-Streifen des aktuellen Suchpfades erreicht werden, dass lediglich die Summenfolgen des neuen R-Streifens bestimmt werden müssen. Vergleichbares gilt bei vertikaler Suche, siehe Abbildung 4.19b. Es wird hierbei sichergestellt, dass jedes Pixel des Bildausschnitts nur einmal für die Berechnung der Streifensummenfolgen herangezogen wird.

Die Ergebnisse mehrerer Suchpfade werden entsprechend der Reihenfolge ihrer Auswertung jeweils ans Ende der eindimensionalen Ergebnisstruktur angefügt. Für die Navigation innerhalb des Ergebnisarrays wird zusätzlich eine Verwaltungsstruktur angelegt. In der Verwaltungsstruktur wird für jeden Suchpfad je ein Eintrag erstellt, der je nach Suchpfadrichtung die entsprechende Bildzeile bzw. Bildspalte des aktuellen Suchpfades, den Index des ersten Segments in der Ergebnisstruktur und die Anzahl der Segmente im Suchpfad speichert. Siehe Tabelle 4.5.

Dadurch wird die Möglichkeit einer schnellen Navigation in der eindimensionalen Ergebnisstruktur der Segmente sichergestellt. Weiterhin bleibt hierdurch der räumliche Bezug der Segmente zueinander erhalten. Dieser ist grundlegend für eine effiziente Verkettung von Segmenten zu Bereichen bzw. für die Gruppierung von Kantenpunkten zu Kantenverläufen. Ohne diesen unmittelbaren räumlichen Bezug müsste stets jedes Segment auf Nachbarschaft zu jedem anderen verglichen werden. Der Zeitaufwand wäre enorm.

Diese Vorgehensweise wurde weiterhin für die Auswertung konvexer 2-D Polygonbereiche verallgemeinert. Der wesentliche Unterschied besteht darin, dass die Anfangs- und Endpunkte eines jeden Suchstreifens jeweils aus den Schnittpunkten mit einem konvexen 2-D Polygonbereich berechnet werden müssen und die Berechnung der Kantenrichtungen aus den L- und R-Streifen nur für definierte Einträge erfolgen darf.



**Abbildung 4.19:** Wiederverwendung bereits berechneter Summenfolgen bei Schrittweite  $s = p + 1$  bzw.  $s = p$  für eine ungerade bzw. gerade Maskengröße a) horizontale Suchpfadrichtung b) vertikale Suchpfadrichtung

#### 4.3.8 Beispiel: Auswertung eines Ausschnitts eines Schwarz-Weiß-Bildes

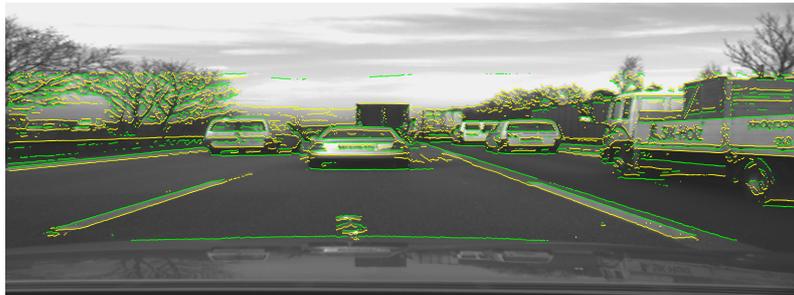
In diesem Abschnitt soll die Leistungsfähigkeit des HOFMANN-Operators anhand eines Beispiels illustriert werden. Hierfür wird ein Halbbild einer Aufnahme einer monochromen  $\frac{1}{2}$ " Kamera mit  $8mm$  Weitwinkelobjektiv, welches eine typische Autobahnszene entsprechend Abbildung 4.20 beinhaltet, innerhalb eines relevanten Bildausschnittes nach Kantenmerkmalen und homogenen Segmenten entlang horizontaler und vertikaler Suchpfade durchsucht.



**Abbildung 4.20:** Originalbild einer typischen Autobahnszene. Der überlagerte weiße Rahmen begrenzt den auszuwertenden Bildausschnitt.

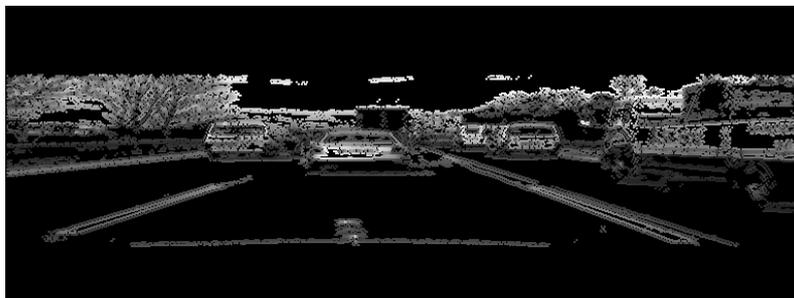
Die Abbildungen 4.21 bis 4.24 zeigen die Ergebnisse einer Bereichsauswertung entlang vertikaler Suchpfade von oben nach unten in Streifen und Streifenabstand  $p$  mit  $pxq$ -

Elementarmaske der Größe  $p = 1$  und  $q = 3$ . In Abbildung 4.21 sind die mit Subpixelgenauigkeit extrahierten gerichteten Kantenpunkte inklusive Kantenrichtung eingezeichnet. Die Art des Helligkeitsübergangs wurde jeweils farblich codiert: hell-dunkel (HD) = gelb, dunkel-hell (DH) = grün. Die gerichteten Kantenpunkte dienen gleichzeitig auch als Segmentgrenzen.



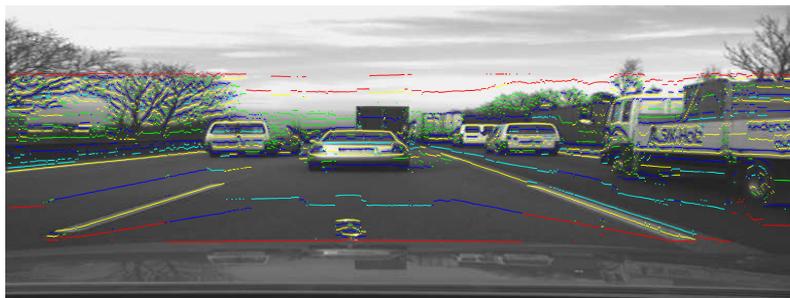
**Abbildung 4.21:** Gerichtete Kantenpunkte bei vertikaler Suchpfadrichtung von oben nach unten mit 1x3-Elementarmaske.

Abbildung 4.22 zeigt die L- und R-Werte bezüglich der extrahierten Kantenpunkte. Diese repräsentieren die lokalen Helligkeiten an den Segmentgrenzen, wohingegen die mittlere Segmenthelligkeit und Helligkeitsänderung im Segment den großflächigeren Helligkeitsbezug wiedergeben.



**Abbildung 4.22:** L- und R-Werte der extrahierten gerichteten Kantenpunkte.

In Abbildung 4.23 sind die Segmentmitten und die mittlere Segmentausrichtung, welche sich als Mittelwert der jeweils begrenzenden Kantenrichtungen ergibt, farblich kodiert. Die Mittelpunkte der Abschlusssegmente am Anfang und Ende jedes Suchpfads sind jeweils rot markiert. In dieser Abbildung ist sehr gut zu erkennen, dass die Segmentgrenzen korrekt extrahiert wurden und die Segmentierung ausgezeichnete Ergebnisse liefert, die auch eine Verkettung benachbarter Segmente zu Bereichen zulässt.



**Abbildung 4.23:** Darstellung der gerichteten Mittelpunkte homogener Segmente mit farblicher Kodierung der unterschiedlichen Helligkeitsübergänge (HD-DH, DH-HD, DH-DH, HD-HD).

Abbildung 4.24 zeigt schließlich die Rekonstruktion des Bildes anhand der mittleren Segmenthelligkeiten und der Helligkeitsänderungen in den homogenen Segmenten. Die hohe Übereinstimmung mit dem ursprünglichen Bild zeigt, dass die wesentliche Bildinformation durch die Bereichsauswertung unter Beibehaltung der Nachbarschaftbeziehungen für darauf aufbauende Verarbeitungsschritte, wie z.B. Verkettungsverfahren, extrahiert und in kompakter Form repräsentiert werden konnte.



**Abbildung 4.24:** Rekonstruktion des Bildes durch Visualisierung der linearen Beschreibungen des Helligkeitsverlaufs homogener Segmente bei vertikaler Suche mit  $1 \times 3$ -Elementarmaske.

Zum Vergleich wurden in den Abbildungen 4.25 und 4.26 verschiedene Elementarblockgrößen ( $4 \times 4$ ,  $8 \times 8$ ) verwendet bei ansonsten unveränderten Parametern des Bildoperators. Es ist zu erkennen, dass die Ergebnisse der Bildauswertung (mit Ausnahme der abnehmenden Auflösbarkeit der Bildmerkmale) unabhängig von der Maskengröße sind. Das bedeutet, es konnte erreicht werden, dass die Rauschschwelle für alle Maskengrößen einheitlich festgelegt werden kann, um bei gegebener Auflösbarkeit gleiche Ergebnisse zu erzielen. Damit ist eine zentrale Bedingung für eine transparente Bildauswertung erfüllt. Die Maskengröße kann an die Größe eines erwarteten Merkmals angepasst werden, wo-



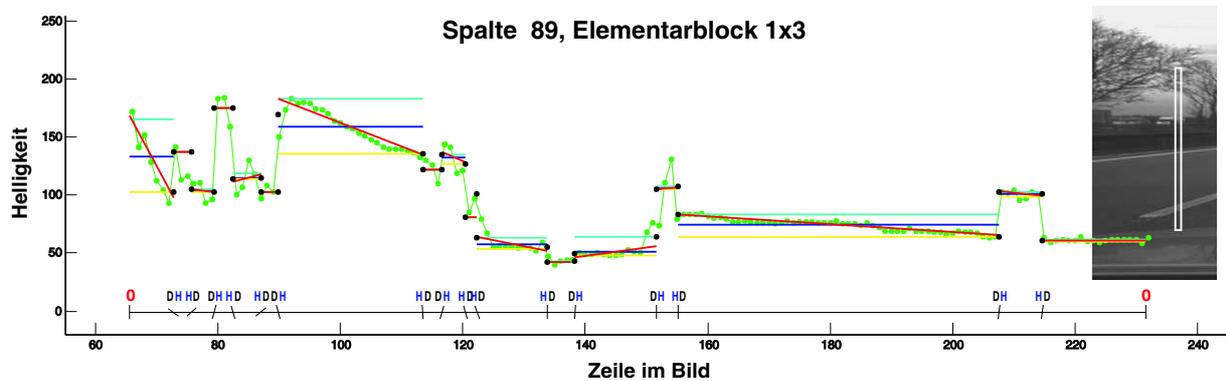
**Abbildung 4.25:** Rekonstruktion des Bildes bei Verwendung einer 4x4-Elementarmaske und Schrittweite 4 zwischen den vertikalen Suchpfaden.



**Abbildung 4.26:** Rekonstruktion des Bildes bei Verwendung einer 8x8-Elementarmaske und Schrittweite 8 zwischen den vertikalen Suchpfaden.

bei die Rauschschwelle global für das Bild bzw. lokal für einen Bildbereich einheitlich festgelegt werden kann.

In Abbildung 4.27 ist exemplarisch der Helligkeitsverlauf entlang eines einzelnen vertikalen Suchpfades dargestellt. In ihr sind der Helligkeitsverlauf = grün , die Approximation durch



**Abbildung 4.27:** Segmentierung eines Helligkeitsverlaufs entlang eines vertikalen Suchpfades. Zur Rechten ist der zugrunde liegende vertikale Suchstreifen im Bild dargestellt.

eine Gerade = rot, der Segmentmittelwert = blau, die minimale Elementarblockhelligkeit je Segment = gelb, die maximale Elementarblockhelligkeit je Segment = türkis und die L- und R-Werte je Segmentgrenze jeweils als schwarze Punkte dargestellt. Weiterhin sind in

der Zeile des Helligkeitswertes null die Segmentlängen, die Kantenrichtungen und die Art der Helligkeitsübergänge (0 = nicht definiert, HD, DH) eingezeichnet. Zur Rechten ist der Suchstreifen weiß umrandet hervorgehoben. Die Abbildung soll nochmals die kompakte Form der extrahierten Information verdeutlichen.

In den Abbildungen 4.28 und 4.29 ist die Rekonstruktion des Bildausschnittes bei horizontaler Suche mit 1x3-Elementarmaske und Streifenabstand 1 dargestellt. In Abbildung 4.28



**Abbildung 4.28:** Rekonstruktion des Bildes bei Verwendung einer 1x3-Elementarmaske und Schrittweite 1 zwischen den horizontalen Suchpfaden.

ist zu erkennen, dass selbst die extrem flachen Spurmarkierungen in der linken Bildhälfte richtig segmentiert werden. Die Kantenrichtungen werden korrekt bestimmt.

In Abbildung 4.29 ist der Wert für die Selektivität von  $0^\circ$  auf  $45^\circ$  erhöht worden, wodurch alle Segmentgrenzen mit einer Kantenrichtung kleiner  $\pm 45^\circ$  relativ zu Suchpfadrichtung unterdrückt werden. Es ist zu erkennen, dass die flachen und extrem flachen Spurmarkierungen nicht mehr als Segmentgrenzen zugelassen werden und lediglich vertikale Kanten als Segmentgrenzen dienen.



**Abbildung 4.29:** Rekonstruktion des Bildes bei Verwendung einer 1x3-Elementarmaske und Schrittweite 1 zwischen den horizontalen Suchpfaden mit Selektivität= $45^\circ$ .

Als ein letztes Beispiel ist in Abbildung 4.30 die vertikale Auswertung eines Bildausschnittes mit einer 1x2-Elementarmaske dargestellt. Es ist zu erkennen, dass durch die im Vergleich zu Abbildung 4.21 kleinere Maske noch etwas feinere Details aufgelöst werden können, wobei im Nahbereich vereinzelt auch Helligkeitsübergänge als Segmentgrenzen detektiert werden, die dem Pixelrauschen zugeordnet werden könnten.



**Abbildung 4.30:** Vertikale Auswertung eines Bildausschnitts mit  $1 \times 2$ -Elementarmaske und Schrittweite eins: oben) gerichtete Kantenpunkte, Mitte links) L- und R-Werte, Mitte rechts) gerichtete Segmentmittelpunkte, unten links) lineare Rekonstruktion, unten rechts) horizontale Verkettung gerichteter Kantenpunkte entsprechend Abschnitt 4.3.13.1.

Elementarmaske	Schrittweite	Anzahl der Segmente	Rechenzeit [ms]
1x3	1	12235	18,21
1x3	3	4074	6,54
3x3	3	3991	6,26
3x3	6	2014	3,37
2x2	2	7054	9,72
2x2	4	3486	5,24
4x4	4	2701	4,56
2x2	8	1706	2,65
4x4	8	1319	2,46
8x8	8	937	2,52

**Tabelle 4.6:** Rechenzeitbedarf und Anzahl der Segmente bei vertikaler Suche mittels HOFMANN-Operator in Abhängigkeit von der Elementarmaske und der Schrittweite.

### 4.3.9 Rechenzeitbedarf des HOFMANN-Operators

Die Auswertung eines [170x760]-Pixel großen Bildbereichs eines Halbbildes entsprechend Abbildung 4.20 mit unterschiedlichen Elementarmasken und Schrittweiten benötigt bei vertikaler Suchpfadrichtung auf einer Pentium4 CPU mit 2400 MHz Taktfrequenz eine Rechenzeit in [ms] entsprechend Tabelle 4.6.

Der Tabelle ist weiterhin zu entnehmen, dass bei gleicher Schrittweite bei kleinerer Maske die Anzahl der resultierenden Segmente größer ist. Das liegt daran, dass durch eine kleinere Maske feinere Strukturen aufgelöst werden können.

### 4.3.10 Statistik der photometrischen Eigenschaften eines Bildbereichs aus der Beschreibung homogener Segmente

Entsprechend Tabelle 4.4 stehen nun auch für eine Analyse der Bildhelligkeit eine Vielzahl an Informationen zur Verfügung, welche für eine Anpassung der Parameter der Bildoperatoren an die Lichtverhältnisse genutzt werden können. Für jeden Bildstreifen können aus den Ergebnissen der Segmentierung charakteristische Kenngrößen extrahiert werden. Angenommen der Helligkeitsverlauf eines Bildstreifens kann durch  $n_S$  Segmente nachgebildet werden, dann ergibt sich der mittlere Helligkeitswert  $b_S$  über alle Segmente  $i$  der Länge  $l_i$  und der mittleren Helligkeit  $b_i$  aus

$$b_S = \frac{\sum_{i=1}^{n_S} (l_i \cdot b_i)}{\sum_{i=1}^{n_S} l_i}. \quad (4.6)$$

Für einen größeren Bereich  $G$ , der in  $n_G$  Bildstreifen unterteilt wurde, gilt dann

$$b_G = \frac{\sum_{j=1}^{n_G} (\sum_{i=1}^{n_{Sj}} (l_{ij} \cdot b_{ij}))}{\sum_{j=1}^{n_G} (\sum_{i=1}^{n_{Sj}} l_{ij})}. \quad (4.7)$$

Der Wert für  $b_S$  bzw.  $b_G$  unterscheidet sich vom Mittelwert der mittleren Helligkeiten, denn dieser ergibt sich aus

$$b_{MeanS} = \frac{\sum_{i=1}^{n_S} b_i}{n_S} \text{ bzw. } b_{MeanG} = \frac{\sum_{j=1}^{n_G} (\sum_{i=1}^{n_{Sj}} b_{ij})}{\sum_{j=1}^{n_G} n_{Sj}}. \quad (4.8)$$

Der absolute Minimalwert und Maximalwert der Helligkeit aller Segmente eines Bildstreifens kann aus

$$Min_{MinS} = \min_{i=1}^{n_S} (min_i) \text{ bzw. } Max_{MaxS} = \max_{i=1}^{n_S} (max_i) \quad (4.9)$$

bestimmt werden. Für einen größeren Bereich gilt

$$Min_{MinG} = \min_{j=1}^{n_G} (\min_{i=1}^{n_{Sj}} (min_{ij})) \text{ bzw. } Max_{MaxG} = \max_{j=1}^{n_G} (\max_{i=1}^{n_{Sj}} (max_{ij})). \quad (4.10)$$

Aus der Differenz ergibt sich der absolute Hub der Helligkeit eines Bildstreifens zu

$$Hub_S = Max_{MaxS} - Min_{MinS} \quad (4.11)$$

bzw. der Hub eines ganzen Bildbereichs zu

$$Hub_G = Max_{MaxG} - Min_{MinG}. \quad (4.12)$$

Der Hub eines einzelnen Segments ergibt sich aus der Differenz des Minimal- und Maximalwertes eines Segments  $Hub_i = max_i - min_i$ . Der mittlere Hub innerhalb eines Streifens bzw. Bildbereiches ergibt sich dann aus

$$Hub_{MeanS} = \frac{\sum_{i=1}^{n_S} Hub_i}{n_S} \text{ bzw. } Hub_{MeanG} = \frac{\sum_{j=1}^{n_G} (\sum_{i=1}^{n_{Sj}} Hub_{ij})}{\sum_{j=1}^{n_G} n_{Sj}}. \quad (4.13)$$

Soll der Maximalwert bzw. Minimalwert weniger sensitiv gegenüber vereinzelt Ausreißern im Helligkeitsverlauf sein, kann der Minimalwert bzw. Maximalwert über die Mittelwerte  $b_i$  aller Segmente verwendet werden:

$$b_{MinS} = \min_{i=1}^{n_S} (b_i) \text{ bzw. } b_{MaxS} = \max_{i=1}^{n_S} (b_i) \quad (4.14)$$

bzw. für einen größeren Bereich

$$b_{MinG} = \min_{j=1}^{n_G} (\min_{i=1}^{n_{Sj}} (b_{ij})) \text{ bzw. } b_{MaxG} = \max_{j=1}^{n_G} (\max_{i=1}^{n_{Sj}} (b_{ij})) \quad (4.15)$$

Je nachdem, ob der Mittelwert des Bildstreifens näher am Minimalwert bzw. Maximalwert liegt, ist der Bildstreifen demnach eher dunkler bzw. heller. Der praktische Nutzen dieser

Zweiteilung ist für die Selektion eines Merkmals aus der Menge von Merkmalen allerdings eher gering.

Ist man an einer Statistik der helleren zu den dunkleren Segmenten interessiert, kann man z.B. zunächst die mittlere Helligkeit  $b_G$  über alle Segmente errechnen. Dann wählt man aus der Menge aller Segmente diejenigen Segmente aus, deren mittlere Helligkeiten  $b_i$  größer bzw. kleiner  $b_G$  sind. Für diese Untermengen kann dann eine (mit den jeweiligen Segmentlängen gewichtete) mittlere Helligkeit  $b_{hellG}$  für die helleren Segmente bzw.  $b_{dunkelG}$  für die eher dunkleren Segmente bestimmt werden.

Für diese Untermengen mag es dann z.B. wiederum interessant sein, die Minimal- und Maximalwerte der Minimal- bzw. Maximalwerte über alle Segmente dieser Untermengen zu bestimmen, um daraus den Hub der jeweiligen Untermenge zu erhalten...

Eine näherungsweise Unterteilung der Menge der Segmente nach Helligkeit in eher helle, eher dunklere und Segmente mit mittlerer Helligkeit erhält man dann z.B., indem als obere Schwelle

$$S_{hell} = \frac{b_G + b_{hellG}}{2} \quad (4.16)$$

bzw. als untere Schwelle

$$S_{dunkel} = \frac{b_G + b_{dunkelG}}{2} \quad (4.17)$$

verwendet wird.

Als zusätzliche statistische Kenngrößen können weiterhin der Mittelwert, Minimalwert und Maximalwert aller L- bzw. R-Werte errechnet werden oder es können der Mittelwert, Minimalwert und Maximalwert der absoluten Differenz zwischen R- und L-Wert bestimmt werden. Eine weitere interessante Größe stellt das Verhältnis von L-Wert zu R-Wert dar. Für die Zuordnung von Merkmalen zwischen Bildern verschiedener Kameras sind relative Helligkeitsmaße ein wichtiges Kriterium, da sich aufgrund unterschiedlicher Blendenstellungen, Belichtungszeiten und Verstärkungsfaktoren die absoluten Helligkeiten eines Merkmals sicher unterscheiden werden. Diesbezüglich wird im Abschnitt 4.3.11 eine Möglichkeit zur Bestimmung einer Abbildungsvorschrift zur näherungsweisen Transformation von Bildhelligkeiten vorgestellt, wodurch auch der Vergleich von absoluten Helligkeitswerten möglich wird.

Dann kann z.B. noch die mittlere, minimale, maximale Segmentlänge eines Streifens bzw. Bereichs bestimmt werden...

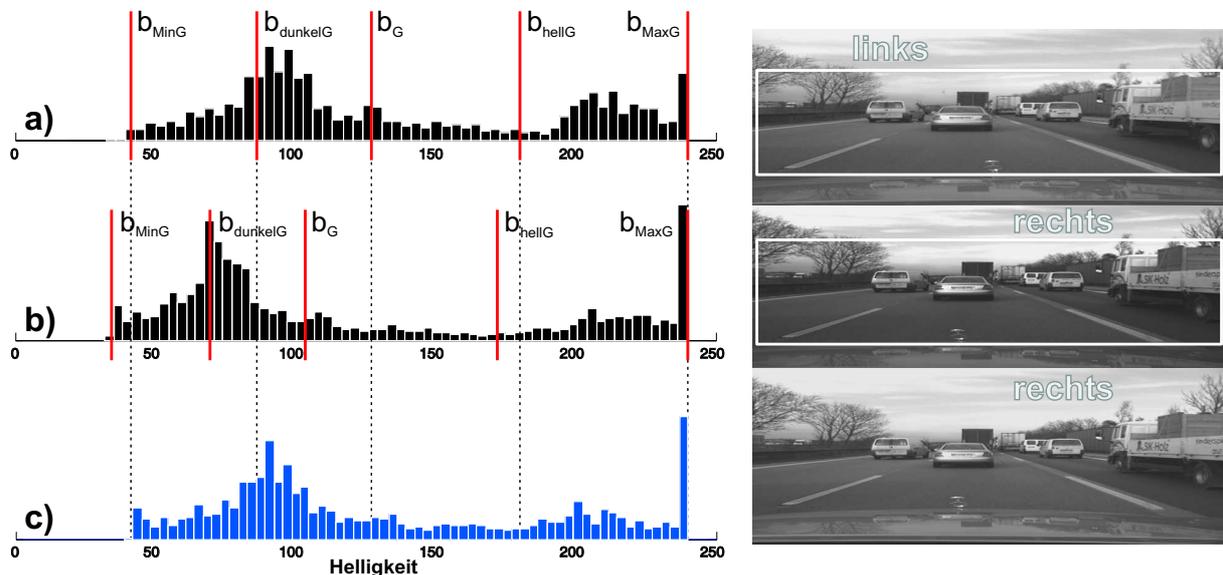
Welche Größen sinnvoll für eine Auswertung sind, muss dabei von Fall zu Fall entschieden werden. Die Aufzählung soll hier lediglich die Möglichkeiten verdeutlichen, welche durch das Wissen über die Beschaffenheit homogener Segmente gegenüber einer reinen Kantenextraktion eröffnet werden.

Die Statistik eines kleineren Bildausschnitts bzw. einzelnen Bildstreifens sollte dabei stets in Relation zur Statistik eines größeren Bildausschnitts betrachtet werden. Neben der Bestimmung von Kenngrößen eines Bildausschnitts zum aktuellen Zeitpunkt kann es dabei auch sinnvoll sein, die Veränderung bestimmter Größen z.B. mittels gleitendem Mittelwert über der Zeit zu betrachten.

### 4.3.11 Beispiel: Ermittlung einer näherungsweise Abbildungsvorschrift für Helligkeiten

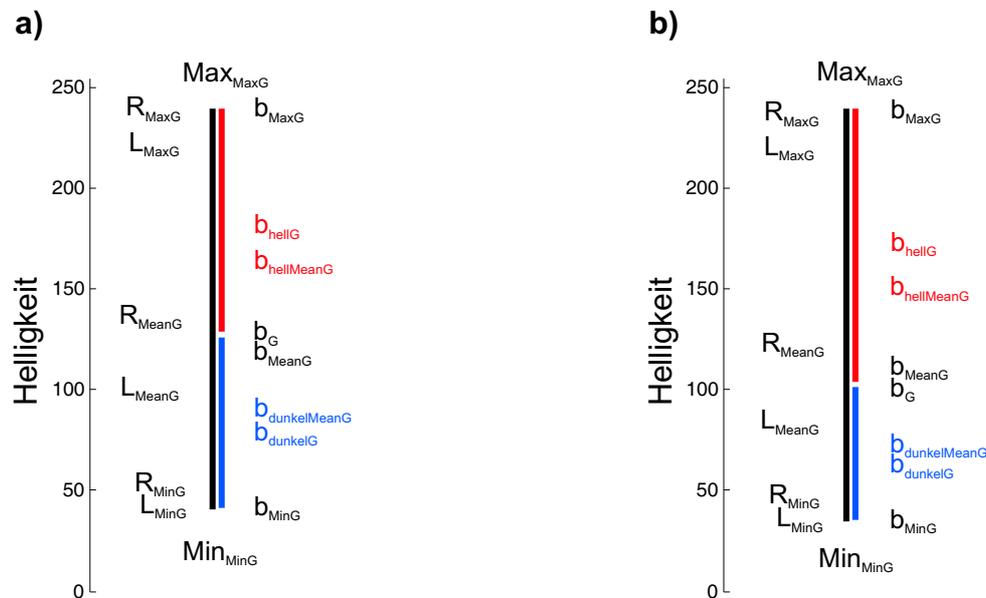
Dieser Abschnitt soll anhand eines sehr einfachen Beispiels eine Anwendung zur Ausnutzung der statistischen Kenngrößen illustrieren. Gegeben sind zwei Bilder einer Stereokameraanordnung, wobei sich aufgrund verschiedener Blendeneinstellungen, eventuell abweichender Verschlusszeiten und automatischer Kameraverstärkungen deutlich unterschiedliche Helligkeitsverteilungen ergeben haben. Die beiden Bilder stellen aufgrund der näherungsweise parallelen optischen Achsen und gleicher Brennweiten die gleiche Szene dar. Demnach kann davon ausgegangen werden, dass die Segmentierungen eines Bildausschnittes in beiden Bildern bis auf die absoluten Helligkeiten vergleichbare Ergebnisse liefern.

Die Abbildungen 4.31a bzw. 4.31b zeigen das linke bzw. rechte Bild einer Stereoanordnung und die dazugehörigen Histogramme ihrer Pixelhelligkeiten. Aus den Bildern und anhand der Histogramme ist zu erkennen, dass das rechte Bild im Vergleich zum linken Bild deutlich dunkler ist. Nun sei eine Abbildungsvorschrift gesucht, die den Vergleich der Helligkeiten z.B. zur Linken und zur Rechten von Kantenmerkmalen im linken und rechten Kamerabild erlaubt. Das Ergebnis einer solchen Abbildungsvorschrift ist in Abb. 4.31c zum Vergleich mit Abb. 4.31a und 4.31b dargestellt, soll aber erst später näher betrachtet werden.



**Abbildung 4.31:** Bilder unterschiedlicher Helligkeit einer parallelen Stereoanordnung mit den entsprechenden Histogrammen der Helligkeit: a) linkes Bild b) rechtes Bild c) an die Helligkeitsverteilung des linken Bildes angepasstes rechtes Bild nach der Helligkeitstransformation

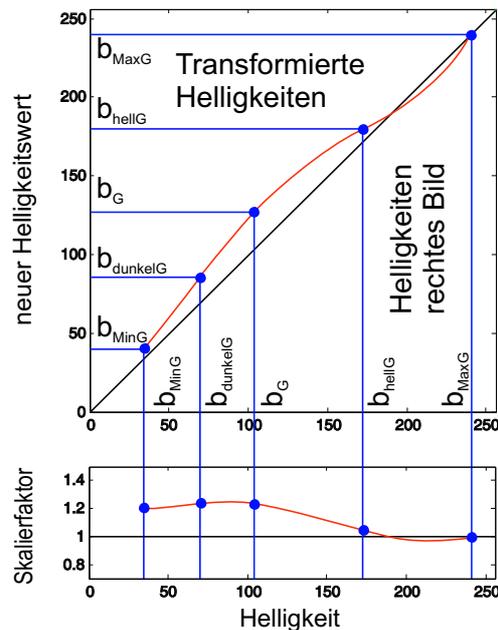
Zunächst werden die charakteristischen photometrischen Kenngrößen des in den Abbildungen 4.31a und 4.31b weiß umrandeten interessanten Bereichs in beiden Bildern durch Auswertung benachbarter Bildstreifen ermittelt. Abbildung 4.32 zeigt die daraus errechneten statistischen photometrischen Kenngrößen.



**Abbildung 4.32:** Statistische photometrische Kenngrößen a) des linken Stereobildes b) des rechten Stereobildes

Beide Bilder sind bei den sehr hellen Intensitäten in der Sättigung, wodurch die Abbildungsvorschrift für Helligkeiten im oberen Helligkeitsbereich verfälscht wird. Die Abbildungsvorschrift soll anhand von fünf Stützstellen bestimmt werden. Als Stützstellen dienen jeweils  $b_{MinG}$ ,  $b_{dunkelG}$ ,  $b_G$ ,  $b_{hellG}$  und  $b_{MaxG}$  des linken bzw. rechten Bildausschnittes. Der Verlauf zwischen den fünf Stützstellen wird mittels eines Polynoms 4.-ter Ordnung approximiert. Abbildung 4.33 zeigt sowohl die resultierende Abbildungsvorschrift für die Transformation der Helligkeiten, als auch den notwendigen Skalierfaktor, der für die Transformation je Helligkeit notwendig ist. Die Stützstellen sind durch blaue Punkte markiert.

In Abbildung 4.31c ist das Ergebnis der Transformation dargestellt. Es ist zu erkennen, dass durch die Transformation die Helligkeitsverteilung des rechten der des linken Bildes angepasst werden konnte. Diese Abbildungsvorschrift stellt lediglich eine grobe Näherung dar, zeigt aber, dass selbst durch diese einfache Vorgehensweise die Auswertung der Bildinformation z.B. für die Zuordnung von Merkmalen erleichtert werden kann.



**Abbildung 4.33:** Approximation der Abbildungsvorschrift für die Helligkeiten des rechten Bildes entsprechend der Helligkeitsverteilung des linken Bildes innerhalb des Bereichs  $[b_{MinG}; b_{MaxG}]$ .

#### 4.3.12 Zuordnung korrespondierender Merkmale

Das Ziel der vorhergehenden Abschnitte bestand darin, eine möglichst exakte, aussagekräftige Beschreibung eines Bildstreifens bzw. einer Menge paralleler Bildstreifen zu generieren, um damit die Möglichkeit zu schaffen, die Selektion aufgrund der Merkmalsbeschreibung und damit die Zuordnung von Merkmalen für eine Interpretation zuverlässiger zu machen. Eine Zuordnung korrespondierender Merkmale wird benötigt

- für die Verkettung benachbarter Merkmale zu Kantenverläufen oder zu Bereichen,
- für die Selektion eines Merkmals zur Bestimmung der Abweichung bezüglich einer erwarteten Position,
- für die Bestimmung eines Merkmalsflusses aus aufeinanderfolgenden Bildern einer Bildfolge,
- und um aus den Positionen in den Bildern eines Stereokamerapaares die Disparität korrespondierender Merkmale zu bestimmen, um durch Triangulation die Entfernung des zugrundeliegenden Objekts im Raum zu errechnen.

Die Fülle an Information bezüglich eines Segments bzw. einer Segmentgrenze erleichtert den für eine Zuordnung notwendigen Vergleich d.h. die Quantifizierung einer Ähnlichkeit

erheblich. Durch die richtungssensitive Kantenextraktion werden Segmentgrenzen genauer und zuverlässiger detektiert. Die lineare Beschreibung eines Helligkeitsverlaufs wird gegenüber dem Originalbild nicht durch irgendwelche richtungsabhängigen Spaltenmittlungen verfälscht und die L- und R-Werte geben sehr gut die lokale Helligkeit an einer Kante wieder.

Je nachdem, ob die homogenen Segmente oder ihre Segmentgrenzen als die Merkmale betrachtet werden, können unterschiedliche Kriterien für die Quantifizierung einer Ähnlichkeit herangezogen werden.

Für den Fall, dass eine Segmentgrenze das gesuchte Merkmal darstellt, werden die Art des Helligkeitsübergangs, die Position, die Kantenrichtung und das Verhältnis von L- und R-Wert verglichen. Zusätzlich können die mittleren Helligkeiten und Segmentlängen der anschließenden Segmente dazu verwendet werden, um ein Merkmal im Kontext benachbarter Merkmale zu betrachten.

Für den Fall, dass ein homogenes Segment als das gesuchte Merkmal betrachtet wird, werden die mittlere Segmenthelligkeit, die mittlere Segmentrichtung, die Segmentlänge, die Art der Helligkeitsübergänge an den Grenzen, die L- und R-Werte der begrenzenden gerichteten Kantenpunkte und eventuell die mittleren Helligkeiten der angrenzenden Segmente verglichen. Nachdem es durchaus vorkommen kann, dass sich z.B. in benachbarten Bildstreifen ein langes Segment im ersten Streifen aus zwei Segmenten im zweiten Streifen zusammensetzt, sollte ein Verkettungsverfahren für Segmente in der Lage sein, diese Möglichkeit in Betracht zu ziehen. Entsprechend sollte ein Verkettungsverfahren für gerichtete Kantenpunkte zu Kantenverläufen einzelne Lücken bei Bedarf überbrücken können, sofern die restlichen beschreibenden Parameter eine konsistente Interpretation zulassen.

Durch die Verringerung von Fehlzuordnungen wird die Genauigkeit und Robustheit der Zustandsschätzung gesteigert.

### **4.3.13 Gruppierung von gerichteten Kantenpunkten zu Kantenverläufen**

Die Bereichsauswertung mittels des HOFMANN-Operators entsprechend Abschnitt 4.3.7 liefert zum einen eine Liste von Segmenten mit begrenzenden gerichteten Kantenelementen, zum anderen wird eine Verwaltungsstruktur generiert, die eine schnelle Navigation unter Beibehaltung der Nachbarschaftsbeziehungen innerhalb der eindimensionalen Ergebnisdatenstruktur des in Suchstreifen zerlegten Bereichs ermöglicht. In der Verwaltungsstruktur wurde für jeden Suchpfad je ein Eintrag angelegt, der je nach Suchpfadrichtung die entsprechende Bildzeile bzw. Bildspalte des aktuellen Suchpfades, den Index des ersten Segments in der Ergebnisdatenstruktur und die Anzahl der Segmente im Suchpfad beinhaltet.

Als Beispiel soll hier die Verkettung von Kantenelementen bei vertikaler Suchpfadrichtung entsprechend der Bereichsauswertung nach Abbildung 4.21 näher erläutert werden. Die

Verkettung erfolgt analog zur Bereichsauswertung in Suchpfadrichtung von oben nach unten und von links nach rechts im Bild.

#### 4.3.13.1 Lineare Verkettung von gerichteten 2D-Kantenpunkten

Beginnend mit dem ersten Eintrag in der Ergebnisdatenstruktur wird kollinear zur Kantenpunkttrichtung im Abstand der Suchpfade nach einem Eintrag im Ergebnisintervall des benachbarten Suchstreifens gesucht. Für die Ermittlung einer Korrespondenz müssen die Eigenschaften eines Kandidatenkantenpunktes den gemittelten Eigenschaften der bereits verketteten Kantenpunkte ähnlich sein. Für die Bestimmung einer Ähnlichkeit können z.B. Distanzmaße wie die *Mahalanobis*-Distanz berechnet werden, die eine Gewichtung der Beiträge unterschiedlicher Parameter bei der Ermittlung der Distanz erlaubt. Mittels eines festgelegten Schwellenwertes kann dann eine Entscheidung über die Ähnlichkeit getroffen werden. Ein anderer Weg besteht darin, in Abhängigkeit von den zu vergleichenden Parametern Ähnlichkeitsintervalle zu definieren, wobei nur diejenigen Kantenpunkte als ähnlich bezüglich einer Gruppe bereits verketteter Punkte angesehen werden, deren beschreibende Parameter innerhalb der Intervalle liegen.

Findet sich ein ähnlicher Kantenpunkt, werden mittels der Gleichungen (A.20), (A.26) und (A.30) die Steigung  $a$ , der Achsabschnitt  $b$  und die Varianz  $Var$  der Abweichung der Kantenpunkte von der Ausgleichsgeraden als Maß für die Qualität der Ausgleichsgeraden berechnet. Weiterhin wird der neu verkettete Kantenpunkt als 'benutzt' markiert, so dass er nachfolgend bei der Verkettung anderer Kantenpunkte als Kandidat nicht mehr betrachtet werden muss und damit Rechenzeit eingespart werden kann.

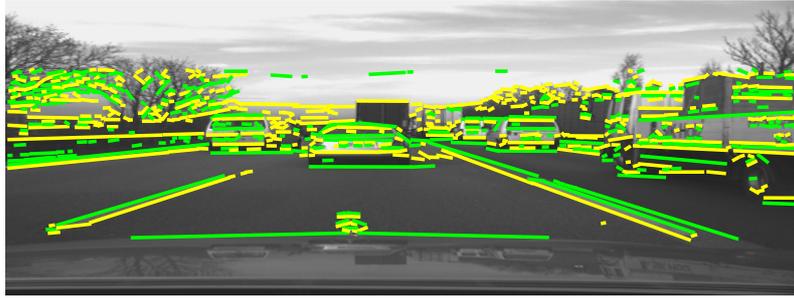
Findet sich im Ergebnisintervall eines Suchstreifens kein geeigneter Kandidat, dann können bis zu *AnzahlLuecken* Lücken durch das Verfahren überbrückt werden, bis das Verkettungsverfahren für den aktuellen Kantenverlauf als abgeschlossen betrachtet wird und mit dem nächsten noch unbenutzten Kantenpunkt als neuen Startpunkt eines Kantenverlaufs mit dem Verkettungsverfahren wieder begonnen wird. Das Verfahren endet, sobald kein Kantenpunkt mehr als Startpunkt oder Kandidat zur Verkettung zur Verfügung steht.

Ergebnis der Verkettung ist eine Liste von linearen Kantenverläufen die durch die Parameter entsprechend Tabelle 4.7 beschrieben werden. In Abbildung 4.34 ist die Verkettung der in Abschnitt 4.3.8 aus einem Grauwertbild in vertikalen Suchpfaden extrahierten gerichteten Kantenpunkte dargestellt. Ein weiteres Beispiel zur Gruppierung von gerichteten Kantenpunkten zu Kantenverläufen findet sich im Ergebnisteil in der Anwendung zur Grabenerkennung in Kapitel 8.2.6.

Bei Bedarf wäre es möglich, auch die Indices der verketteten Kantenpunkte, die den Zugriff auf die Einträge in der Ergebnisstruktur des HOFMANN-Operators ermöglichen, in einer Liste zu speichern, um zu einem späteren Zeitpunkt weitere beschreibende Parameter bestimmen zu können bzw. das Wissen über die Nachbarschaften verfügbar zu halten.

**Tabelle 4.7:** Datenstruktur für die Ergebnisse einer linearen Verkettung

uBegin, vBegin	Anfangspunkt im Bildkoordinatensystem
uEnd, vEnd	Endpunkt im Bildkoordinatensystem
du, dv	Richtungsvektor im Bildkoordinatensystem
AngleEdge	Winkel im Bildkoordinatensystem
a	Steigung im Bildkoordinatensystem
b	je nach Suchpfadrichtung u- bzw. v-Abschnitt im Bildkoordinatensystem
Len	Länge des linearen Kantenverlaufs
MeanL	mittlerer L-Wert
MeanR	mittlerer R-Wert
MeanSegPos	mittlere Segmentlänge in Suchpfadrichtung
MeanSegNeg	mittlere Segmentlänge entgegen der Suchpfadrichtung
NrPoints	Anzahl der verketteten Kantenpunkte
su, sv	Summe der u- bzw. v-Koordinaten der verketteten Kantenpunkte
suu, svv	Summe der Quadrate der u- bzw. v-Koordinaten der verketteten Kantenpunkte
suv	Summe der gemischten Terme
SumNorm	Summe der Gewichte bei variabler Gewichtung der verketteten Kantenpunkte



**Abbildung 4.34:** Lineare Verkettung gerichteter Kantenpunkte.

Die lineare Verkettung lässt Kantenpunkte nur dann als Kandidaten zu, wenn ihr orthogonaler Abstand bezogen auf den linearen Gesamtverlauf der Ausgleichsgeraden innerhalb eines Epsilonschlauches liegt, der durch einen erlaubten Schwellenwert für die Varianz  $Var$  festgelegt wird. Möchte man eine Gruppierung von Kantenpunkten entlang eines beliebigen glatten kontinuierlichen Verlaufs, so führt dieses Vorgehen nicht zum Ziel, da das Verfahren z.B. bei einem konstant gekrümmten Verlauf nach Überschreiten der Epsilon-schwelle abbricht. Im nächsten Abschnitt wird hierfür eine Erweiterung des Verkettungsverfahrens zur Gruppierung von Punktemengen durch lokale Verkettung beschrieben.

#### 4.3.13.2 Gruppierung von Kantenpunkten zu Punktemengen durch lokale Verkettung

Der Unterschied zum Verfahren des vorhergehenden Abschnitts besteht darin, dass nun ein Kandidatenkantenpunkt nicht mehr bezüglich der Gesamtmenge an bereits gruppierten Punkten bewertet wird, sondern lediglich bezüglich der in einem nachgeführten lokalen Fensters gelegenen Punkte die Ähnlichkeitsbedingungen erfüllen muss. Die Vorgehensweise für das lokale Fenster ist dagegen identisch zum vorhergehenden Abschnitt. Erfüllt ein neuer Kantenpunkt die Bedingungen zur Gruppierung, dann wird das lokale Fenster nachgeführt, so dass der in Verkettungsrichtung am weitesten entfernte Kantenpunkt aus der lokalen Verkettung herausfällt. Das Gruppierungsverfahren terminiert, sobald kein geeigneter Kantenpunkt mehr für die Gruppierung gefunden werden kann. Für die Gesamtmenge der gruppierten Punkte werden am Ende die Parameter entsprechend Tabelle 4.7 berechnet. Da als Kriterium für einen Abbruch der Gruppierung der Epsilonschlauch nur auf das lokale Fenster angewandt wird, kann und darf die Varianz der Abweichung aller gruppierten Kantenpunkte von der resultierenden Ausgleichsgerade größer sein.

#### 4.3.13.3 Gruppierung von homogenen Segmenten zu Bereichen

Nach denselben Prinzipien wie in 4.3.13.1 und 4.3.13.2 können auch Segmente bei ähnlichen Helligkeitsverläufen zu homogenen Bereichen verknüpft werden. Beispielsweise werden bei der Bestimmung potentieller Fremdfahrzeuge aufgrund des Schattens unter einem

Fahrzeug gerichtete Kantenpunkte zu Linien und die daran anschließenden Segmente zu Bereichen verknüpft. Eine Schwierigkeit besteht darin, eine Beschreibung zu finden, welche die räumlichen Bezüge der Bereiche zueinander erhält.

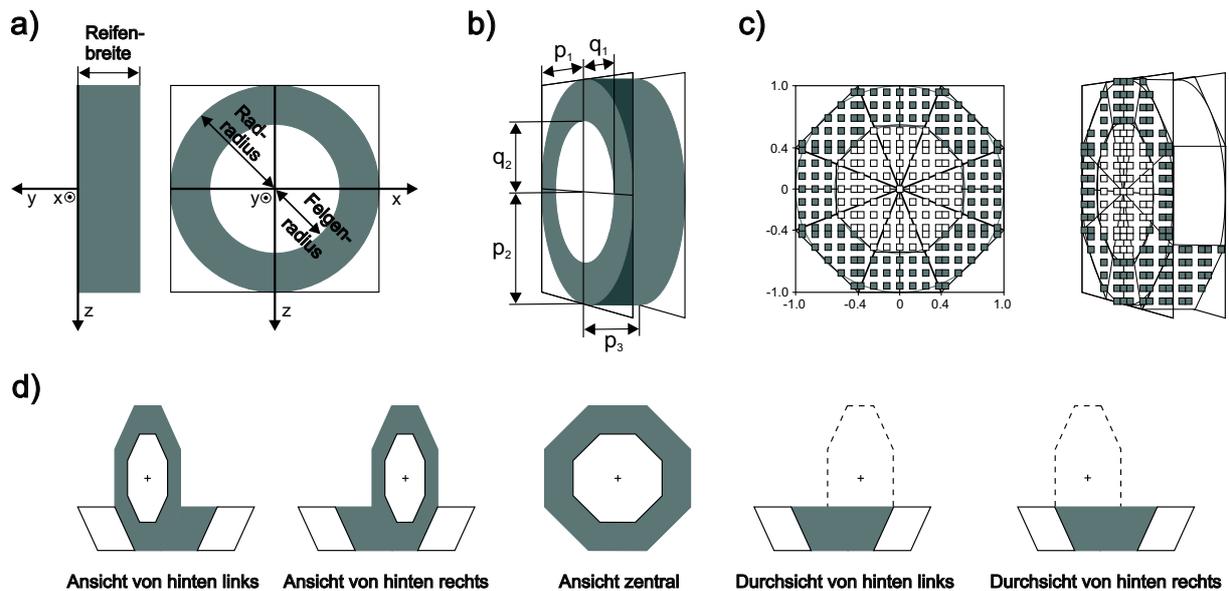
## 4.4 Explizite Suche nach Subobjekten mittels generisch erzeugter Muster am Beispiel der Raderkennung

Die Erkennung komplexer Subobjekte mittels Streifenoperatoren, wie z.B. in Abschnitt 4.1 oder 4.3 vorgestellt, erweist sich bei hohem Detaillierungsgrad und gleichzeitig geringer Auflösung eines Subobjekts im Bild häufig als äußerst schwierig. Die Parametrisierungen der Streifenoperatoren könnten zwar mittels generischer Formmodelle koordiniert werden, die Zusammenschau der Messergebnisse wäre allerdings sehr mühsam, sobald die Dimensionen der zu extrahierenden Merkmale eines Subobjekts an der Grenze oder sogar unterhalb der möglichen Auflösung der Streifenoperatoren liegen.

Beispielsweise besitzen Radkappen oder Alufelgen durch das Design ihrer Speichen einen hohen Detaillierungsgrad. Die Fülle der möglichen Designs führt dazu, dass der Aufwand für eine detaillierte Modellierung in keinem Verhältnis zum Nutzen steht. Weiterhin sind die Details für die prinzipielle Aufgabe der Raderkennung eher unwesentlich. Deshalb wird hier eine sehr einfache Beschreibung der photometrischen Eigenschaften gewählt: Der Reifen des Rades sei dunkel und die Radkappe oder Alufelge eher hell.

In solch einem Fall macht es Sinn, entfernungs- und aspektabhängig eine angepasste Pixelmaske zu erzeugen, die die wesentlichen Charakteristika des Subobjekts repräsentiert, um möglichst gut mit dem erwarteten Erscheinungsbild eines Subobjekts im Bild zu korrelieren. Mittels generischer Formmodelle in 3-D wird diese durch Projektion ins Bild für den jeweils aktuellen Blickwinkel erzeugt, wobei wiederum aus verschiedenen Modellen je nach der Klasse der Ansicht und Sichtbarkeit das passende Modell ausgewählt wird. Solch eine an die Erwartung angepasste Korrelation hat zum Ziel die genaue Position innerhalb eines Signalverlaufs, hier Suchpfads, aufgrund einer Maximumsuche zu bestimmen, ohne dem einzelnen Pixel innerhalb einer Maske besonderes Gewicht zu geben. Die Summe der Übereinstimmungen ist ausschlaggebend, wobei die Nachbarschaftbeziehungen im Pixelmuster implizit berücksichtigt werden.

Das Rad soll hier als Beispiel eines Subobjekts dienen. Abbildung 4.35a zeigt die typischen Parameter für die Modellierung eines Rades: die Reifenbreite, den Radius des Rades und der Felge. Bei den nachfolgenden Überlegungen werden die erwarteten Maße der Felge ausgehend von den Maßen des Rades über das Verhältnis der Radien zwischen Felge und Rad gebildet. Das Koordinatensystem des Subobjekts sei kollinear mit dem übergeordneten Hauptobjekt. Sein Ursprung befinde sich auf der Außenhülle des Hauptobjektes, d.h. hier auf dem das Fremdfahrzeug einhüllenden Quader. Näheres zur Formmodellierung von Fremdfahrzeugen findet sich in Kapitel 5.2.1.



**Abbildung 4.35:** Aspektabhängige Erzeugung von Pixelmustern zur Raderkennung:  
a) generisches Radmodell in 3-D,  
b) generisches Radmodell nach Projektion ins Bild,  
c) Approximation eines Rades durch ein Achteck und Festlegung der hellen und dunklen Pixelpositionen des Musters relativ zum Fahrzeug-äußeren Radmittelpunkt,  
d) fünf verschiedene Klassen von parametrisierbaren Radmustern entsprechend den Aspektbedingungen und den Pixellängen  $p_1$ ,  $p_2$  und  $p_3$  im Bild

In Abbildung 4.35b ist die Projektion eines Rades aufgrund seiner aktuellen Position im Raum und der aktuellen Ansicht der Kamera auf das Subobjekt dargestellt. Aus der generischen 3-D Beschreibung werden durch die Projektion explizite Pixelmaße im Bild erzeugt. Das Erscheinungsbild des Rades kann hier mittels der in der Abbildung eingezeichneten Parameter  $p_1$  und  $p_2$  für die Länge der Hauptachsen der Radellipse und  $p_3$  für die Reifenbreite beschrieben werden. Die Parameter  $q_1$  und  $q_2$  ergeben sich näherungsweise durch Multiplikation der Parameter  $p_1$  und  $p_2$  mit dem Verhältnis der Radien zwischen Felge und Rad.

Abbildung 4.35c zeigt die Approximation der Gestalt eines Rades durch ein Achteck. Die Knickstellen des Polygons sind jeweils bei 40% und 100% des Radradius. Das Pixelmuster unterteilt sich in Pixelpositionen, die hellen Bereichen zugeordnet sind, und Pixelpositionen, die dunklen Bereichen zugeordnet sind. Die Pixelpositionen werden stets relativ zum Ursprung des Subobjektkoordinatensystems im Bild angegeben.

Die maximale Anzahl der Pixel im generischen Muster kann begrenzt werden. Die Schrittweite jeweils entlang der Koordinatenachsen ergibt sich aus dem Quotienten der erwarteten Ausdehnung zur erlaubten Auflösung, wobei nur Schrittweiten größer gleich eins zugelassen werden. Dies ist in der Abbildung 4.35c angedeutet. Bei Schrittweiten größer

eins ergibt sich eine Unterabtastung des erwarteten Erscheinungsbildes. Hierdurch wird gewährleistet, dass unabhängig von der erwarteten Ausdehnung die für die Korrelation benötigte Rechenzeit begrenzt bleibt und dennoch die Nachbarschaftsbeziehungen ausgenutzt werden.

Die Korrelation erfolgt entlang eines Suchpfads. Relativ zu den Positionen entlang des Suchpfads werden jeweils die Helligkeitswerte aller 'hellen' bzw. 'dunklen' Pixelpositionen im Bild aufaddiert und durch die jeweilige Anzahl an 'hellen' bzw. 'dunklen' Positionen dividiert. Die Differenz der beiden normierten Summen entspricht der Maskenantwort an einer bestimmten Position des Suchpfads. Der Verlauf der Maskenantwort wird anschließend nach einem Maximum durchsucht. Das Maximum entspricht dabei der vermuteten Position des Subobjekts entlang des Suchpfads.

Das hier implementierte Verfahren geht zunächst von tendenziell horizontalen Suchpfaden für die Korrelation aus. An der gefundenen vermuteten Subobjektposition im Suchpfad wird anschließend noch eine Optimierung entlang eines vertikalen Suchpfads durchgeführt. Die Fußpunkte der Räder sind wichtige Merkmale für die Bestimmung der Ausrichtung und der Entfernung eines Fremdfahrzeugs.

In Abbildung 4.35d sind verschiedene Klassen an parametrisierbaren Radmustern je nach Aspektklasse und erwarteter Sichtbarkeit dargestellt. Zusätzlich zur Abbildung 4.35c sind bei einigen der Masken links und rechts 'helle' Pixelpositionen ergänzt worden. Diese Bereiche sind zwar physikalisch keine Bestandteile des Subobjekts, stellen aber erfahrungsgemäß ein charakteristisches Merkmal dieser Subobjekte dar. Für die Ansichten von vorne können die dargestellten Masken spiegelbildlich verwendet werden. Näheres zur Klassifizierung der Aspektbedingungen findet sich im Abschnitt 5.2.2.

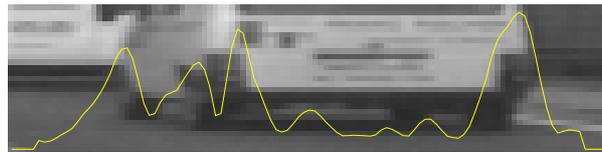
## 4.5 Suche nach vertikalen Häufungen unspezifischer Helligkeitsübergänge innerhalb eines horizontalen Bildstreifens

Für die Detektion von vertikalen Häufungen unspezifischer Helligkeitsübergänge innerhalb eines horizontalen Bildstreifens wurde in [THOMANEK 96] ein Bildoperator ähnlich einer ternären Maske vorgeschlagen, der als neuen Kerngedanken die Betragsbildung der zeilenweise berechneten Maskenantworten vor ihrer spaltenweisen Summation vorsieht. Der Operator - im folgenden als *Betragsoperator* bezeichnet - liefert ein robustes Merkmal zur Bestimmung der Position einer äußeren vertikalen Berandung eines Fahrzeugs. Die Korrelation  $K$  an der Position  $(u, v)$  im Bildkoordinatensystem eines Grauwertbildes  $B$  ergibt sich bei Verwendung einer Maske mit einer Nullspalte, einer halben Maskenhöhe  $h$  und einer Länge  $b$  des vor- bzw. nachlaufenden Teils der Maske aus

$$K(u, v) = \sum_{i=-h}^h \left| \sum_{j=1}^b (B(u+j, v+i) - B(u-j, v+i)) \right|. \quad (4.18)$$

Die Position  $(u, v)$  wird entlang eines tendenziell horizontalen Suchpfads variiert und die Korrelationen  $K(u, v)$  ermittelt. Die Parameter der Korrelationsmaske und des Suchpfads werden dabei anhand der erwarteten Position und Ausdehnung eines gesuchten vertikalen Helligkeitsübergangs im Bild festgelegt. Anschließend wird der Signalverlauf nach seinem Maximum durchsucht. In der Regel entspricht die Position des Maximums bei ausreichender Nähe zur erwarteten Position der gesuchten Position der äußeren Berandung.

Die Abbildung 4.36 zeigt beispielhaft den Signalverlauf der Korrelation  $K$  bei einer Maskenlänge  $b = 5$  und der Höhe  $h = \frac{\text{Zeilen}-1}{2}$  bezüglich des Bildausschnitts. Das Ver-



**Abbildung 4.36:** Verlauf der Korrelation bei Verwendung des Betragsoperators

fahren kann sehr einfach auch für die Suche nach horizontalen unspezifischen Helligkeitsübergängen adaptiert werden.



# Kapitel 5

## Umfeldwahrnehmung mittels integraler räumlich-zeitlicher Modelle

Ein wesentliches Ziel der maschinellen Wahrnehmung besteht darin, eine interne Repräsentation der realen Welt aufzubauen und diese entsprechend den Veränderungen der realen Welt zu aktualisieren. Zur Lösung einer bestimmten Aufgabe wird diese Repräsentation in der Regel auf eine endliche Menge für relevant erachteter Objekte beschränkt, für die einem Erkennungsexperten a priori Wissen in Form von Modellvorstellungen bezüglich ihrer Gestalt und Dynamik beigebracht wurde. Dieses Wissen wird zu Beginn des Erkennungsvorgangs für die initiale Detektion von Objekten genutzt und anschließend bei der Verfolgung der Objekthypothesen für die Eingrenzung des Suchraums verwendet.

### 5.1 4-D Ansatz

Die Grundlage der Arbeiten zur visuellen Wahrnehmung am ISF stellt der 4-D Ansatz nach [MEISSNER 82], [MEISSNER&DICKMANN 83], [DICKMANN 87], [WÜNSCHE 87], [DICKMANN&GRAEFE 88] dar. Das Kürzel 4-D soll hierbei die Modellierung und Schätzung der Form von Objekten in 3-D und ihrer Dynamik, d.h. Lage und Veränderung der Lage im Raum, über der Zeit (1-D) zum Ausdruck bringen. Der zentrale Gedanke besteht in der Modellierung der Kontinuität der beschreibenden Parameter in Raum und Zeit, wodurch das Wissen über das *Hier-und-Jetzt* unter Berücksichtigung der statistischen Kenngrößen der Messfehler und der auf das System einwirkenden Störgrößen je nach Qualität der Modellvorstellung in einem mehr oder weniger ausgedehnten räumlich-zeitlichen Fenster extrapoliert werden kann.

Durch Projektion einer extrapolierten Modellvorstellung in den Sensordatenraum, welche der Erwartung an eine veränderte Lage von Merkmalen im Datenbereich eines Sensors ent-

spricht, ist es möglich, den Suchraum für die Suche nach Merkmalen einzuschränken. Damit wird die Gefahr von Fehlzuordnungen verringert und gleichzeitig werden die vorhandenen Rechenressourcen sehr effizient genutzt. Die Abweichung zwischen der Erwartung und der Messung wird zur Korrektur der Modellvorstellung und damit zur Nachführung der Objektmodelle entsprechend den Veränderungen der realen Welt verwendet.

### 5.1.1 Rekursive Zustandsschätzung

Die Schätzung der Modellparameter bzw. Zustandsgrößen erfolgt dabei rekursiv, d.h. der aktuelle Schätzwert beinhaltet jeweils alle seit dem Start der rekursiven Schätzung durch die Messungen eingebrachte Information. Im Vergleich zu nicht rekursiven Schätzverfahren kann deshalb eine Speicherung zeitlich zurückliegender Messwerte entfallen. Jede Messung wird durch ihren Wert und das als bekannt angenommene Messrauschen charakterisiert. Dieses sei mittelwertfrei und normalverteilt. Je kleiner das Messrauschen, desto größer ist das „Vertrauen“ in die Messung.

Für die rekursive Zustandsschätzung wird ein *Erweitertes Kalman Filter (EKF)* verwendet. Dieses stellt eine Erweiterung des linearen Kalman Filters auf nichtlineare zeitdiskrete Systeme dar. Ein nichtlineares Messmodell und eine nichtlineare Systemgleichung werden hierzu bezüglich der aktuellen Schätzung linearisiert. Für die Zustandsprädiktion werden dabei die nichtlineare Systemgleichung und für die Bestimmung der erwarteten Messwerte aufgrund des prädizierten Zustandsvektors das nichtlineare Messmodell verwendet. In [WÜNSCHE 87] und [MYSLIWETZ 90] werden die Vor- und Nachteile verschiedener Varianten des Kalman Filters hinsichtlich ihrer numerischen Stabilität und Recheneffizienz diskutiert, die dann zur Auswahl der seit Jahren am ISF genutzten Variante des EKF nach [BIERMAN 77] geführt hat. Eine kurze Beschreibung der prinzipiellen Funktionsweise des Erweiterten Kalman Filters findet sich im Anhang B.

### 5.1.2 Messmodell

Die Vorgehensweise nach dem 4-D Ansatz dient der Fusion von Messdaten eines oder mehrerer Sensoren, um daraus *eine* konsistente Repräsentation der Objekte bzw. Subjekte der Welt zu erhalten. Die Verbindung zur realen Welt wird mittels der im System vorhandenen Sensorik hergestellt. Hierbei ist es von zentraler Bedeutung, die Abbildungsvorschrift von Merkmalen im Raum auf die Sensordaten gut zu kennen. Denn erst durch ihre Kenntnis wird die Interpretation der Sensordaten ermöglicht. Eine solche Abbildungsvorschrift wird *Messmodell* (bezüglich eines Sensors) genannt. Die Qualität des Messmodells ist wesentlich für den erreichbaren Grad der Übereinstimmung zwischen der internen Vorstellung und der Realität.

Bei der visuellen Wahrnehmung ist dieses Wissen der Abbildungseigenschaften als Ergebnis eines Kalibriervorgangs im Kameramodell einer jeden Kamera im System abrufbar. Der Kalibriervorgang wird im Kapitel 3.3.1ff beschrieben. Für andere Sensoren gilt im

Prinzip dasselbe. Mittels entsprechender Kalibrierverfahren werden die Parameter eines geeigneten Sensormodells so bestimmt, dass sich eine möglichst exakte Beschreibung der Abbildung auf die Sensordaten ergibt. Solche Messmodelle sind dabei häufig stark nicht-linear, wodurch eine Inversion in der Regel rechenintensiv bzw. oft nur aufgrund stark einschränkender Annahmen und Näherungen überhaupt durchführbar ist. Deshalb wird versucht, diese möglichst zu vermeiden.

Anstatt durch Inversion des Messmodells aus der Messung direkt auf eine Schätzgröße schließen zu wollen, wird beim 4-D Ansatz die Vorstellung über ein Objekt zu einem Zeitpunkt  $t_{k-1}$  (z.B. eine initiale oder geschätzte Gestalt, Position und Dynamik eines Objekts im Raum) unter Ausnutzung des Wissens über die Dynamik eines Objekts für den aktuellen Messzeitpunkt  $t_k$  extrapoliert, mittels des Messmodells in den Sensordatenraum abgebildet und dort mit den aktuellen Sensordaten verglichen. Die Abweichung von den gemessenen Sensordaten wird zur Korrektur der Vorstellung, d.h. z.B. der geschätzten Gestalt, Position und Dynamik im Raum, hergenommen. Die rechenintensive bzw. z.T. nicht durchführbare Inversion des Messmodells wird damit in eine Modellanpassung umgewandelt.

### 5.1.3 Generisches Formmodell und Dynamikmodell

Jedes Objekt besitzt ein *generisches Formmodell* und ein *Dynamikmodell*. Dabei bedeutet „generisch - die Gattung betreffend“ [DUDEN 01].

Das Formmodell dient zur Beschreibung der Gestalt eines Objekts. Es verfügt über verschiedene Freiheitsgrade zur Anpassung des Modells an ein zu vermessendes reales Objekt. In Kapitel 5.2.1 wird das generische Formmodell am Beispiel des Pkw-Formmodells im Vorgriff auf Kapitel 7 beschrieben. In Kapitel 5.4 wird eine Kurzübersicht zur Modellierung der Fahrspur und zur Modellierung der Eigenbewegung relativ zur Fahrspur gegeben, und in Abschnitt 8.2.7 wird die Modellierung eines rechteckigen Grabens als Beispiel eines negativen Hindernisses beschrieben.

Das Dynamikmodell dient zur Beschreibung der Position und der Veränderung der Position eines Objekts relativ zu seinem Bezugskordinatensystem über der Zeit. In ihm können typische Verhaltensmuster mittels typischer Zustandsgrößen in entsprechenden Systemgleichungen repräsentiert werden. In Abschnitt 5.5 werden die Koordinatensysteme zur visuellen Fahrspur- und Objekterkennung beschrieben.

Das jeweilige Bezugskordinatensystem eines Objekts wird auch sein *Vaterkoordinatensystem* genannt. Dem entsprechend ist aus der Sicht des Vaterkoordinatensystems das Objektkordinatensystem ein *Sohnkoordinatensystem*. Ein Vater kann dabei mehrere Söhne, jedoch ein Sohn stets nur einen Vater besitzen.

### 5.1.4 Numerische Bestimmung der partiellen Ableitung des Messmodells

Unter Verwendung dieser objektspezifischen räumlich-zeitlichen Modelle wird eine interne Repräsentation der realen Welt den Veränderungen der realen Welt nachgeführt. Die rechnerinternen Modellvorstellungen erlauben es, eine Vorhersage über die Veränderung der Objektkonstellation zu errechnen.

Die Merkmale im Raum werden unter Verwendung der jeweiligen Messmodelle auf den Datenbereich eines Sensors projiziert. Durch Variation der beschreibenden Parameter der Objektmodelle und Beobachtung der Abweichung im Datenbereich eines Sensors wird der erwartete Einfluss des variierten Parameters auf die Sensordaten bestimmt. Die Division durch die Variation ergibt eine numerische Näherung der partiellen Ableitung für den jeweiligen Eintrag in der Jacobi-Matrix (der partiellen Ableitung  $\mathbf{C}_k$  des Messmodells).

Diese geht in Gleichung (B.7) in die Berechnung der Kalmanverstärkung  $\mathbf{K}_k$  ein, welche die Anteile der Residuen festlegt, die im Korrekturschritt jedem Element des Zustandsvektors zugeschlagen werden.

### 5.1.5 Festlegung der Filterparameter

Unter der Annahme, dass die den Zustandsgrößen überlagerten Störgrößen stochastisch unabhängig voneinander sind, besitzt die Kovarianzmatrix des Systemrauschens  $\mathbf{Q}$  Diagonalform. Ihre Einträge sind dann idealerweise so zu wählen, dass die Quadratwurzeln der Diagonalelemente den erwarteten Streuungen der modellierten Zustandsgrößen um die tatsächlichen physikalischen Werte entsprechen.

Die Gewichtung der Messung gegenüber der Modellvorstellung wird anschaulich über das Verhältnis zwischen einer angenommenen Kovarianz des Systemrauschens  $\mathbf{Q}$  bzw.  $q_i$  für die  $i$ -te von  $n$  Zustandsgrößen und einer bekannten oder angenommenen Messfehlerkovarianz  $\mathbf{R}$  festgelegt.

Ein kleiner Wert für  $q_i$  drückt demnach ein großes Vertrauen in die durch das Systemmodell prädizierte  $i$ -te Zustandsgröße aus und verringert damit den Einfluss der Messung auf die Schätzung. D.h. das Folgeverhalten einer Schätzgröße gegenüber der Messung wird umso träger, je kleiner (allerdings größer null) ein  $q_i$ -Wert gewählt wird. Hingegen wird durch einen großen Wert der Einfluss der Messung bei der Nachführung der Zustandsgröße erhöht. Wird  $q_i$  zu klein gewählt, dann kann das Filter bei der Zustandsinnovation der Änderung im Messwert nicht folgen, das Residuum wird von Schätzzyklus zu Schätzzyklus steigen. Das heißt, die Erwartung entfernt sich immer weiter vom tatsächlichen Wert - Schätzung und Messung divergieren. Die Einstellung der Filterparameter wird als *Filtertuning* bezeichnet.

In der Regel wird zunächst davon ausgegangen, dass alle Pixel eines Bildes das gleiche Sensorrauschen aufweisen. Abhängig vom Messverfahren, das für die Extraktion eines

Merkmals zum Einsatz kommt, ist es jedoch sinnvoll, zusätzlich den unterschiedlichen Genauigkeiten der Verfahren Rechnung zu tragen. Dies erfolgt, indem bei der *sequentiellen Innovation* das einer Messung zugeordnete Messrauschen an das verwendete Verfahren angepasst wird. Für die sequentielle Innovation muss die Voraussetzung erfüllt sein, dass die Messungen  $y_1, \dots, y_m$  des Messvektors  $\mathbf{y}_k$  unabhängig voneinander sind. Weitere Anmerkungen hierzu finden sich im Anhang B.2.

## 5.2 Repräsentation von a priori Expertenwissen zur Wahrnehmung

Das a priori Expertenwissen beinhaltet das Wissen über die Gestalt eines Objekts in 3-D, die Art seiner Merkmale, ihre relative Lage, Ausrichtung und Ausdehnung und die Verfahren zu ihrer Extraktion und zu ihrer Selektion. Weiterhin stellt das Wissen über die Dynamik eines Objektes und seine Repräsentation innerhalb eines Umfeldmodells aufgrund der prinzipiellen Anordnung und der Abhängigkeiten der beschreibenden Koordinatensysteme modelliertes a priori Expertenwissen dar.

Die Gestalt eines Objektes wird dabei häufig implizit mit der Beschreibung der Einhüllenden durch ein detailliertes Gittermodell gleichgesetzt. Allerdings sind viele dieser Objektanten häufig gar nicht beobachtbar, zum anderen sind nicht alle Merkmale zwingend vom Typ Kante. Der Schwerpunkt der Beschreibung liegt bei einigen Merkmalen z.B. auf den photometrischen Eigenschaften. Es können zwar für jedes Objekt oder Subobjekt Gittermodelle generiert werden, allerdings gibt es für bestimmte Merkmale effizientere Methoden diese zu extrahieren, wie am Beispiel des Subobjekts Rad bereits in Abschnitt 4.4 erläutert wurde.

Der Detaillierungsgrad eines Formmodells hängt von den Anforderungen an eine eventuell durchzuführende Klassifikation ab. Für das Verfolgen einer Objekthypothese sind häufig nur wenige spezielle, aber signifikante Merkmale zuverlässig zu detektieren. Die Klassifikation steht nicht im Vordergrund dieser Arbeit, deshalb wird im folgenden Abschnitt ein einfaches Formmodell eines Pkws als Beispiel eines generischen Formmodells im Vorgriff auf Kapitel 7 präsentiert.

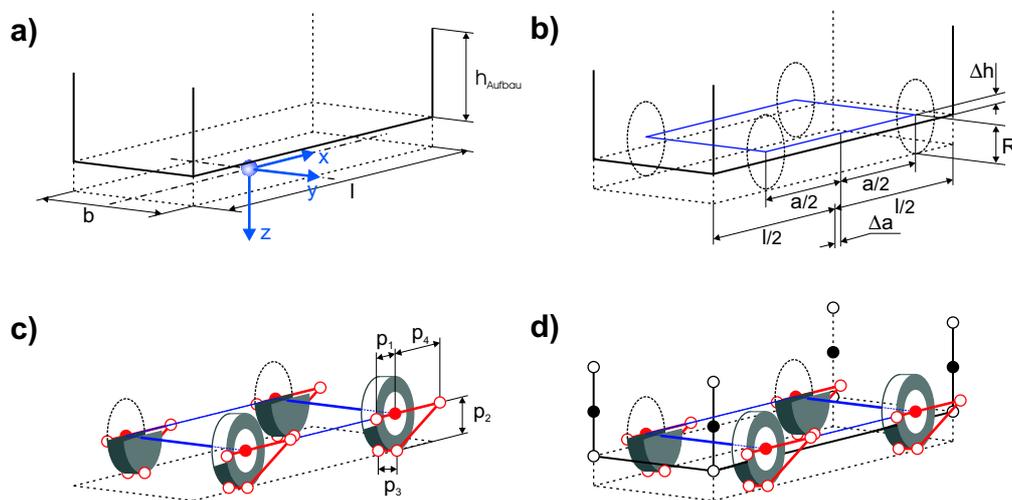
### 5.2.1 Generisches Formmodell eines Fremdfahrzeugs

In den Abbildungen 5.1a bis 5.1d ist der Aufbau des Formmodells für Personenkraftwagen beschrieben. Das Fahrzeugkoordinatensystem wird entsprechend Abbildung 5.1a festgelegt. Es ist ein rechtshändiges orthogonales Koordinatensystem entsprechend der Luftfahrtnorm [LN9300/90] mit der z-Achse in Richtung des Schwerevektors normal zur Bodenebene. Die x-Achse verläuft parallel zur Fahrzeuginnenachse vom Heck zur Fahrzeugfront ausgerichtet und die y-Achse bei Heckansicht von links nach rechts. Der Ur-

sprung des Koordinatensystems liegt in der Bodenebene im Schnittpunkt der Symmetrieachsen in Längs- und Querrichtung. In diesem Koordinatensystem wird die Form und Lage von Merkmalen in 3-D modelliert.

Es wird ein einhüllender Quader definiert, der mittels der Parameter Länge  $l$  und Breite  $b$  direkt parametrisiert werden kann. Aus der Summe der Höhe des Fahrzeugaufbaus  $h_{Aufbau}$  und dem Radius der Räder  $R$  abzüglich der Abweichung des Achshöhe von der Unterkante des Aufbaus  $\Delta h$  ergibt sich die Höhe des Quaders  $h = h_{Aufbau} - \Delta h + R$ .

Weiterhin werden ein Parameter  $a$  für den Achsabstand und  $\Delta a$  für die Abweichung zur Fahrzeugmitte in Längsrichtung eingeführt. Der Achsabstand  $a$  und die Abweichung  $\Delta a$  werden proportional zur Fahrzeuglänge  $l$  definiert, so dass bei einer Anpassung der Fahrzeuglänge die Relationen erhalten bleiben. In den Abbildungen 5.1a und 5.1b sind die entsprechenden Formparameter eingetragen.



**Abbildung 5.1:** Aufbau des generischen Formmodells eines Fremdfahrzeugs: a) Ausmaße  $l$ ,  $b$  und Aufbauhöhe  $h_{Aufbau}$ , b) Achsabstand  $a$ , Abweichung von der Symmetrieachse  $\Delta a$  und  $\Delta h$  zwischen Radachshöhe und Aufbau, c) 3-D  $\rightarrow$  2-D Parametrisierung für die Generierung eines Radmusters in Abhängigkeit vom erwarteten Erscheinungsbild und d) vollständiges 3-D Modell mit Messpunkten auf den vertikalen Quaderkanten

In Abbildung 5.1c sind um die Räder herum 3-D Punkte eingezeichnet. Gefüllte Kreise symbolisieren dabei das erwartete Zentrum eines zu vermessenden Merkmals. Kreise mit weißem Kern dienen zur Parametrisierung des Bildverarbeitungsoperators zur Erkennung von Rädern. Dazu werden die 3-D Punkte in das Bild der jeweils aktuell auszuwertenden Kamera projiziert und daraus die 2-D Pixelabstände  $p_1$ ,  $p_2$  und  $p_3$  bestimmt. Dann wird in Abhängigkeit von den Aspektbedingungen entsprechend Abschnitt 5.2.2 ein geeignetes Radmodell nach Abschnitt 4.4 ausgewählt und unter Verwendung dieses Radmodells das Korrelationsmuster entsprechend dem erwarteten Erscheinungsbild des Rades mit den

Parametern  $p_1$ ,  $p_2$  und  $p_3$  erzeugt. Die beiden ins Bild projizierten Punkte, die die Strecke  $p_4$  begrenzen, legen den halben 2-D Suchbereich bzw. den erlaubten Toleranzbereich zur Linken und zur Rechten ausgehend vom Zentrum fest. Der Suchpfad verläuft dabei von links nach rechts im Bild, kollinear zur Strecke  $p_4$ .

In Abbildung 5.1d sind zusätzlich an den Ecken des Fahrzeugaufbaumodells 3-D Punkte definiert, welche die Parametrisierung der Bildauswertung zur Verfolgung von vertikalen Kanten steuern. Hierbei stellt der gefüllte Kreis wiederum das Zentrum des zu vermessenden Merkmals dar. An den Ecken des Quaders kommt in Abhängigkeit von der Sichtbarkeit und den Aspektbedingungen der Betragsoperator nach Abschnitt 4.5 zum Einsatz. Die vertikale Begrenzung des Suchbereichs erfolgt anhand der Positionen der ins Bild projizierten 3-D Parametrierpunkte. Die horizontale Ausdehnung wird jeweils explizit so errechnet, wie ein in Metern vorgegebener Toleranzbereich orthogonal zum Sehstrahl durch Projektion ins Bild Pixel überdecken würde.

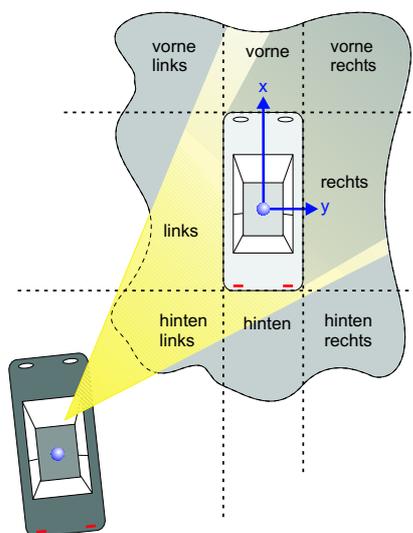
## 5.2.2 Klassifizierung der Aspektbedingung

Nach [DUDEN 01] bedeutet der Begriff Aspekt „... <lat.; >das Hinsehen<>: 1. Blickwinkel, Blickpunkt, Betrachtungsweise, Gesichtspunkt ...“. Die Klassifizierung der Aspektbedingungen legt hierbei Bereiche fest, in denen die Blickwinkel bzw. Blickpunkte auf ein Objekt auf entsprechende Klassen von Aspekten abgebildet werden. Eine derartige Klassifizierung der Aspektbedingungen der Sensoren gegenüber den repräsentierten Objekten wird beispielsweise durchgeführt, um eine zielgerichtete Auswahl von Formmodellen, z.B. für ein Subobjekt Rad, zu ermöglichen oder die Zustandsschätzung in Abhängigkeit von der Beobachtbarkeit bestimmter relevanter Merkmale zu steuern. Eine Aspektbedingung wird dabei immer zwischen einem bestimmten Sensor und einem bestimmten Objekt dadurch ermittelt, dass die Position des Sensors im Objektkoordinatensystem klassifiziert wird.

Die in der Abbildung 5.2 dargestellte Unterteilung ist speziell für Objekte mit grob quaderförmiger Einhüllender konzipiert, wobei bei der Klassifikation von einer Einbauhöhe der Sensoren innerhalb des Bereichs zwischen Bodenebene und Oberkante des Objektquaders ausgegangen wird, so dass Ansichten „von oben“ und „von unten“ unberücksichtigt bleiben können.

Bei der Klassifizierung nach Abbildung 5.2 sind die Klassen aufgrund der sichtbaren Seiten des Quaders definiert. Diese lassen eine entsprechende Fallunterscheidung bezüglich der sichtbaren Objektseitenflächen zu, welches z.B. bei der Festlegung von Verdeckungspolygonen in Abschnitt 5.2.4 ausgenutzt wird.

Die Klassifizierung entsprechend der 8-Teilung ist z.B. für den Fall des Aspekts „hinten links“ nicht ausreichend, um zu entscheiden, ob die Schätzung der Länge  $l$  oder der Breite  $b$  besser konditioniert ist. Zur Steuerung der Innovation muss objektklassenspezifisch definiert werden, unter welcher Ansicht und aufgrund welcher Merkmalkombinationen die Innovation welcher Zustandsgrößen sinnvoll und erlaubt ist.



**Abbildung 5.2:** Klassifizierung der Aspektbedingungen zwischen Objekt und Sensor aufgrund der Position des Sensors im Objektkoordinatensystem: 8-Teilung

### 5.2.3 Sichtbarkeit von 3-D Messpunkten im Bild

Zur Bestimmung der Sichtbarkeit von 3-D Messpunkten im Bild werden verschiedene Kriterien angewandt. Das erste Kriterium überprüft zunächst alle 3-D Messpunkte dahingehend, ob ihre  $x$ -Komponenten im Kamerakoordinatensystem einen größeren Wert als die Brennweite der Kamera besitzen, wobei der Ursprung des Kamerakoordinatensystems im Projektionszentrum liegt. Das heißt, es wird geprüft, ob jeder Messpunkt vor der Projektionsebene im richtigen Halbraum liegt und damit theoretisch sichtbar ist.

Die Punkte sind zusammengefasst zu Kanten. Sobald mindestens ein Punkt einer Kante theoretisch sichtbar ist, wird die Kante als theoretisch sichtbar markiert. Mehrere in einer Ebene liegende Kanten legen wiederum Flächen fest. Eine Fläche ist theoretisch sichtbar, wenn mindestens eine ihrer Kanten theoretisch sichtbar und der Normalenvektor der Fläche in die Halbebene weist, in der auch das Projektionszentrum liegt. Ist mindestens eine Fläche theoretisch sichtbar, dann sei das übergeordnete Objekt theoretisch sichtbar. In Abbildung 5.3 sind die Orientierungen aller modellierten Kanten des Fahrzeugmodells eingezeichnet. Der Umlaufsinn der die Flächen begrenzenden Kanten ist so zu berücksichtigen, dass der Normalenvektor der Flächen stets vom Objekt nach außen zeigt.

Als zweites Kriterium wird für theoretisch sichtbare Objekte überprüft, ob alle theoretisch sichtbaren Messpunkte auch im Bildbereich liegen. Hierfür wird die Projektion der 3-D Messpunkte in die Bildebene durchgeführt und kontrolliert, ob sie innerhalb des Bildbereichs liegen. Die Sichtbarkeit der Kanten und Flächen wird entsprechend angepasst.

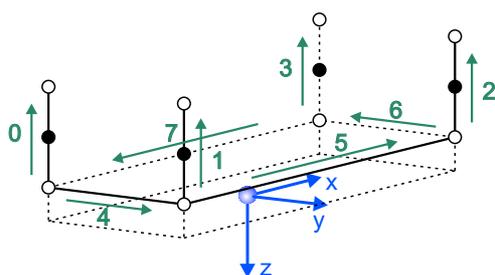


Abbildung 5.3: Orientierungen der Kanten im Fahrzeugkoordinatensystem

### 5.2.4 Verdeckungen

Als drittes Kriterium ist zu überprüfen, ob Verdeckungen vorliegen. Hierfür werden für jedes Objekt in Abhängigkeit von den Aspektbedingungen entsprechend der 8-Teilung aus Abbildung 5.2a umschreibende Polygone in 3-D definiert. Dies ist in Abbildung 5.4a, 5.4b und 5.4c beispielhaft für die Ansichten von „hinten rechts“, „hinten“ und „hinten links“ dargestellt.

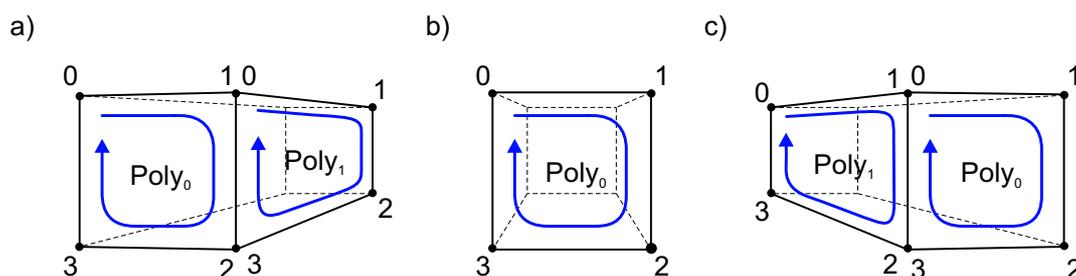


Abbildung 5.4: Festlegung von Verdeckungspolygonen in 3-D. Entsprechend dem eingezeichneten Umlaufsinn beschreibt das Polygon  $Poly_0$  je nach Ansicht stets das Heck bzw. die Front eines Quaders und das Polygon  $Poly_1$  die linke bzw. rechte Seitenfläche.

Für jedes Polygon wird außerdem der horizontale und vertikale Winkelbereich im aktuellen Kamerakoordinatensystem und die mittlere Entfernung vom Projektionszentrum bestimmt. Mittels der Entfernung- und Winkelbereichsinformation wird entschieden, ob das aktuell auszuwertende Objekt durch ein anderes Objekt voll oder teilweise verdeckt wird. Bei vollständiger Verdeckung wird das Objekt als nicht sichtbar markiert. Bei teilweiser Verdeckung werden alle theoretisch sichtbaren Messpunkte überprüft, ob ihre Projektionen innerhalb des ins Bild projizierten 2-D Bereichs eines verdeckenden Polygons liegen. Ist dies der Fall wird ein entsprechender Messpunkt als nicht sichtbar markiert.

Durch die Festlegung von 3-D Polygonen können auch Subobjekte des Eigenfahrzeugs, welche zu Verdeckungen führen könnten, wie z.B. die Motorhaube, die A-Holme oder

der Mercedesstern modelliert und bei der Verdeckungsanalyse berücksichtigt werden. Bei der derzeitigen Implementierung muss die Randbedingung erfüllt sein, dass die Projektionen der 3-D Polygone ins Bild zu konvexen 2-D Polygonen führen. Ansonsten würde die Möglichkeit bestehen, dass die Überprüfung, ob ein Messpunkt innerhalb eines 2-D Polygonbereichs liegt oder nicht, fehlerhaft ist.

### 5.2.5 Aspektabhängige Beobachtbarkeit von Zustandsgrößen

Die Beobachtbarkeit bestimmter Zustandsgrößen hängt direkt mit der Messbarkeit der für eine Zustandsschätzung relevanten Merkmale bzw. Kombinationen von Merkmalen zusammen. Das heißt, nur wenn mindestens die minimal notwendige Anzahl an relevanten Merkmalen oder Merkmalkombinationen messbar ist, kann und darf die Zustandsinnovation der entsprechenden Zustandsgröße durchgeführt werden.

Beispielsweise macht es nur dann Sinn den Abstand zwischen der vorderen und der hinteren Achse eines Fremdfahrzeugs zu schätzen, wenn bei Seitenansicht sowohl die Position des Vorderrads als auch des Hinterrads bestimmt werden konnten. Bei Heck- oder Frontalansicht sollte die Schätzung dieses Formparameters unterbunden bzw. eingefroren werden.

Es ist zu überlegen, ob in diesem Fall auch das Vertrauen in die Schätzung, repräsentiert durch die entsprechenden Einträge in der Fehlerkovarianzmatrix, bis zur nächsten möglichen Zustandsinnovation konstant gehalten wird oder ob das Vertrauen entsprechend der steigenden Unsicherheit über der Zeit abnehmen soll.

Mit dem Wissen, dass es sich z.B. bei den Formparametern um die Beschreibung eines starren Körpers handeln soll, ist es vertretbar anzunehmen, dass das letzte beste Wissen sich über der Zeit nicht verschlechtert. Für dynamische Größen wie z.B. die Position, Geschwindigkeit und Ausrichtung sollte demgegenüber das Vertrauen in die Schätzwerte ohne Zustandsinnovation abnehmen.

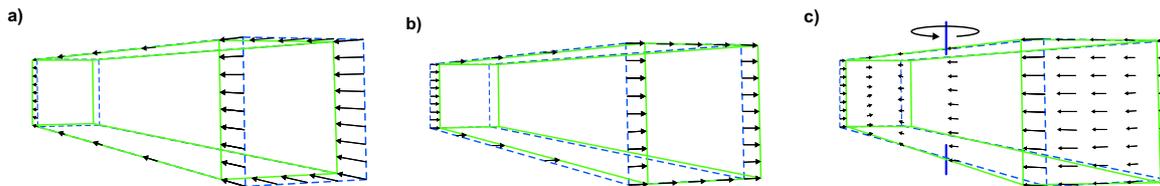
## 5.3 Berücksichtigung unspezifischer Merkmale bei der Zustandsschätzung

Im Folgenden werden zwei Ansätze zur Berücksichtigung vergleichsweise unspezifischer Merkmale bei der Zustandsschätzung beschrieben. Es handelt sich dabei zum einen um die erwartungsbasierte Bestimmung des Merkmalfusses entlang der Bildzeilen, herrührend aus der dominanten Objektbewegung, und zum anderen um einen Ansatz zur erwartungsbasierten Stereoauswertung. Im Wesentlichen wird nur die Annahme getroffen, dass alle Merkmale in erster Näherung auf den Seitenflächen eines umschreibenden Quaders liegen und die zu vergleichenden Bildbereiche annähernd die gleiche Größe besitzen.

### 5.3.1 Erwartungsbasierte Bestimmung des Merkmalflusses

Es wird hier angenommen, dass ein Merkmalfluss im Bild hauptsächlich auf zwei Ursachen zurückzuführen ist. Entweder ein Objekt, auf dem sich Merkmale befinden, bewegt sich relativ zur Kamera oder die Kamera verändert ihre Position und/oder Blickrichtung. Hier sollen zunächst nur Objekte betrachtet werden, deren grundlegende Form sich nicht ändert und damit im Wesentlichen als starre Körper behandelt werden können. Beiträge zum Merkmalfluss aufgrund von Spiegelungen in den Scheiben oder auf lackierten Oberflächen sowie Schattenwurf, z.B. von einer Schilderbrücke, und sonstige Änderungen der Beleuchtungsbedingungen werden hier als vernachlässigbar betrachtet, auch wenn sie das öfteren nicht sind. Weiterhin seien die internen Parameter der Kamera unveränderlich.

Das heißt, ein Merkmalfluss im Bild sei entweder durch Eigenbewegung der Kamera und/oder durch Positionsänderungen der Objekte im Raum hervorgerufen. Solange die Lage eines Objekts im Kamerakoordinatensystem unverändert bleibt, soll dieses demnach keinen optischen Fluss erzeugen. Die Abbildungen 5.5 a), b) und c) veranschaulichen den Merkmalfluss durch Veränderung der Fremdfahrzeugposition in Längsrichtung, in Querrichtung und um die vertikale Achse am Schwerpunkt.



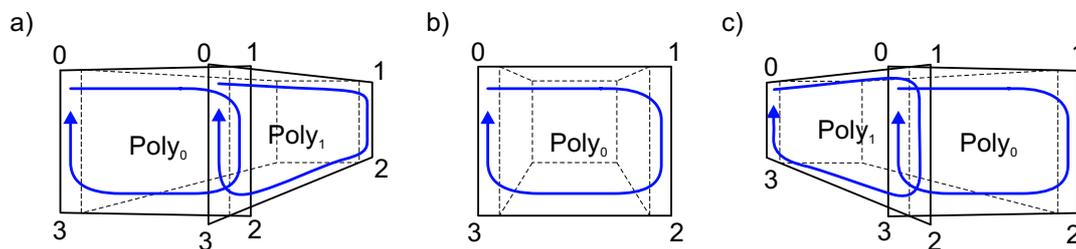
**Abbildung 5.5:** Veranschaulichung des Merkmalflusses bei Veränderung der Fremdfahrzeugposition a) in  $x$ , b) in  $y$  und c) in der Ausrichtung  $\Psi$ .

Eine Trennung der Einflüsse herrührend aus einer beliebigen Eigenbewegung der Kamera und beliebiger Bewegung von Objekten im Raum ist dabei ohne Kenntnis der Lage der Objekte im Raum nicht möglich. Deshalb wird häufig in der Literatur nur der Fall einer statischen Kamera untersucht und versucht, durch Zuordnungen von signifikanten 2-D Merkmalen im Bild (z.B. Ecken) die Bewegungsvektoren zu extrahieren, gleichförmige Bewegungen zu segmentieren, um daraus Objekthypothesen zu generieren. Oder es wird versucht, bei bekannter Eigenbewegung der Kamera durch Bewegungstereo die Lage der Objekte im Raum zu bestimmen. Durch eine zusätzliche Bewegung eines Objekts im Raum wird die Entfernungsbestimmung durch Stereo-Triangulation entsprechend fehlerbehaftet, da die Disparität zwischen einer Merkmalposition im Bild zum Zeitpunkt  $t_k$  und zum Zeitpunkt  $t_{k+1}$  um den Anteil verfälscht ist, um den sich die Projektion ins Bild durch die Bewegung des Objekts verändert hat.

Bei bekannter Eigenbewegung der Kamera vom Zeitpunkt  $t_k$  zum Zeitpunkt  $t_{k+1}$  kann bei bekannter Lage eines Objekts im Kamerakoordinatensystem zum Zeitpunkt  $t_k$  die

erwartete Position zum Zeitpunkt  $t_{k+1}$  relativ zur Kamera bestimmt werden. Die Differenz der Projektionen ins Bild legt den erwarteten Merkmalfluss durch Eigenbewegung fest. Die Abweichung des gemessenen Merkmalflusses vom erwarteten Merkmalfluss stellt den Anteil dar, der durch die Objektbewegung hervorgerufen wird.

Die Form eines Fahrzeugs wird in der Modellvorstellung grob durch einen umschreibenden Quader angenähert, wobei die sichtbaren Seiten des Quaders explizit durch konvexe Polygone in 3-D beschrieben werden. Um für die Berechnung des Merkmalflusses auch die Randbereiche des Quaders mit einbeziehen zu können, werden entsprechend Abbildung 5.6 erweiterte Polygone in 3-D definiert. Durch Projektion dieser Polygone ins Bild werden diejenigen Bereiche festgelegt, welche aussagekräftige Merkmale eines Fahrzeugs enthalten sollten.



**Abbildung 5.6:** Festlegung erweiterter 3-D Polygone.

Neben der Lage und Ausrichtung relativ zu einem Referenzkoordinatensystem wird für jedes Fremdfahrzeug unter Verwendung geeigneter Dynamikmodelle die Geschwindigkeitskomponente entlang der Längsachse geschätzt, welche über die Ausrichtung in die Geschwindigkeitskomponenten in x- und y-Richtung des Referenzkoordinatensystem umgerechnet werden kann. Demnach ist es möglich, neben dem Anteil der Eigenbewegung auch die erwartete Objektbewegung dem erwarteten Merkmalfluss hinzuzufügen, so dass die verbleibende Abweichung nun den Anteil darstellt, der im Idealfall durch eine veränderte Dynamik in der Objektbewegung hervorgerufen wurde. Diese Abweichung wiederum kann zur Nachführung der die Objektbewegung beschreibenden Zustandsgrößen verwendet werden.

Die interne Repräsentation der Objekte der „realen Welt“ erfolgt via Szenenbaum relativ zum mitbewegten Fahrspurkoordinatensystem. Die Bewegung des Eigenfahrzeugs wird relativ zu diesem Referenzobjekt bestimmt, so dass hierdurch eine implizite Trennung zwischen Eigenfahrzeug- und Fremdfahrzeugbewegung durchgeführt wird. Die Interpretation der Information über die reale Welt erfolgt Ego-zentriert von nah nach fern z.B. beim Aufbau des Wissens über das Umfeld und bei der Verdeckungsanalyse. Bei exakt bekannter Eigenbewegung und ideal modellierter und geschätzter Objektbewegung würde demnach die Abweichung zwischen dem erwarteten Merkmalfluss und dem gemessenen Merkmalfluss null sein.

Die dominante Bewegungskomponente eines Fahrzeugs ist in Längsrichtung. Eine Querbewegung oder ein Gierwinkel können bei Normalbetrieb nur im Zusammenhang mit einer Längsbewegung aufgebaut werden. Darum wird lediglich der Längsanteil des Merkmalflusses für die Innovation zugelassen. Die Veränderungen der Ausrichtung und der Querablage werden anhand der äußeren Berandungen und der Positionen der Räder bestimmt.

Für jedes Objekt werden für jeden betrachteten Bildsensor für alle 2-D Polygonbereiche die Ergebnisse der horizontalen Bereichsauswertung mittels des HOFMANN-Operators für den aktuellen und den zurückliegenden Messzeitpunkt gespeichert. Die Ergebnisse der aktuellen Messung werden dann im nächsten Messzyklus zur „letzten“ Messung, und der Speicherbereich der Messergebnisse der bis dato letzten Messung wird durch die dann aktuelle Messung überschrieben. Der Zugriff auf die Speicherbereiche erfolgt über Zeiger, so dass durch einen zyklischen Tausch der Zeiger der Verwaltungsaufwand minimal ist.

Zur Zuordnung der Bildmerkmale zwischen der aktuellen und der vorhergehenden 2-D Polygonbereichsauswertung werden zum einen die photometrischen Eigenschaften ausgenutzt, zum anderen wird für jedes Merkmal eine Statistik angelegt, welche als wesentliche Größen die Anzahl der erfolgreichen Zuordnungen und den gleitenden Mittelwert des berechneten Merkmalflusses beinhaltet. Das Ziel der Statistik liegt darin, nur robust und damit wiederholt erkannte Merkmale für die spätere Zustandsinnovation zuzulassen.

Zunächst wird mittels der Eckpunkte der 2-D Polygonbereiche der aktuellen und der vorhergehenden Messung der mittlere zu erwartende Merkmalfluss bestimmt. Alternativ könnte für jedes Merkmal durch lineare Interpolation der individuelle erwartete Merkmalfluss bestimmt werden. Dies ist etwas aufwendiger, allerdings auch genauer. Dieser dient als Referenzwert, relativ zu dem eine Korrespondenz innerhalb eines erlaubten Toleranzbereichs liegen darf. Der Toleranzbereich wird in Abhängigkeit vom erwarteten Merkmalfluss festgelegt, wobei ein erlaubter Minimalwert fix vorgegeben wird.

Zu Beginn der Flussberechnung werden dann alle Merkmale jedes Suchstreifens der aktuellen Messung mit den Merkmalen des jeweils zugeordneten Streifens der vorhergehenden Messung anhand der photometrischen Eigenschaften (Art des Helligkeitsübergangs, mittlere Segmenthelligkeit links und rechts der gerichteten Kanten, L- und R-Wert ) und der Nähe zur erwarteten Position aufgrund des erwarteten Merkmalflusses auf Korrespondenz überprüft. Kann eine Korrespondenz gefunden werden, wird dieses Merkmal als „bereits zugeordnet“ markiert und in der Ergebnisstruktur der aktuellen Messung der Zähler für die Anzahl der Zuordnungen dieses Merkmals um 1 erhöht. Diejenigen Merkmale, bei denen es gelungen ist, über mehrere Zyklen hinweg eine Korrespondenz zu finden, und deren Zähler demnach eine vorgegebene Mindestanzahl überschreitet, werden als robuste Merkmale betrachtet und für die Zustandsinnovation zugelassen.

Die entsprechenden Zeileneinträge in der **C**-Matrix des Kalmanfilters werden numerisch bestimmt. Durch Variation der dominanten Zustandsgröße, Bestimmung der Veränderung der Positionen der Eckpunkte der 2-D Polygone im Bild und Division durch die Variation ergibt sich eine numerische Näherung für die entsprechenden Einträge in den Zeilen der Jacobi-Matrix bezüglich der Eckpunkte. In den Spalten derjenigen Zustandsgrößen,

die nicht variiert wurden, werden Nullen eingetragen. Durch lineare Interpolation werden daraus die jeweiligen Zeileneinträge für jede beliebige Merkmalposition innerhalb des aktuellen 2-D Polygons ermittelt.

### 5.3.2 Erwartungsbasierte Stereoauswertung bei divergenten optischen Achsen

Unter der Voraussetzung, dass die optischen Achsen der Stereokameraanordnung zumindest näherungsweise in einer Ebene liegen (d.h. die Differenz der Nickwinkel sei  $|\Delta\theta| < 0.5^\circ$ ), die Rollwinkel der Sensorchips der Kameras bezüglich dieser Ebene null sind und die internen Kameraparameter der beiden Kameras annähernd gleich sind, wodurch z.B. eine Skalierung der Bildpositionen aufgrund eines Brennweitenunterschieds entfallen kann, können die Ergebnisse der 2-D Polygonauswertungen, die für die Bestimmung des Merkmalfusses notwendig waren, für die Entfernungsbestimmung mittels Stereotriangulation genutzt werden. Das heißt, es ist keine nochmalige Bildauswertung notwendig.

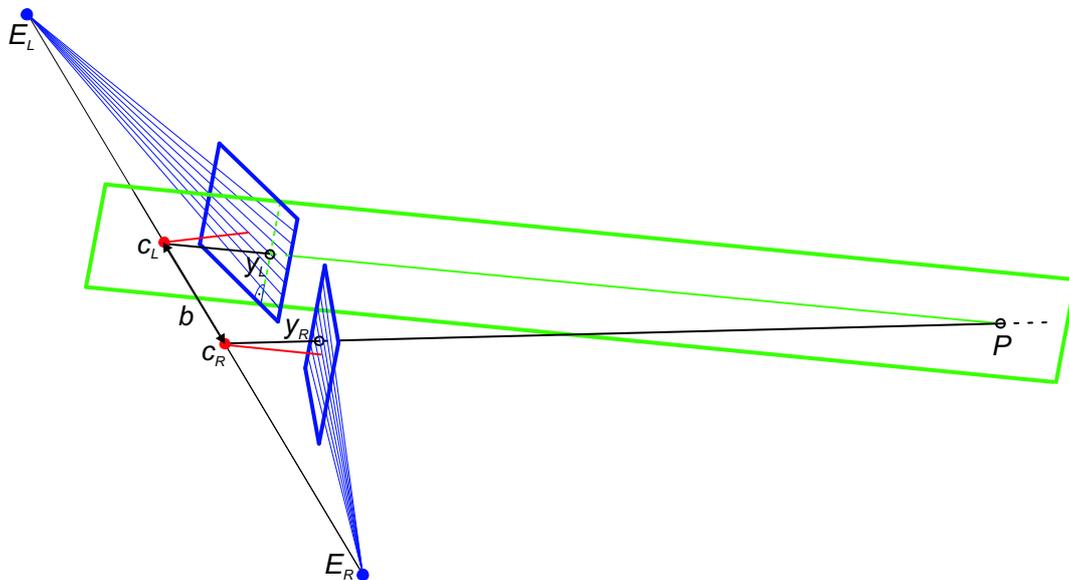
Der Abstand der Projektionszentren  $c_L$  und  $c_R$  der linken bzw. rechten Stereokamera, die sogenannte *Stereobasis*, sei  $b$  mit  $b > 0$ . Weiterhin muss überprüft werden, ob sich ein gerade betrachtetes Objekt innerhalb des Überlappungsbereichs des Sichtfelds der Kameraanordnung befindet. Diese Überprüfung kann aufgrund der 2-D Polygonbereiche in den Bildern oder anhand der 3-D Winkelbereiche erfolgen. Ob es sich um eine parallele oder divergente Kameraanordnung handelt, ist dabei vollkommen unerheblich.

Sind die Voraussetzungen erfüllt, kann eine Stereointerpretation zur Entfernungsbestimmung durchgeführt werden. Hierfür wird in der Regel die sogenannte *Epipolarbedingung* ausgenutzt, welche die 2-dimensionale Suche nach einer Korrespondenz auf eine eindimensionale Suche entlang sogenannter *Epipolarlinien* reduziert. Der Epipolarlinie bezüglich eines Bildpunkts der linken Kamera entspricht dabei der ins Bild der rechten Kamera projizierte Sehstrahl durch diesen Bildpunkt und  $c_L$ . Alle 3-D Punkte dieses Sehstrahls werden auf den gleichen Bildpunkt des linken Kamerabildes abgebildet und die Projektionen ins rechte Kamerabild können nur auf der zugehörigen Epipolarlinie liegen. Deshalb kann die Suche nach einer Korrespondenz für jeden Bildpunkt der linken Kamera auf eine eindimensionale Suche entlang der zugehörigen Epipolarlinie in der rechten Kamera beschränkt werden.

Bei paralleler Stereoanordnung und vernachlässigbarer Linsenverzerrung verlaufen die Epipolarlinien parallel zu den Bildzeilen. Im Gegensatz dazu schneiden sich bei divergenter Anordnung und vernachlässigbarer Linsenverzerrung die Epipolarlinien je Kamera in je einem Punkt - einem sogenannten *Epipol*. Die Epipole liegen auf der Verbindungsgeraden durch die beiden Projektionszentren  $c_L$  und  $c_R$  der Kameras und seien hier mit  $E_L$  und  $E_R$  bezeichnet. Bei paralleler Stereoanordnung liegen sie im Unendlichen.

Die Epipolarlinien ergeben sich als Schnittgeraden der beiden Projektionsebenen mit allen Ebenen, die die Verbindungsgerade durch die Projektionszentren  $c_L$  und  $c_R$  beinhalten.

Einige dieser Epipolarlinien sind unter anderem in Abbildung 5.7 eingezeichnet. Bei nicht vernachlässigbarer Linsenverzerrung werden aus den Epipolarlinien „Epipolarcurven“ im Bild. Dies kann man sich anschaulich so vorstellen, dass die Projektion einer Geraden im Raum z.B. bei radialer Verzerrung nicht mehr auf eine Projektionsebene erfolgt, sondern auf eine mehr oder weniger gewölbte Projektionsfläche. Durch Rektifikation der Bilder mittels inverser Kameramodelle, welche bildlich gesprochen die Pixel einer gewölbten Projektionsfläche in eine Projektionsebene transformiert, ergeben sich daraus wieder Epipolarlinien. Eine solche Rektifikation, bei der jede Pixelposition transformiert werden muss, ist selbst bei Verwendung vorab berechneter Look-Up-Tabellen zeitlich aufwendig und erhöht keinesfalls die im Bild enthaltene Information.



**Abbildung 5.7:** Veranschaulichung der Vorgehensweise zur Stereoauswertung bei divergenter Kameraanordnung

Um diese Rektifikation zu vermeiden, kann in endlichen (eventuell beliebig kleinen) lokalen Bereichen um die Positionen eines korrespondierenden Merkmalpaars in den Bildern der linken bzw. rechten Stereokamera in erster Näherung von der Parallelität und Kolinearität der Epipolarlinien mit den Bildzeilen ausgegangen werden. Die korrespondierenden 2-D Polygonbereiche eines Objekts im Bild der linken und rechten Stereokamera werden nun als solche lokale Bereiche aufgefasst, wodurch aufgrund der vorausgesetzten Identität der internen Kameraparameter durch direkten Vergleich der extrahierten Bildmerkmale ohne Skalierung die Suche nach Korrespondenzen erfolgen kann.

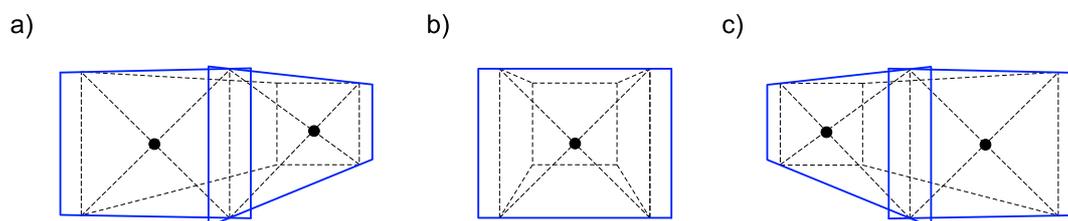
Die Zuordnung zwischen den Merkmalen zweier Polygonbereiche erfolgt anhand der photometrischen Eigenschaften. Die *Disparität*  $d = y_L - y_R$  bezeichnet dabei die Differenz der horizontalen Bildkoordinaten korrespondierender Bildmerkmale. Für die Positionsbestimmung in 3-D werden nur solche Korrespondenzen zwischen Merkmalen zugelassen, deren

Disparität innerhalb eines Konfidenzintervalls liegt. Hierfür werden die Positionen von 3-D Testpunkten relativ zum erwarteten Mittelpunkt eines 3-D Polygons um einen Wert, der entfernungsabhängig festgelegt wird, variiert und die zu erwartenden Disparitäten bzw. das Konfidenzintervall bestimmt. Beispielsweise könnten solche 3-D Testpunkte bei Heckansicht in  $x$ -Richtung je nach Entfernung um  $\pm 1m$  (im Nahbereich) bis  $\pm 5m$  (im Fernbereich) bezüglich des Hecks variiert und daraus das erlaubte Konfidenzintervall ermittelt werden.

Es seien  $y_L$  und  $y_R$  die horizontalen Bildkoordinaten zweier korrespondierender Merkmale im Bild der linken bzw. rechten Kamera. Die 3-D Positionsbestimmung erfolgt durch Berechnung des Schnittpunkts  $P$  zwischen dem Sehstrahl durch  $y_R$  und  $c_R$  und der Ebene, die den Sehstrahl durch  $y_L$  und  $c_L$  beinhaltet und parallel zur vertikalen Ausrichtung des linken Kamerabildes verläuft. Aufgrund dieser Vorgehensweise ist auch bei minimal windschiefen korrespondierenden Sehstrahlen (aufgrund nicht vermeidbarer noch so kleiner Ungenauigkeiten bei der Montage bzw. Kalibrierung der Kameras) gewährleistet, dass stets die Position im Raum bestimmt werden kann. In Abbildung 5.7 ist dies dargestellt.

Bei Vorhandensein einer Mindestanzahl an Stereozuordnungen können je 2-D Polygon eine mittlere 3-D Position, die Standardabweichungen und der maximale Hub in  $x$ - und  $y$ -Richtung bestimmt werden. In einem Nachverarbeitungsschritt könnten weiterhin Fehlzuzuordnungen bzw. 3-D Punkte, die zu weit von der mittleren 3-D Position entfernt sind, als Ausreißer detektiert und für die Berechnung gesperrt werden.

Die mittlere 3-D Position einer Seitenansicht ( $Poly_1$ ) wird dabei zur Korrektur der  $y$ -Position relativ zum Mittelpunkt des jeweiligen 3-D Polygons relativ zum Fremdfahrzeugschwerpunkt ins Kalmanfilter eingespeist und die mittlere 3-D Position einer Heck- bzw. Frontansicht ( $Poly_0$ ) zur Korrektur der  $x$ -Position verwendet. Der Mittelpunkt wird hier stellvertretend für einen beliebigen Punkt innerhalb der jeweiligen Polygonebene benutzt. Die Mittelpunkte der 3-D Polygone sind unter verschiedenen Ansichten in Abbildung 5.8 dargestellt.



**Abbildung 5.8:** Mittelpunkte der 3-D Polygone bei Ansicht a) von hinten rechts b) von hinten und c) von hinten links.

## 5.4 Kurzübersicht zur Fahrspurmodellierung und zur Modellierung der relativen Eigenposition

Die horizontale Fahrbahnkrümmungsdynamik der Eigenfahrspur (*OwnLane*) wird mittels einer mitgeführten, über einen definierten Vorausschaubereich gemittelten Klothoide nach [DICKMANN 88] modelliert. Die Krümmung einer Klothoide in der Vorausschauentfernung bzw. Lauflänge  $l$  entlang der Klothoide errechnet sich dabei aus der Krümmung am Fußpunkt der Klothoide  $c_{0h}$  und der Krümmungsänderung  $c_{1h}$  zu

$$c_h(l) = c_{0h} + c_{1h} \cdot l. \quad (5.1)$$

Der Gierwinkel der Tangente an die Klothoide für die Lauflänge  $l$  gegenüber der  $x$ -Achse des Fahrspurkoordinatensystems, der sogenannte *Bahnazimutwinkel*  $\chi$ , wird durch Integration des Krümmungsverlaufs bestimmt:

$$\chi(l) = \int_0^l c_h(\lambda) d\lambda = c_{0h} \cdot l + c_{1h} \cdot \frac{l^2}{2}. \quad (5.2)$$

Daraus folgt für kleine Krümmungswerte  $|c_{0h}|$  und  $|c_{1h}|$ , wie sie z.B. auf Autobahnen und gut ausgebauten Landstraßen üblich sind,

$$x(l) = \int_0^l \cos(\chi(\lambda)) d\lambda \approx l \quad (5.3)$$

und

$$y(l) = \int_0^l \sin(\chi(\lambda)) d\lambda \approx \frac{c_{0h}}{2} \cdot l^2 + \frac{c_{1h}}{6} \cdot l^3. \quad (5.4)$$

Bei entsprechend großer gewollter Vorausschau kann es eventuell notwendig werden, weitere Terme der Taylorreihenentwicklungen in den Näherungen zu berücksichtigen. Dies gilt natürlich insbesondere für enge Kurven, wie sie z.B. in Ein- und Ausfahrten auf Autobahnen auftreten. In der Regel wird jedoch die gefahrene Geschwindigkeit in Abhängigkeit von der Krümmung einer befahrenen Kurve so gewählt, dass die Näherungen (5.3) und (5.4) für die dann notwendige Vorausschau ausreichend genau sind. Die Extrapolation eines geschätzten Krümmungsverlaufs über den Vorausschaubereich hinaus ist im Übrigen nur eingeschränkt empfehlenswert.

Orthogonal zu dieser Skelettlinie der Fahrspur wird in Abhängigkeit von der Lauflänge  $l$  im Abstand  $\pm \frac{(b_0 + l \cdot b_1)}{2}$  nach Spurmarkierungen gesucht, wobei  $b_0$  die Spurbreite am Fußpunkt der Klothoide und  $b_1$  der Spurbreitenänderung bezeichnen. Die Beschreibung der Eigenfahrspur wurde von [MYSLIWETZ 90] um ein Vertikalkrümmungsmodell erweitert. Dabei wird dem horizontalen Krümmungsverlauf analog ein vertikaler Krümmungsverlauf in Abhängigkeit der Lauflänge  $l$  überlagert. Durch [BEHRINGER 96] wurde im vertikalen Krümmungsmodell die Idee eines Relaxationsterms eingeführt, der die Größen  $c_{0v}$  und  $c_{1v}$  über der Zeit zwangsweise gegen null (dem erwarteten Normalzustand) führt.

Als Querdynamikmodell des Eigenfahrzeugs wird das lineare ebene Einspurmodell nach [MITSCHKE 72], [DONGES 78] verwendet, für das von [MYSLIWETZ 90] weitere Vereinfachungen durchgeführt wurden. Das Fahrzeugquerdynamikmodell liefert u.a. als wesentliche Größen die Ablage und Ausrichtung des Eigenfahrzeugs gegenüber der Eigenfahrspur. Das diskretisierte Zustandsraummodell kann in [MYSLIWETZ 90] im Anhang A3 nachgelesen werden. Darauf aufbauend wurde die Eigenzustandsschätzung von [BEHRINGER 96] um einen Schätzer für den Nickwinkel einer Kamera gegenüber der Bodenebene ergänzt, welcher bei fester Einbaulage der Kamera auch als Nickwinkel des Fahrzeugaufbaus (*Ego-Body*) interpretiert werden kann. Dieser Nickwinkel wird in [SIEDERSBERGER 03] dazu verwendet, die niederfrequente Drift der inertialen Nickwinkelschätzung zu kompensieren.

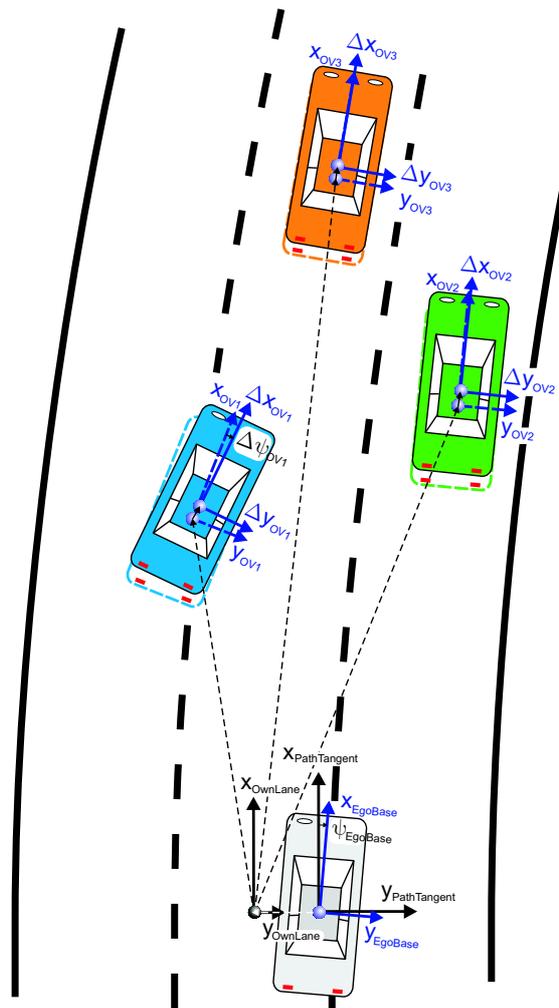
## 5.5 Koordinatensysteme zur visuellen Fahrspur- und Objekterkennung

Der Ursprung des Fahrspurkoordinatensystems ist identisch mit dem Fußpunkt des Fahrspurverlaufs und befindet sich bezüglich der in der Bodenebene liegenden Normalen zur Fahrspur stets auf gleicher „Höhe“ mit dem Ursprung des Eigenfahrzeug-Basiskoordinatensystems (*EgoBase*). Dieser resultiert aus der Projektion des Fahrzeugschwerpunkts in die Bodenebene. Die  $x$ -Achse des *EgoBase*-Koordinatensystems ist parallel zur Fahrzeuglängsachse. Alle Koordinatensysteme sind *rechtshändige* Koordinatensysteme entsprechend der Luftfahrtnorm [LN9300/90].

Parallel zum Fahrspurkoordinatensystem wurde von [LÜTZELER 02] ein Zwischenkoordinatensystem, das sogenannte *PathTangent*-Koordinatensystem eingeführt, dessen Ursprung mit der Projektion des Fahrzeugschwerpunkts in die Bodenebene zusammenfällt. Siehe Abbildung 5.9. Seine Translation gegenüber dem Fahrspurkoordinatensystem dient der Beschreibung der Ablage des Egofahrzeugs gegenüber der Fahrspur. Die Rotation gegenüber dem *EgoBase*-Koordinatensystem repräsentiert die Ausrichtung des Ego gegenüber der Tangente an die Fahrspur am Fußpunkt der gemittelten Klothoidenbeschreibung.

Prinzipiell gibt es zwei mögliche Vorgehensweisen zur Berücksichtigung der Eigenbewegung in der Beschreibung der Relativlagen von Objekten gegenüber dem Eigenfahrzeug. Für den Fall, dass alle Objekte unmittelbar relativ zum Eigenfahrzeug beschrieben werden, müssen bei allen Objekten explizit die je Zyklus durch Eigenbewegung hervorgerufenen Anteile abgezogen werden. Alternativ dazu kann ein gemeinsames Referenzobjekt für eine implizite Trennung zwischen Eigenbewegung und Objektbewegung verwendet werden.

Beispielsweise kann das Fahrspurkoordinatensystem als solch ein Referenzobjekt dienen. Verändert sich die relative Lage des Eigenfahrzeugs gegenüber der Fahrspur, ist diese



**Abbildung 5.9:** Draufsicht auf einen Ausschnitt der Koordinatensystemhierarchie.

Änderung implizit bei allen Objekten, die relativ zum Referenzobjekt beschrieben wurden, berücksichtigt. Für den Fall, dass z.B. in der Detektionsphase zeitweilig kein Referenzobjekt zur Verfügung steht, kann als Rückfalloption stets die explizite Korrektur der Eigenbewegung betrachtet werden.

In der Anwendung entsprechend Kapitel 7 wird den Fremdfahrzeugen ein von Zeitschritt zu Zeitschritt nachgeführtes Koordinatensystem unterlagert, welches als Akkumulator für die Änderungen in der Fremdfahrzeugposition und Ausrichtung dient. Dies ist u.a. in Abbildung 5.9 dargestellt. Das Ziel hierbei besteht darin, ein möglichst einfaches Dynamikmodell zu verwenden, welches die Kopplung zwischen Längs- und Querbewegung beinhaltet. Das heißt, eine Änderung der Querablage relativ zum Referenzkoordinatensystem soll nur aufgrund einer von null verschiedenen Ausrichtung und einer Längsbewegung möglich sein. Es wurde hier ein Newtonmodell für die Schätzung der Position und

Geschwindigkeit in Fremdfahrzeuglängsrichtung und ein weiteres Newtonmodell für die Schätzung der Ausrichtung und Gierrate gegenüber dem Referenzkoordinatensystem verwendet. Details bezüglich des Newtonmodells für die Längsdynamik finden sich in Kapitel 6.2 in [RIEDER 00] bzw. in Kapitel 4.1.1 in [THOMANEK 96].

Zu Beginn eines jeden neuen Zyklus wird für jedes Fremdfahrzeug die erwartete Lageänderung des  $\Delta OV$ -Koordinatensystems durch Prädiktion der Fremdfahrzeugbewegung ermittelt. Das unterlagerte  $OV$ -Positionskoordinatensystem wird danach durch Akkumulation der Änderungen nachgeführt und die Änderungen im  $\Delta OV$ -Koordinatensystems zu null gesetzt. Der Grund für diese Vorgehensweise liegt darin, dass dadurch die numerische Bestimmung der Jacobimatrixeinträge ohne Umrechnung durch direkte Variation der jeweiligen Lageparameter beschrieben im  $OV$ -Koordinatensystem erfolgen kann, d.h. durch Variation der  $x$ -Koordinate in Fahrzeuglängsrichtung und des Gierwinkels am Fremdfahrzeugschwerpunkt.

# Kapitel 6

## Hybrider adaptiver Tempomat

Derzeit werden große Anstrengungen zur Entwicklung von Fahrerassistenz- und Komfortsystemen unternommen. Beispiele für solche Systeme sind z.B. ein Spurverlassenswarnsystem, ein Stauassistent oder ein adaptiver Tempomat (engl. ACC - **A**daptive **C**ruise **C**ontrol), der die aktuell gefahrene Geschwindigkeit an den Verkehrsfluss in der eigenen Fahrspur bzw. im aufgrund des Lenkeinschlags voraussichtlich befahrenen Bereich, dem sogenannten *Fahrschlauch* (engl. *driving tube*), anpasst.

In Zusammenarbeit mit einer Zulieferfirma der Automobilindustrie wurde am ISF ein System entwickelt, das ein radarbasiertes ACC System, ähnlich den derzeit bereits im Handel erhältlichen Systemen, mit dem am ISF entwickelten EMS-VISION System kombiniert. Das Radarsystem stand hierbei lediglich als Black Box zur Verfügung, welches eine Liste von Radarobjekten und eine unter Verwendung der ABS-Sensorik berechnete tiefpassgefilterte Krümmung eines Fahrschlauchs lieferte. Für diese Kombination wurde die Bezeichnung Hybrid Adaptive Cruise Control (HACC) System gewählt.

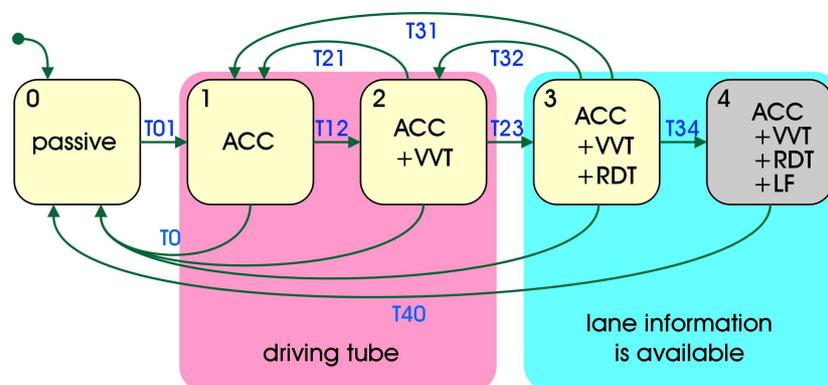
### 6.1 Funktionsbeschreibung

Das HACC System sei ein Komfortsystem, welches für den Einsatz auf Autobahnen und autobahnähnlich ausgebauten Straßen mit weißen Spurmarkierungen auf beiden Seiten der Fahrspuren konzipiert ist. In dieser klar definierten Domäne seien die erwarteten Objekte auf Straßenfahrzeuge beschränkt. Das eigene Fahrzeug (EGO) soll dabei vom Fahrer manuell gelenkt werden, nur die Längsregelung soll automatisiert durch das HACC System erfolgen. Hierfür muss der Fahrer eine Wunschgeschwindigkeit vorgeben. Der Computer regelt die Geschwindigkeit des Fahrzeugs dann derart, dass der Sicherheitsabstand zu vorausfahrenden Fremdfahrzeugen (OtherVehicles (OVs)) eingehalten wird und die Differenz zwischen Wunsch- und aktueller Geschwindigkeit möglichst klein ist. Es obliegt der Verantwortung des Fahrers, die Fahrspur zu wählen und zu entscheiden, ob überholt werden soll oder nicht. Wenn beispielsweise ein anderes Fahrzeug vor dem Eigenfahrzeug

fährt, dessen Geschwindigkeit kleiner als die Wunschgeschwindigkeit ist, dann bremst das EGO unter Einhaltung des Sicherheitsabstandes ab und folgt mit der Geschwindigkeit des vorausfahrenden Fahrzeugs. Die Beschleunigungen bzw. Verzögerungen werden durch die Längsregelung dabei so gewählt, dass es für den Passagier stets komfortabel bleibt. Der Sicherheitsabstand zu vorausfahrenden Fremdfahrzeugen soll dabei niemals kleiner als eine vorgegebene Zeitlücke (z.B.  $1.6s$ ) mal der aktuellen Geschwindigkeit sein. Das HACC System ist kein Sicherheitssystem. Der menschliche Fahrer muss sich der Verkehrssituation stets bewusst sein. Er ist rechtlich verantwortlich für alle Aktionen des Fahrzeugs. Der Fahrer kann zu jeder Zeit das HACC System durch Betätigen des Gaspedals bzw. der Bremse überstimmen. Dadurch dass das HACC System kein Sicherheitssystem darstellen soll, ist die maximal mögliche Verzögerung, die das System kommandieren kann, auf  $-2.5 \frac{m}{s^2}$  begrenzt. Das impliziert, dass es dem HACC System auch nicht erlaubt ist, eine Notbremsung durchzuführen.

## 6.2 Systemdesign mit skalierbarer Systemleistung

Das System wurde so konzipiert, dass es auf unterschiedlichen Leistungsstufen operieren kann. Dies ist in Abbildung 6.1 dargestellt. Im initialen Systemzustand ist die Längsregelung passiv, der menschliche Fahrer regelt die Geschwindigkeit und Fahrtrichtung selbst.



**Abbildung 6.1:** Skalierbare Leistung des HACC Systems: ACC = radarbasiertes Adaptive Cruise Control, VVT = visuelle Objekterkennung (engl. *Vehicle Validation and Tracking*), RDT = Fahrspurerkennung (engl. *Road Detektion and Tracking*), LF = vollautonomes Spurfolgen (engl. *Lane Follow*). Diese Abbildung ist aus [HOFMANN ET AL. 01].

Der Übergang zur ersten Stufe der Systemleistung erfolgt, indem das konventionelle, rein radarbasierte ACC aktiviert wird. Die Entscheidung, ob ein anderes Fahrzeug möglicherweise für die Längsregelung relevant ist oder nicht, d.h. ob es vor dem eigenen Fahrzeug

in der eigenen Fahrspur mit einer Geschwindigkeit kleiner als die eigene Wunschgeschwindigkeit fährt, wird unter Zuhilfenahme des Fahrschlauchs getroffen. Dieser verläuft im Nahbereich stets tangential zur Fahrzeuglängsachse. Seine Krümmung wird aus den Relativgeschwindigkeiten der vier Räder unter Verwendung der ABS-Sensorik berechnet. Das System hat auf dieser Leistungsstufe keine Kenntnis über die tatsächliche Lage und den Verlauf der Fahrspuren.

Eine zweite Leistungsstufe wird dadurch erreicht, dass die OV-Hypothesen, die durch das radarbasierte ACC generiert wurden, mittels visueller Objekterkennung validiert und ihre relative Lage zum EGO und ihre Dimensionen bestimmt werden. Bei erfolgreicher Validierung werden die OVs in den Szenenbaum eingefügt und veröffentlicht. Alle Objekte werden mittels Bildverarbeitung nach dem 4-D Ansatz verfolgt. Auch auf der zweiten Leistungsstufe liegt keine Information über die Lage und den Verlauf einer aktuell befahrenen Fahrspur vor.

Auf der dritten Leistungsstufe wird zusätzlich die Detektion und Verfolgung der eigenen Fahrspur und die Zuordnung der validierten OVs zur einer Fahrspur (der eigenen oder den beiden anschließenden Fahrspuren) durchgeführt.

Eine vierte Leistungsstufe könnte das vollständig autonome Folgen der aktuell befahrenen Fahrspur sein. Dies war nicht im Fokus dieses Projekts, stellt aber eine Standardfunktionalität des EMS-Vision Systems dar.

Die gewünschte Leistungsstufe wird vom Fahrer über eine Mensch-Maschine-Schnittstelle (engl. *Human Machine Interface* (HMI)) vorgegeben. In Abbildung 6.1 stehen die Transitionen  $T_{xy}$  zwischen den verschiedenen Leistungsstufen für Bedingungen, die für einen Übergang von einer Leistungsstufe zu einer anderen erfüllt sein müssen.  $T_{01}$  überprüft, ob das Radarmodul betriebsbereit zur Ausführung der radarbasierten ACC-Funktionalität ist. Die Transition  $T_{12}$  verifiziert, dass die Bildverarbeitungsmodule gestartet und der VVT Prozess Daten liefert. Transition  $T_{23}$  überprüft, ob die automatische Fahrspurdetektion abgeschlossen ist und der RDT Prozess die Fahrspur trackt.

Wenn der RDT Prozess die Verfolgung der Fahrspur abbricht und eine automatische Neudetektion durchgeführt wird, während sich das System in der Leistungsstufe 3 befindet, so wechselt dieses automatisch via  $T_{32}$  in die Leistungsstufe 2 bis  $T_{23}$  wieder erfüllt ist. Wenn die Witterungs- oder Beleuchtungsbedingungen für die Bildauswertung ungenügend sind (Dies wird nicht explizit, sondern nur implizit über die Anzahl und den Kontrast der Merkmale ermittelt.), wechselt das System über die Transitionen  $T_{31}$  oder  $T_{21}$  auf Leistungsstufe 1. Wenn das Radarsystem nicht aktiv ist bzw. keine Daten liefert oder der Fahrer das System überstimmt, geht das System stets über  $T_0$  in den passiven Zustand.

Wenn die gewünschte Leistungsstufe 4 ist, darf das System mittels  $T_{34}$  in das autonome Spurfolgen wechseln und wird den autonomen Modus via  $T_{40}$  erst dann wieder verlassen, wenn das radarbasierte ACC inaktiv werden sollte, der Fahrer das System durch Betätigen der Bremse oder Lenkeingriff überstimmt oder die Spurverfolgung eine Neuinitialisierung startet.

Dabei gibt es noch erheblichen Forschungsbedarf bezüglich der Fragestellung, wie ein menschlicher Fahrer nach längerer vollautonomer Fahrt dazu gebracht werden kann, die Verkehrssituation schnell richtig zu erfassen oder alternativ auch bei vollautonomer Fahrt dem Verkehrsgeschehen stets zu folgen, so dass eine sichere Übergabe an den menschlichen Fahrer überhaupt erfolgen kann. Die Entscheidung auf volle Autonomie zu verzichten und dem Fahrer das Lenken zu überlassen, erfordert, dass er stets dem Verkehrsgeschehen folgen muss.

### 6.3 Gegenüberstellung der Eigenschaften eines Radars und eines visuellen Sensors

Der Reiz einer Kombination eines radarbasierten ACC Systems mit einem Vision-System liegt in dem Bestreben, sowohl die Unzulänglichkeiten eines reinen radarbasierten ACC Systems wie auch eines reinen Bildverarbeitungssystems kompensieren zu können. Ein rein radarbasiertes ACC Systems hat beispielsweise folgende Unzulänglichkeiten:

- Reflektionen des ausgesendeten Radarsignals an Leitplanken können zu fehlerhaften Objekthypothesen und damit zu Fehlalarmen führen.
- Zwei Fahrzeuge, welche nebeneinander mit annähernd der gleichen Geschwindigkeit fahren, sind nur schwer zu unterscheiden und erscheinen häufig nur als *ein* Hindernis, welches auch nur *einer* Fahrspur zugeordnet wird. Das bedeutet, ein Fahrzeug oder Motorrad, das sich mit annähernd der gleichen Geschwindigkeit neben einem Lkw befindet, kann für das Radar unsichtbar sein.
- Aufgrund der geringen Winkelauflösung eines Radarsystems ist die Bestimmung der lateralen Position eines Fremdfahrzeugs relativ zum Eigenfahrzeug deutlich ungenauer als die Bestimmung der Entfernung in Längsrichtung. Entsprechend kann keine Aussage über die tatsächlichen Dimensionen eines potentiellen Hindernisses gemacht werden.
- Die eigene Position und die relativen Positionen der OV's können aufgrund fehlenden Wissens nicht in Bezug zu den realen Fahrspuren gesetzt werden. Dadurch wird die Entscheidung, ob ein Hindernis relevant ist oder nicht, deutlich erschwert. Dies gilt insbesondere für größere Objektentfernungen.
- Die Algorithmik des hier verwendeten radarbasierten ACCs unterdrückt Gegenverkehr und Objekthypothesen, deren absolute Geschwindigkeit kleiner als ein Schwellenwert (z.B.  $< 10 \frac{m}{s}$ ) ist. Ein derartiges konventionelles ACC System ist für Stop-and-Go Verkehr ungeeignet.

- Der verwendete 77 GHz Long-Range-Radar besitzt einen sehr kleinen horizontalen Öffnungswinkel, so dass Einscherer im Nahbereich nicht bzw. in der Regel zu spät erkannt werden.

Andererseits

- sind die Sensoreigenschaften eines Radarsystems im Wesentlichen unabhängig von den Witterungs- und Beleuchtungsbedingungen
- und die Bestimmung der Entfernungen und der relativen Geschwindigkeiten der OV's ist sehr präzise.

Die Vorteile eines Visionsystems liegen in der Möglichkeit

- einer genauen Bestimmung der lateralen Position von OV's relativ zum Eigenfahrzeug.
- Es können die Dimensionen anderer Fahrzeuge bzw. der entsprechenden einhüllenden Quader ermittelt werden. Hierdurch kann eine Klassifikation der Hindernisse aufgrund ihrer Größe durchgeführt werden und eine modellbasierte Vorhersage des Fremdfahrzeugverhaltens rückt in den Bereich des Möglichen.
- Es kann die Lage und der Verlauf der eigenen Fahrspur detektiert und über der Zeit verfolgt werden.

Als Konsequenz daraus

- können die Formparameter der eigenen Fahrspur extrahiert werden.
- kann die Position des EGO relativ zur eigenen Fahrspur bestimmt werden.
- kann ein Spurwechsel in Abhängigkeit vom Gierwinkel und der Ablagenänderung des Eigenfahrzeugschwerpunkts gegenüber der Fahrspur erkannt werden.
- können die Positionen anderer Fahrzeuge relativ zur eigenen Fahrspur bestimmt werden.

Von Nachteil ist,

- dass die Bildverarbeitungsverfahren stark von den Witterungs- und Beleuchtungsverhältnissen abhängig sind.
- dass die Entfernungsbestimmung mittels eines rein Vision-basierten ACC Systems schwierig ist, da diese Information durch die perspektivische Projektion verloren geht. Deshalb ist es im Vergleich wesentlich schwieriger, genaue Werte für die realitiven Geschwindigkeiten von OV's zu schätzen.

Radar und Vision haben demnach komplementäre Sensoreigenschaften. Eine Kombination der beiden Sensortypen sollte zu einer verbesserten Leistung des Gesamtsystems führen. Dies konnte exemplarisch mit dem Experimentalfahrzeug VaMP der UniBwM bei Versuchsfahrten im realen Straßenverkehr gezeigt werden.

## 6.4 Verwendete Hardware

Als Experimentalplattform wurde einer der Versuchsträger des ISF, ein Mercedes 500SEL entsprechend Abbildung 6.2, verwendet. Dieser Versuchsträger für autonome Mobilität und Rechnersehen im Pkw - VaMoRs-P, kurz VaMP - wurde für das HACC Projekt mit einem 77 GHz Long-Range-Radarsystem (LRR) ausgerüstet, welches entsprechend Abbildung 6.3 a) in der vorderen Stoßstange unterhalb des Nummernschildes mittig eingebaut wurde. Das Radarsystem verwendet eine Radarkeule mit einem horizontalen Öffnungs-

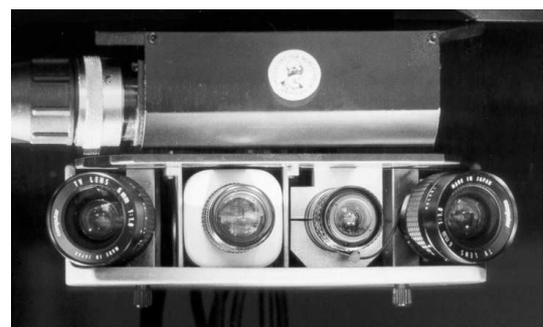


Abbildung 6.2: Versuchsträger für autonome Mobilität im Pkw - VaMP

winkel von  $\pm 4^\circ$ . Es ist in der Lage, die relativen Geschwindigkeiten und Entfernungen zu OV's in einem Bereich von  $2m$  bis  $130m$  mit einer Genauigkeit von ca.  $\pm 0.5m$  (Nahbereich) bis  $\pm 1.5m$  (Fernbereich) zu messen. Auf dem Markt sind inzwischen z.B. von der Firma



a) 77 GHz Long-Range-Radar



b) Ein-Achsen-Kameraplattform

Abbildung 6.3: Einbaulagen a) des Radarsensors und b) der Kameraplattform

Bosch auch Radarsysteme mit größerem Öffnungswinkel z.B.  $\pm 8^\circ$  und z.B. 3 Radarkeulen erhältlich, welches die horizontale Winkelauflösung verbessert. Die prinzipiellen Vorzüge

und Unzulänglichkeiten behalten jedoch unverändert ihre Gültigkeit. Das radarbasierte ACC Modul verwendet die Daten der ABS Sensorik, um daraus die Krümmung des Fahr-schlauchs zu berechnen. Der Fahr-schlauch entspricht aus der Sicht des rein radarbasierten ACCs dem besten Wissen über die erwartete Fahrtrajektorie des Eigenfahrzeugs.

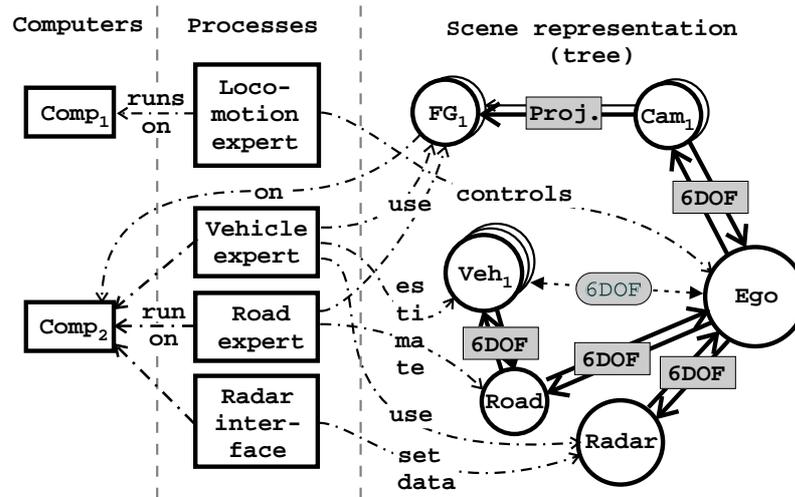
Das System ist in der Lage, das vor dem Fahrzeug befindliche Umfeld mittels mehrerer Ka-meras zu beobachten, welche auf einer gierbaren Ein-Achsen-Plattform entsprechend Ab-bildung 6.3 b) montiert sind. Von dieser multifokalen Kameraanordnung, für die der Name MARVEYE (Multi-focal active/reactive Vehicle Eye) entsprechend [DICKMANN 95] ge-prägt wurde, werden für dieses Projekt nur die Videodaten einer hochsensitiven Schwarz-Weiß-Kamera (S/W)(dritte Kamera von links) und das Helligkeitssignal einer 3-Chip Farbkamera (zweite von links) ausgewertet. Die Plattform war in diesem Projekt nicht aktiv. Diese bifokale Kameraanordnung ist mit einem Normalobjektiv (an der Grenze zum Weitwinkel) auf der S/W-Kamera mit einem horizontalen Öffnungswinkel von  $\pm 22^\circ$  und einem Teleobjektiv auf der 3-Chip Farbkamera mit einem horizontalen Öffnungswinkel von  $\pm 5.5^\circ$ . Für die Bildauswertung werden Halbbilder mit einer Auflösung von 768x286 Pixeln im 40ms Takt verwendet.

Für die Bildverarbeitungsverfahren wurde nur einer der insgesamt drei im System verfügbaren Bildverarbeitungsrechner verwendet. Auf diesem Computer (Comp2) liefen der VVT, der RDT und das Radar Interface-Modul. Der Radarprozess ist eine Schnittstel-le zu der Radar-Hardware, die dem System lediglich als Black Box zur Verfügung stand. Die Aktuatoren erhielten ihre Kommandos vom Fahrzeugsteuerungs PC (Comp1), auf dem das Modul für die Fahrzeugregelung lief.

## 6.5 Datenbasis des HACC Systems

Der prinzipielle Aufbau eines Szenenbaums mittels homogener Koordinatentransformatio-nen ist in Abschnitt 3.1 beschrieben worden. In Abbildung 6.4 werden die Verknüpfungen zwischen dem Szenenbaum, der Prozessliste und den im HACC System repräsentierten Computern veranschaulicht.

Beispielsweise werden die Positionen und die Bewegungsgrößen anderer Fahrzeuge relativ zum Koordinatensystem der aktuellen Eigenfahrspur geschätzt. Dieses Koordinatensy-tem liegt in der Mitte der Fahrspur, tangential zur Spur und bewegt sich entlang des Fahrspurverlaufs gleichförmig mit dem Eigenfahrzeugschwerpunkt. Die Position und Be-wegung des Eigenfahrzeugs wird ebenfalls relativ zum Eigenfahrspurkoordinatensystem beschrieben. Das Eigenfahrspurkoordinatensystem dient als Referenzkoordinatensystem für eine „kooperierende“ Schätzung der Eigen- und Fremdfahrzeugpositionen, wodurch eine implizite Trennung der jeweiligen Bewegungsanteile vollzogen wird.



**Abbildung 6.4:** Datenbasis des HACC Systems aus [HOFMANN ET AL. 01]:  $FG_1$  = Framegrabber<sub>1</sub>,  $Cam_1$  = Kamera<sub>1</sub>,  $Veh_1$  = Fremdfahrzeug<sub>1</sub> bzw.  $OV_1$ , Proj. = nicht invertierbare perspektivische Projektion, 6DOF = homogene Koordinatentransformation mit sechs Freiheitsgraden (6 degrees of freedom)

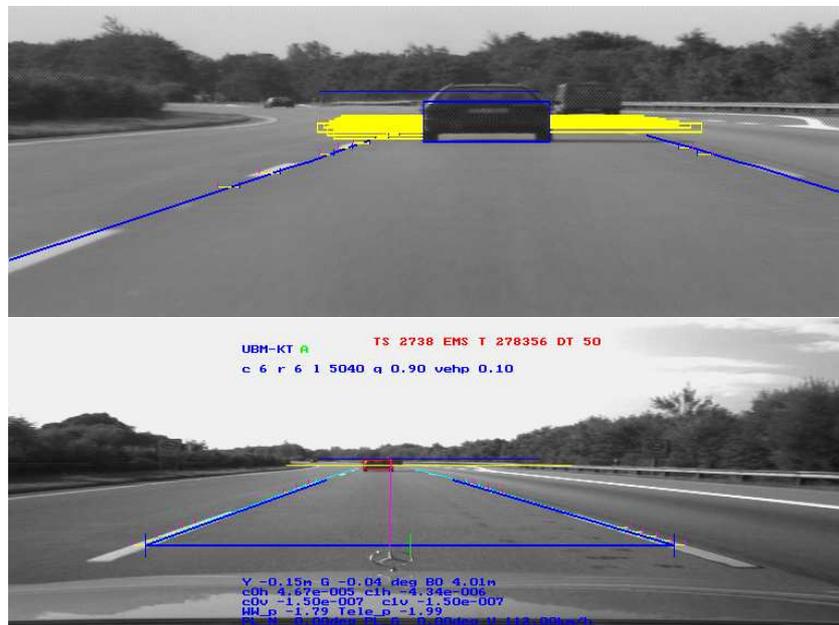
## 6.6 Verfahren zur Erkennung und Verfolgung der Fahrspur

Das Modul zur Fahrspurerkennung und -verfolgung durchsucht die Grauwertbilder entlang horizontaler Suchstreifen nach Spurmarkierungen. Hierfür wird mittels ternärer Masken entsprechend Abschnitt 4.1 nach charakteristischen Dunkel-Hell-Dunkel-Helligkeitsübergängen gesucht, welche an erwarteten Positionen eine erwartete Breite und Orientierung besitzen. Diese Spurmarkierungen werden von Zeitschritt zu Zeitschritt unter Verwendung der 3-D Modellvorstellung über die Form und Lage der Fahrspur durch Projektion ins jeweilige Bild berechnet. Die Abweichungen zwischen den gemessenen und den erwarteten Positionen werden für die Korrektur der 3-D Modellvorstellung verwendet.

Zur Initialisierung des Trackingvorgangs muss jedoch zunächst die Detektion der Fahrspur durchgeführt worden sein. Dazu muss eine geeignete Helligkeitsschwelle als Parameter für die Bildverarbeitungsoperatoren ermittelt werden, welche die Extraktion der Dunkel-Hell-Dunkel-Kombinationen zuverlässig gewährleistet. Beginnend mit einer sehr hohen Rauschschwelle wird diese sukzessive verringert, bis die Anzahl der gefundenen Kombinationen ausreichend ist und der Wert der Schwelle noch größer als ein erlaubter minimaler Wert ist. Sobald eine ausreichende Anzahl an Merkmalkombinationen extrahiert werden konnte, werden diese zu Linien verknüpft und daraus die initialen Werte für die Spurbreite, Ablage und die Ausrichtung relativ zum Eigenfahrzeug ermittelt. Die Detektion der Fahrspur erfolgt nur anhand des Bildes der Nahbereichskamera, denn auf Autobahnen

und ähnlichen Straßen kann im Nahbereich ( $6 - 30m$ ) der Einfluß der Spurkrümmung vernachlässigt werden. Während der automatischen Fahrspurdetektion dient die Lage relativ zum Fahrschlauch als Kriterium für die Beurteilung der Relevanz eines Fremdfahrzeugs. Die Fahrspurgeometrie wird mittels des gemittelten mitbewegten Klothoidenmodells nach [DICKMANN 88] beschrieben.

Die initialen Werte der Fahrspurdetektion werden als Startwerte für die Schätzung der Form- und Lageparameter entsprechend dem 4-D Ansatz nach Kapitel 5 verwendet. In Abbildung 6.5 ist die Fahrspur im Bild der Telekamera (oben) und im Bild der Nahbereichskamera (unten) visualisiert. Die Formparameter des 3-D Fahrspurmodells sind die horizontale und vertikale Krümmung, die Spurbreite und ihre Änderungen. Die Lageparameter umfassen die horizontale Ablage, die Ausrichtung relativ zur Fahrzeuglängsachse und den Nickwinkel des Fahrzeugaufbaus relativ zur Fahrspur. Siehe auch Abschnitt 5.4.



**Abbildung 6.5:** Verkürzen von Suchfenstern aufgrund einer Überlappung mit dem Verdeckungsbereich eines OV's.

Durch sukzessive Vergrößerung der Vorausschauentfernung von nah ( $< 40m$ ) nach fern ( $< 100m$ ) wird der Krümmungsverlauf der realen Fahrspur zuverlässig durch die Modellvorstellung nachgebildet. Für das Bild der Nahbereichskamera werden die Suchfenster innerhalb einer Vorausschauentfernung von  $6 - 40m$  und für das Bild der Telekamera im Bereich von  $30 - 100m$  in 3-D definiert und durch Projektion in die jeweiligen Bilder die erwarteten Positionen in Bildkoordinaten bestimmt. Falls die Anzahl der aus dem Telebild extrahierten Spurmarkierungsmerkmale über mehrere Zyklen eine bestimmte minimal geforderte Anzahl unterschreitet, wird der Vorausschaubereich der Telekamera verkürzt und

anschließend von Zyklus zu Zyklus wieder sukzessive von nah nach fern erweitert. Dies erhöht die Robustheit der Spurverfolgung.

Bevor die Merkmalextraktion gestartet wird, werden zunächst alle Suchfenster dahingehend überprüft, ob ein Fahrzeug eine erwartete Spurmarkierung verdeckt, d.h. die erwartete Position einer Spurmarkierung innerhalb des 2-D Verdeckungsbereichs eines Fahrzeugs liegt. In diesem Fall wird das Suchfenster deaktiviert. Für den Fall, dass es zwischen dem Verdeckungsbereich und einem Suchfenster lediglich eine Überlappung gibt, wird das Suchfenster entsprechend verkürzt und nur wenn die resultierende Suchpfadlänge anschließend unzureichend ist, deaktiviert. In Abbildung 6.5 sind die aufgrund einer Überlappung verkürzten Suchfenster zur Extraktion der Spurmarkierungen gelb eingezeichnet. Die restlichen Suchfenster wurden für eine verbesserte Übersichtlichkeit ausgeblendet.

## 6.7 Verfahren zur Fahrzeugdetektion, Hypothesengenerierung und deren Verfolgung

Die Regelung der Geschwindigkeit in Fahrzeuginnenrichtung erfordert, dass das HACCC System alle potentiellen Hindernisse detektiert. Im HACCC System werden neue Fahrzeughypothesen allerdings nur auf Initiative des Radarsystems, nach entsprechender Validierung durch Vision generiert. Im ACC-Radarmodul laufen hierfür Signalverarbeitungsverfahren, welche Reflektionen mit ähnlicher Entfernung, relativer Geschwindigkeit und Amplitude zu Objekten gruppieren. Das Radarsystem generiert daraus alle  $60ms$  eine Liste potentieller Fahrzeuge.

Für diese Radarmessungen wird zunächst untersucht, ob sie bereits existierenden Hypothesen im Szenenbaum zugeordnet werden können. Dies wird erreicht, indem zunächst entsprechend Abbildung 6.6 um jedes OV ein Konfidenzbereich definiert wird und danach alle Radarobjekte diesen zugeordnet werden, sofern sie innerhalb der Konfidenzbereiche liegen. Siehe dazu auch [RIEDER 00].

Die bereits existierenden Fahrzeughypothesen werden bezüglich ihrer Entfernung vom Eigenfahrzeug sortiert. Danach werden die horizontalen Winkelbereiche berechnet, die von jedem Fremdfahrzeug verdeckt werden. Diejenigen Radarmessungen, die bis dato noch nicht zugeordnet werden konnten, werden dahingehend überprüft, ob sie in einem der Verdeckungsbereiche liegen. Ist dies der Fall werden sie verworfen. Verbleibende Radarmessungen, welche weder einem existierenden Fremdfahrzeug noch einem Verdeckungsbereich zugeordnet werden konnten, sind Kandidaten für neue Fahrzeughypothesen. Diese werden nun mittels Vision überprüft. Falls die Validierung erfolgreich ist, wird eine neue Fremdfahrzeughypothese in die Szenenrepräsentation eingefügt.

Wird eine Hypothese für mehrere Zyklen weder vom Radar noch durch Vision bestätigt bzw. aktualisiert, wird sie aus dem Szenenbaum entfernt. Alle Fahrzeughypothesen im Szenenbaum werden mittels 4-D Bildauswertung verfolgt.

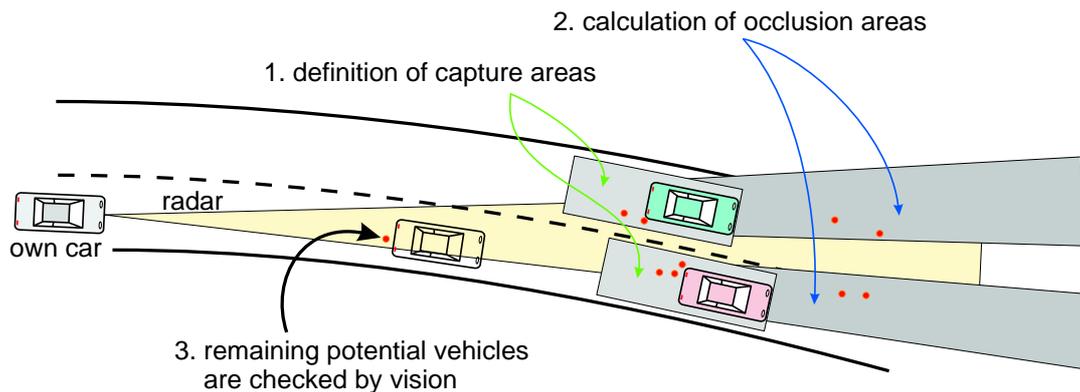


Abbildung 6.6: Generierung von Fahrzeughypothesen aus [HOFMANN ET AL. 01].

An der Position eines Kandidaten für ein potentielles Hindernis wird ein Quadermodell initialisiert, um damit die Form des potentiellen Fahrzeugs zu umschreiben. Die Orientierung des Quaders im Raum wird als parallel zur Orientierung der Fahrspur in der entsprechenden Entfernung angenommen.

In Abhängigkeit von der Ausrichtung des Quaders relativ zum Ego wird dann entweder die Länge oder die Breite des umschreibenden Quaders geschätzt. Weiterhin werden die Ablage und Quergeschwindigkeit und die Position, Geschwindigkeit und Beschleunigung in Längsrichtung bezüglich des Fahrspurkoordinatensystems am Fußpunkt der mitgeführten gemittelten Klothoide mittels EKF für jedes OV geschätzt.

Wie bereits in Abschnitt 6.5 erwähnt, werden sowohl die Positionen und Positionsänderungen der OVs als auch die Eigenbewegung relativ zum Fahrspurkoordinatensystem geschätzt, wodurch eine implizite Trennung zwischen der Eigenbewegung und den Fremdfahrzeugbewegungen resultiert, welche die Robustheit während des Trackings von Fremdfahrzeugen erhöht.

## 6.8 Auswahl des relevanten Fahrzeugs für die Fahrzeuglängsregelung

Die Auswahl eines für die Längsregelung relevanten Objekts erfordert eine Analyse der Objektkonstellation, welche wiederum die Kenntnis der aktuellen Positionen der Objekte und ihrer Bewegungsvektoren in 3-D voraussetzt. Die Entscheidung über die Relevanz eines Objekts erfolgt dabei unter Verwendung der impliziten Annahme, dass Fahrzeuge sich die meiste Zeit innerhalb der Fahrspuren bewegen, durch Zuordnung der Fahrzeuge zu den Fahrspuren.

Ein rein radarbasiertes ACC System (nach Abschnitt 6.2 Leistungsstufe 1) kann für die Beurteilung der Relevanz lediglich auf die Kenntnis des Fahrschlauchs zurückgreifen, dessen einziger variabler Parameter - die Krümmung - unter Verwendung der ABS Sensorik

aus den relativen Geschwindigkeiten der vier Räder berechnet wird. Der Fahrschlauch ist dabei vollkommen unabhängig von der Lage und vom Verlauf der Fahrspur. Dies kann den Abbildungen 6.7 und 6.8 entnommen werden.

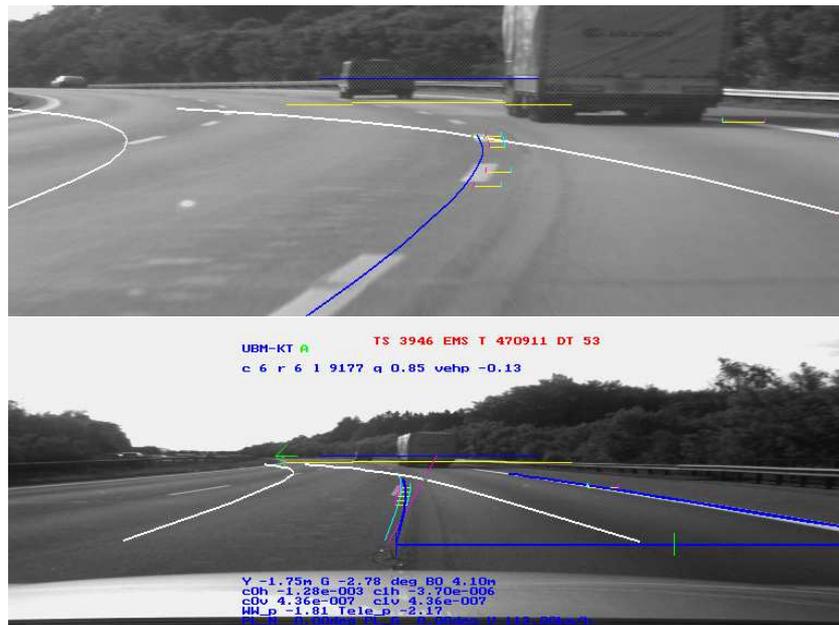


**Abbildung 6.7:** Visualisierung des Fahrschlauchs (weiße Linien) und der Fahrspur (dunkelblaue Linien). Das Eigenfahrzeug fährt nahe der rechten Fahrspurmarkierung, jedoch noch innerhalb der eigenen Fahrspur. Der im Telebild sichtbare Kleinbus würde zum relevanten Fahrzeug werden und damit einen Fehlalarm darstellen.

Querbewegungen innerhalb einer Fahrspur resultieren in einer sich ständig ändernden Krümmung des Fahrschlauchs. Um diesen Effekt zu mildern, wird die Berechnung der Krümmung des Fahrschlauchs in aller Regel mittels eines Tiefpasses gefiltert. Dies hat zur Folge, dass die Krümmung des Fahrschlauchs bei starken Lenkwinkeländerungen dem tatsächlich gefahrenen Kreisbogen entweder naheilt oder überschwingt.

Die Relevanzentscheidung unter Verwendung des Fahrschlauchs kann deshalb speziell in größerer Entfernung leicht zu Fehlalarmen führen. Wenn beispielsweise das Eigenfahrzeug ein anderes Fahrzeug in einer Linkskurve überholt, kann dieses zum relevanten Objekt werden, wenn sich aufgrund der Schwankungen des Lenkwinkels des Eigenfahrzeugs innerhalb der Eigenfahrspur die Krümmung des Fahrschlauchs ändert. Siehe auch Abbildung 6.7.

Abbildung 6.8 zeigt den Verlauf des Fahrschlauchs und des Fahrspurmodells während eines Spurwechsels. Der Abbildung ist zu entnehmen, dass eine Relevanzentscheidung aufgrund des Fahrschlauchs in größerer Entfernung definitiv falsch sein würde.



**Abbildung 6.8:** Visualisierung des Fahrschlauchs und der Fahrspur während eines Spurwechsels

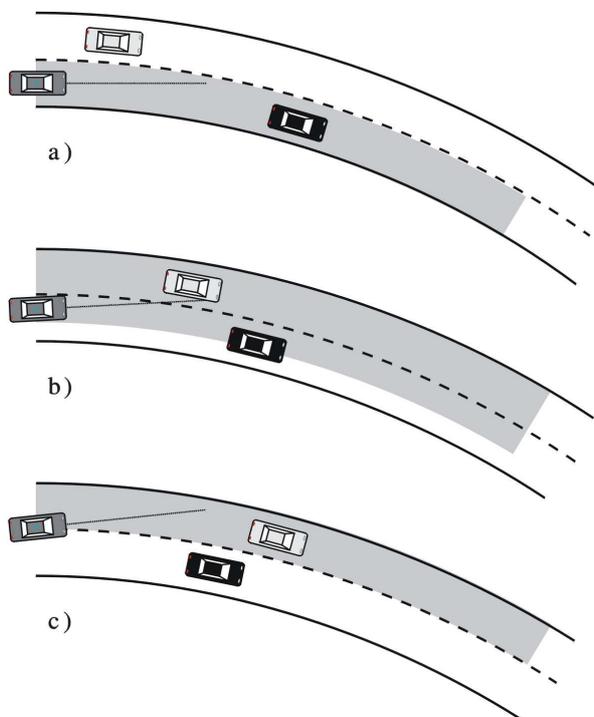
Im Gegensatz hierzu ermöglicht die visuelle Fahrspurerkennung und -verfolgung die Positionsbestimmung des Eigenfahrzeugs relativ zur Fahrspur mit allen sechs Freiheitsgraden und ferner die Bestimmung der Formparameter der Fahrspur. Für die Zuordnung von Fremdfahrzeugen zu den Fahrspuren und die Entscheidung über ihre Relevanz können drei Fälle unterschieden werden: Das Eigenfahrzeug fährt innerhalb der aktuellen Eigenfahrspur,

1. ohne dass ein Spurwechsel angezeigt oder erkannt wird.
2. wobei ein Spurwechsel nach links angenommen bzw. durch Setzen des linken Blinkers angezeigt wird.
3. wobei ein Spurwechsel nach rechts angenommen bzw. durch Setzen des rechten Blinkers angezeigt wird.

Die Detektion eines Spurwechsels kann durch Überwachung der Ausrichtung und der Ablage des Eigenfahrzeugs gegenüber der Fahrspur erfolgen. Sobald die in einer Vorausschauentfernung von z.B. 10m vorhergesagte Ablage des EGOs größer als 60% der Spurbreite ist, wird angenommen, dass sich das Eigenfahrzeug gerade im Spurwechsel befindet.

Im ersten Fall ist der Relevanzbereich identisch mit der aktuellen Eigenfahrspur entsprechend Abbildung 6.9a. Ein anderes Fahrzeug wird genau dann zur eigenen Fahrspur

zugeordnet, wenn der Betrag der Querablage zur Fahrspurskelettlinie kleiner als die halbe Fahrspurbreite ist. Falls ein anderes Fahrzeug bereits der Eigenfahrspur zugeordnet wurde, bleibt es ihr zugeordnet, solange der Betrag der Querablage kleiner als die halbe Fahrspurbreite plus die halbe Fahrzeugbreite ist.



**Abbildung 6.9:** Relevanzbereich als Funktion der Querablage

Während eines Spurwechsels ist es vernünftig, die Eigenfahrspur solange weiter zu tracken, wie die Projektion des Schwerpunkts des Eigenfahrzeugs in die Bodenebene innerhalb der bisherigen Eigenfahrspur liegt und erst dann die getrackte Fahrspur entsprechend zu wechseln. Allerdings wird die Relevanzentscheidung bezüglich der aktuellen Fahrspur nur unbefriedigende Ergebnisse liefern, denn das EGO wird ja innerhalb kürzester Zeit die Fahrspur wechseln. Während und unmittelbar nach dem Spurwechsel sollte die Regelung der Geschwindigkeit des Eigenfahrzeugs zusätzlich in Abhängigkeit der Fahrzeuge in der Zielfahrspur erfolgen.

Normalerweise wird ein Überholmanöver (Fall 2) ausgeführt, um danach schneller als in der derzeitigen Fahrspur fahren zu können. Während des dafür notwendigen Beschleunigungsvorgangs und Ausschermanövers akzeptieren menschliche Fahrer häufig eine Unterschreitung des für gewöhnlich gewählten Sicherheitsabstandes. Das bedeutet, die Erhöhung der Geschwindigkeit geht dabei auf Kosten des Sicherheitsabstandes gegenüber einem vorausfahrenden Fahrzeug in der derzeitigen Fahrspur. Das Setzen des linken Blinkers könnte als Indikator für den Beginn einer Beschleunigung dienen, indem dem System erlaubt wird, die Zeitlücke zum vorausfahrenden Fahrzeug zu verkleinern.

Gleichzeitig sollte der Relevanzbereich auf die Zielfahrspur erweitert werden. Seine Breite innerhalb der aktuellen Fahrspur sollte sukzessive als Funktion der Querablage verringert werden. Siehe Abbildung 6.9b. Solange die Projektion des Fahrzeugschwerpunkts auf die Bodenebene innerhalb der aktuellen Fahrspur liegt, sollte sich die Breite des Relevanzbereichs über den Bereich, der vom Eigenfahrzeug überdeckt wird, und über die Zielfahrspur erstrecken. Der erweiterte Relevanzbereich verläuft dabei parallel zur Fahrspurskelettlinie. Sobald der Schwerpunkt des Eigenfahrzeugs die Spurmarkierung überschritten hat, wird aus der Zielfahrspur die aktuelle Fahrspur und der Relevanzbereich wird identisch mit der aktuellen Fahrspur. Siehe Abbildung 6.9c. Der Fall 3 entspricht näherungsweise dem Fall 2, nur dass die Verkleinerung des Sicherheitsabstands nicht erlaubt wird.

Hiernach werden alle OV's entsprechend ihrer Entfernung zum Eigenfahrzeug sortiert und das nächstgelegene OV innerhalb der aktuellen Eigenfahrspur wird zum relevanten Zielobjekt für die Längsregelung.

## 6.9 Experimentelle Ergebnisse

In Abbildung 6.10 wurde ein manueller Spurwechsel von der mittleren Fahrspur zur rechten Fahrspur durchgeführt. In den linken Hälfte sind jeweils die Bilder der Nahbereichskamera und in der rechten Hälfte die der Telekamera dargestellt. Nur innerhalb des relativ kleinen Sichtbereichs des Radars, der im Vergleich zum horizontalen Öffnungswinkel der Telekamera noch etwas kleiner ist, werden Fremdfahrzeughypothesen generiert.

Das oberste Bild zeigt das Eigenfahrzeug innerhalb der mittleren Fahrspur beim Folgen eines vorausfahrenden Fahrzeugs mit dessen Geschwindigkeit und entsprechendem Sicherheitsabstand mittels automatisierter Längsregelung. Das relevante Fahrzeug wurde mit einem roten Rechteck markiert, während andere ebenfalls getrackte Fahrzeuge durch ein hellgraues Rechteck markiert wurden.

Das zweite Bild der Sequenz zeigt das Eigenfahrzeug unmittelbar vor Überschreiten der rechten Fahrspurmarkierung durch den Fahrzeugschwerpunkt. Das relevante Fahrzeug ist noch immer das vorausfahrende Fahrzeug in der mittleren Fahrspur.

Das dritte Bild visualisiert das Spurwechselmanöver unmittelbar nach Überschreiten der rechten Fahrspurmarkierung durch den Fahrzeugschwerpunkt. Das relevante Fahrzeug wurde vom vorausfahrenden Fahrzeug in der mittleren Spur zum vorausfahrenden Fahrzeug in der rechten Fahrspur, welche nun die neue Eigenfahrspur ist, gewechselt. Dies kann der Abbildung aufgrund des eingezeichneten blauen Fahrspurverlaufs und der roten Markierung des relevanten Fremdfahrzeugs entnommen werden.

Das vierte Bild zeigt die Annäherung des Eigenfahrzeugs an das vorausfahrende relevante Fahrzeug mittels der ACC Funktionalität. Dem System wurde im Übrigen nicht beigebracht, dass auf deutschen Autobahnen nach § 5 Abs. 1 StVO geboten ist: „Es ist links zu überholen.“, bzw. dass es auch Ausnahmen gibt: „Bei Fahrzeugschlangen auf mehreren Fahrstreifen für eine Richtung, bei Stau oder zäh fließendem Verkehr (wenn links höchstens Tempo 60 km/h gefahren wird), dann ist maximal 20 km/h schnelleres Rechtsüberholen erlaubt.“

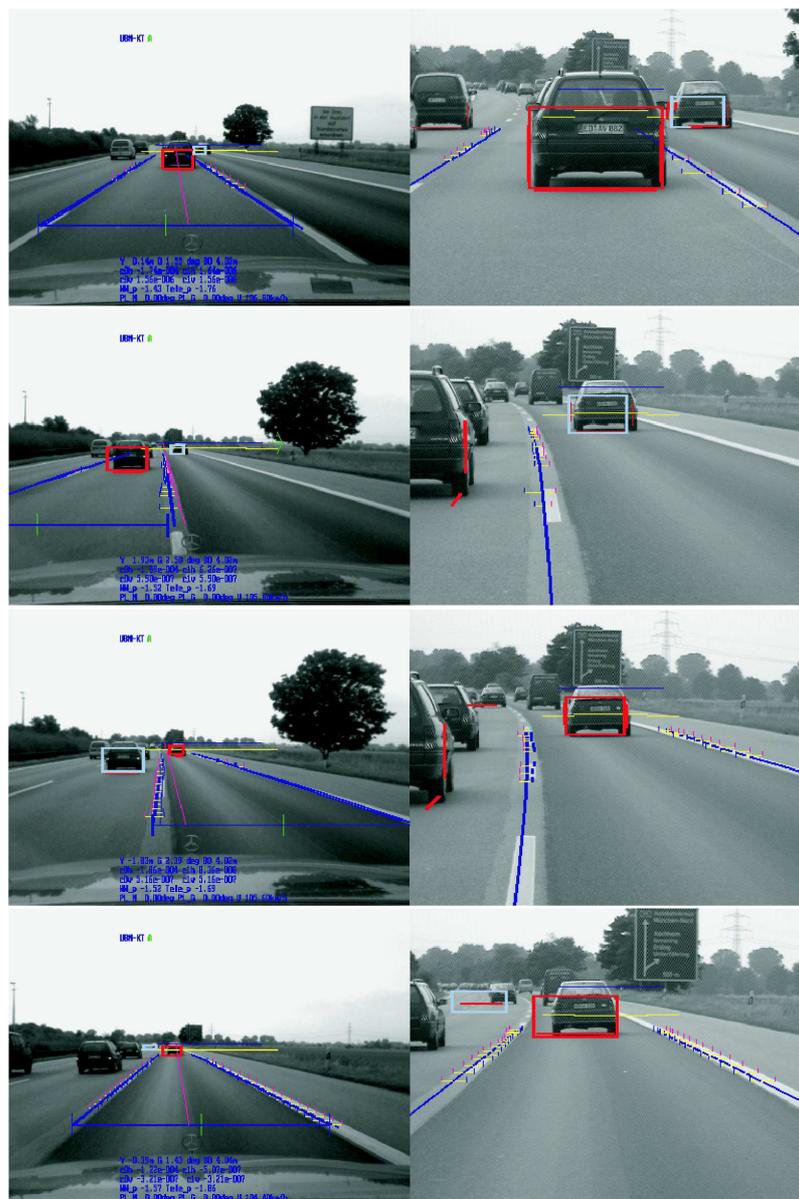


Abbildung 6.10: Spurwechsel nach rechts

## 6.10 Zusammenfassung und Ausblick

Die Kombination eines Radarsensors mit visueller Sensorik führt zu einem System mit verbesserten Leistungsmerkmalen, welches in der Lage ist, unterschiedliche Funktionen gleichzeitig zur Verfügung zu stellen. Durch die z.T. komplementären Sensoreigenschaften der beiden Sensortypen können die jeweiligen Unzulänglichkeiten der Sensoren verringert bzw. ausgeglichen werden, um insgesamt eine robustere ACC Funktionalität zu gewähr-

leisten.

Die Abhängigkeit der visuellen Wahrnehmung von den Witterungs- und Beleuchtungsbedingungen stellt dabei noch immer ein Problem dar. Für Allwettertauglichkeit bei beliebigen Beleuchtungsverhältnissen besteht noch erheblicher Forschungsbedarf. Im HACC System wurde ansatzweise versucht, die Witterungsbedingungen und Beleuchtungsverhältnisse über die Anzahl und den Kontrast der detektierten Merkmale zu beurteilen, um daraus implizit auf die Wahrnehmungsleistung zu schließen, um gegebenenfalls beispielsweise die Reinitialisierung der Fahrspur zu starten.

Das Wissen über den eigenen Zustand und die Objekte der Umwelt ist in einer zentralen Wissensbasis repräsentiert und steht damit allen Experten gleichzeitig zur Verfügung. Das HACC System ist in der Lage eine geeignete Leistungsstufe in Abhängigkeit vom Systemzustand bzw. der aktuellen Leistungsfähigkeit der Experten zu wählen.

Zur Verbesserung der Robustheit bei der Verfolgung von Fremdfahrzeugen hat dabei erheblich die implizite Trennung zwischen der Eigenbewegung und den Fremdfahrzeugbewegungen unter Verwendung der Fahrspur als Referenzkoordinatensystem beigetragen. Bezüglich der Fahrspurerkennung ist als Fazit zu ziehen, dass auf Autobahnen bei großen Vorausschauentfernungen ( $> 70m$ ) in der Modellierung der Fahrspur die Berücksichtigung der vertikalen Fahrspurkrümmung speziell auf hügeliger Strecke für eine robustes Tracking unabdingbar ist. Zur Verbesserung der Robustheit hat sich weiterhin die Überwachung der Anzahl der im Fernbereich detektierten Spurmarkierungsmerkmale als Kriterium für die Zuverlässigkeit der Fahrspurverfolgung bewährt. Für den Fall, dass sich zu wenige Merkmale finden, wird die Vorausschauentfernung prozentual verkürzt und dann sukzessive wieder von nah nach fern aufgebaut. Dass die Berücksichtigung von Verdeckungen, die durch andere Fahrzeuge hervorgerufen sind, die Robustheit verbessert, versteht sich von selbst.

Ist nur die Lage des Eigenfahrzeugs relativ zur Fahrspur und die prinzipielle Fahrspurgeometrie von Interesse, dann kann die Fahrspurerkennung auf den Nahbereich beschränkt und damit auf die Modellierung der vertikalen Krümmung verzichtet werden. Auf Autobahnen könnte bei Vorausschauentfernungen kleiner  $30m$  sogar vollständig auf die horizontale Krümmungsschätzung verzichtet werden [LÜTZELER 02].

Durch das im Vergleich zum rein radarbasierten ACC System erweiterte Wissen über die Umwelt kann das HACC System neben der verbesserten ACC Funktionalität weitere Funktionen bieten, die im Wesentlichen auf der Kenntnis der Fahrspur gründen. Beispielsweise kann ohne großen Aufwand die Funktionalität eines Spurverlassenswarnsystems in Abhängigkeit vom Setzen des Blinkers und unter Verwendung des Wissens über die Position und Ausrichtung des Eigenfahrzeugs relativ zur Eigenfahrspur gezeigt werden. Weiterhin stellt das längs- und quergeregelte Folgen bzgl. eines vorausfahrenden Fahrzeugs entlang der Eigenfahrspur eine Standardfunktionalität dar. Diese erfährt allerdings im HACC System eine Einschränkung dadurch, dass die Signalvorverarbeitung des alleinig Hypothesen initiierenden radarbasierten ACC Systems Fremdfahrzeuge mit Geschwindigkeiten kleiner z.B.  $10\frac{m}{s}$  ausblendet. Dadurch, dass im HACC System keine zum

Radarsystem alternative visuelle bottom-up Fahrzeugdetektion implementiert wurde, ist damit das HACC System für Stop-and-Go Verkehr ungeeignet. Es werden beispielsweise keine stehenden Hindernisse erkannt.

Weiterhin ist der horizontale Öffnungswinkel des Radarsystems für eine Detektion von Einscherern und Überholern zu gering. Diese sind aber speziell im Bereich niedriger Geschwindigkeiten wichtig. Die Erfahrungen mit der bifokalen HACC Kameraanordnung haben weiterhin gezeigt, dass auch der horizontale Öffnungswinkel von  $\pm 22^\circ$  der Nahbereichskamera für eine frühzeitige Einscherer- und Überholerdetektion bei niedrigeren Geschwindigkeiten nicht ausreichend ist.

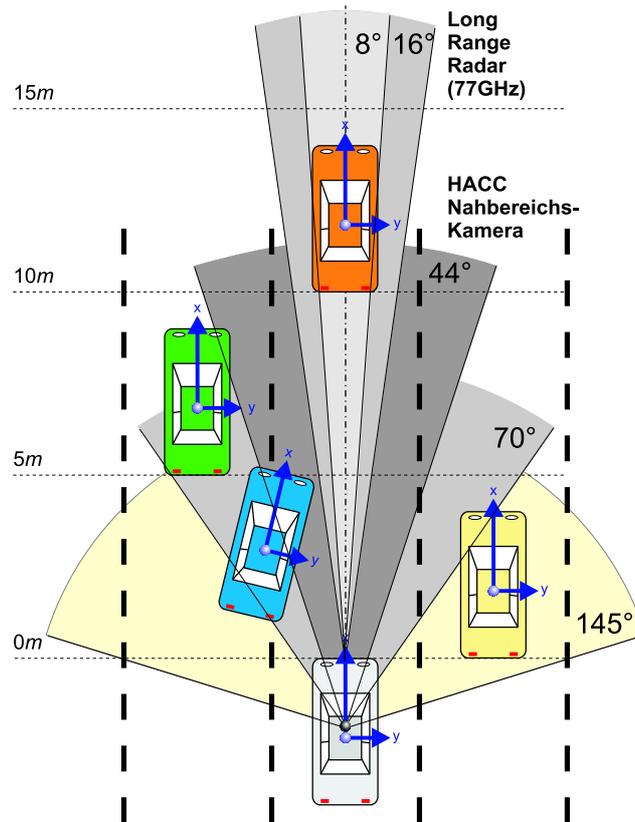
# Kapitel 7

## Erkennung von Fremdfahrzeugen im Nahbereich

Unabhängig von der spezifischen Realisierung können zunächst gewisse Anforderungen an eine Umfeldwahrnehmung und an einen zu erkundenden Sichtbereich aufgrund einer zu realisierenden Funktionalität herausgearbeitet werden. Am Beispiel der visuellen Objekterkennung soll hier ein Beitrag zur Erkennung von Fremdfahrzeugen im Nahbereich bei niederen Geschwindigkeiten erbracht werden, wobei der Fokus in diesem Kapitel ausschließlich auf dem Bereich vor dem Fahrzeug liegen soll.

Nachfolgend werden zunächst einige Bezeichnungen definiert: Unter *mitbewegtem Verkehr* seien alle Fahrzeuge zu verstehen, welche mit an den Verkehrsfluß angepasster Geschwindigkeit einer Fahrspur folgen. Ein *Überholer* gegenüber einem Bezugsobjekt sei ein Fremdfahrzeug, das relativ zum Bezugsobjekt auf einer benachbarten Fahrspur schneller an einem Bezugsobjekt vorbeifährt. Ein *Einscherer* gegenüber einem Bezugsobjekt sei ein Fremdfahrzeug, das aus einer zur Eigenfahrspur des Bezugsobjekts benachbarten Fahrspur auf dessen Fahrspur wechselt. Demgegenüber sei ein *Ausscherer* ein Fremdfahrzeug, das aus der Eigenfahrspur des Bezugsobjekts auf eine benachbarte Fahrspur wechselt.

Bei niederen Geschwindigkeiten, z.B. auf Autobahnen im Stau, können komplexe Situationen auftreten, für die die Umfeldwahrnehmung eines rein radarbasierten ACC oder des im vorhergehenden Kapitel beschriebenen HACCC nicht ausreichend ist. Dies liegt zum einen daran, dass die derzeit im Handel befindlichen Long-Range-Radar Systeme (LRR) einen sehr kleinen horizontalen Öffnungswinkel besitzen. Beispielsweise hat das im HACCC System verwendete LRR einen horizontalen Öffnungswinkel von  $\pm 4^\circ$  oder ein LRR System der Firma Bosch (das derzeit bei Audi in Serie verwendet wird) einen horizontalen Öffnungswinkel von  $\pm 8^\circ$ . Wie in Abbildung 7.1 zu erkennen ist, sind bei niederen Geschwindigkeiten die Sichtbereiche der LRR-Systeme im Nahbereich zu klein, um mitbewegten Verkehr in den Nachbarspuren, Überholer oder potentielle Einscherer zu detektieren. Deren frühzeitige Erkennung ist allerdings für eine robuste Erfassung und Interpretation einer Fahrsituation zur Gewährleistung eines sicheren Stop&Go Betriebs notwendig.



**Abbildung 7.1:** Darstellung einer Stausituation und Visualisierung der Sichtbereiche von Sensoren mit unterschiedlichen horizontalen Öffnungswinkeln zur Umfeldwahrnehmung

Zum anderen ist es mit einem Radarsystem nicht möglich, die Ausmaße eines potentiellen Hindernisses zu bestimmen, wodurch die Klassifikation relevanter Objekte stark eingeschränkt ist. Die Klassifikation erfolgt nur über die Bewegung. Das heißt, es werden angehaltene Objekte und bewegte Objekte als potentielle Hindernisse zugelassen. Stehende Hindernisse, also Objekte, welche sich aus der Sicht der Wahrnehmung bisher noch nicht bewegt haben, werden bewusst ausgeblendet. Sehr anschaulich verdeutlicht diese Problematik das häufig angeführte Beispiel einer Coladose auf der Fahrbahn. Ein Radarsystem ist nicht in der Lage zwischen einer Coladose und einem stehenden Fahrzeug (z.B. am Stauende) zu unterscheiden.

## 7.1 Anforderungen an eine Umfeldwahrnehmung bei niederen Geschwindigkeiten im Nahbereich

Beim HACCC System werden allein vom LRR System potentielle Fremdfahrzeughypothesen generiert, welche dann durch die Bildverarbeitung verifiziert oder verworfen werden.

Das heißt, nur innerhalb des relativ kleinen Sichtbereichs des LRR können beim HACC überhaupt Fremdfahrzeuge detektiert werden. Für höhere Geschwindigkeiten ist diese Vorgehensweise ausreichend, da hier in der Regel im Nahbereich keine Einscherer zu erwarten sind.

Demgegenüber ist bei niederen Geschwindigkeiten davon auszugehen, dass andere Fahrzeuge teilweise sehr spontan einen Spurwechsel durchführen, um einen geringfügigen Geschwindigkeitsvorteil zu erhalten. Dabei wird nicht selten sehr knapp vor dem Eigenfahrzeug ein- oder ausgeschert. Dies geschieht häufig gerade dann, wenn in einer Fahrspur der Verkehr zum Stehen kommt bzw. die Gefahr dazu besteht und erwartet wird, dass eine andere Fahrspur einen höheren Verkehrsfluss besitzt.

Um Einscherer überhaupt erkennen zu können, ist deshalb der Sichtbereich durch zusätzliche Sensoren zu erweitern. Dies kann z.B. durch 24 GHz Short-Range-Radar Sensoren (SRR) mit 45° Öffnungswinkel in leicht divergenter Anordnung (z.B. 20° Überlappungsbereich) erfolgen, wobei diese wiederum die prinzipielle Problematik der Erkennung stehender Hindernisse nicht lösen können. Eine weitere Möglichkeit wäre der Einsatz von Laserscannern oder Multibeam Lasersystemen, deren Vorteil in einer sehr guten aktiven Entfernungsbestimmung entlang definierter Raumrichtungen in einer oder mehreren Ebenen besteht.

Allerdings bieten sowohl Lasersensoren als auch Radarsensoren keine Möglichkeit zur Bestimmung der Positionen der Fahrspuren relativ zum Eigenfahrzeug und demnach auch keine Möglichkeit zur Zuordnung von Fremdfahrzeugen zu Fahrspuren. Die Zuordnung von Fahrzeugen zu Fahrspuren und die Veränderung ihrer relativen Lage stellt aber eine wichtige Informationsquelle für die Interpretation einer Fahrsituation mit anschließender Verhaltensentscheidung dar.

Für eine globale Bestimmung der Lage und Geometrie von Straßen wird bisweilen versucht, das a priori Wissen digitaler Straßenkarten mittels Global Positioning System (GPS) nutzbar zu machen. Allerdings können diese Bestrebungen keinen Ersatz dafür bieten, dass es derzeit und vermutlich auch in Zukunft unabdingbar sein wird, zusätzlich eine lokale Bestimmung der Lage und Geometrie des Straßenbereichs durchzuführen und seine Befahrbarkeit (Existenz von Baustellen, Schutz bezüglich veralteter oder fehlerhafter digitaler Karten) zu überprüfen. Hierfür werden Kameras eingesetzt. Aus diesem Grund bietet es sich an, bereits vorhandene Sensorik auch für die Erkennung von Fremdfahrzeugen zu verwenden.

Hierfür sind Verfahren zur monokularen Detektion und Verifikation von Fremdfahrzeugen zu entwickeln. Es sind alternative Informationsquellen, wie z.B. stereobasierte Verfahren und Verfahren zur Bestimmung des Merkmalflusses zu integrieren.

Die Verfahren sollen zur Erkennung von stehenden Hindernissen, mitbewegten Objekten, Einscherern und Überholern dienen. Hierfür soll ein möglichst großes Sichtfeld zur möglichst frühen Detektion ausgewertet werden. Zur frühzeitigen Erkennung von Einscherern und Ausscherern wären außerdem Verfahren wünschenswert, welche die Möglichkeit zur Bestimmung der Ausrichtung von Fremdfahrzeugen eröffnen.

Bei niederen Geschwindigkeiten entsprechend Abbildung 7.1 erfordert eine Erkennung von Einscherern im Nahbereich einen horizontalen Sichtbereich in der Größenordnung von mindestens  $44^\circ$ . Je nach Einschergeschwindigkeit kann es dabei allerdings aufgrund von Einschwingzeiten oder Integrationszeiten der Erkennungsalgorithmik zur Vermeidung einer Kollision bereits zu spät sein. Verdeckungen durch die Motorhaube erschweren die visuelle Erkennung relevanter Merkmale zusätzlich. Zur Steigerung der Erkennungsleistung ist entweder die Kombination mit weiteren Sensoren ratsam, die eine visuelle Objektverifikation triggern, oder eine weitere Vergrößerung des Sichtbereichs sinnvoll, um durch zeitlich-räumliche Integration von Bildmerkmalen frühzeitig Indizien für potentielle Hindernisse akkumulieren zu können, um dann rechtzeitig entsprechende Hypothesen instanzieren zu können.

Ein solch vergrößerter Sichtbereich, der auch die Möglichkeit zur Detektion von Überholern beinhalten würde, deren Heck gegenüber der eigenen Fahrzeugfront nur  $5m$  entfernt liegt, hätte z.B. einen horizontalen Öffnungswinkel von ca.  $70^\circ$ . Siehe Abbildung 7.1. Mit aktiver Blickrichtungssteuerung für einen Gierbereich von  $\pm 37,5^\circ$ , so wie es im Versuchsfahrzeug VaMP (Abbildung 6.2 und 6.3b) prinzipiell möglich ist, könnte dann ein horizontaler Sichtbereich von  $145^\circ$  verwirklicht werden, der bei einer Einbaulage hinter der Windschutzscheibe am Ort des inneren Rückspiegels die Erkennung von Fahrzeugen ermöglicht, deren Heck in der Nebenspur auf der Höhe der Front des Eigenfahrzeugs liegt. In Abbildung 7.1 ist dieser Sichtbereich als gelber Winkelbereich dargestellt.

## 7.2 Verfahren zur monokularen Detektion von Längsverkehr im Nahbereich

Bei Tag und entsprechenden Witterungsbedingungen stellt der dunkle Bereich unterhalb eines Fahrzeugs bzw. der Schattenwurf eines Fahrzeugs ein sehr markantes Merkmal für das Vorhandensein eines Fahrzeugs dar [THOMANEK 96]. Ein großer Vorteil dieses Merkmals besteht darin, dass aus der Lage des Schattens im Bild unter der Annahme einer ebenen Fahrbahn bzw. bei bekanntem vertikalen Krümmungsverlauf unter Verwendung des Messmodells die Entfernung in 3-D errechnet werden kann, welche ja infolge der perspektivischen Projektion ins Bild verloren gegangen ist. Ähnliches gilt für die Berührungspunkte der Räder von Fahrzeugen mit der Fahrbahn.

Zur Detektion des dunklen Bereichs wird deshalb zunächst innerhalb eines interessanten Bildausschnitts in vertikalen Suchpfaden von oben nach unten nach Helligkeitsübergängen von einem dunklen Schatten zur helleren Straße gesucht. Bei bekannter Fahrbahn kann der interessante Bildausschnitt auf den Bereich der Projektion der Fahrbahn ins Bild beschränkt werden.

Die *notwendige Vorausschau* ergibt sich aus dem Bremsweg, der bei der aktuell gefahrenen Geschwindigkeit bei einer vorgegeben maximalen Verzögerung notwendig ist, um vor einem stehenden Hindernis sicher anzuhalten. Einen oberen Grenzwert für den interessanten Bereich im Bild stellt stets der Horizont im Bild dar.

Infolge der endlichen Auflösung einer Kamera können Merkmale einer bestimmten Ausdehnung nur bis zu einer *möglichen Vorausschau* aufgelöst, d.h. vermessen werden. Die entsprechenden Überlegungen werden am Beispiel der Erkennung eines negativen Hindernisses nochmals ausführlich in Abschnitt 8.2.1 erläutert.

Falls der Bereich der Fahrbahn nicht bekannt ist, bietet es sich an, den interessanten Bereich auf den Bereich unterhalb des Horizonts bis zur möglichen Vorausschau und auf jeweils eine benachbarte Fahrspur typischer Breite zur Linken und zur Rechten zu begrenzen. Dies erfolgt unter der Annahme, dass sich das Eigenfahrzeug mittig in der Eigenfahrspur befindet.

### 7.2.1 Detektion von Fremdfahrzeugen durch explizite Suche nach dem Schatten unter dem Fahrzeugaufbau und nach Rädern

Für die Detektion von Pkw-Fremdfahrzeugen wurde hier eine explizite Suche nach dem Schatten unter dem Fahrzeugaufbau und nach den Rädern durchgeführt. Nachfolgend werden die Verfahrensschritte detailliert erläutert:

**Schritt 1:** Ähnlich wie in [THOMANEK 96] mit dem Kantendetektor wird hier mittels des HOFMANN-Operators im interessanten Bildausschnitt in vertikalen Suchpfaden von oben nach unten nach Helligkeitsübergängen zwischen einem dunklen Schatten und dem helleren Straßenbelag gesucht.

**Schritt 2:** Die gefundenen gerichteten Kantenpunkte werden entsprechend Abschnitt 4.3.13.2 zu Linien gruppiert.

**Schritt 3:** Unter Verwendung der bekannten Einbaulage der Kamera und der Annahme einer ebenen Fahrbahn im Nahbereich werden die zu den 2-D Linien korrespondierenden 3-D Linien in der Bodenebene bestimmt.

**Schritt 4:** Hiernach werden die 3-D Linien bezüglich ihrer Lage im Raum gefiltert. Dabei wird für Fahrzeuge in einem normalen Betriebszustand angenommen, dass ihre Längsachsen näherungsweise kollinear zum Fahrspurverlauf ausgerichtet sind. Die Schattenkante des Hecks in die Bodenebene sei demnach näherungsweise orthogonal zum Fahrspurverlauf. Deshalb werden nur solche 3-D Linien als potentielle Schattenkanten zugelassen, deren Ausrichtung relativ zur Fahrspurausrichtung bzw., falls keine Fahrspurinformation vorhanden ist, relativ zur Längsachse des Eigenfahrzeugs ( $> 45^\circ$  und  $< 135^\circ$ ) bzw. ( $< -45^\circ$  und  $> -135^\circ$ ) sind.

**Schritt 5:** Die Entfernung der Linien in  $x$ -Richtung vom Schwerpunkt des Eigenfahrzeugs sei  $> 4m$  und  $< 40m$ . Die erlaubte  $y$ -Position hängt von der Anzahl der vorhandenen Fahrspuren zur Linken und zur Rechten der Eigenfahrspur ab. Falls

keine Information über Nebenspuren verfügbar ist, werden nur 3-D Linien zugelassen deren  $y$ -Koordinaten der Anfangs- und Endpunkte im Intervall  $-6m < y < 6m$  liegen.

**Schritt 6:** Die Länge einer potentiellen Schattenkante in 3-D sei  $> 0,5m$  und  $< 2,4m$ .

**Schritt 7:** In Abhängigkeit von der Lage eines Linienmittelpunkts in 3-D bzw. der Ansicht relativ zur Kamera wird nun eine Fallunterscheidung durchgeführt, um zum einen entsprechend dem erwarteten Erscheinungsbild entfernungsabhängig Radmuster entsprechend Abschnitt 4.4 zu generieren und zum anderen um die Suchstrategie an die erwartete Ansicht in 3-D anzupassen.

**Fall 1: Ansicht von hinten rechts** (Fahrzeug in linker Spur relativ zur Eigenfahrspur): Zunächst wird nach dem Rad hinten rechts im Bereich des Endpunkts einer 3-D Linie gesucht. Falls diese Suche erfolgreich ist, wird nach einem Rad vorne rechts gesucht. Falls diese Suche nicht erfolgreich war, wird das Rad hinten rechts eventuell das Rad vorne rechts. Darum wird in typischem Abstand entgegen dem Fahrspurverlauf nach dem Rad hinten rechts gesucht. Falls diese Suche erfolgreich war, wird nach dem Rad hinten links gesucht. Falls diese Suche erfolgreich war, wird als weiterer Plausibilitätstest nach einem Nummernschild gesucht. Die Erkennung eines Nummernschilds funktioniert allerdings derzeit nur bei hellem Nummernschild auf dunklem Grund. In Abhängigkeit von den Messergebnissen wird eine Hypothese instanziiert oder verworfen. Eine Hypothese wird verworfen, wenn bei Pkws nicht wenigstens zwei Räder gefunden werden konnten.

Der Schwerpunkt einer Pkw-Hypothese wird dabei entweder aus dem rechten Vorder- und Hinterrad oder aus dem linken und rechten Hinterrad unter Verwendung Pkw-typischer initialer Werte für die Ausdehnung berechnet. Alternativ könnte bei bekannter Spurweite (Abstand zwischen linkem und rechtem Hinterrad in 3-D) nach den Außenkanten eines Pkws gesucht werden und über die Länge der Diagonale auf die Länge des Fahrzeugs geschlossen werden.

**Fall 2: Ansicht von hinten:** Falls die Linie eine typische Länge zwischen  $1,2m$  und  $2,0m$  hat, wird am Anfangspunkt der 3-D Linie nach dem linken Hinterrad und am Endpunkt der 3-D Linie nach dem rechten Hinterrad gesucht. Ist die Länge kleiner, dann sollte überprüft werden (Konzept), ob der Schatten zu einem Motorrad gehört. Hier könnte der Betragsoperator zur Bestimmung der Außenkanten zum Einsatz kommen [THOMANEK 96]. Falls es sich um kein Motorrad handelt, dann wird am Anfangspunkt nach dem linken Hinterrad und in einem typischen Abstand z.B.  $1,6m$  nach dem rechten Hinterrad bzw. am Endpunkt nach dem rechten Hinterrad und davon ausgehend in einem typischen Abstand nach dem linken Hinterrad gesucht.

**Fall 3: Ansicht von hinten links** (Fahrzeug in rechter Spur relativ zur Eigen-

fahrspur): Der Fall 3 besitzt bezüglich der Fremdfahrzeughängsachse die spiegelbildliche Logik zu Fall 1.

Für weitergehende Plausibilitätstests könnten noch die äußeren Berandungen oder die A-Holme herangezogen werden. Weiterhin könnte die Symmetrie eines potentiellen Fahrzeughecks bzw. einer potentiellen Fahrzeugfront überprüft werden. Falls eine Symmetrie vorliegt, ist das ein gutes Indiz für deren tatsächliche Existenz.

Mittels der ermittelten Berührungspunkte der Räder mit dem Boden besteht die Möglichkeit, auf die Ausrichtung eines Fahrzeugs relativ zur Fahrspur bzw. zum Eigenfahrzeug zu schließen.

**Schritt 8:** Ein neues potentiell Fahrzeug wird dann mit der Liste bereits bestehender Fahrzeughypothesen verglichen. Dies geschieht, indem bezüglich jedes Schwerpunkts einer jeden bestehenden Hypothese ein Fangbereich z.B. mit  $x \pm 2m$  und  $y \pm 1m$  definiert wird. Liegt der Schwerpunkt eines neuen potentiellen Fahrzeugs innerhalb eines solchen Fangbereichs, wird es der entsprechenden bestehenden Hypothese zugeordnet, d.h. es wird keine neue Hypothese instanziiert.

**Schritt 9:** Falls keine Übereinstimmung mit bestehenden Hypothesen gefunden wird, wird eine neue Hypothese instanziiert und initialisiert. Falls die aktuelle Eigenfahrspur im Szenenbaum repräsentiert ist, dient diese als Referenz relativ zu der die neue Hypothese eingegangen wird, ansonsten wird die neue Hypothese relativ zum EgoBase-Koordinatensystem im Szenenbaum erzeugt.

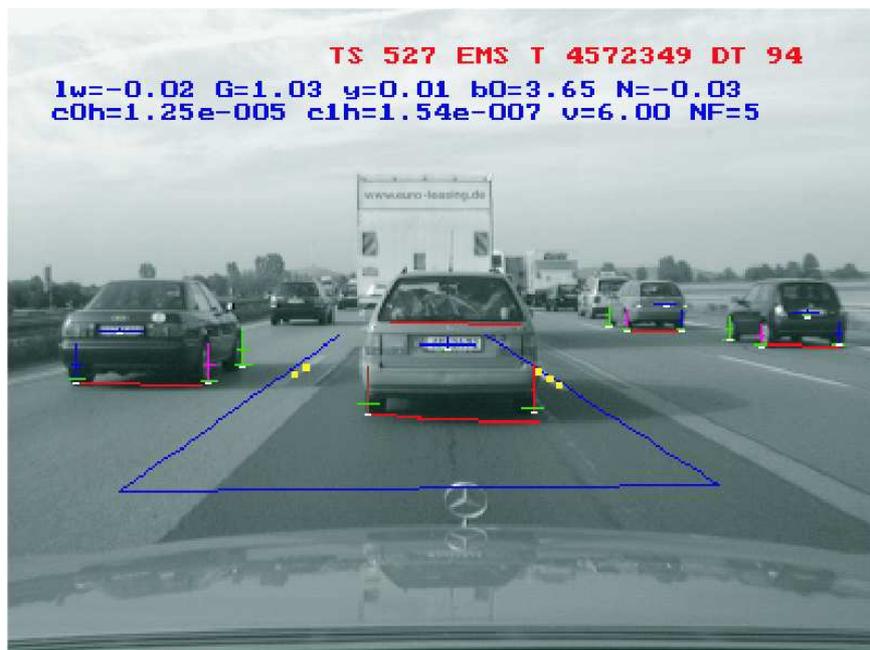
**Schritt 10:** Für jede neue Hypothese wird nun noch ein Schätzer instanziiert, initialisiert und gestartet.

In Abbildung 7.2 sind die dunklen Schattenkanten als rote Linien eingezeichnet. Weiterhin sind die detektierten Räder als Kreuz der Hauptachsen der Radellipsen dargestellt. Die Breite des grünen Querbalkens visualisiert die erwartete Reifenbreite. Die Pixeldifferenz in vertikaler Richtung zwischen diesem grünen Balken und dem weißen Berührungspunkt mit der Bodenebene entspricht dem vertikalen Delta, das nach der tendenziell horizontalen Suche durch vertikale Optimierung gefunden wurde.

Die hinteren dem Eigenfahrzeug zugewandten Räder wurden lila dargestellt, die vorderen inneren Räder grün und die hinteren äußeren Räder dunkelblau. Beim direkt vorausfahrenden Zielfahrzeug wurde das linke hintere Rad dunkelrot und das rechte hintere Rad hellrot eingezeichnet. Außerdem sind in der Abbildung die detektierten Nummernschilder durch blaue Kreuze zur Veranschaulichung ihrer Länge und Breite markiert.

## 7.2.2 Generierung einer Objekthypothese aufgrund des optischen Merkmalflusses zur Überholerdetektion

Erste Untersuchungen zur Überholerdetektion durch Extraktion eines positiven Merkmalflusses von dunkel-hell hell-dunkel Folgen in einem erweiterten Sichtbereich mittels



**Abbildung 7.2:** Suche nach Rädern im Bereich dunkler Schatten zur Detektion von Fremdfahrzeugen

divergenter Kameraanordnung wurden durch [RIEDER 00] durchgeführt. Bei dichtem Verkehr, so wie er in Stausituationen bei niederen Geschwindigkeiten vorkommt, lieferten die Algorithmen allerdings nur unbefriedigende Ergebnisse, da im mitbewegten Verkehr die Trennung zwischen den einzelnen Fahrzeugen nicht möglich war.

Ein weiterer Ansatz (Konzept) könnte darin bestehen, dass der erwartete Fluss von Merkmalen der Bodenebene aufgrund der Eigenbewegung bestimmt wird. Abweichungen gegenüber der Erwartung ergeben sich dann im Idealfall (keine Fehlzuordnung) nur durch Objekte, die sich entweder selbst bewegen oder aus der Bodenebene hervorstehen. Dabei sollte bei bekannter Lage und Ausdehnung der Fahrbahn nur im Bereich der Fahrbahn der Merkmalfluss bestimmt werden. Die Untersuchung dieses Detektionsansatzes bleibt weiteren Arbeiten überlassen.

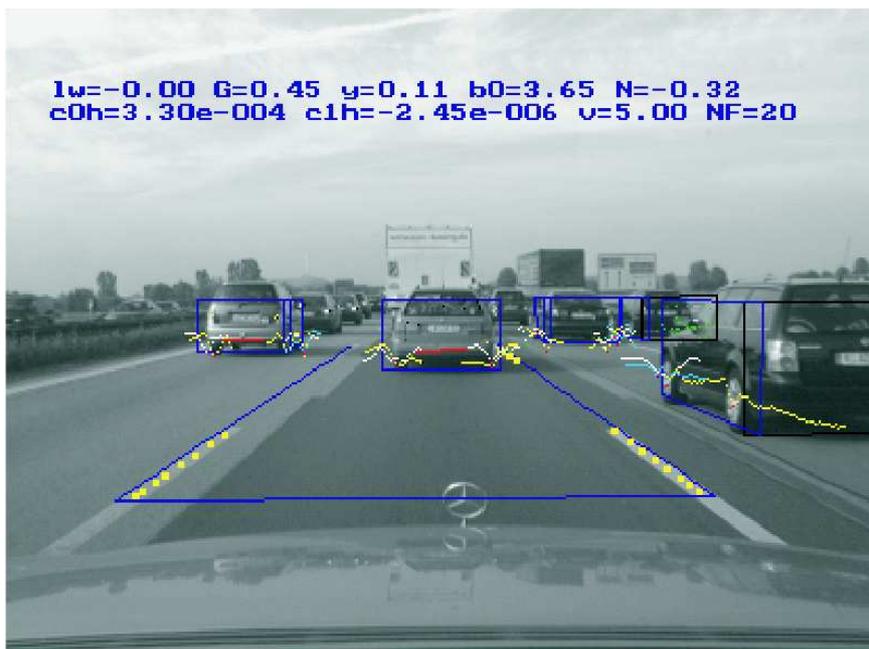
### 7.3 Verfolgung einer Objekthypothese

In Kapitel 5.2.1 wurde am Beispiel des Pkw-Formmodells die Modellierung von Objekten mittels generischer Modelle erläutert. Es wurde die Parametrisierung der Bildverarbeitung am Beispiel der Raderkennung und der Bestimmung der äußeren Berandung mittels Betragsoperator beschrieben. Die Generierung der Radmuster erfolgt dabei in Abhängigkeit von den Aspektbedingungen entsprechend Kapitel 5.2.2, der Entfernung in 3-D und

den Abbildungseigenschaften der Kamera-Objektiv-Kombination. Die Messung im Bild erfolgt in Abhängigkeit von der Sichtbarkeit der Messpunkte. Hierfür wurde entsprechend Abschnitt 5.2.3 zunächst die theoretische Sichtbarkeit bestimmt und durch die Analyse gegenseitiger Verdeckungen entsprechend Abschnitt 5.2.4 ergänzt.

Die Bereiche der Verdeckungspolygone wurden weiterhin für die Bestimmung des horizontalen Merkmalflusses verwendet, die zur Stützung der Zustandsschätzung durch unspezifische Merkmalkonstellationen entsprechend Kapitel 5.3.1 dienen. Ergänzend hierzu wurde bei divergenter Anordnung in Kapitel 5.3.2 eine Möglichkeit zur Einbringung expliziter Stereoinformation beschrieben.

Abbildung 7.3 zeigt als blaue Rechtecke die Bereiche der ins Bild projizierten Verdeckungspolygone, in denen der Merkmalfluss bestimmt wird. Am rechten Rand sind zwei schwarze Verdeckungsbereiche eingezeichnet. Das schwarze Rechteck im Vordergrund wurde als ungültig markiert, weil es teilweise außerhalb des Bildbereichs liegt. Das dahinterliegende weiter entfernte Rechteck ist schwarz, weil es durch ein näher gelegenes Rechteck zum Teil verdeckt wird. Es ist zu erkennen, dass die Messpunkte des verdeckten Objekts innerhalb des verdeckenden Polygons als grüne Punkte markiert sind. Diese sollen verdeutlichen, dass sie als gesperrte Messpunkte eingetragen wurden und deshalb dort keine Messung erfolgt.



**Abbildung 7.3:** Verfolgen von Fremdfahrzeughypothesen. Schwarze Polygone werden durch andere Polygone verdeckt oder liegen zum Teil außerhalb des Bildbereichs.

In Abbildung 7.3 sind weiterhin die Signalverläufe der Korrelationen der Radmuster (gelb und cyan) bzw. des Betragsoperators (weiß) entlang der Suchpfade im Bild eingezeichnet,

wobei die roten Punkte im gelben, cyan oder weißen Farbverlauf die gefundene Position eines Hinterrades, respektive Vorderrades und einer äußeren Berandung darstellen.

## 7.4 Zusammenfassung

In Abschnitt 7.2.1 wurde ein Verfahren zur Detektion von Pkw-Fremdfahrzeugen aufgrund des dunklen Bereichs unter dem Fahrzeugaufbau und der Räder in Abhängigkeit von den Aspektbedingungen vorgestellt. Die Suche nach den Rädern erfolgt mittels generischer Radmuster, welche in Abhängigkeit von der Entfernung und der Ansicht an das erwartete Erscheinungsbild angepasst werden. Es wurde eine einfache Suchlogik realisiert, welche in Abhängigkeit von den Suchergebnissen die Suchfenster für die Suche nach weiteren Merkmalen parametrisiert und daraus Hypothesen generiert. Über die Positionen der Berührungspunkte der Räder mit der Bodenebene besteht zum einen die Möglichkeit, auf die Entfernung in 3-D zu schließen und zum anderen die Ausrichtung eines Fremdfahrzeugs relativ zum Eigenfahrzeug und zu den Fahrspuren zu bestimmen.

# Kapitel 8

## Erkennung negativer Hindernisse

Die Detektion von Hindernissen und die Bestimmung ihrer Ausdehnung, Position und Ausrichtung relativ zum eigenen Fahrzeug ist beim Querfeldeinfahren unbedingt notwendig für eine sichere Fahrt. Es können hierbei positive Hindernisse, wie z.B. Steine, Büsche und Bäume, also Objekte, welche relativ zur gerade befahrenen Bodenoberfläche aus dieser hervorstehen, und negative Hindernisse, wie z.B. Schlaglöcher oder Gräben, auftreten. In diesem Kapitel soll der Schwerpunkt auf der Erkennung und Vermessung eines negativen Hindernisses in Form eines Grabens liegen. Diese Aufgabe stellt aufgrund der sehr ungünstigen, entfernungsabhängigen Aspektbedingungen eine besondere Herausforderung dar.

Bei rein monokularer Bildauswertung ist die Klassifizierung, ob es sich bei einem Objekt um ein positives oder negatives Hindernis handelt bzw. ob es überhaupt ein Hindernis ist, äußerst schwierig. In einem Grauwertbild oder Farbbild können unter schlechten Beleuchtungsbedingungen z.B. ein quer liegender Baumstamm, ein Graben oder der Schatten eines Baumes bei tiefstehender Sonne bezüglich Form und Helligkeit nicht zu unterscheiden sein. Aus diesem Grund wird hier beim Querfeldeinfahren für die Erkennung von Hindernissen eine Kombination aus monokularer und stereoskopischer Bildauswertung gewählt.

Das Wahrnehmungssystem verwendet dabei eine von der Firma Sarnoff entwickelte Bildverarbeitungshardware für Stereo-Bildauswertung in Echtzeit. Diese Hardware liefert sowohl ein Grauwertbild als auch das zugehörige Disparitätenbild, wodurch Probleme bei der Datenfusion zwischen Grauwertbild und Disparitätenbild vermieden bzw. verringert werden. Zu jedem Bildpunkt der linken Kamera, welcher einem Bildpunkt der rechten Kamera zugeordnet werden konnte, wird ein Disparitätenwert relativ zur sogenannten *Horopterebene* ermittelt, der durch Triangulation bei bekannter Stereobasis in einen Entfernungswert umgerechnet werden kann. Näheres hierzu ist in Abschnitt 8.1.3 zu finden. Bildmerkmale, welche mittels Kantenoperatoren aus dem Grauwertbild bestimmt werden, stellen meist eine Obermenge für die Merkmale dar, welche aus dem Disparitätenbild zu bestimmen sind. Eine Hypothesengenerierung unter Verwendung von Grauwertkanten kann deshalb zumeist früher, das heißt, mit größerer Distanz zu einem potentiellen

Hindernis erfolgen, denn das Signal-Rausch-Verhältnis ist günstiger als das des Disparitätenbildes. Im Nahbereich dagegen kann der umgekehrte Effekt beobachtet werden, denn im Grauwertbild werden dann mehr und mehr irrelevante Kantenpunkte extrahiert.

Auf dem derzeitigen Entwicklungsstand werden bei der Suche nach negativen Hindernissen zunächst aus dem Grauwertbild in einem durch eine sinnvolle Vorausschau festgelegten interessanten Bereich entlang vertikaler Suchpfade gerichtete Kantenpunkte extrahiert. Anschließend wird für jeden gerichteten Kantenpunkt im Disparitätenbild überprüft, ob eine signifikante Diskontinuität bzw. ein signifikanter Übergang in der Disparität festzustellen ist, der auf eine Höhenänderung schließen lässt. Näheres hierzu findet sich im Abschnitt 8.2.4.

In Abschnitt 8.2.5 werden die bestätigten Kantenpunkte zu 2D-Linien verkettet. Wobei das Verkettungsverfahren nach Kapitel 4.3.13.2 zum Einsatz kommt.

Die 2D-Linien werden dann auf Parallelität und Kolinearität hin überprüft, gegebenenfalls zusammengefasst und daraus eine Hypothese generiert. Dieses Vorgehen erlaubt es, zwischen Schatten und negativen Hindernissen zu unterscheiden (siehe Abschnitt 8.2.6 und 8.2.7). Unter einer Hypothese wird hier eine konkrete parametrisierte Repräsentation eines potentiellen realen Objekts im Rechner verstanden. Diese Repräsentation besteht aus einem generischen Formmodell und einem Lagemodell mit sechs Freiheitsgraden. Die mittels Kalmanfilterung zu schätzenden Parameter dieser Modelle werden als Zustandsgrößen bezeichnet.

Sobald ein potentielles Hindernis detektiert wurde, wird überprüft, ob es einer bereits bestehenden Hypothese zugeordnet werden kann. Falls ja, werden die Form- und Lageparameter dieser Hypothese mit den neuen Messdaten innoviert. Falls nein, wird eine neue Hypothese generiert. Diese wird dann zunächst über mehrere Videozyklen verfolgt, bevor sie als ein neues Objekt im allgemeinen Szenenbaum veröffentlicht wird. Dies bewirkt, dass die Wahrnehmung unempfindlicher gegenüber falschen Hypothesen wird. Nach der Veröffentlichung stehen die Daten eines Objekts auch anderen Experten zur Verfügung. Dies wird in Kapitel 8.3 erläutert.

Basierend auf den aktuellen Form- und Positionsschätzungen wird die Sichtbarkeit jedes hypothetisierten Objektes im aktuellen Kamerabild überprüft. Wird ein als sichtbar erwartetes hypothetisiertes Objekt über einen längeren Zeitraum nicht bestätigt, so wird die Länge der korrespondierenden Hypothese zunächst verkürzt und danach, falls weiterhin keine Bestätigung möglich ist, die Hypothese gelöscht. Details hierzu finden sich in Abschnitt 8.3.2.

Die Innovation einer Hypothese erfolgt ebenfalls in Abhängigkeit von der erwarteten Sichtbarkeit eines hypothetisierten Objekts. Sind nur Teilbereiche der Projektion einer Hypothese im Bild sichtbar, so sind nicht alle Zustandsgrößen direkt bestimmbar. Die Vorgehensweise zur Steuerung der Innovation wird in Abschnitt 8.3.3 beschrieben.

Im Anschluss wird in Kapitel 8.4 das Zusammenspiel zwischen Wahrnehmung, Blickrichtungssteuerung und Fahrzeugführung erläutert.

Zur gezielten Blickzuwendung legt der Wahrnehmungsexperte Aufmerksamkeitsbereiche fest, die die „Bedürfnisse“ der Wahrnehmung berücksichtigen. In Abhängigkeit von der Verhaltensentscheidung kann die zentrale Entscheidungsinstanz dann durch Parametrisierung des Wahrnehmungsexperten zwischen fünf verschiedenen Wahrnehmungsmodi wählen. Die Wahrnehmungsmodi und die Festlegung der Aufmerksamkeitsbereiche werden in Abschnitt 8.4.1 beschrieben.

In Abschnitt 8.5 werden dann die Ergebnisse einer realen Versuchsfahrt diskutiert.

## 8.1 Hardware

Die Hardware, welche bei der Erkennung negativer Hindernisse zum Einsatz kommt, kann man im Wesentlichen in drei Teilbereiche unterteilen. Der erste Teilbereich umfasst die eingesetzten Kameras, die Kameraanordnung und die verwendeten Objektive. Der zweite Teilbereich beinhaltet die für die Steuerung der Blickrichtung notwendige Zwei-Achsen Kamera-Plattform. Und der dritte Teilbereich bezieht sich auf die eingesetzte Hardware zur Echtzeit Stereobildauswertung.

### 8.1.1 Kamerakonfiguration

Für die Grabenerkennung werden als Bildsensoren zwei parallel angeordnete  $\frac{1}{2}$ ” Schwarz-Weiß-Fingerkameras der Firma Teli Typ CS 3150C mit C-Mount Adaptern und C-Mount Objektiven verwendet. Die Stereobasis  $b$  beträgt  $27\text{cm}$ . Als Objektive für die beiden Kameras der Stereokonfiguration werden zwei identische Objektive der Firma Cosmicar mit  $f = 16\text{mm}$  Brennweite verwendet. Aufgrund der großen Brennweite können Linsenverzerrungen vernachlässigt und das klassische Lochkameramodell verwendet werden. Eine schematische Darstellung der Stereoanordnung zeigt Abbildung 8.1. Die Projektionszentren der beiden Kameras befinden sich in den Punkten  $c_L$  und  $c_R$ .

Bei paralleler Kameraanordnung gilt, dass die Epipolarlinien mit den Bildzeilen zusammenfallen. Ein Raumpunkt  $P$  mit dem Abstand  $x$  zur Basislinie wird mit  $y$ -Koordinate  $y_L$  in der linken und  $y_R$  in der rechten Bildebene abgebildet. Die Disparität  $d$  der Projektion eines Raumpunktes im rechten Bild relativ zur Projektion in die linke Bildebene ergibt sich demnach zu

$$d = y_L - y_R \quad (8.1)$$

bzw.  $y_R = y_L - d$ .

Bei bekannter Entfernung  $x$ , Brennweite  $f$  und Stereobasis  $b$  kann die Disparität auch unter Anwendung des Strahlensatzes aus

$$d = f \frac{b}{x} \quad (8.2)$$

bestimmt werden. Im Umkehrschluß kann bei bekannter Disparität  $d$ , Brennweite  $f$  und Stereobasis  $b$  die Entfernung  $x$  errechnet werden. Abbildung 8.2 zeigt eine Fotografie der Kameraanordnung im Versuchsfahrzeug VAMORS. Das parallele Stereokamerapaar ist durch Pfeile markiert.

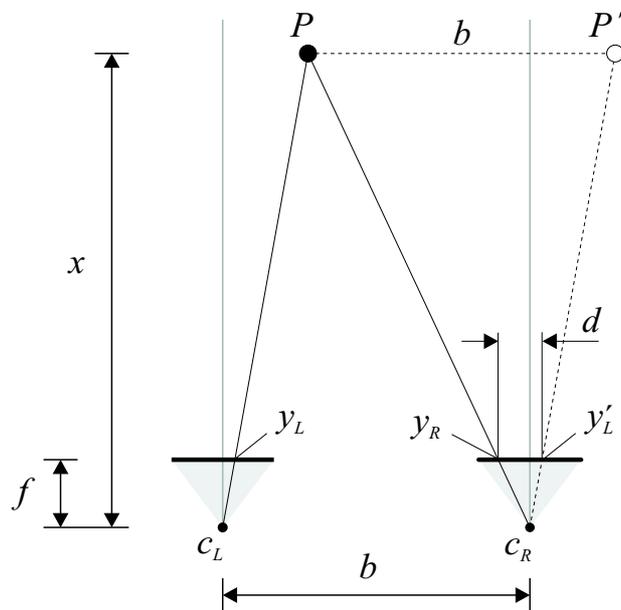


Abbildung 8.1: Stereokameraanordnung mit parallelen optischen Achsen

### 8.1.2 Aktive Zwei-Achsen-Kameraplattform



Abbildung 8.2: Fotografie der Kameraplattform mit paralleler Stereokameraanordnung

Wie in Abbildung 8.2 zu sehen, sind die Kameras des EMS-Vision Systems im Versuchsfahrzeug VAMORS auf einer drehbaren Kameraplattform montiert. Diese ist in der Lage,

sowohl die Gierachse als auch die Nickachse auf einen vorgegeben Sollwinkel zu stellen. Im Wesentlichen wird die Kameraplattform dazu benutzt, schnelle Blickzuwendungen, sogenannte *Sakkaden*, und Fixierungen auf ein zu vermessendes Objekt durchzuführen. Innerhalb eines Planungszeitraums der Blickrichtungssteuerung, dessen Dauer je nach Situation mit 3-5 *sec* fest vorgegeben ist, kann nach [PELLKOFER 03] die Plattform 2-3 Sakkaden ausführen, wobei jede Sakkade ca. 400 *msec* benötigt.

### 8.1.3 Acadia-Stereobildverarbeitungshardware

Für die Echtzeit Stereobildauswertung wird eine Spezialhardware der Firma Sarnoff eingesetzt. Hierbei handelt es sich um eine PC-PCI-Einsteckkarte mit der Typbezeichnung *Acadia*. Diese ist in Abbildung 8.3 zu sehen. Als Basisrechner dient ein handelsüblicher Single-Pentium III 850 MHz unter Windows NT 4.0.

Die Acadia-Stereoalgorithmen rechnen zunächst das rechte Kamerabild mittels einer affinen Transformation auf das linke Kamerabild um. Diese affine Transformation ist im Idealfall im Prinzip nichts anderes, als eine Beschreibung für die Ebene relativ zu der die Summe aller Disparitäten minimal ist - die sogenannte *Horopterebene*. Für den Fall, dass diese mit der Bodenebene zusammenfällt, bedeutet das anschaulich, dass die Punkte der Bodenebene eine relative Disparität null besitzen und diejenigen, welche höher oder tiefer liegen, eine positive bzw. negative Disparität erhalten. Im Disparitätenbild erscheinen deshalb bezogen auf die Bodenebene Erhöhungen heller und Vertiefungen dunkler, wobei einer relativen Disparität von null der 8-Bit Grauwert 127 entsprechen soll. Dies ist beispielsweise der Abbildung 8.5 zu entnehmen. Für weitere Details zur Acadia und als Einstieg in die zugrunde liegenden Basisalgorithmen können die Veröffentlichungen [BURT 83], [BURT 95] oder [VAN DER WAL 00] dienen.



**Abbildung 8.3:** Acadia PC-PCI-Einsteckkarte der Firma Sarnoff zur Echtzeit Stereobildauswertung

## 8.2 Prinzipielle Vorgehensweise zur Erkennung negativer Hindernisse

In den nachfolgenden Kapiteln 8.2.1 bis 8.2.7 werden die zur Erkennung und Vermessung eines negativen Hindernisses notwendigen Schritte erläutert. Als Beispiel dient eine Schwarz-Weiß-Aufnahme der linken Kamera des Stereosystems entsprechend Abbildung 8.4 und das dazugehörige Disparitätenbild eines Grabens aus Abbildung 8.5. Im unteren Drittel des Schwarz-Weiß-Bildes wurde ein künstlicher Schatten eingefügt, indem in diesem Bereich alle Helligkeitswerte halbiert wurden. Dieser künstliche Schatten wird dazu dienen, die Fähigkeit zur Unterscheidung zwischen einem Schatten und einer echten Vertiefung zu erklären.

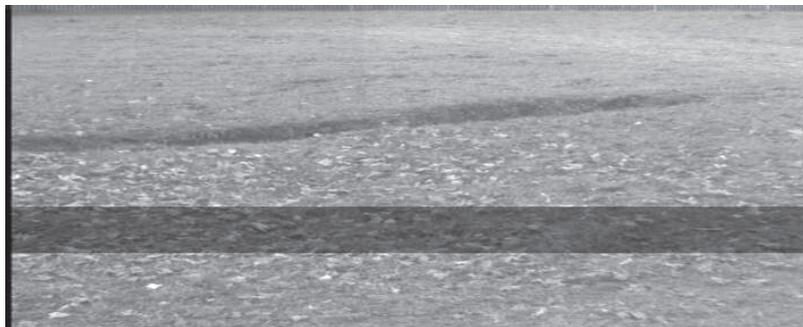


Abbildung 8.4: Schwarz-Weiß-Aufnahme eines Grabens



Abbildung 8.5: Disparitätenbild eines Grabens

### 8.2.1 Festlegung eines interessanten Bereichs für die Bildauswertung

Bei der Durchführung einer autonomen Fahrmission ist das Gebiet, durch das eine zukünftige Fahrtrajektorie verlaufen soll, auf Befahrbarkeit zu überprüfen. Prinzipiell handelt es sich bei diesem Gebiet um ein dreidimensionales Raumvolumen. Dieses hat im Idealfall die

Form eines Schlauchs, der minimal die Ausmaße des Eigenfahrzeugs in Höhe und Breite besitzt.

Bevor nun mit der Bildauswertung begonnen werden kann, ist der für die Bildverarbeitung interessante Bereich festzulegen. Dies erfolgt zunächst in 3D. Durch Projektion in die Bildebene ergibt sich dann der interessante Bereich im Bild. Zur Vereinfachung wird angenommen, dass die Bodenoberfläche im Nahbereich des Fahrzeugs durch eine Ebene angenähert werden kann. Unter dem *Vorausschaubereich* versteht man den Bereich, in dem ein gesuchtes Merkmal einer vorgegebenen Größe gerade noch vermessbar, d.h. wahrnehmbar, ist. Der Vorausschaubereich sollte den Bereich der Projektion des Fahrerschlauchvolumens auf die Bodenebene einschließen.

Die *mögliche Vorausschau* wird durch die Wahl der Objektivbrennweite, die Anbringung der Kameras am Fahrzeug und die Auflösung der CCD-Chips der Kameras beeinflusst. Als Randbedingung ist festzulegen, ab welcher Ausdehnung ein Merkmal eines Hindernisses erkannt werden soll. Für die Erkennung eines solchen Merkmals im Bild ist es notwendig, dass die Projektion des Merkmals ins Bild in Suchrichtung mindestens so viele Pixel überdeckt, wie eine halbe Suchmaske groß ist. Legt man z.B. fest, dass die kleinste halbe Maskenbreite drei Pixel betragen soll und z.B. ab einer Größe von 0,25 m Breite eine Vertiefung als Hindernis aufgefasst werden soll, dann ist damit bei bekannten Aspektbedingungen die maximale mögliche Vorausschau für die Wahrnehmung festgelegt.

Die *notwendige Vorausschau* für ein sicheres Anhalten bzw. Ausweichen vor einem Hindernis ergibt sich aus der Geschwindigkeit des Eigenfahrzeugs, denn diese legt bei bekannter Verzögerung den minimalen Bremsweg fest. Ist die mögliche Vorausschau kleiner als die notwendige Vorausschau aufgrund der momentanen Geschwindigkeit, so muss für ein sicheres Fahren die Geschwindigkeit verringert werden.

Unter der Annahme einer ebenen Bodenoberfläche kann die obere begrenzende Bildzeile der Projektion des möglichen Vorausschaubereichs ins Bild näherungsweise dadurch bestimmt werden, dass für jedes Pixel der vertikalen Mittellinie des Bildes die Schnittpunkte der Sehstrahlen durch diese Pixel mit der Bodenebene bestimmt werden. Die Entfernungsdaten werden in einer Look-Up-Tabelle (LUT) gespeichert, welche die Zuordnung zwischen Bildzeile und Entfernung in 3D ermöglicht. Bei Verwendung vertikaler Suchpfade ist nun zu überprüfen, ob jeweils 3 aufeinanderfolgende Bildzeilen weniger als 0,25 m der Bodenebene überdecken. Solange diese Bedingung erfüllt ist, ist die Wahrnehmung in der Lage, solche Merkmale aufzulösen.

Die untere Grenze des notwendigen Vorausschaubereiches hängt vom bisherigen Verlauf des Wahrnehmungsvorgangs ab. Wurde das System gerade aus dem Stand gestartet und ist kein Vorwissen über den direkt vor dem Fahrzeug liegenden Bereich vorhanden, so muss der Vorausschaubereich bereits bei 0m bezogen auf das vordere Ende des Fahrzeugs beginnen. In Abhängigkeit von den zur Verfügung stehenden Kameras und deren Sichtbereichen, ist gegebenenfalls ein Absuchen des Vorausschaubereichs durch aktive Blickrichtungssteuerung notwendig.

Befindet sich das Fahrzeug in Bewegung, so wird in aller Regel davon ausgegangen, dass sich in einem bestimmten Bereich direkt vor dem Fahrzeug kein Hindernis befindet, denn

wenn dort ein Hindernis wäre, so wäre es bereits bei der Untersuchung dieses Bereichs zu einem früheren Zeitpunkt detektiert worden.

Wurde der Nahbereich überprüft, kann die untere Grenze des Vorausschaubereichs sukzessive vergrößert werden, so dass sich auch die obere Grenze des Vorausschaubereichs vergrößert. Solange die notwendige Vorausschau kleiner als die aktuell mögliche Vorausschau ist, darf die Geschwindigkeit erhöht werden. Diejenigen Bildzeilen, die diesen Vorausschaubereich begrenzen, legen den Bereich von Interesse (engl. „Region of Interest“ = ROI) fest. Das heißt, es werden je nach Plattformstellung die oberste und unterste Bildzeile für die Merkmalsextraktion in vertikalen Suchpfaden festgelegt (Abbildung 8.6).

Die Wahrnehmung sollte demnach die Fortbewegung dahingehend beeinflussen, dass sie eine Obergrenze für die erlaubte Geschwindigkeit in Abhängigkeit vom Auflösungsvermögen für Merkmale bzw. von den Wahrnehmungsfähigkeiten festlegt.

Die Breite der ROI im Bild richtet sich nach der Breite des Bereichs gültiger Daten im Disparitätenbild, welcher bedingt durch die Stereoverarbeitungsalgorithmen im Vergleich zur Größe des Grauwertbildes der linken Kamera deutlich kleiner ist. Das Grauwertbild ist in dieser Anwendung 704x288 Pixel groß. Der auswertbare Pixelbereich im Disparitätenbild liegt horizontal näherungsweise im Intervall [80, 650] Pixel und vertikal im Intervall [20, 280] Pixel. Dies wird in Abbildung 8.7 durch den weißen Rahmen verdeutlicht.

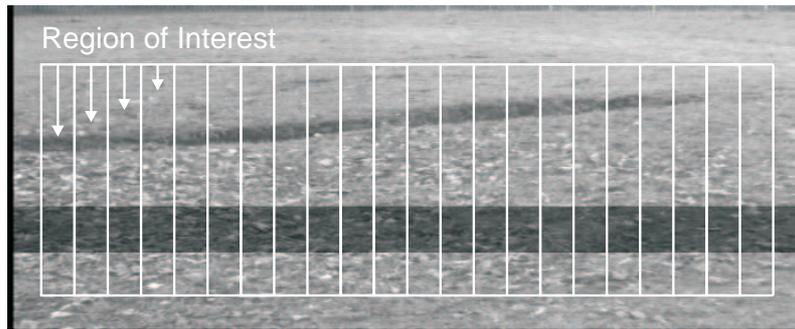
Dadurch dass kein Vorwissen über Vorhandensein, Lage und Ausdehnung von irgendwelchen Hindernissen vorhanden ist, gibt es keinen plausiblen Grund den Detektionsbereich weiter einzuschränken.

## 8.2.2 Adaptive Festlegung der Maskengröße für Bildoperatoren

Die Maskengröße wird stets so gewählt, dass im zentralen Bereich der ROI unabhängig von der Plattformstellung möglichst die gleiche Fläche in 3D überdeckt wird. Unter Verwendung der LUT wird berechnet, wie weit die Projektionen zweier benachbarter Pixel in der Bodenebene auseinanderliegen. Daraus wird dann die Anzahl der Pixel bestimmt, welche notwendig ist, um näherungsweise einen gewünschten Bereich abzudecken. Diese Anzahl legt dann die Größe der halben Bildoperatorenmaske fest. Je kleiner die Maskengröße, desto feinere Helligkeitsübergänge können ermittelt werden. In Abschnitt 8.2.4 in Abbildung 8.9 kann man die je nach Entfernung unterschiedlichen Maskengrößen bei der Überprüfung der Grauwertkantenpunkte im Disparitätenbild sehr schön erkennen.

## 8.2.3 Kantensuche im Grauwertbild

Die Suche nach Grauwertübergängen erfolgt im linken Kamerabild unter Verwendung des HOFMANN-Operators aus Kapitel 4.3. Dabei werden entlang vertikaler Suchpfade, wie in Abbildung 8.6 dargestellt, gerichtete Kantenpunkte extrahiert, sofern sie eine vorgegebene Schwelle für den Helligkeitsübergang überschreiten. Die Ergebnisse der Kantensuche sind



**Abbildung 8.6:** Festlegung des auszuwertenden Bildbereichs



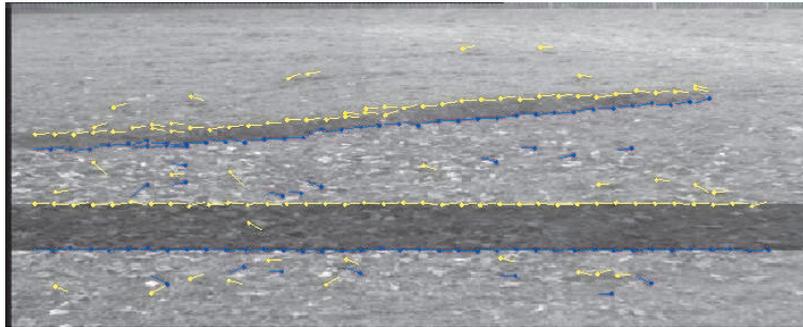
**Abbildung 8.7:** Bereich gültiger Disparitäten im Disparitätenbild

in Abbildung 8.8 dargestellt. Die linksgerichteten Fähnchen stellen einen Übergang von Dunkel nach Hell, die rechtsgerichteten Fähnchen von Hell nach Dunkel dar.

Unter Verwendung der Informationen über die homogenen Grauwertsegmente der Suchpfade ist eine Helligkeitsanalyse entsprechend Abschnitt 4.3.10 realisierbar. Mittels der charakteristischen photometrischen Kenngrößen der Helligkeitswerte innerhalb der ROI kann die Rauschschwelle adaptiv an die aktuellen Beleuchtungsverhältnisse angepasst werden. Diese Helligkeitsanalyse ist derzeit noch nicht in diese Anwendung integriert.

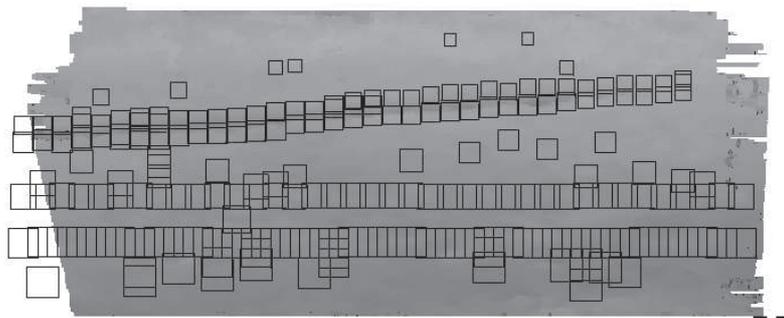
#### 8.2.4 Validierung der Kantenpunkte im Disparitätenbild

Für alle gefundenen gerichteten Kantenpunkte wird im Disparitätenbild überprüft, ob der Grauwertübergang mit einem Disparitätenübergang zusammenfällt. Diejenigen Kantenpunkte, welche dieser Überprüfung genügen, werden als gültig markiert. Die Maskengröße für die Überprüfung des Disparitätenübergangs wird wiederum in Abhängigkeit vom überdeckten Bereich in 3D festgelegt. Das heißt, die Maskengröße für Bildbereiche, welche weit entfernten 3D-Bereichen zugeordnet werden, ist kleiner als für Pixel des Nahbereichs. Abbildung 8.9 zeigt die unterschiedlichen Maskengrößen als schwarze Quadrate. Die als gültig markierten Kantenpunkte sind in Abbildung 8.10 dargestellt. Die Grauwertkanten, welche durch den künstlichen Schatten hervorgerufen wurden, konnten hierbei von den „echten“

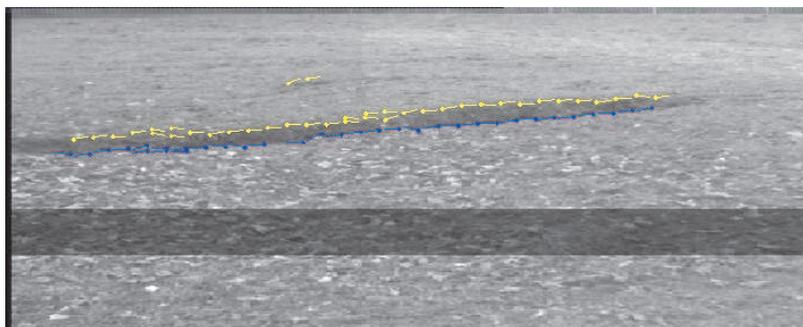


**Abbildung 8.8:** Kantensuche im Grauwertbild

Grabenkantenpunkten getrennt werden. Die Merkmale des künstlichen Schattens bleiben deshalb bei allen weiteren Betrachtungen unberücksichtigt.



**Abbildung 8.9:** Validierung der gerichteten Kantenpunkte



**Abbildung 8.10:** Darstellung der bestätigten Kantenpunkte

### 8.2.5 Verkettung der Kantenpunkte zu Kanten

Die gültigen Kantenpunkte werden unter Verwendung der Informationen über Kantenrichtung, räumliche Nähe und der Art des Helligkeitsübergangs zu Kanten verkettet. Es wird das Verkettungsverfahren aus Kapitel 4.3.13.2 verwendet. Das Ergebnis dieser Verkettung ist eine Liste von Kantensegmenten, wie in Abbildung 8.11 dargestellt.

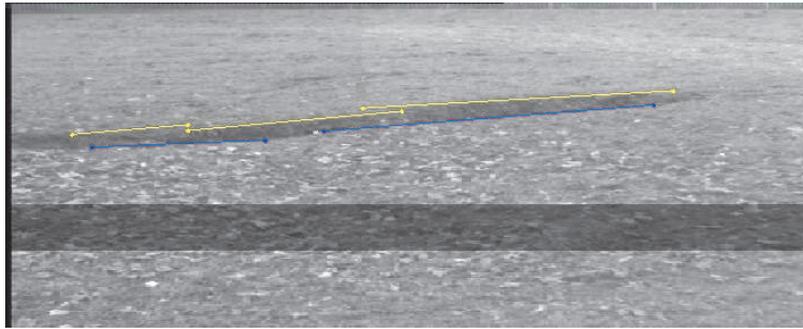


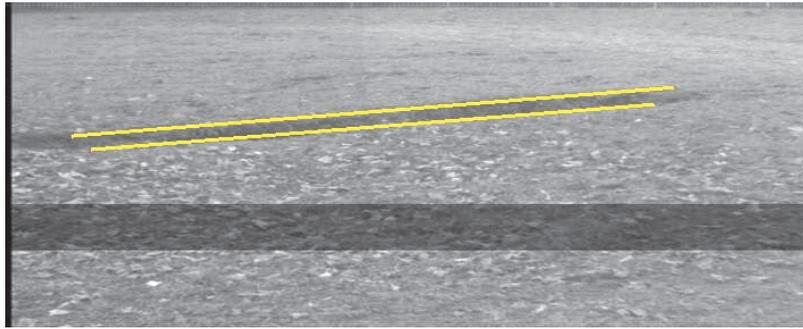
Abbildung 8.11: Verkettung der Kantenpunkte zu Kanten

### 8.2.6 Verknüpfung der Kanten zu Linien

Diese Liste der Kantensegmente wird nun nach möglichst langen hell-dunkel dunkel-hell Kombinationen durchsucht. Es werden dabei Kombinationen gebildet, welche möglichst viele gültige gerichtete Kantenpunkte einschließen. Als Kriterium für die Gültigkeit einer Kombination dienen die Kolinearität gleichartiger Übergänge und die näherungsweise Parallelität der hell-dunkel bezüglich der dunkel-hell Kante. Das Ergebnis dieser Suche sind demnach zwei Linien im Bild, welche u.a. durch ihre Anfangs- und Endpunkte und die Art des Helligkeitsübergangs beschrieben werden. Das Ergebnis der Suche ist in Abbildung 8.12 dargestellt. Hieraus wird dann eine Hypothese für einen potentiellen Graben generiert. Die Mathematik für die Berücksichtigung von Nebenbedingungen bei der Berechnung von Regressionsgeraden kann im Anhang A.5 nachgelesen werden.

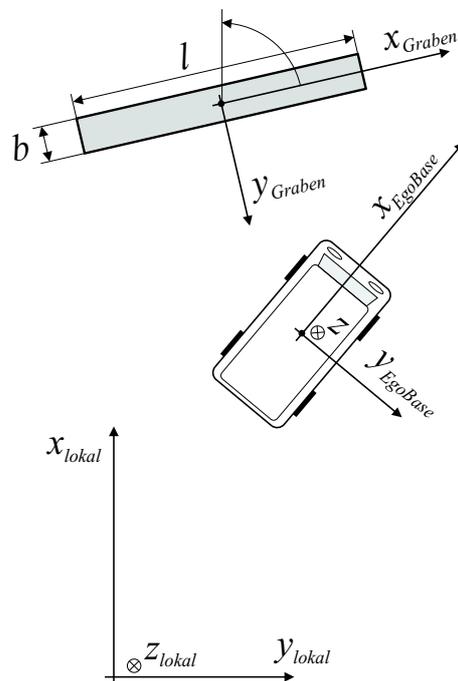
### 8.2.7 Erzeugung einer Grabenhypothese

Zur Modellierung der Form eines Grabens wird ein einfaches in der Bodenebene liegendes Rechteck verwendet. Das Formmodell besitzt als Parameter eine Länge  $l$  und eine Breite  $b$ . Der Ursprung des rechtshändigen Modellkoordinatensystems befindet sich jeweils im Schwerpunkt des Grabens. Die  $x$ -Achse des Koordinatensystems fällt mit der längeren Hauptachse des Grabens zusammen. In dieser Richtung wird auch die Länge des Grabens bestimmt - orthogonal dazu die Breite. Die Beschreibung der Lage eines Grabens erfolgt jeweils relativ zu einem ortsfesten geodätischen Koordinatensystem - im folgenden



**Abbildung 8.12:** Kombination der Kantensegmente zu Linien

als *lokales Koordinatensystem* bezeichnet - , welches zu Beginn eines jeden Missionsabschnitts zur Wegpunktnavigation parallel zur aktuellen Fahrzeugausrichtung instanziiert wird. Die Position des Eigenfahrzeugs relativ zu diesem Koordinatensystem wird mittels inertial-basierter Eigenzustandsschätzung (IBSE) unter Verwendung von GPS, von Winkelratensensoren und einem Odometer stets aktualisiert [SIEDERSBERGER 03]. Das Formmodell eines Grabens und die Relativlage zum lokalen Koordinatensystem und zum Eigenfahrzeug sind in Abbildung 8.13 dargestellt.



**Abbildung 8.13:** Modellierung eines Grabens und Beschreibung der Relativlage zum Eigenfahrzeug

Zur Bestimmung der Modellparameter werden ausgehend von den Ergebnissen aus Abschnitt 8.2.6 zunächst die Eckpunkte der Vorderkante des Grabens in die Bodenebene projiziert und daraus der Gierwinkel des Grabens relativ zum lokalen Koordinatensystem ermittelt. Abbildung 8.14a zeigt in einer Draufsicht die in die Bodenebene projizierte Vorder- und Hinterkante eines potentiellen Grabens. In Abbildung 8.14b ist die Berechnung der Lotfußpunkte der Projektionen der linken und rechten Ecke der Hinterkante auf die Vorderkante des Grabens dargestellt. Mit dem Ortsvektor  $\vec{p}_0$  als Vektor  $\vec{a}$  zum Aufpunkt  $P_0$  und dem Richtungsvektor  $\vec{r} = \vec{p}_1 - \vec{p}_0$  errechnen sich die Ortsvektoren der Lotfußpunkte  $L_2$  und  $L_3$  aus

$$\begin{aligned} k_2 &= \frac{\vec{r} \circ (\vec{p}_2 - \vec{a})}{|\vec{r}|^2} \\ \vec{l}_2 &= \vec{a} + k_2 \cdot \vec{r} \end{aligned} \quad (8.3)$$

bzw.

$$\begin{aligned} k_3 &= \frac{\vec{r} \circ (\vec{p}_3 - \vec{a})}{|\vec{r}|^2} \\ \vec{l}_3 &= \vec{a} + k_3 \cdot \vec{r} \end{aligned} \quad (8.4)$$

Der Abstand von  $P_2$  bzw.  $P_3$  zur Vorderkante wird als  $b_2$  bzw.  $b_3$  bezeichnet. Die Breite  $b$  des Grabens errechnet sich näherungsweise aus dem Mittelwert von  $b_2$  und  $b_3$ .

$$b = \frac{b_2 + b_3}{2} \quad (8.5)$$

Die Skalare  $k_2$  bzw.  $k_3$  nehmen für den Fall, dass der Lotfußpunkt innerhalb der Strecke  $\overline{P_0P_1}$  liegt, einen Wert aus dem Intervall  $[0..1]$  an. Ist der Wert negativ, so liegt der zugehörige Lotfußpunkt rechts von  $P_0$ , ist der Wert größer als 1, dann links von  $P_1$ . Zur Ermittlung der Länge  $l$  des Rechtecks wird die größte Ausdehnung in  $x$ -Richtung gesucht. Hierfür wird unter Verwendung von  $k_2$  entschieden, ob  $P_0$  oder  $L_2$  weiter links liegt, bzw. unter Verwendung von  $k_3$ , ob  $P_1$  oder  $L_3$  den rechten Rand festlegt. Der Abstand zwischen den äußeren Grenzen in  $x$ -Richtung entspricht der Länge  $l$ . Dies ist in Abbildung 8.14c und 8.14d dargestellt.

Aus der Lage der Eckpunkte wird der Schwerpunkt des Grabens bestimmt. Dieser legt den Ursprung für das Grabenkoordinatensystem fest. Somit sind nun alle Form- und Lageparameter des Modells bestimmt. Abbildung 8.15 zeigt die Projektion des Grabenmodells ins Grauwertbild.

Die aktuell aus der Messung generierte Hypothese wird im Folgenden mit *Messhypothese* bezeichnet. Die Messhypothese ist eine virtuelle, das heißt, nicht explizit im Szenenbaum repräsentierte Hypothese. Sie dient zur Überprüfung einer Übereinstimmung zwischen bereits bestehenden Hypothesen und der Messung bzw., falls keine Übereinstimmung gefunden werden konnte, als Vorlage zur Instanzierung einer neuen Hypothese im Szenenbaum. Im folgenden Kapitel wird die Verwaltung mehrerer Grabenhypothesen erläutert.

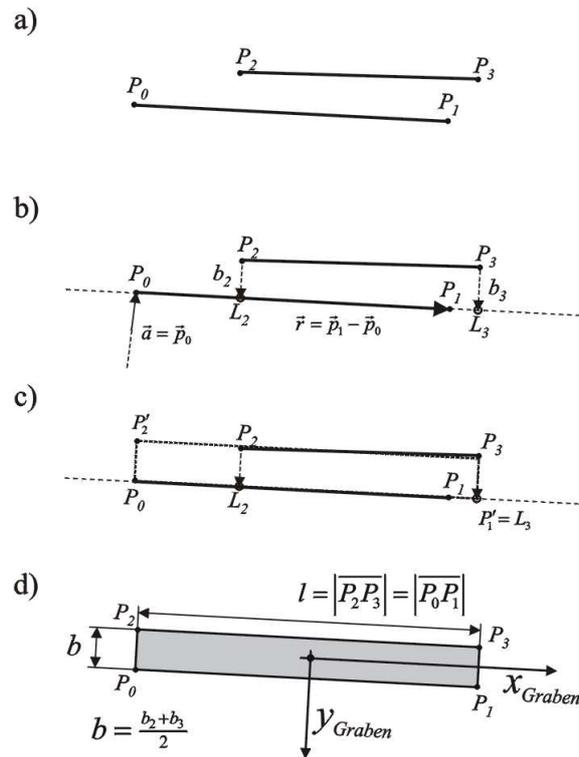


Abbildung 8.14: Bestimmung der Parameter des Grabenmodells

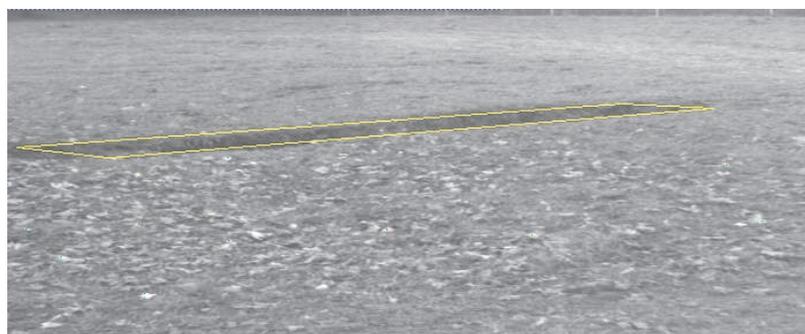


Abbildung 8.15: Visualisierung der Grabenhypothese im Grauwertbild

### 8.3 Hypothesenverwaltung

Jeder Prozess, der Daten mit der zentralen Wissensrepräsentation (Dynamische Objekt-daten Basis - DOB) austauschen will, erzeugt sich zu Prozessbeginn einen zunächst leeren lokalen Szenenbaum. Indem er sich auf bestimmte Objektklassen registriert, die für den Prozess von Bedeutung sind, erhält er lokale Kopien der aktuellen Knoteneinträge

der Objekte des allgemeinen Szenenbaums. Jeweils zu Beginn eines neuen Systemzyklus werden die Daten der allgemeinen Repräsentation mit denen der lokalen Kopie synchronisiert. Dabei werden die Einträge eines Objektes der lokalen Repräsentation nur dann an die allgemeine Repräsentation weitergeleitet, wenn der Szenenknoten eines Objektes als öffentlich markiert wurde.

Das heißt, nur veröffentlichte Objekte stehen auch anderen Prozessen über den allgemeinen Szenenbaum zur Verfügung. Dadurch hat man die Möglichkeit, ein neu detektiertes Objekt zunächst lokal zu verwalten, für eine gewisse Zeit zu verifizieren und erst danach zu veröffentlichen. Dadurch kann die Fehlalarmrate deutlich verringert werden. Allerdings erhöht sich natürlich die Totzeit ein wenig, bis eine Reaktion auf ein detektiertes Objekt erfolgt. Es ist aber in aller Regel möglich, einen guten Kompromiß zwischen einer ausreichend kleinen Totzeit und einer kleinen Fehlalarmrate zu finden.

Der lokale Szenenbaum stellt eine Funktionalität zur Verfügung, die es ermöglicht, dass sich der Prozess für eine interessante Klasse von Szenenobjekten eine Liste von Pointern auf alle Objekte dieser Klasse im aktuellen lokalen Szenenbaum erzeugt. Wenn im Folgenden vom Einfügen einer Hypothese in die Liste der bestehenden Hypothesen gesprochen wird, dann ist damit gemeint, dass eine neue Hypothese relativ zum lokalen geodätischen Koordinatensystem in den Szenenbaum eingefügt worden ist und die Liste der Pointer auf Objekte der Klasse Grabenhypothese (*CNegObstacle*) danach aktualisiert wurde. Entsprechend bedeutet das Löschen einer Hypothese aus der Liste, dass der zugehörige Szenenknoten aus dem Szenenbaum entfernt wurde.

Die lokale Verwaltung der Grabenhypothesen kann grob in drei Bereiche unterteilt werden:

1. Erzeugen einer neuen Grabenhypothese aus der Messhypothese, wenn entweder die Liste der Hypothesen noch leer ist oder keine Übereinstimmung mit einer bestehenden Hypothese gefunden werden konnte.
2. Überprüfung der Übereinstimmung der Messhypothese mit allen bestehenden Hypothesen der Liste und gegebenenfalls Innovation der Zustandsgrößen
3. Verkürzen der Länge bzw. Löschen einer Hypothese bei längerem Ausbleiben einer Bestätigung durch Messung

Für alle drei Bereiche ist eine Beurteilung der Vermessbarkeit bestehender Hypothesen notwendig. Hierfür wird eine Analyse der erwarteten Sichtbarkeit der hypothetisierten Objekte durchgeführt.

### 8.3.1 Sichtbarkeitsanalyse

Es werden hier vier Arten der Sichtbarkeit unterschieden. Die Sichtbarkeit wird dabei daran festgemacht, wie die Projektionen der vier Eckpunkte einer Hypothese relativ zum Bereich gültiger Disparitäten liegen.

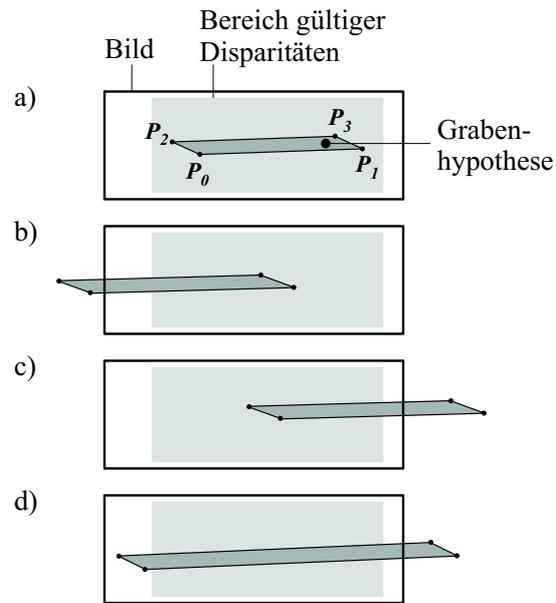


Abbildung 8.16: Analyse der Sichtbarkeit einer Hypothese

**Fall 1:** Alle vier Eckpunkte liegen innerhalb des gültigen Disparitätenbereichs. Das heißt, das zur Hypothese korrespondierende Objekt sollte vollständig sichtbar und damit vermessbar sein. Dies ist in Abbildung 8.16a dargestellt.

**Fall 2:** Der rechte Rand der Hypothese ist vollständig im Bild, der linke Rand nicht. Siehe Abbildung 8.16b.

**Fall 3:** Der linke Rand der Hypothese ist vollständig im Bild, der rechte Rand nicht. Siehe Abbildung 8.16c.

**Fall 4:** Weder linker noch rechter Rand sind vermessbar. Allerdings liegt sowohl ein Teil der Vorderkante als auch ein Teil der Hinterkante im gültigen Bereich der Disparitäten. Siehe Abbildung 8.16d.

In allen anderen Fällen wird das zu einer Hypothese korrespondierende Objekt als nicht sichtbar und damit nicht vermessbar markiert. Im Folgenden ist unter der Sichtbarkeit einer Hypothese die aufgrund der Modellvorstellung erwartete Sichtbarkeit des korrespondierenden realen hypothetisierten Objekts zu verstehen.

### 8.3.2 Verkürzen oder Löschen einer bestehenden Hypothese

Anhand der Sichtbarkeit wird entschieden, ob eine Hypothese durch Messung bestätigt werden müsste oder nicht. Erfolgt eine Bestätigung einer Hypothese, dann wird die Varia-

ble *LastUpdate* auf den aktuellen Systemzykluszähler *EMSCycle* gesetzt. Ist eine Hypothese nicht sichtbar, so wird ein Zähler *NumCyclesNotVisible* für die Anzahl der Zyklen, in denen eine Hypothese nicht sichtbar war, inkrementiert, der verhindert, dass eine nicht sichtbare Hypothese aus dem Szenenbaum gelöscht wird. Sollte eine Hypothese sichtbar sein und die Messung ist nicht erfolgreich, so wird je nach Art der Sichtbarkeit die Hypothese zunächst verkürzt bzw. nach einer gewissen Zeit aus der Liste der Hypothesen gelöscht.

Die Anzahl der relevanten Zyklen  $\Delta Cycle$  seit der letzten Bestätigung einer Hypothese berechnet sich aus der Differenz des aktuellen Systemzykluszählers *EMSCycle* minus dem Zyklus der letzten Bestätigung *LastUpdate* minus der Anzahl *NumCyclesNotVisible*, in der die Hypothese nicht vermessbar war.

In den Fällen 2 und 3 nach Abschnitt 8.3.1 wird die Länge der Hypothese um einen prozentualen Anteil verkürzt und der Schwerpunkt so verschoben, dass der nicht sichtbare linke bzw. rechte Rand der Hypothese jeweils seine Position beibehält. Durch diese Anpassung der Hypothese wird der Zeitpunkt der letzten Bestätigung *LastUpdate* nicht verändert. Es wird die Anzahl der Zyklen  $\Delta Cycle$  mit einem Grenzwert verglichen. Überschreitet  $\Delta Cycle$  den Grenzwert, dann wird die Hypothese aus der Liste gelöscht.

In den Fällen 1 und 4 erfolgt keine Verkürzung der Länge der Hypothese. Hier wird lediglich die Hypothese gelöscht, wenn die Anzahl der Zyklen  $\Delta Cycle$  den Grenzwert überschreitet.

### 8.3.3 Steuerung der Innovation

Bei der Überprüfung, ob eine bestehende Hypothese mit der aktuellen Messhypothese übereinstimmt, werden die Ausrichtung, Entfernung und Breite der Grabenhypothesen miteinander verglichen. Kann eine Übereinstimmung innerhalb eines Toleranzbereichs festgestellt werden, so ist für die Innovation der Zustandsgrößen von Bedeutung, welcher Teil der erwarteten Hypothese im sichtbaren Bildausschnitt zu finden ist.

Das heißt z.B., nur wenn alle vier Eckpunkte einer Hypothese innerhalb des sichtbaren Bildausschnittes liegen, ist die Länge des Grabens aus dem Bild direkt bestimmbar. Ist weder rechter noch linker Rand des Grabens im gültigen Bereich des Bildes, dann kann keine Aussage über die Länge des Grabens getroffen werden. Weiterhin, nur wenn sowohl Vorder- als auch Hinterkante eines Grabens sichtbar sind, kann die Breite geschätzt werden.

Nachfolgend werden zunächst einige Bezeichnungen eingeführt:

- $\vec{p}_{Aerw}$  und  $\vec{p}_{Eerw}$  seien die Ortsvektoren des Anfangs- bzw. Endpunkts der erwarteten Mittellinie eines hypothetisierten Objekts (HO).
- $\vec{p}_{Amess}$ ,  $\vec{p}_{Emess}$ ,  $\vec{p}_{Mmess}$  seien die Ortsvektoren des Anfangs-, End- und Mittelpunktes der Mittellinie der Messhypothese.
- $\vec{p}_{Lot}$  sei der Ortsvektor des Lotfußpunktes.

Weiterhin seien

- $x_{erw}$  - der erwartete  $x$ -Wert des Schwerpunkts eines HOs
- $y_{erw}$  - der erwartete  $y$ -Wert des Schwerpunkts eines HOs
- $l_{erw}$  - die erwartete Länge eines HOs
- $b_{erw}$  - die erwartete Breite eines HOs
- $\psi_{erw}$  - der erwartete Gierwinkel eines HOs
- $x_{mess}$  - der gemessene  $x$ -Wert der Messhypothese
- $y_{mess}$  - der gemessene  $y$ -Wert der Messhypothese
- $l_{mess}$  - die gemessene Länge der Messhypothese
- $b_{mess}$  - die gemessene Breite der Messhypothese
- $\psi_{mess}$  - der gemessene Gierwinkel der Messhypothese.

Die Residuen ergeben sich aus der Differenz der anhand der Sichtbarkeit angepassten Messdaten  $x_{in}$ ,  $y_{in}$ ,  $l_{in}$ ,  $b_{in}$  und  $\psi_{in}$  und der Erwartung

$$\begin{aligned}
 x_{res} &= x_{in} - x_{erw} \\
 y_{res} &= y_{in} - y_{erw} \\
 l_{res} &= l_{in} - l_{erw} \\
 b_{res} &= b_{in} - b_{erw} \\
 \psi_{res} &= \psi_{in} - \psi_{erw}.
 \end{aligned} \tag{8.6}$$

Es sei

$$\vec{r}_{0mess} = \frac{\vec{p}_{Emess} - \vec{p}_{Amess}}{|\vec{p}_{Emess} - \vec{p}_{Amess}|}. \tag{8.7}$$

der auf die Länge 1 normierte Richtungsvektor der Messhypothese, der vom Anfang der Mittellinie  $\vec{p}_{Amess}$  in Richtung des Endpunktes  $\vec{p}_{Emess}$  zeigt.

Entsprechend den vier Fällen aus Abschnitt 8.3.1 sind bei der Innovation einer Hypothese ebenfalls vier Fälle zu unterscheiden.

**Fall 1:** Es liegen alle Eckpunkte im gültigen Bereich des Bildes.

Die Innovation der Hypothese, der die Messhypothese zugeordnet wurde, kann mit den Residuen, die direkt aus der Differenz der Parameter  $x_{mess}$ ,  $y_{mess}$ ,  $l_{mess}$ ,  $b_{mess}$ ,  $\psi_{mess}$  der Messhypothese und den erwarteten Werten bestimmt werden, erfolgen. Es gilt

$$\begin{aligned}
 x_{in} &= x_{mess} \\
 y_{in} &= y_{mess} \\
 l_{in} &= l_{mess} \\
 b_{in} &= b_{mess} \\
 \psi_{in} &= \psi_{mess}
 \end{aligned}$$

**Fall 2:** Nur der rechte Rand liegt im gültigen Bildbereich.

Die Parameter  $b_{mess}$ ,  $\psi_{mess}$  der Messhypothese können direkt für die Innovation genutzt werden.

$$\begin{aligned} b_{in} &= b_{mess} \\ \psi_{in} &= \psi_{mess} \end{aligned}$$

$x_{in}$ ,  $y_{in}$ ,  $l_{in}$  sind aus der Differenz zwischen der Erwartung der Lage der Hypothese und der Lage der Messhypothese durch Festhalten der Position des jeweils unsichtbaren Randes errechenbar. In Abbildung 8.17a ist dies schematisch dargestellt. Zunächst wird der Lotfußpunkt  $P_{Lot}$  des Endpunkts der Messhypothese  $P_{Emess}$  auf die erwartete Mittellinie der Hypothese gebildet. Als Nebenprodukt fällt bei dieser Berechnung entsprechend Gleichung 8.3 bzw. 8.4 wieder ein Skalar  $k$  an, welcher für die Bestimmung der Länge  $l_{in}$  benötigt wird. Für die Länge der angepassten Messhypothese gilt

$$l_{in} = k \cdot l_{erw}, \quad (8.8)$$

wobei das Inkrement  $dl = (k - 1) \cdot l_{erw}$  ist (Abbildung 8.17a).

Der Schwerpunkt  $(x_{in}, y_{in})$  der angepassten Messhypothese ergibt sich aus

$$\begin{bmatrix} x_{in} \\ y_{in} \end{bmatrix} = \vec{p}_{Emess} - \frac{l_{in}}{2} \cdot \vec{r}_{0mess}. \quad (8.9)$$

Dieser wird also ausgehend vom Endpunkt der Mittellinie der Messhypothese  $P_{Emess}$  entgegen der Orientierung der Mittellinie mit halber Länge  $l_{in}$  verschoben. Die Länge wird um die volle Differenz  $dl$  verlängert bzw. verkürzt, je nachdem ob  $k > 1$  oder  $k < 1$  ist. Die Länge  $l_{in}$  der Messhypothese und ihr Schwerpunkt werden demnach stets so angepasst, dass der nicht sichtbare Rand der zugeordneten Hypothese näherungsweise seine erwartete Position in Längsrichtung beibehält.

**Fall 3:** Nur der linke Rand liegt im gültigen Bildbereich.

Dies ist in Abbildung 8.17b dargestellt. Es wird der Lotfußpunkt  $P_{Lot}$  des Anfangspunkts der Messhypothese  $P_{Amess}$  auf die erwartete Mittellinie der Hypothese gebildet. Der Rest stimmt mit der Vorgehensweise von Fall 2 überein. Die Länge der angepassten Messhypothese errechnet sich aus

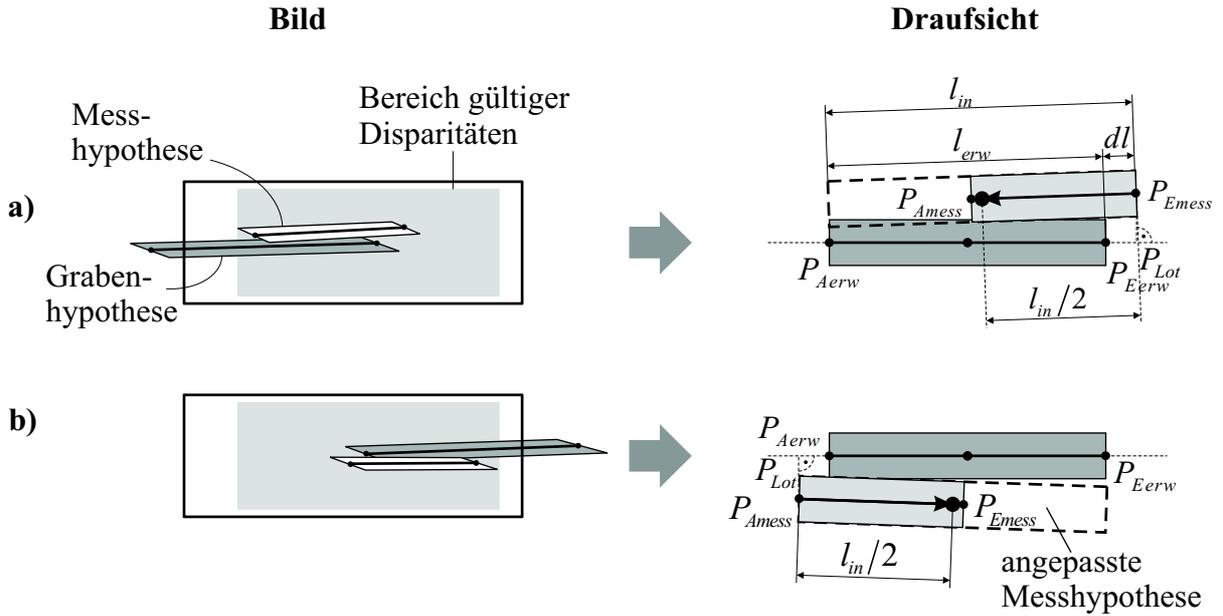
$$l_{in} = (1 - k) \cdot l_{erw}, \quad (8.10)$$

wobei  $dl = -k \cdot l_{erw}$  ist. Der Schwerpunkt  $(x_{in}, y_{in})$  der angepassten Messhypothese ergibt sich aus

$$\begin{bmatrix} x_{in} \\ y_{in} \end{bmatrix} = \vec{p}_{Amess} + \frac{l_{in}}{2} \cdot \vec{r}_{0mess}. \quad (8.11)$$

Weiterhin gelte

$$\begin{aligned} b_{in} &= b_{mess} \\ \psi_{in} &= \psi_{mess}. \end{aligned}$$



**Abbildung 8.17:** Anpassung der Messhypothese

**Fall 4:** Weder rechter noch linker Rand sind vermessbar. Allerdings liegt sowohl ein Teil der erwarteten Vorderkante als auch ein Teil der erwarteten Hinterkante der Hypothese im gültigen Bereich.

In diesem Fall ist die Länge der Grabenhypothese nicht beobachtbar, das heißt, die Innovation der Zustandsgröße  $l$  wird nicht durchgeführt und  $l_{in}$  ist nicht definiert. Die Größen  $b_{mess}$  und  $\psi_{mess}$  der Messung können direkt genutzt werden. Die Größen  $x_{in}, y_{in}$  können durch die Normalkomponente an eine Entfernungsänderung angepasst werden, indem der Schwerpunkt normal zur Ausrichtung der Hypothese verschoben wird. Es wird der Lotfußpunkt  $P_{Lot}$  des Mittelpunkts der Messhypothese  $P_{Mmess}$  auf die erwartete Mittellinie der Hypothese gebildet. Die Normalkomponente zur Anpassung der Schwerpunktposition  $(x_{in}, y_{in})$  folgt aus  $\vec{p}_{norm} = \vec{p}_{Mmess} - \vec{p}_{Lot}$ .

$$\begin{bmatrix} x_{in} \\ y_{in} \end{bmatrix} = \begin{bmatrix} x_{exp} \\ y_{exp} \end{bmatrix} + \vec{p}_{norm} \quad (8.12)$$

$$\begin{aligned} b_{in} &= b_{mess} \\ \psi_{in} &= \psi_{mess} \end{aligned}$$

In allen anderen Fällen wird die jeweilige Hypothese als nicht vermessbar deklariert.

Aus der Differenz der Werte  $x_{in}, y_{in}, l_{in}, b_{in}, \psi_{in}$  und der Erwartung werden die Residuen  $x_{res}, y_{res}, l_{res}, b_{res}, \psi_{res}$  entsprechend Gleichung (8.6) gebildet und damit das Kalmanfilter zur Innovation der Zustandsgrößen gespeist.

## 8.4 Zusammenspiel von Wahrnehmung, Blickrichtungssteuerung und Fahrzeugführung

Bevor eine Hypothese als Objekt für andere Experten öffentlich gemacht wird und das Wissen darüber via DOB kommuniziert wird, muss eine Hypothese zunächst über mehrere Videozyklen hinweg bestätigt worden sein. Hierfür wird die Liste aller bisherigen potentiellen Gräben durchlaufen und mit der aktuellen Messhypothese verglichen. Kann keine Übereinstimmung gefunden werden, wird eine neue Hypothese im lokalen Szenenbaum erzeugt. Kann eine Übereinstimmung bzw. Zuordnung der Messung mit der Erwartung gefunden werden, dann werden die Zustandsgrößen dieses potentiellen Grabens innoviert. Ab einer bestimmten Anzahl an Bestätigungen wird dann eine Hypothese veröffentlicht, d.h. das Wissen über das Objekt wird anderen Experten sehr früh zugänglich gemacht, obwohl sich dieses durch eine weitere Annäherung des Eigenfahrzeugs an das Objekt in aller Regel noch weiter verbessert.

Die Veröffentlichung der Hypothese ist allerdings notwendig, weil ansonsten die Situationsanalyse für die Blickrichtungssteuerung dem bestätigten Objekt keine Aufmerksamkeit (Relevanz) schenken würde. Für eine gezielte Verbesserung der Wahrnehmungsleistung durch Blickrichtungssteuerung ist dies allerdings unabdingbar.

Für die Steuerung der Blickrichtung werden für alle wahrzunehmenden Objekte oder Teile von Objekten Aufmerksamkeitsbereiche durch den Wahrnehmungsexperten definiert. Diese dienen zur Planung der Blickrichtung unter der Randbedingung einer begrenzten Anzahl an Sakkaden für eine zeitlich begrenzte Planungsperiode [PELLKOFER 03]. In Kapitel 8.4.1 werden die Aufmerksamkeitsbereiche genauer erläutert.

Die Situationsanalyse für die Blickrichtungssteuerung ist nach [PELLKOFER 03] Teil der zentralen Entscheidungsinstanz. Hier wird festgelegt, ob ein Objekt relevant ist für die Blickrichtungssteuerung oder nicht. Diese Relevanz kann auch nicht durch BDGA (Behavior Decision for Gaze and Attention) festgelegt werden, da diese keinen Überblick über die zu erfüllende Mission haben sollte. Die Kompetenz von BDGA ist also darauf beschränkt, einen Ablauf festzulegen, wie innerhalb eines Planungszeitraums die gestellte Aufgabe, nämlich  $n$  durch einen Relevanzwert gewichtete Objekte möglichst effizient zu fixieren. Im Konfliktfall, d.h. wenn z.B. der zeitliche Rahmen nicht ausreicht, um alle relevanten Objekte zu fixieren und eine minimale Verweildauer zu gewährleisten, kann dieser nur durch die zentrale Entscheidungsinstanz aufgelöst werden, und das auch nur dann, wenn es möglich ist, die Anzahl der relevanten Objekte zu reduzieren. Ist dies aufgrund der Situation nicht möglich, dann kann solch ein Konflikt nur durch Vergrößerung des Planungszeitraums beseitigt werden.

Relevant sind beim Querfeldeinfahren solche Objekte, welche sich zumindest zum Teil innerhalb des Fahrschlauchs befinden. Nur wenn ein Objekt relevant ist, haben auch seine Aufmerksamkeitsbereiche eine Relevanz. Diese Aufmerksamkeitsbereiche sind alle von der Klasse CVisObjBase abgeleitet. Die Blickrichtungssteuerung - ihr entspricht der Prozess GazeControl [PELLKOFER 03] - erzeugt sich eine Liste von Verweisen auf alle Objekte

der Klasse CVisObjBase im Szenenbaum. Ein Rückschluß von den Aufmerksamkeitsbereichen auf die zugrundeliegenden hypothetisierten Objekte ist hier nicht mehr möglich. Der Algorithmus zur Blickrichtungssteuerung versucht je nach Vorgabe mit maximal 2 oder 3 Sakkaden alle Aufmerksamkeitsbereiche in der fest vorgegebenen Planungsperiode abzubilden. Gelingt dies nicht, wird das am wenigsten relevante Objekt derzeit einfach unberücksichtigt gelassen.

Weiterhin wird eine Hypothese nach der Veröffentlichung zunächst als „*wird gerade vermessen*“ markiert, damit die Entscheidungsinstanz nicht sofort beim Auftauchen einer Hypothese ein neues Verhalten z.B. Ausweichen anstößt, sondern der Blickrichtungssteuerung noch die Zeit lässt, gezielt die vom Wahrnehmungsexperten definierten Aufmerksamkeitsbereiche mittels Sakkaden zu fixieren. Sobald die Vermessung beendet ist, was in Abhängigkeit von der Häufigkeit der wahrgenommen Merkmale - hier speziell der linken und rechten Berandung - erfolgt, wird die Hypothese dann als „*vermessen*“ markiert.

Solange diese Vermessung nicht abgeschlossen ist, d.h. das Objekt noch nicht „*vermessen*“ wurde, bewegt sich das Eigenfahrzeug in der Regel auf das Objekt zu. Bei einer Geschwindigkeit des Eigenfahrzeugs von z.B.  $v = 5 \frac{m}{s}$  und einer ersten Detektion eines negativen Hindernisses in ca. 20 m Abstand zum Hindernis bleiben bei einer minimalen notwendigen Entfernung von 12 m - 15 m, um je nach Lage des Objektes ausweichen zu können, gerade mal 1 - 1,6 sec für die Vermessung der Hypothese übrig. Dies ist eine sehr geringe Zeitspanne, um eine genauere Vermessung der Hypothese mittels gezielter Blickzuwendung durchführen zu können, da bereits eine Sakkade ca. 400msec benötigt.

Wenn der Abstand zu einem potentiellen Hindernis Gefahr läuft, zu klein zu werden, um bei aktueller Geschwindigkeit noch ausweichen zu können, wird deshalb die Fähigkeit des Anhaltens aktiviert [SIEDERSBERGER 03]. Hierdurch erhält die Wahrnehmung die Zeit, um bei gezielter Blickzuwendung auf die linke und rechte Berandung des Objekts dieses genauer zu vermessen. Wird das Hindernis schließlich als „*vermessen*“ markiert und ist der Abstand zum Hindernis noch ausreichend, um aus dem Stand heraus ein Ausweichmanöver starten zu können, dann wird die entsprechende Fähigkeit aktiviert. Falls dies nicht möglich ist, so verharret das Fahrzeug derzeit im sicheren Zustand, d.h. es bleibt stehen. Ansonsten müsste das Fahrzeug, um den Abstand für ein Ausweichen ausreichend zu vergrößern, autonom rückwärts fahren können, was bedingt durch die Art der Automatikgetriebe mit den Fahrzeugen der UBM nicht realisierbar ist [SIEDERSBERGER 03].

#### **8.4.1 Wahrnehmungsmodi zur Steuerung der Aufmerksamkeitsbereiche**

Die Wahrnehmung von Objekten erfolgt in der Regel anhand bestimmter objektspezifischer Merkmale und Gruppen von Merkmalen. Für die Wahrnehmung einzelner Merkmale wurde in Abschnitt 8.2.1 bereits eine Randbedingung festgelegt. Es wurde definiert, ab welcher Ausdehnung ein Merkmal mit der kleinsten Bildoperatorensuchmaske prinzipiell

detektierbar sein soll. Aufgrund dieser Randbedingung wird je nach den Aspektbedingungen der maximal mögliche Vorausschaubereich bestimmt.

Für die Wahrnehmung von Gruppen von Merkmalen bzw. für die Gruppierung von Einzelmerkmalen zu Merkmalgruppen gelten natürlich auch Randbedingungen. In der Anwendung zur Grabenerkennung wurde eine solche Randbedingung derart festgelegt, dass die horizontale Ausdehnung der Projektion der Grabenhypothese bzw. eines Teils der Grabenhypothese ins Bild z.B. mindestens die Hälfte der Pixel des gültigen Bildbereichs ausfüllen soll. Damit wird sichergestellt, dass die Algorithmen zur Gruppierung von Einzelmerkmalen zu Merkmalgruppen ausreichend Einzelmerkmale zur Verfügung haben sollten. Die geforderte minimale Anzahl an Pixeln wird durch die Variable  $N_{HorPixelMin}$  festgelegt. Weiterhin sei  $N_{HorPixelMax}$  die Anzahl der Pixel in einer Zeile.

Damit die zweite Randbedingung durch gezielte Blickzuwendung erfüllt werden kann, legt der Wahrnehmungsexperte für jede Objekthypothese Aufmerksamkeitsbereiche in 3-D fest. Diese Aufmerksamkeitsbereiche werden relativ zum Objektkoordinatensystem beschrieben und mit einer Logik versehen, die der Blickrichtungssteuerung Alternativen an die Hand gibt. Die Erfüllung der Randbedingung durch die Blickrichtungssteuerung erfolgt dabei implizit, denn die Blickrichtungssteuerung weiß von den „Bedürfnissen“ der Wahrnehmung nichts außer den Aufmerksamkeitsbereichen.

Für die Festlegung der Größe der Aufmerksamkeitsbereiche ist die bereits verbal formulierte Bedingung in Abhängigkeit von der Objektentfernung in 3-D in ein Längenmaß umzurechnen. Aus der Breite des CCD-Chips kann mittels Strahlensatz die geforderte minimale Anzahl an Pixeln im gültigen Bildbereich  $N_{HorPixelMin}$  in eine minimale Länge  $l_{min}$  in 3-D umgerechnet werden.

Mit  $s_{max}$  gleich der horizontalen Größe des CCD-Chips, der euklidischen Entfernung  $x$  des Schwerpunkts eines hypothetisierten Objekts zum Ursprung des Kamerakoordinatensystems und der Brennweite  $f$  errechnet sich die minimale Länge des geforderten Bereichs in Metern aus

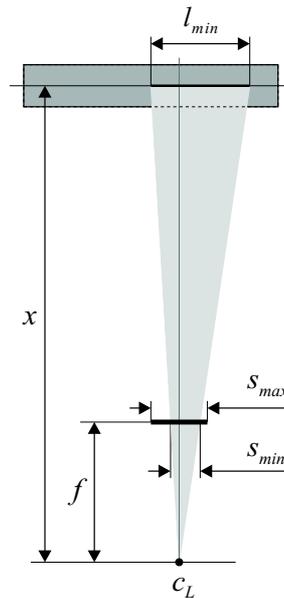
$$l_{min} = s_{min} \cdot \frac{x}{f}, \quad (8.13)$$

wobei

$$s_{min} = \frac{N_{HorPixelMin}}{N_{HorPixelMax}} \cdot s_{max} \quad (8.14)$$

ist. In Abbildung 8.18 wird die Berechnung von  $l_{min}$  mittels Strahlensatz noch einmal veranschaulicht.

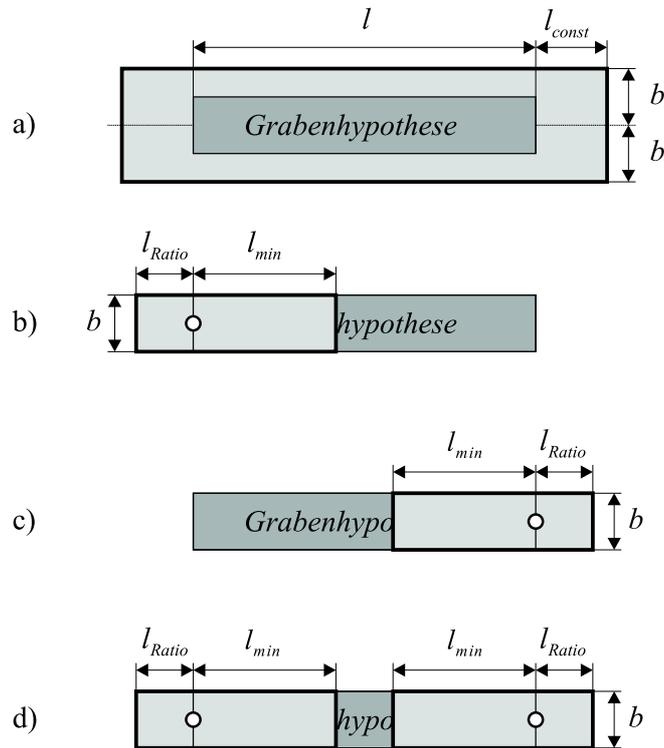
Der Wahrnehmungsexperte stellt fünf verschiedene Wahrnehmungsmodi zur Verfügung. Diese werden in Abhängigkeit von der Verhaltensentscheidung der zentralen Entscheidungsinstanz von BDGA durch die Parameterisierung der Wahrnehmungsfähigkeit vorgegeben. Die unterschiedlichen Wahrnehmungsmodi legen jeweils unterschiedliche Kombinationen an Aufmerksamkeitsbereichen fest. Nachfolgend werden die fünf Wahrnehmungsmodi vorgestellt:



**Abbildung 8.18:** Erforderliche Ausdehnung eines Aufmerksamkeitsbereichs für die Wahrnehmung

**FixPointLeft** Der Wahrnehmungsmodus **FixPointLeft** stellt eine Oder-Kombination aus zwei Aufmerksamkeitsbereichen dar. Eine Oder-Kombination legt fest, dass die „Bedürfnisse“ des Wahrnehmungsexperten erfüllt sind, wenn entweder der eine oder der andere Aufmerksamkeitsbereich im Bild sichtbar ist. Der erste Aufmerksamkeitsbereich beinhaltet dabei die gesamte Grabenhypothese, wie in Abbildung 8.19a zu sehen. Ein zweiter Aufmerksamkeitsbereich schließt den linken Rand der Grabenhypothese ein. Diese ist in Abbildung 8.19b dargestellt. Die Aufmerksamkeitsbereiche sind jeweils etwas größer als durch die Größe der Grabenhypothese bzw. die minimale Länge  $l_{min}$  vorgegeben, um zum Rand des gültigen Bereichs etwas Abstand zu lassen. Der Aufmerksamkeitsbereich aus Abbildung 8.19a hat z.B. die doppelte Breite der geschätzten Breite  $b$  des Grabens und eine um  $2 \cdot l_{const}$  größere Länge als die geschätzte Länge. Der konstante Wert  $l_{const}$  wurde z.B. auf 1,5m gesetzt. Der Aufmerksamkeitsbereich aus Abbildung 8.19b hat dagegen die gleiche Breite  $b$  wie die Hypothese, allerdings erstreckt er sich vom linken Rand ausgehend noch um  $l_{Ratio} = k \cdot l_{min}$  mit  $k = 0,3$  nach links und um  $l_{min}$  nach rechts. In der Abbildung 8.19 wurde der jeweils interessante Fixationspunkt mit einem schwarzen Kringel mit weißer Füllung markiert.

**FixPointRight** Entsprechendes gilt für den Wahrnehmungsmodus **FixPointRight**. Dieser stellt wiederum eine Oder-Kombination aus zwei Aufmerksamkeitsbereichen dar. Der erste ist identisch mit dem ersten des Wahrnehmungsmodus **FixPointLeft**, der zweite ergibt sich durch Spiegelung des linken Aufmerksamkeitsbereichs am Schwer-



**Abbildung 8.19:** Wahrnehmungsmodi:

- I) FixPointLeft kombiniert die Aufmerksamkeitsbereiche a) und b)
- II) FixPointRight kombiniert die Aufmerksamkeitsbereiche a) und c)
- III) LeftAndRight kombiniert die Aufmerksamkeitsbereiche a) und d)

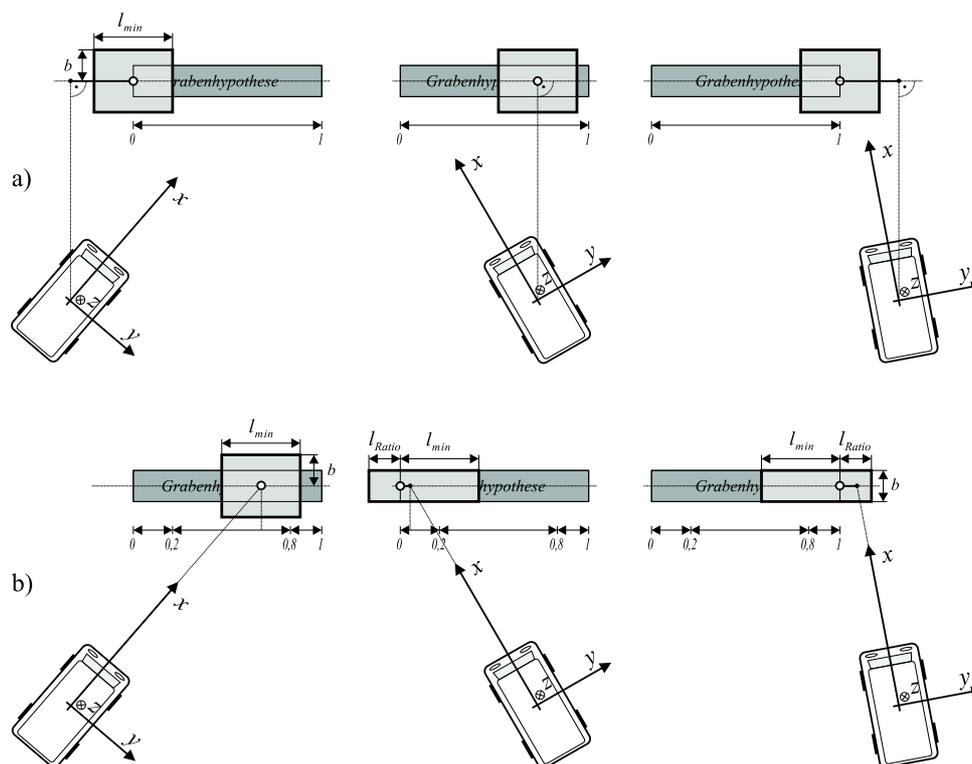
punkt der Grabenhypothese. Siehe Abbildung 8.19c.

**LeftAndRight** Der Wahrnehmungsmodus **LeftAndRight** kombiniert dagegen drei Aufmerksamkeitsbereiche. Der erste Aufmerksamkeitsbereich aus Abbildung 8.19a wird dabei mit einer „Sowohl-Als-Auch“-Kombination nach Abbildung 8.19d Oderverknüpft. Das heißt, der Blickrichtungssteuerung stehen zwei Alternativen zur Verfügung, entweder lässt sich der Aufmerksamkeitsbereich eins ins Bild abbilden oder es wird abwechselnd auf den linken und rechten Aufmerksamkeitsbereich fixiert, damit die Wahrnehmung die Vermessungsaufgabe erfüllen kann.

**FixPointNearest** Ein weiterer Modus ist **FixPointNearest**. Dieser stellt wiederum eine Kombination aus zwei Aufmerksamkeitsbereichen dar. Der erste entspricht Abbildung 8.19a. Der zweite errechnet sich jeweils in Abhängigkeit des Lotfußpunktes des Schwerpunkts des Eigenfahrzeugs auf die Mittellinie der Grabenhypothese. Je nachdem, wo der Lotfußpunkt zu liegen kommt, wird somit immer der nächstgelegene Punkt auf der Mittellinie als Zentrum des Aufmerksamkeitsbereichs gewählt.

Dieser stellt gleichzeitig den nächstgelegenen Punkt auf der Mittellinie relativ zum Schwerpunkt des Eigenfahrzeugs dar. Der Aufmerksamkeitsbereich hat eine Breite von  $2 \cdot b$  und eine Länge  $l_{min}$ . Dies ist in Abbildung 8.20a dargestellt.

**FixPointEgoAxis** Im Gegensatz zum Wahrnehmungsmodus **FixPointNearest** wird bei **FixPointEgoAxis** nicht der Lotfußpunkt, sondern der Schnittpunkt der Fahrzeuglängsachse mit der Mittellinie der Grabenhypothese als Kriterium für die Lage des interessanten Fixationspunktes für den Aufmerksamkeitsbereich gewählt. Weiterhin wird eine Fallunterscheidung durchgeführt. Liegt der Schnittpunkt innerhalb des Intervalls 20%-80% der Länge der Mittellinie, dann wird der Aufmerksamkeitsbereich symmetrisch zum Schnittpunkt mit Länge  $l_{min}$  und Breite  $2 \cdot b$  definiert. Liegt der Schnittpunkt links davon, dann wird ein Aufmerksamkeitsbereich entsprechend Abbildung 8.19b verwendet, liegt er rechts davon, dann entsprechend Abbildung 8.19c. Dies ist in Abbildung 8.20b dargestellt.



**Abbildung 8.20:** Wahrnehmungsmodi:

IV) **FixPointNearest** kombiniert den Aufmerksamkeitsbereich a) mit Abb. 8.19a

V) **FixPointEgoAxis** kombiniert den Aufmerksamkeitsbereich b) mit Abb. 8.19a

## 8.4.2 Parametrisierung der Wahrnehmung

In Abhängigkeit von der Verhaltensentscheidung der zentralen Entscheidungsinstanz wird durch die Entscheidungsinstanz von BDGA durch die Parametrisierung der Wahrnehmung vorgegeben, welcher Wahrnehmungsmodus aktiv sein soll. Dieser legt wiederum fest, wie die Aufmerksamkeitsbereiche liegen sollen. Daraus ergibt sich, ob vornehmlich der linke Rand (beim Linksausweichen), der rechte Rand (beim Rechtsausweichen), der zentrale Bereich (beim Anhalten vor dem Graben) oder der linke und rechte Rand (genaues Vermessen) ausgewertet wird. Der Wahrnehmungsexperte legt also die Aufmerksamkeitsbereiche entsprechend der Vorgabe so fest, dass ein gutes Wahrnehmungsergebnis zu erwarten ist.

## 8.4.3 Anknüpfung an das Fähigkeitsnetz

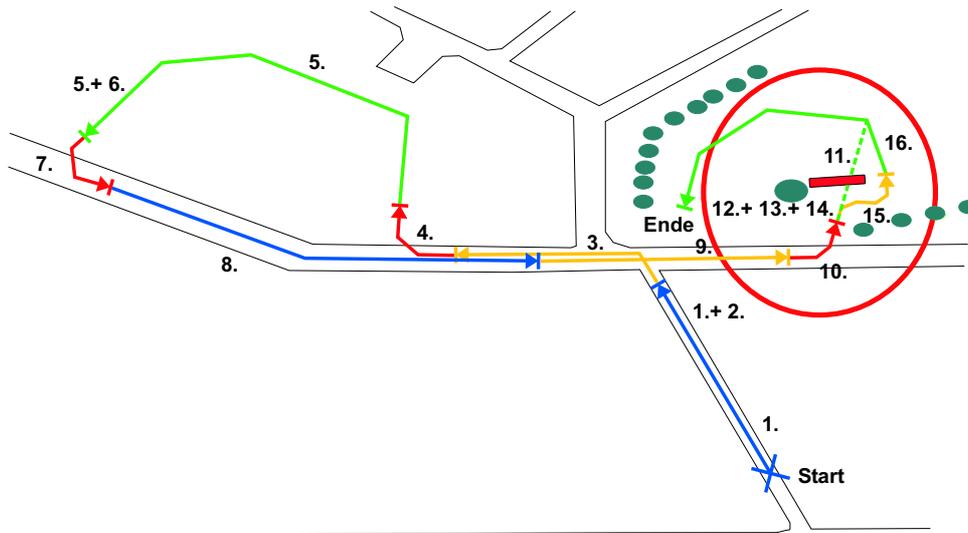
Die Parametrisierung der Wahrnehmung zur Erkennung negativer Hindernisse erfolgt über das Fähigkeitsnetz. Details zum Fähigkeitsnetz finden sich in [PELLKOFER 03] und [SIEDERSBERGER 03]. Hierfür wird beim Systemstart ein entsprechender Fähigkeitsknoten im Fähigkeitsnetz erzeugt, der die statischen Abhängigkeiten zwischen den Fähigkeiten wiedergibt. Eine Fähigkeit ist im System erst dann aktivierbar, sobald alle Voraussetzungen bzgl. der Hardware und unterlagerter Fähigkeiten erfüllt sind und sich ein entsprechender Experte im Fähigkeitsnetz auf einen entsprechenden Fähigkeitsknoten registriert hat, der die Fähigkeit zur Verfügung stellt, d.h. mit Leben erfüllt. Beim Start einer Fähigkeit wird eine Instanz einer Fähigkeit erzeugt, mit ihrem korrespondierenden Fähigkeitsknoten verknüpft, via Fähigkeitsnetz von der zentralen Entscheidungsinstanz parametrisiert und anschließend gestartet. Prinzipiell können parallel mehrere Instanzen der gleichen Fähigkeit aktiv sein, sofern sie nicht exklusiven Zugriff auf bestimmte Ressourcen benötigen. Während der Ausführung der Fähigkeit können die Parameter der Fähigkeit verändert werden und die Fähigkeit auch wieder beendet werden.

## 8.5 Ergebnisse einer realen Versuchsfahrt

Anhand einer Querfeldein-Versuchsfahrt werden nun einige Ergebnisse anschaulich erläutert. Die Fahrt war ein Teil des Szenarios der Abschlussdemonstration zum Forschungsprojekt **Intelligente Fahrzeugfunktionen Phase 3 (IFF3)** zur autonomen Fahrzeugführung, welche am 23.10.01 vor internationalem Publikum auf dem Versuchsgelände der Universität der Bundeswehr in Neubiberg gezeigt wurde. Abbildung 8.21 zeigt die Stationen der Demonstrationsfahrt.

Wesentliche Stationen der Demonstrationsfahrt vom 23.10.01 waren:

1. Folgen des Verlaufs eines Feldwegs
2. Erkennen und Vermessen einer von der Missionsplanung erwarteten Querstraße
3. Abbiegen nach Links auf die Querstraße



**Abbildung 8.21:** Plan der IFF3-Abschlussdemonstration

4. Verlassen der Straße nach rechts
5. Querfeldeinfahren mit GPS-Wegpunktenavigation
6. Erkennen und Vermessen einer von der Missionsplanung erwarteten Querstraße
7. Auffahren auf eine Straße
8. Straße folgen
9. Überqueren einer Kreuzung
10. Verlassen der Straße nach links
11. Querfeldeinfahren mit GPS-Wegpunktenavigation
12. Erkennen eines negativen Hindernisses
13. Anhalten vor dem Hindernis

Die hier gezeigten Ergebnisse beziehen sich auf die Stationen 11 bis 13 der Demonstration, wobei in Erweiterung der Demonstrationsfahrt vom 23.10.01 zusätzlich zur

12. Erkennung eines negativen Hindernisses und
13. dem Anhalten vor einem negativen Hindernis,
14. eine genauere Vermessung eines Grabens durch gezielte Blickzuwendung,
15. Anfahren und Ausweichen an einem Hindernis und danach
16. Fortsetzung der GPS-Wegpunktenavigation

verwirklicht wurden. Der hier interessierende Teilbereich ist in Abbildung 8.21 durch einen Kreis hervorgehoben.

### 8.5.1 Erläuterungen zum Ausweichen an einem Graben anhand einer Bildfolge

Die Bildfolgensequenz der Abbildungen 8.22 bis 8.32 soll die wesentlichen Elemente zur Realisierung einer autonomen Querfeldeinfahrt mit Ausweichen an einem Graben illustrieren. Leider konnte die Visualisierung der Grabenhypothesen nicht direkt im Bild der gleichen Kamera erfolgen, welche für die Messung verwendet wurde. Die hier gezeigten Bilder stammen von einer  $\frac{1}{3}$ " drei Chip-Farbkamera, welche mit einem  $12\text{mm}$  Objektiv ausgestattet ist. Dies entspricht einem  $16\text{ mm}$  Objektiv bei  $\frac{1}{2}$ " Kameras. Die Öffnungswinkel sind demnach mit den Öffnungswinkeln der Kameras, welche für die Messung verwendet werden, vergleichbar. Durch die exakte Repräsentation der Relativlagen der Kameras im Szenenbaum und eine gute Kalibrierung ist gewährleistet, dass die Projektion der Hypothesen ins Bild zum Bild passen. Allerdings entsteht durch die Kommunikation der Objektdaten auf einen anderen Rechner eine Totzeit von einem Zyklus. Ein weiterer Zyklus geht durch die Tatsache verloren, dass die Visualisierung stets im vorletzten Videobild erfolgt, da das aktuelle Bild für Messungen reserviert ist. Das heißt, dass die Visualisierung im Bild stets mit einer Totzeit von mindestens 2 Zyklen behaftet ist. Gerade bei schnellen Blickzuwendungen kann hierdurch eine z.T. deutliche Diskrepanz zwischen der Visualisierung einer Hypothese und dem realen Bild entstehen. Dies bedeutet allerdings nicht, dass die Schätzgrößen derart fehlerbehaftet sind. Während einer Sakkade werden keine Messungen im Bild durchgeführt.

Abbildung 8.22 zeigt das Anfahren eines GPS-Wegpunktes (Station 11) durch das Eigenfahrzeug nach Verlassen der Straße (Station 10). In den Bildern links oben befindet sich ein Baum und der Schatten des Baums. Im mittleren Bereich des Bildes sind mehrere hellere und dunklere Bereiche zu sehen. Es werden keine falschen Hypothesen erzeugt.



**Abbildung 8.22:** Querfeldeinfahren mit aktivierter Hinderniserkennung

In Abbildung 8.23 wird im linken und mittleren Bild für das menschliche Auge langsam ein Graben erkennbar. Der Wahrnehmungsmodus für die Definition von Aufmerksamkeitsbereichen ist zu Beginn der Default-Modus `FixpointEgoAxis`. Bevor eine Hypothese veröffentlicht wird, muss sie eine bestimmte Anzahl an Bestätigungen erfahren (z.B. willkürlich 15). Das Bild in der Mitte ist genau 15 Zyklen vor dem Bild rechts aufgenommen, wobei das Bild rechts ein Bild vor der Veröffentlichung der Hypothese ist.



**Abbildung 8.23:** Annähern an einen Graben

In Abbildung 8.24 ist die Projektion der Grabenhypothese ins Bild nach der Veröffentlichung dargestellt. Ihr Zustand ist auf „*wird gerade vermessen*“ gesetzt. Wird die Grabenhypothese von der zentralen Entscheidungsinstanz als relevant für die Missionsdurchführung erachtet, dann wird für eine genauere Vermessung des Grabens der Wahrnehmungsmodus durch die zentrale Entscheidungsinstanz auf **LeftAndRight** gesetzt. Alle vier Eckpunkte der Visualisierung liegen im Bild (wobei der hier sichtbare Bereich nicht ganz dem der Messung entspricht). Das heißt, die Hypothese ist vollständig sichtbar. Durch die weitere Annäherung an den Graben verschwindet der linke Rand aus dem Bild, so dass die Bestimmung der Länge des hypothetisierten Objekts nur noch eingeschränkt erfolgt.



**Abbildung 8.24:** Veröffentlichung der Hypothese

Die geringfügigen Abweichungen der Visualisierung bezüglich des realen Bildes rühren zum einen von der bereits angesprochenen Totzeit von zwei Zyklen her und zum anderen aus der Vereinfachung, dass der Nahbereich des Eigenfahrzeugs durch eine Ebene angenähert und Nickwinkelschwankungen aufgrund von Bodenunebenheiten vernachlässigt wurden.

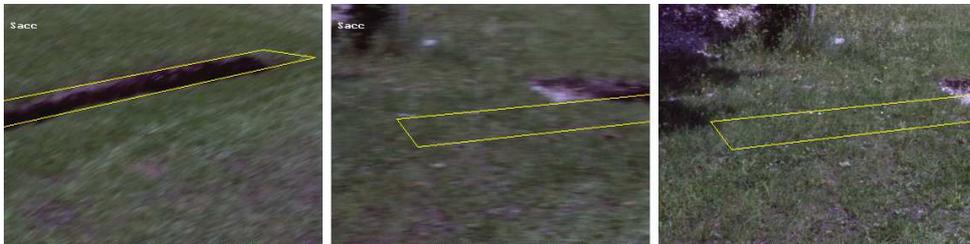
Unterschreitet der Abstand zum Hindernis eine notwendige Entfernung fürs Ausweichen, dann erhält die Fahrzeugführung den Auftrag anzuhalten. In Abbildung 8.25 ist das Ein- und Ausnicken des Eigenfahrzeugs deutlich zu erkennen.

In Abbildung 8.26 ist im linken und mittleren Bild die Sakkade auf die erwartete Position des linken Randes der Hypothese dargestellt. In der linken oberen Ecke wird während einer Sakkade das Kürzel „Sacc“ ausgegeben. Während der Sakkade werden keine Messungen im Bild durchgeführt. Nach der Sakkade ist in Abbildung 8.26 rechts zu erkennen, dass



**Abbildung 8.25:** Anhalten vor Graben mit Ein- und Ausnicken

an der erwarteten Position des linken Randes dieser nicht durch die Messung bestätigt werden kann.



**Abbildung 8.26:** Sakkade nach links

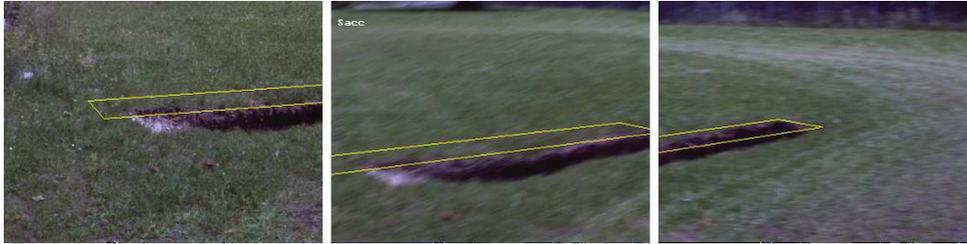
Abbildung 8.27 zeigt die Verkürzung der Länge der Hypothese, wobei der unsichtbare rechte Rand seine Position beibehält. Während der Verkürzung wird je Zyklus ein prozentualer Anteil von der Länge der Hypothese abgezogen.



**Abbildung 8.27:** Verkürzen der Hypothese und Nachführung der Blickrichtung auf linken Rand

Gleichzeitig mit der Verkürzung wird die Blickrichtung entsprechend der Position des erwarteten linken Randes nachgeführt. Im rechten Bild von Abbildung 8.27 kann schließlich die Erwartung durch Messung bestätigt werden und es erfolgt eine Anpassung der Hypothese durch Innovation der Modellparameter.

Nachdem nun der linke Rand vermessen wurde, erfolgt in Abbildung 8.28 eine Blickzuwendung nach rechts, um zu überprüfen, ob der rechte Rand noch korrekt vermessen ist.



**Abbildung 8.28:** Blickzuwendung auf rechten Rand

In Abbildung 8.29 wird nochmals eine Blickzuwendung auf den linken Rand durchgeführt. Sobald die Anzahl der bestätigten Messungen der linken und rechten Berandung einen Grenzwert überschritten hat, wird die Hypothese als „vermessen“ markiert. Die zentrale Entscheidungsinstanz legt hiernach fest, ob das Eigenfahrzeug weiterhin stehen bleiben oder ob links bzw. rechts am Hindernis vorbei ausgewichen werden soll.

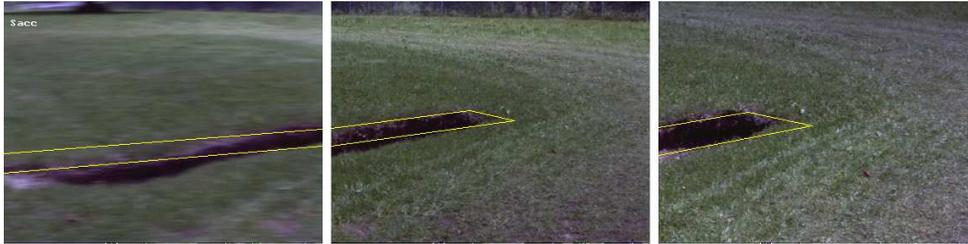


**Abbildung 8.29:** Blickzuwendung auf linken Rand; Hypothese wird als „vermessen“ markiert

Bei dieser Versuchsfahrt wurde entschieden, dass rechts ausgewichen werden soll. Deshalb wird der Wahrnehmungsmodus durch BDGA auf `FixpointRight` gesetzt, wodurch die Blickrichtungssteuerung wieder den rechten Rand fixiert. Siehe Abbildung 8.30. Die Fahrzeugführung erhält den Auftrag anzufahren (Bild 8.30 rechts) und unter Verwendung der ermittelten Grabenparameter eine Ausweichtrajektorie zu bestimmen und abzufahren [SIEDERSBERGER 03].

Bei dieser Fahrt treten wieder Aufbauschwankungen auf, die hier nicht durch einen inertialen Stabilisierungsmodus für die Blickrichtung ausgeglichen wurden. In Abbildung 8.31 sind diese Störungen bei der Blickfixierung auf den rechten Rand während des Ausweichmanövers zu sehen.

Sobald die Grabenhypothese keine Überschneidung mehr mit dem Fahrschlauch hat, verlieren das Objekt und damit die Aufmerksamkeitsbereiche ihre Relevanz. Es erfolgt eine



**Abbildung 8.30:** Nach der Entscheidung rechts auszuweichen, Blickzuwendung auf den rechten Rand und Anfahren



**Abbildung 8.31:** Ausweichen mit Fixierung auf rechte Berandung

Sakkade auf den geplanten Fahrspurverlauf und der Wahrnehmungsmodus für negative Hindernisse wird auf die Default-Einstellung `FixpointEgoAxis` zurückgesetzt. Dies ist in Abbildung 8.32 dargestellt.



**Abbildung 8.32:** Blickzuwendung auf Fahrweg

## 8.5.2 Verlauf der Parametrisierung der Wahrnehmung

Der Verlauf der Parametrisierung der Wahrnehmung wird in Abbildung 8.33 nochmals gezeigt. Ohne relevante Grabenhypothese ist der Wahrnehmungsmodus `FixpointEgoAxis` eingestellt. Bei Veröffentlichung einer Hypothese (erste große runde Markierung) und sofern diese relevant ist, wird zum Wahrnehmungsmodus `LeftandRight` gewechselt. Sobald die Hypothese als „vermessen“ markiert wurde (zweite große runde Markierung),

wird je nach Verhaltensentscheidung bei „Stehen bleiben“, „links Ausweichen“ bzw. „rechts Ausweichen“ der Wahrnehmungsmodus auf `FixpointEgoAxis`, `FixpointLeft` oder `FixpointRight` gesetzt. Verliert die Hypothese ihre Relevanz, da keine Überschneidung mehr mit dem Fahrschlauch existiert, wird der Wahrnehmungsmodus auf `FixpointEgoAxis` zurückgesetzt.

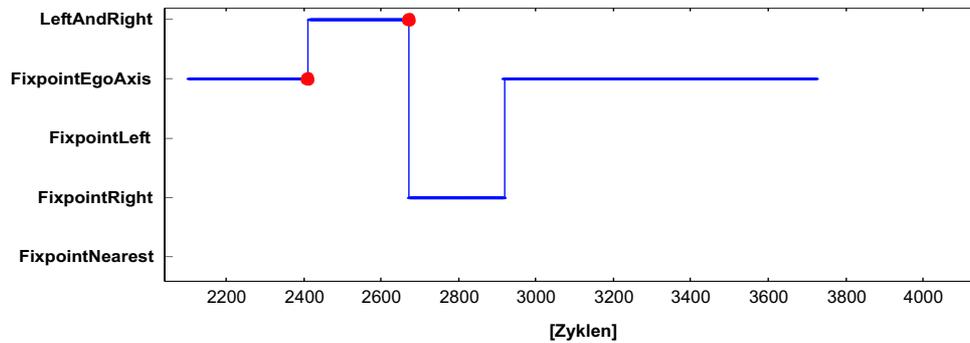


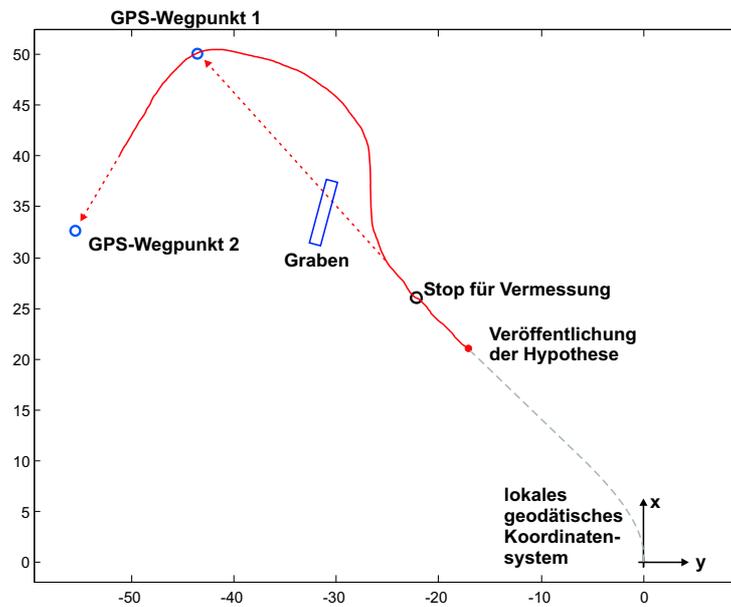
Abbildung 8.33: Parametrisierung der Wahrnehmung

### 8.5.3 Vogelperspektive des Ausweichmanövers an einem Graben

In Abbildung 8.34 ist diese Versuchsfahrt noch einmal in einer Draufsicht auf die Bodenebene dargestellt. Die gestrichelte Linie zeigt ausgehend vom lokalen geodätischen Koordinatensystem den Verlauf der Schwerpunktsposition des Eigenfahrzeugs während der Annäherung auf den GPS-Wegpunkt 1. Der Anfangspunkt der durchgezogenen Linie stellt den Ort des Schwerpunktes des Eigenfahrzeugs zum Zeitpunkt der Veröffentlichung der Hypothese dar. Der schwarze Kringel markiert den Ort, an dem das Eigenfahrzeug zur genaueren Vermessung des Grabens stehen blieb. Der weitere Verlauf der durchgezogenen Linie illustriert die Trajektorie des Ausweichmanövers mit anschließendem Übergang zurück zur Wegpunktnavigation. Details zur Planung der Ausweichtrajektorie sind in [SIEDERSBERGER 03] nachzulesen. Die Position und Ausdehnung des eingezeichneten Grabens entspricht exakt den durch Bildverarbeitung bestimmten Parametern zu dem Zeitpunkt, an dem die Hypothese „*vermessen*“ gesetzt wurde.

### 8.5.4 Diskussion des Verlaufs der Modellparameter bei der Grabenerkennung

Die Abbildung 8.35 zeigt den Verlauf der Länge, der Breite und der Sichtbarkeit der Grabenhypothese ab dem Zeitpunkt der Veröffentlichung bis zu dem Zeitpunkt, an dem die Hypothese „*vermessen*“ gesetzt wird. Ein bestimmter Zeitpunkt ist dabei gleichbedeutend



**Abbildung 8.34:** Vogelperspektive der Fahrtrajektorie im lokalen geodätischen Koordinatensystem

mit einem bestimmten Wert des Zykluszahlers, der auf der Abszisse der Grafik aufgetragen ist. In der Abbildung sind die von der Blickrichtungssteuerung durchgeführten Sakkaden als dunkelgraue Balken markiert. Die Breite eines Balkens korrespondiert mit der Dauer der Sakkade in Zyklen. Weiterhin sind drei interessante Bereiche besonders hervorgehoben. Im ersten Bereich ist zu sehen, dass während der Annäherung an den Graben kurzzeitig weder der linke noch der rechte Rand der Grabenhypothese im Bild sichtbar waren. Dadurch konnte zu diesem Zeitpunkt aus der Messung keine Aussage über die Länge des Grabens gemacht werden. Im oberen Teil der Grafik ist zu sehen, dass während dieser Zeitspanne keine Innovation der Länge durchgeführt wurde, was sich darin zeigt, dass die ermittelte Länge konstant bleibt.

Der zweite Bereich markiert die Zeitspanne des Anhaltens bzw. des Ein- und Ausnickens. Nachdem in dieser Anwendung der Nickwinkel nicht explizit modelliert bzw. inertial gemessen wurde, kommt es hier zu einer Verschlechterung der geschätzten Modellparameter. Nach der ersten Sakkade sollte der linke Rand der Grabenhypothese sichtbar sein (vergleiche die Abbildungen 8.26 und 8.27). Nachdem keine Bestätigung durch die Messung erfolgt, wird im dritten hervorgehobenen Bereich die geschätzte Länge des Grabens verkürzt, wobei die Breite konstant bleibt, da keine Innovation durchgeführt wird. Die von der aktiven Kameraplattform eingestellten Winkel der Nick- und Gierachse zur Fixierung eines Aufmerksamkeitsbereichs (hier: des linken Randes) werden automatisch beim Erstellen der homogenen Transformationsmatrizen via Szenenbaum berücksichtigt.

Danach konnte die erwartete Position des linken Randes durch die Messhypothese bestätigt werden. Die Modellparameter werden durch Innovation der Zustandsgrößen

durch das Kalmanfilter aktualisiert. Im weiteren zeitlichen Verlauf konvergieren die Modellparameter gegen die Werte des realen Grabens. Dieser hat ungefähr eine Länge von  $6m$  und eine Breite von  $1,1m$ .

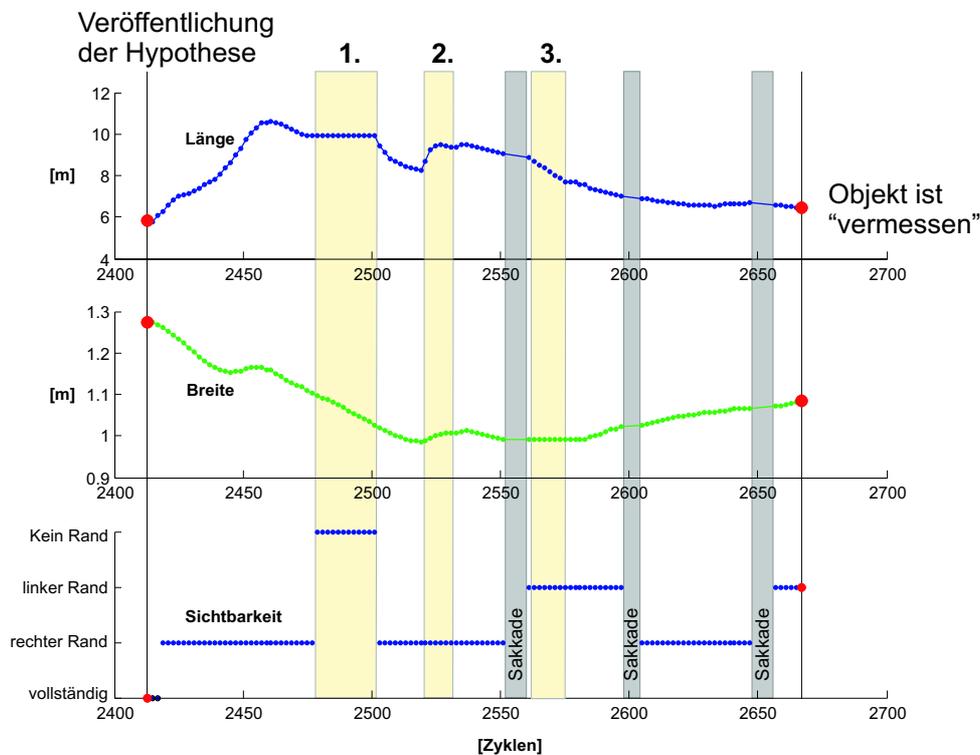
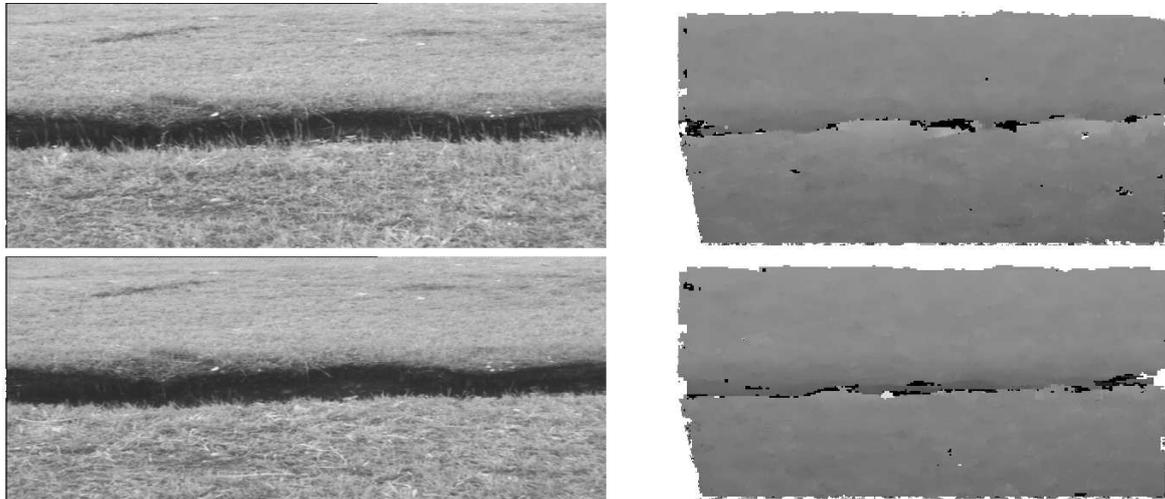


Abbildung 8.35: Datenplot der Formparameter und der Sichtbarkeit der Hypothese

### 8.5.5 Anmerkungen zur Stereohardware

Die Abbildung 8.36 zeigt im linken oberen bzw. linken unteren Teilbild jeweils eine Aufnahme eines Grabens mit und ohne vereinzelte Grashalme an der Vorderkante des Grabens. Im rechten Teil der Abbildung sind die zu diesen Aufnahmen zugehörigen Disparitätenbilder, die durch die Stereohardware ermittelt wurden, dargestellt. Im Disparitätenbild rechts oben ist im Vergleich zum Disparitätenbild rechts unten zu erkennen, dass die eigentlich dominante Vertiefung (negative Disparität) an den Stellen der vereinzelten Grashalme durch das Wattebausch-artige Aufblähen der positiven Disparitäten zum Teil überdeckt wird. Dadurch wird die Verifikation von Grauwertkanten im Disparitätenbild deutlich erschwert, da die Position einer Grauwertkante in diesem Fall nicht unbedingt mit einem Übergang zu negativen Disparitäten zusammenfällt. Weiterhin soll die Abbildung 8.36 verdeutlichen, dass das Stereobildverarbeitungssystem sehr empfindlich gegenüber „Verunreinigungen“ wie z.B. einzelnen Grashalmen reagiert und dass da-

mit eine Unterscheidung zwischen einem festen Hindernis und einem „Pseudo“-Hindernis, welches vom Eigenfahrzeug leicht überfahren werden kann, nicht möglich erscheint.



**Abbildung 8.36:** Grauwertbild und Disparitätenbild mit und ohne vereinzelte Grashalme

## 8.6 Zusammenfassung der Ergebnisse

Eingebunden in eine komplexe Mission konnte die bottom-up Detektion negativer Hindernisse unter Verwendung aktiver Blickzuwendung mit daran anschließendem Ausweichmanöver gezeigt werden. Im Gegensatz zu Anwendungen, welche die Bestätigung des Vorhandenseins und anschließend die Lagebestimmung eines erwarteten Objekts verfolgen, wobei eine Erwartung der Lage und Ausdehnung eines gesuchten Objekts durch die Missionsplanung vorgegeben wird (Beispiel: Querstraßenerkennung [LÜTZELER 02]), erfolgt hier die Detektion negativer Hindernisse ohne Vorwissen, ob überhaupt ein Hindernis vorhanden ist und welche Lage und Ausdehnung es besitzt.

Der interessante Bereich (engl. *Region of Interest* (ROI)) für die Bildauswertung wird in Abhängigkeit vom Auflösungsvermögen der Kamera-Bildoperator Kombination für ein gesuchtes Objektmerkmal festgelegt. Für die Erkennung werden zwei unterschiedliche Bildverarbeitungsverfahren fusioniert. Die Bestimmung interessanter Bildbereiche innerhalb der ROI erfolgt Kanten-basiert aus dem Grauwertbild. Die Verifikation der Relevanz für die Hypothesengenerierung geschieht anhand des Disparitätenbildes des Stereosystems. Durch Verkettung gerichteter Kantenpunkte werden aus verrauschten Kantenpunktmenngen dominante Strukturen extrahiert. Die Maskengrößen für die Kantenpunktextraktion und für die Verifikation im Disparitätenbild werden entfernungsabhängig an die Größe eines gesuchten Objektmerkmals im Bild angepasst.

Die Wahrnehmung für die Erfüllung der Vermessungsaufgabe ist mit der Fahrzeugführung und der Blickrichtungssteuerung abgestimmt. Durch ein gezieltes Anhalten in ausreichender Entfernung zu einem Hindernis wird der Blickrichtungssteuerung die Zeit gegeben, die von der Wahrnehmung vorgegebenen Aufmerksamkeitsbereiche zu fixieren. Die Aufmerksamkeitsbereiche für die Objektmerkmale werden nach den „Bedürfnissen“ der Wahrnehmung dimensioniert und positioniert.

Es wird eine Sichtbarkeitsanalyse aller bestehenden Hypothesen durchgeführt. Unter Berücksichtigung der Ergebnisse dieser Analyse erfolgt entweder eine Steuerung der Innovation der Zustandsgrößen eines Objekts im Kalmanfilter, wenn die Messhypothese einer bestehenden Hypothese zugeordnet werden konnte, oder die Länge der Hypothese wird verkürzt, wenn keine Bestätigung durch Messung möglich ist, bzw. gelöscht, wenn längere Zeit keine Bestätigung stattgefunden hat.

Es wurde ein Mechanismus zur abgestuften Bereitstellung von Hypothesendaten eingeführt. Zunächst wird eine Hypothese nur lokal verwaltet. Nach einer bestimmten Anzahl an Bestätigungen wird sie dann veröffentlicht, wobei sie zunächst als „wird gerade vermessen“ gesetzt wird. Nach einer genaueren Vermessung wird sie dann als „vermessen“ markiert. Erst hiernach wird z.B. aus den Objektdaten ein Ausweichkurs ermittelt und ein Ausweichmanöver gestartet.

Das Konzept zur Grabenerkennung wurde in realen Versuchsfahrten verifiziert.

# Kapitel 9

## Zusammenfassung

Die Wahrnehmung des Fahrzeugumfeldes zur Detektion von Objekten und zur Bestimmung der eigenen Position relativ zu anderen Objekten ist für die Realisierbarkeit intelligenter autonomer Fahrfunktionen grundlegend. Ein Ziel dieser Arbeit bestand deshalb darin, Bildverarbeitungswerkzeuge zur Steigerung des Informationsgewinns aus der Bildfolgenanalyse zu entwickeln.

Diese Bildverarbeitungswerkzeuge (Operatoren) wurden so konzipiert, dass sie sowohl für die Detektion von Objekten als auch fürs Verfolgen bestehender Objekthypothesen verwendet werden können. Hierfür wurden in den Kapiteln 4.2 und 4.3 zwei neue Verfahren zur Segmentierung von Helligkeitsverläufen entwickelt. Dabei wird entlang horizontaler bzw. vertikaler Bildstreifen nach Helligkeitsübergängen gesucht, wobei diese in der Regel die Grenzen homogener Helligkeitsbereiche bilden.

Die Beschreibung eines Bildstreifens gewinnt hierdurch gegenüber den bisherigen Verfahren durch die Fülle der extrahierten Information deutlich an Aussagekraft, wodurch auch verbesserte Möglichkeiten zur Selektion und Gruppierung von Merkmalen zur Verfügung stehen. Die Zuordnung eines gemessenen Bildmerkmals bezüglich einer aktuellen Vorstellung kann dabei umso zuverlässiger erfolgen, je aussagekräftiger die extrahierte Bildbeschreibung zur Unterscheidung gegenüber gleichartigen Merkmalen ist.

Die entwickelten Ansätze und Konzepte wurden exemplarisch in verschiedenen Anwendungen in den beiden Versuchsfahrzeugen für autonome Mobilität und Rechnersehen (VAMORS und VAMP ) implementiert und erfolgreich getestet. Unter anderem wurde in Kapitel 6 ein hybrides ACC System (HACC) in Form einer Kombination eines radarbasierten ACC mit visueller Fahrspur- und Objekterkennung realisiert, das in Abhängigkeit von der von den Wahrnehmungsmodulen zur Verfügung gestellten Information auf unterschiedlichen Leistungsstufen arbeiten kann.

Es wurden ferner in Kapitel 5.3 und 7.2 Strategien und in Kapitel 4.4 Bildverarbeitungsoperatoren zur rein visuellen Detektion und Verfolgung von Pkws im Nahbereich für ein Stauassistenzsystem entwickelt und auf Videosequenzen verifiziert. Die Detektion von

Pkws bei seitlicher Ansicht stellt dabei eine besondere Herausforderung dar und konnte mittels aspektabhängiger generischer Muster entsprechend Kapitel 5.2.1 und 4.4 zur Detektion von Rädern gezeigt werden.

Als ein Teil einer größeren vollautonomen Mission mit querfeldein GPS-Wegpunktnavigation konnte in Kapitel 8 die Erkennung eines negativen Hindernisses mit anschließendem Ausweichmanöver unter Verwendung der entwickelten Bildverarbeitungsverfahren und aktiver Blickrichtungssteuerung in realen Versuchsfahrten demonstriert werden. Das Zusammenspiel zwischen Wahrnehmung, Blickrichtungssteuerung und Fahrzeugführung wurde dabei speziell am Beispiel der Erkennung eines negativen Hindernisses erläutert. Hierfür wurden vom Wahrnehmungsexperten relevante Aufmerksamkeitsbereiche in 3-D definiert, welche für die Erfüllung der Wahrnehmungsaufgabe und für die Interaktion mit der Blickrichtungssteuerung notwendig waren.

Als Rahmen für die Entwicklung der Fahrerassistenzanwendungen diente das EMS-VISION System des ISF der UniBwM.

# Anhang A

## Mathematische Grundlagen

### A.1 Ausgleichsrechnung

Die Bestimmung einer Ausgleichsgeraden bzw. einer Ausgleichsebene durch eine Punktmenge erfordert die Lösung eines linearen Gleichungssystems. Handelt es sich dabei um ein überbestimmtes Gleichungssystem, d.h. existieren mehr Gleichungen als Unbekannte, dann kann nur eine Näherungslösung ermittelt werden. Häufig wird als Kriterium zur Minimierung eines überbestimmten Gleichungssystems das minimale Quadrat der Fehler der Punktmenge gegenüber der Näherungslösung verwendet.

Es sei

$$\mathbf{y} = \mathbf{A} \cdot \mathbf{p} \quad (\text{A.1})$$

ein überbestimmtes lineares Gleichungssystem mit  $\mathbf{y} \in \mathbb{R}^n$ ,  $\mathbf{p} \in \mathbb{R}^m$  und  $\mathbf{A} \in \mathbb{R}^{n \times m}$  mit  $n > m$ . Weiterhin sei

$$\text{Rang}(\mathbf{A}) = m.$$

Daraus folgt,  $\mathbf{A}^T \mathbf{A}$  ist nicht singulär, symmetrisch und positiv definit.

Für die  $i$ -te Gleichung ( $i \in [1..n]$ ) des Gleichungssystems ergibt sich ein Fehler

$$e_i = y_i - [\mathbf{A} \cdot \mathbf{p}]_i. \quad (\text{A.2})$$

Die zu minimierende Summe der Fehlerquadrate errechnet sich aus

$$J = \sum_{i=1}^n e_i^2 = \mathbf{e}^T \mathbf{e} = (\mathbf{y} - \mathbf{A} \cdot \mathbf{p})^T (\mathbf{y} - \mathbf{A} \cdot \mathbf{p}). \quad (\text{A.3})$$

Als notwendige Bedingung für ein Minimum muss die erste Ableitung von  $J$  nach  $\mathbf{p}$  null sein, d.h.

$$\frac{\partial J}{\partial \mathbf{p}} = 2(\mathbf{A}^T \mathbf{A} \cdot \mathbf{p} - \mathbf{A}^T \cdot \mathbf{y}) \stackrel{!}{=} 0. \quad (\text{A.4})$$

Mit  $\mathbf{A}^T \mathbf{A}$  positiv definit ist die Bedingung

$$\frac{\partial^2 J}{\partial \mathbf{p}^2} = 2(\mathbf{A}^T \mathbf{A}) > 0 \quad (\text{A.5})$$

für ein Minimum erfüllt. Das bedeutet, dass sich aus Gleichung (A.4) die Summe der kleinsten Fehlerquadrate für

$$\mathbf{p} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{y} \quad (\text{A.6})$$

ergibt. Der Ausdruck  $(\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T$  entspricht hierbei der Moore-Penrose-Pseudoinversen.

## A.2 Bestimmung einer Ausgleichsgeraden durch eine Punktemenge

Die Gleichung (A.6) wird nun zur Bestimmung einer Ausgleichsgeraden einer Punktemenge angewandt. Als Punktemenge diene eine Menge von  $(x_i, y_i)$ -Wertepaaren mit  $i = 1..n$  Elementen. Es wird vorausgesetzt, dass die  $x_i$ -Werte

- aufeinanderfolgend sind und
- äquidistante Abstände voneinander besitzen,

d.h. sowohl *eindeutig* als auch *ohne Lücke* im Verlauf sind. Weiterhin soll hier der Spezialfall betrachtet werden, dass dieser äquidistante Abstand  $x_{i+1} - x_i = 1$  ist und  $x_i \rightarrow i$  mit  $i = 1..n$  als Indexvektor einer Zahlenfolge  $Y$  betrachtet werden kann.

Eine solche Zahlenfolge ist z.B. ein gemittelter Grauwertverlauf eines Bildausschnitts eines Grauwertbildes. Der Bildausschnitt besteht z.B. aus 3 aufeinanderfolgenden Zeilen eines Bildes, startet bei Spalte 100 und endet bei Spalte 199. D.h., der Ausschnitt hätte eine Größe von 3x100 Pixeln. Der gemittelte Grauwertverlauf dieses Bildausschnittes ergibt sich durch Summation über die Spalten der Höhe  $h = 3$  und Division durch die Spaltenhöhe. Die gemittelten Werte  $y_i$  zu den Werten  $x_i$  folgen aus den Werten  $y_{ji}$  des Bildausschnitts durch

$$y_i = \frac{\sum_{j=1}^h (y_{ji})}{h}. \quad (\text{A.7})$$

Eine Ausgleichsgerade  $\bar{y}_i = f(x_i)$  ist durch die Parameter  $a$  und  $b$  eindeutig festgelegt.

$$\bar{y}_i = a \cdot x_i + b \quad (\text{A.8})$$

Der Parameter  $a$  ist die *Steigung* der Geraden, der Parameter  $b$  legt den Schnittpunkt der Ausgleichsgeraden mit der  $y$ -Achse fest.

Das Koordinatensystem sei nun so definiert, dass  $b$  dem Mittelwert der Ausgleichsgeraden entspricht. Das bedeutet, dass die Indizierung so verändert werden muss, dass die neue Indizierung symmetrisch zur Mitte der Zahlenfolge und damit der Ursprung des Koordinatensystems mittig zur gesuchten Ausgleichsgeraden liegt. Der Index  $x_i \rightarrow i$  mit  $i = 1..n$  wird deshalb umgewandelt in einen symmetrischen Index:

$$w_i \rightarrow i - \frac{n+1}{2} \quad (\text{A.9})$$

d.h.  $w_i \in [1 - \frac{n+1}{2}, n - \frac{n+1}{2}]$ . Für ein symmetrisches  $w$  gilt:

$$\sum_{i=1}^n w_i = 0 \quad (\text{A.10})$$

Für eine Zahlenfolge der Länge  $n = 8$  ergibt sich z.B. der Indexvektor:

$$\mathbf{w} = [-3.5, -2.5, -1.5, -0.5, +0.5, +1.5, +2.5, +3.5]^T \quad (\text{A.11})$$

Die Ausgleichsgerade folgt als Lösung des Gleichungssystems:

$$\begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} w_1 & 1 \\ \vdots & \vdots \\ w_n & 1 \end{bmatrix} \cdot \begin{bmatrix} a \\ b \end{bmatrix} \quad (\text{A.12})$$

$$\begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} (1 - \frac{n+1}{2}) & 1 \\ (2 - \frac{n+1}{2}) & 1 \\ \vdots & \vdots \\ (n - \frac{n+1}{2}) & 1 \end{bmatrix} \cdot \begin{bmatrix} a \\ b \end{bmatrix} = \mathbf{A}_n \cdot \begin{bmatrix} a \\ b \end{bmatrix} \quad (\text{A.13})$$

und daraus unter Verwendung von Gleichung (A.6)

$$\begin{bmatrix} a \\ b \end{bmatrix} = (\mathbf{A}_n^T \mathbf{A}_n)^{-1} \mathbf{A}_n^T \cdot \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \quad (\text{A.14})$$

mit

$$\mathbf{A}_n^T \mathbf{A}_n = \begin{bmatrix} \sum_{i=1}^n (w_i^2) & \sum_{i=1}^n w_i \\ \sum_{i=1}^n w_i & n \end{bmatrix} \quad (\text{A.15})$$

$$\Delta = |\mathbf{A}_n^T \mathbf{A}_n| = \sum_{i=1}^n (w_i^2) \cdot n - \left( \sum_{i=1}^n w_i \right)^2 \quad (\text{A.16})$$

$$(\mathbf{A}_n^T \mathbf{A}_n)^{-1} \cdot \Delta = \begin{bmatrix} n & -\sum_{i=1}^n w_i \\ -\sum_{i=1}^n w_i & \sum_{i=1}^n (w_i^2) \end{bmatrix} \quad (\text{A.17})$$

und

$$\mathbf{A}_n^T \cdot \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^n (w_i \cdot y_i) \\ \sum_{i=1}^n y_i \end{bmatrix}. \quad (\text{A.18})$$

Die Steigung der Ausgleichsgeraden berechnet sich dann zu

$$a = \frac{n \cdot \sum_{i=1}^n (w_i \cdot y_i) - \sum_{i=1}^n w_i \cdot \sum_{i=1}^n y_i}{\Delta}. \quad (\text{A.19})$$

Mit Gleichung (A.10) folgt für ein symmetrisches  $w$

$$a = \frac{\sum_{i=1}^n (w_i \cdot y_i)}{\sum_{i=1}^n (w_i^2)}. \quad (\text{A.20})$$

Die Summe der gemischten Terme soll im folgenden mit

$$swy_n := \sum_{i=1}^n (w_i \cdot y_i) \quad (\text{A.21})$$

bezeichnet werden. Der Ausdruck im Nenner von Gleichung (A.20) lässt sich unter Verwendung der Formeln für Potenzsummen vereinfachen.

Es gilt

$$sww_n := \sum_{i=1}^n (w_i^2) = \frac{1}{12} \cdot n \cdot (n+1) \cdot (n-1) \quad (\text{A.22})$$

unter Verwendung von Gleichung (A.9) mit

$$\sum_{i=1}^n i = \frac{n \cdot (n+1)}{2} \quad (\text{A.23})$$

$$\sum_{i=1}^n (i^2) = \frac{n \cdot (n+1) \cdot (2n+1)}{6}. \quad (\text{A.24})$$

D.h.

$$a = \frac{swy_n}{sww_n}. \quad (\text{A.25})$$

Der y-Abschnitt bzw. der Mittelwert  $b$  berechnet sich aus

$$b = \frac{-\sum_{i=1}^n w_i \cdot \sum_{i=1}^n (w_i \cdot y_i) + \sum_{i=1}^n (w_i^2) \cdot \sum_{i=1}^n y_i}{\Delta}. \quad (\text{A.26})$$

Und für  $b$  folgt mit Gleichung (A.10), wie gefordert,

$$b = \frac{\sum_{i=1}^n y_i}{n}. \quad (\text{A.27})$$

## A.3 Varianz der Abweichung der Werte einer Zahlenfolge zur Ausgleichsgeraden

Als ein Maß für die Qualität der Approximation der Wertefolge durch die Ausgleichsgerade  $\bar{y}_i = a \cdot w_i + b$  kann die Varianz der Differenz der tatsächlichen Werte zur Ausgleichsgeraden bestimmt werden. Für die Differenz gilt

$$e_i := y_i - \bar{y}_i = y_i - a \cdot w_i - b \quad (\text{A.28})$$

mit  $i = 1..n$ . Es sei

$$\mathbf{e} := [e_1, e_2, \dots, e_n]. \quad (\text{A.29})$$

Dann gilt unter Verwendung des Verschiebungssatzes für Varianzen und mit den Gleichungen (A.28), (A.27), (A.25) und (A.10)

$$\begin{aligned} \text{Var}(\mathbf{e}) &:= \frac{\sum_{i=1}^n e_i^2}{n} - \left( \frac{\sum_{i=1}^n e_i}{n} \right)^2 \\ &= \frac{\sum_{i=1}^n (e_i + b)^2}{n} - \left( \frac{\sum_{i=1}^n (e_i + b)}{n} \right)^2 \\ &= \frac{\sum_{i=1}^n (y_i - a \cdot w_i)^2}{n} - \left( \frac{\sum_{i=1}^n y_i - a \cdot \sum_{i=1}^n w_i - \sum_{i=1}^n b}{n} + b \right)^2 \\ &= \frac{\sum_{i=1}^n (y_i^2 - 2a \cdot y_i \cdot w_i + a^2 w_i^2)}{n} - \left( \underbrace{\frac{\sum_{i=1}^n y_i}{n}}_{=b} - a \cdot \underbrace{\frac{\sum_{i=1}^n w_i}{n}}_{=0} - \underbrace{\frac{\sum_{i=1}^n b}{n}}_{=b} + b \right)^2 \\ &= \frac{\sum_{i=1}^n y_i^2}{n} - 2a \cdot \frac{\sum_{i=1}^n (y_i \cdot w_i)}{n} + a^2 \frac{\sum_{i=1}^n w_i^2}{n} - b^2 \\ &= \frac{\sum_{i=1}^n y_i^2}{n} - 2 \cdot \frac{swy_n}{sww_n} \cdot \frac{swy_n}{n} + \frac{swy_n^2}{n \cdot sww_n} - b^2 \\ &= \frac{\sum_{i=1}^n y_i^2}{n} - \frac{swy_n^2}{n \cdot sww_n} - b^2. \end{aligned} \quad (\text{A.30})$$

Die Varianz ist demnach sehr einfach unter Verwendung bereits vorhandener Zwischenwerte zu bestimmen.

## A.4 Verknüpfung zweier Ausgleichsgeraden zu einer

Das Ziel der Verknüpfung ist das Errechnen einer resultierenden Ausgleichsgeraden und eines Qualitätsmaßes aus den Parametern zweier unterlagerter Ausgleichsgeraden, ohne

**Tabelle A.1:** Zuordnung der Indizes für die Verknüpfung zweier Ausgleichsgeraden

$i_1$	1	...	$n_1$		1	...	$n_2$		$i_2$
$\mathbf{w}_1$	$(1 - \frac{n_1+1}{2})$	...	$(n_1 - \frac{n_1+1}{2})$		$(1 - \frac{n_2+1}{2})$	...	$(n_2 - \frac{n_2+1}{2})$		$\mathbf{w}_2$
$i_3$	1	...	$n_1$		$(n_1 + 1)$	...	$(n_1 + n_2)$		
$\mathbf{w}_3$	$(1 - \frac{n_1+n_2+1}{2})$	...	$(n_1 - \frac{n_1+n_2+1}{2})$		$(n_1 + 1 - \frac{n_1+n_2+1}{2})$	...	$(n_1 + n_2 - \frac{n_1+n_2+1}{2})$		

dass auf die Werte der Zahlenfolgen explizit zugegriffen werden muss. Die resultierende Zahlenfolge, die durch die Vereinigung der beiden Zahlenfolgen entstehen würde, aus denen die unterlagerten Ausgleichsgeraden ermittelt wurden, soll wiederum eindeutig und ohne Lücke sein, d.h. die beiden Zahlenfolgen müssen benachbart sein. Die Indizierung der resultierenden Zahlenfolge soll symmetrisch sein. Deshalb müssen die beschreibenden Parameter der beiden Eingangsdatenfolgen auf die neue Indizierung umgerechnet werden. Die fürs Verknüpfen notwendigen beschreibenden Parameter der unterlagerten Ausgleichsgeraden sind die Längen  $n_1$  und  $n_2$  der Zahlenfolgen, die Summen  $sy_1$  und  $sy_2$  der  $y_i$ -Werte, die Summen  $sy_1y_1$  und  $sy_2y_2$  der Quadrate der  $y_i$ -Werte und die Steigungen  $a_1$  und  $a_2$ . Es gelte

$$sy_1 := \sum_{i=1}^{n_1} y_{i_1}, \quad sy_2 := \sum_{i=1}^{n_2} y_{i_2}, \quad sy_1y_1 := \sum_{i=1}^{n_1} (y_{i_1}^2) \quad \text{und} \quad sy_2y_2 := \sum_{i=1}^{n_2} (y_{i_2}^2). \quad (\text{A.31})$$

#### A.4.1 Verknüpfung zweier Zahlenfolgen beliebiger Länge

Den Indizes  $i_1, \mathbf{w}_1$  der ersten Zahlenfolge und  $i_2, \mathbf{w}_2$  der zweiten Zahlenfolge können entsprechend Tabelle A.1 Indizes  $i_3, \mathbf{w}_3$  der neuen Zahlenfolge zugeordnet werden. Die symmetrische Neuindizierung nach  $\mathbf{w}_3$  kann als Koordinatentransformation in ein parallel zur  $x$ -Achse verschobenes Koordinatensystem interpretiert werden. Die  $y$ -Abschnitte  $b_1$  und  $b_2$  haben demnach im neuen gemeinsamen Koordinatensystem andere Werte. Die Werte der Steigungen  $a_1$  und  $a_2$  bleiben dagegen unverändert gültig, da sie bezüglich einer Translation ihrer Koordinatensysteme invariant sind.

Für die Berechnung der Steigung  $a_{n_1n_2}$  und der Varianz  $Var_{n_1n_2}$  wird die Summe  $swy_{n_1n_2}$  der gemischten Terme bezogen auf den neuen Indexvektor  $\mathbf{w}_3$  benötigt. Deshalb werden zunächst die gemischten Terme  $swy_{n_1merge}$  und  $swy_{n_2merge}$  für den neuen Indexvektor bestimmt. Der Trick besteht darin, dass die Steigungen  $a_1$  und  $a_2$  vom Koordinatensystem unabhängig sind und über Gleichung (A.19) mit (A.16) eine Beziehung zwischen den Summen der gemischten Terme und den Steigungen besteht. Die Summe der gemischten Terme  $swy_{n_1n_2}$  ergibt sich dann aus der Summe der nach  $\mathbf{w}_3$  umgerechneten Terme  $swy_{n_1merge}$  und  $swy_{n_2merge}$ .

Nach Gleichung (A.16) folgt

$$\Delta_1 = s_{12} \cdot n_1 - s_{11}^2 \quad (\text{A.32})$$

$$\Delta_2 = s_{22} \cdot n_2 - s_{21}^2 \quad (\text{A.33})$$

mit

$$s_{11} := \sum_{i=1}^{n_1} \left( i - \frac{n_1 + n_2 + 1}{2} \right) = -\frac{1}{2} \cdot n_1 \cdot n_2 \quad (\text{A.34})$$

$$s_{12} := \sum_{i=1}^{n_1} \left( \left( i - \frac{n_1 + n_2 + 1}{2} \right)^2 \right) = \frac{1}{12} \cdot n_1 \cdot (-1 + n_1^2 + 3n_2^2) \quad (\text{A.35})$$

$$s_{21} := \sum_{i=1}^{n_2} \left( n_1 + i - \frac{n_1 + n_2 + 1}{2} \right) = \frac{1}{2} \cdot n_1 \cdot n_2 = -s_{11} \quad (\text{A.36})$$

$$s_{22} := \sum_{i=1}^{n_2} \left( \left( n_1 + i - \frac{n_1 + n_2 + 1}{2} \right)^2 \right) = \frac{1}{12} \cdot n_2 \cdot (-1 + 3n_1^2 + n_2^2). \quad (\text{A.37})$$

Nach Gleichung (A.19) gilt

$$a_1 = \frac{n_1 \cdot swy_{n1merge} - s_{11} \cdot sy_1}{\Delta_1} \quad (\text{A.38})$$

$$a_2 = \frac{n_2 \cdot swy_{n2merge} - s_{21} \cdot sy_2}{\Delta_2}. \quad (\text{A.39})$$

Daraus folgt für die auf den neuen Indexvektor umgerechneten Summen der gemischten Terme

$$swy_{n1merge} = \frac{\Delta_1 \cdot a_1 + s_{11} \cdot sy_1}{n_1} \quad (\text{A.40})$$

$$swy_{n2merge} = \frac{\Delta_2 \cdot a_2 + s_{21} \cdot sy_2}{n_2}. \quad (\text{A.41})$$

Nun gilt

$$swy_{n1n2} = swy_{n1merge} + swy_{n2merge}. \quad (\text{A.42})$$

Es sei entsprechend Gleichung (A.22)

$$sww_{n1n2} = \frac{1}{12} \cdot (n_1 + n_2) \cdot (n_1 + n_2 + 1) \cdot (n_1 + n_2 - 1). \quad (\text{A.43})$$

Nach Gleichung (A.25) gilt

$$a_{n1n2} = \frac{swy_{n1n2}}{sww_{n1n2}}. \quad (\text{A.44})$$

Nach Gleichung (A.27) folgt

$$b_{n1n2} = \frac{sy_1 + sy_2}{n_1 + n_2}. \quad (\text{A.45})$$

Aus Gleichung (A.30) ergibt sich

$$Var_{n1n2} = \frac{sy_1 + sy_2}{n_1 + n_2} - \frac{swy_{n1n2}^2}{(n_1 + n_2) \cdot sww_{n1n2}} - b_{n1n2}^2. \quad (\text{A.46})$$

### A.4.2 Verknüpfen von Zahlenfolgen gleicher Länge

Beim Verknüpfen von Zahlenfolgen gleicher Länge können einige Gleichungen deutlich vereinfacht werden. Das Verknüpfen von Zahlenfolgen gleicher Länge tritt z.B. beim Aufbau einer Segmentierungspyramide nach Kapitel 4.2 auf. Es soll gelten:

$$n_1 := n_2 := \frac{n_3}{2} \quad (\text{A.47})$$

Aus den Gleichungen (A.34) bis (A.37) folgt mit (A.47) :

$$s_{11} := \sum_{i=1}^{\frac{n_3}{2}} \left( i - \frac{n_3 + 1}{2} \right) = -\frac{1}{8} \cdot n_3^2 \quad (\text{A.48})$$

$$s_{12} := \sum_{i=1}^{\frac{n_3}{2}} \left( \left( i - \frac{n_3 + 1}{2} \right)^2 \right) = \frac{1}{24} \cdot n_3 \cdot (n_3^2 - 1) \quad (\text{A.49})$$

$$s_{21} := \sum_{i=1}^{\frac{n_3}{2}} \left( \frac{n_3}{2} + i - \frac{n_3 + 1}{2} \right) = \frac{1}{8} \cdot n_3^2 = -s_{11} \quad (\text{A.50})$$

$$s_{22} := \sum_{i=1}^{\frac{n_3}{2}} \left( \left( \frac{n_3}{2} + i - \frac{n_3 + 1}{2} \right)^2 \right) = \frac{1}{24} \cdot n_3 \cdot (n_3^2 - 1) = s_{12} \quad (\text{A.51})$$

Mit den Gleichungen (A.32) und (A.33) folgt daraus

$$\begin{aligned} \Delta_1 &= s_{12} \cdot \frac{n_3}{2} - s_{11}^2 = \\ &= \frac{1}{48} \cdot \left( \frac{n_3^4}{4} - n_3^2 \right) \end{aligned} \quad (\text{A.52})$$

$$\Delta_2 = s_{22} \cdot \frac{n_3}{2} - s_{21}^2 = s_{12} \cdot \frac{n_3}{2} - s_{11}^2 = \Delta_1. \quad (\text{A.53})$$

Nach den Gleichungen (A.40) und (A.41) gilt

$$swy_{n1merge} = \frac{\Delta_1 \cdot a_1 + s_{11} \cdot sy_1}{\frac{n_3}{2}} \quad (\text{A.54})$$

$$\begin{aligned} swy_{n2merge} &= \frac{\Delta_2 \cdot a_2 + s_{21} \cdot sy_2}{\frac{n_3}{2}} = \\ &= \frac{\Delta_1 \cdot a_2 - s_{11} \cdot sy_2}{\frac{n_3}{2}}. \end{aligned} \quad (\text{A.55})$$

Hieraus ergibt sich die Summe der gemischten Terme  $swy_{n_3}$  als Funktion von  $n_3, a_1, a_2, sy_1$  und  $sy_2$ . Es gilt

$$\begin{aligned} swy_{n_3} &= swy_{n_1merge} + swy_{n_2merge} = \\ &= \frac{2}{n_3} \cdot [\Delta_1 \cdot (a_1 + a_2) + s_{11} \cdot (sy_1 - sy_2)] = \\ &= \frac{1}{96} \cdot n_3 \cdot [(n_3^2 - 4) \cdot (a_1 + a_2) - 24 \cdot (sy_1 - sy_2)]. \end{aligned} \quad (\text{A.56})$$

Es sei entsprechend Gleichung (A.22)

$$sww_{n_3} = \frac{1}{12} \cdot n_3 \cdot (n_3 + 1) \cdot (n_3 - 1). \quad (\text{A.57})$$

Nach Gleichung (A.25) gilt

$$a_{n_3} = \frac{swy_{n_3}}{sww_{n_3}}. \quad (\text{A.58})$$

Nach Gleichung (A.27) folgt

$$b_{n_3} = \frac{sy_1 + sy_2}{n_3}. \quad (\text{A.59})$$

Aus Gleichung (A.30) ergibt sich

$$Var_{n_3} = \frac{sy_1 + sy_2}{n_3} - \frac{swy_{n_3}^2}{n_3 \cdot sww_{n_3}} - b_{n_3}^2. \quad (\text{A.60})$$

### A.4.3 Rekursive Bestimmung einer Ausgleichsgeraden

Angenommen der Anfang eines Segments wäre bekannt, nur das Ende muss erst noch bestimmt werden. Das heißt, die Größe  $n$  eines Segments wird sukzessive von 1 beginnend erhöht, bis das Ende des Segments gefunden ist. Nach Abschnitt A.4.1 lässt sich diese Problemstellung als Verknüpfung eines Segments der Größe  $n_1$  mit einem Segment der Größe  $n_2 = 1$  zu einem Segment der Größe  $n$  interpretieren. Demnach gilt  $n_1 = n - 1$ . Mit den Gleichungen (A.40) bis (A.43) gilt

$$swy_{n_1n_2} = \frac{\Delta_1 \cdot a_1 + s_{11} \cdot sy_1}{n - 1} + \frac{\Delta_2 \cdot a_2 + s_{21} \cdot sy_2}{1} \quad (\text{A.61})$$

$$sww_{n_1n_2} = \frac{1}{12} \cdot n \cdot (n + 1) \cdot (n - 1). \quad (\text{A.62})$$

Mit den Gleichungen (A.32) bis (A.37) gilt

$$s_{11} = -\frac{1}{2} \cdot (n - 1) = -s_{21} \quad (\text{A.63})$$

$$s_{12} = \frac{1}{12} \cdot (n-1) \cdot ((n-1)^2 + 2) \quad (\text{A.64})$$

$$s_{22} = \frac{1}{4} \cdot (n-1)^2. \quad (\text{A.65})$$

Mit  $sy_1 = \sum_{i=1}^{n-1} y_i$  und  $sy_2 = y_n$  ergibt sich unter Verwendung von Gleichung (A.45)

$$b_{[n]} = \frac{sy_1 + y_n}{n} = \frac{\sum_{i=1}^n y_i}{n} \quad (\text{A.66})$$

und mit  $a_1 = a_{[n-1]}$ ,  $a_2 = 0$  aus (A.44) mit (A.61) und (A.62)

$$a_{[n]} = \frac{a_{[n-1]} \cdot (n^3 - 3 \cdot n^2 + 2 \cdot n) + 6 \cdot (y_n \cdot (n-1) - \sum_{i=1}^{n-1} y_i)}{n \cdot (n+1) \cdot (n-1)}$$

bzw.

$$a_{[n]} = a_{[n-1]} \cdot \frac{(n-2)}{(n+1)} + 6 \cdot \frac{y_n - b_{[n]}}{(n+1) \cdot (n-1)} \quad (\text{A.67})$$

als optimale Lösung für die Ausgleichsgerade mit kleinstem Fehlerquadrat, wobei  $a_{[0]} = 0$  und  $a_{[1]} = 0$  sei. Die Varianz  $Var_{[n]}$  kann bei Bedarf nach Gleichung (A.46) bestimmt werden.

An dieser Stelle sei nachfolgend noch eine experimentell ermittelte rekursive Näherungslösung angegeben, welche zwar nicht die optimale, aber dennoch eine sehr gute Näherung der Ausgleichsgeraden liefert. Bei Versuchen mit linearen Zahlenfolgen, denen ein im Verhältnis zur Länge 5 – 10 %-iges Rauschen überlagert wurde, waren die Abweichungen zur optimalen Lösung selbst bei kurzen Zahlenfolgen meist kleiner 1 – 2 %. Der Vorteil dieser Näherung besteht darin, dass gerade bei längeren Segmenten deutlich weniger Rechenoperationen im Vergleich zu Gleichung (A.67) durchzuführen sind. Mit  $b_{[n]}$  entsprechend Gleichung (A.66) und  $dy_{[0]} = 0$  sei

$$dy_{[n]} = dy_{[n-1]} + (y_n - b_{[n]}). \quad (\text{A.68})$$

Die Näherung für die Steigung  $a_{[n]}$  ergibt sich daraus durch Normierung von  $dy_{[n]}$  auf das Ergebnis der Rekursion  $dx_{[n]}$  des Indexes  $x_i \rightarrow i$  mit

$$\begin{aligned} dx_{[n]} &= dx_{[n-1]} + \left(n - \frac{\sum_{i=1}^n i}{n}\right) = \sum_{j=1}^n \left(j - \frac{\sum_{i=1}^j i}{j}\right) = \\ &= \sum_{j=1}^n \left(j - \frac{j \cdot (j+1)}{2j}\right) = \sum_{j=1}^n \left(\frac{1}{2} \cdot (j-1)\right) = \frac{1}{4} \cdot n \cdot (n-1) \end{aligned} \quad (\text{A.69})$$

$$a_{[n]} \approx \frac{dy_{[n]}}{\frac{1}{4} \cdot n \cdot (n-1)}, \quad (\text{A.70})$$

wobei  $dx_{[0]} = 0$  sei.

Das Verfahren kann sehr einfach auch auf 2-D und 3-D Punktemengen erweitert werden und wurde bereits zur Bestimmung von 3-D Ausgleichsgeraden durch 3-D Punktemengen versuchsweise eingesetzt. Bei ungeordneten Punktemengen muss zumindest die Hauptrichtung der Verteilung der 2-D und 3-D Punkte bekannt sein, damit mittels Fallunterscheidung die Differenz zwischen neuem Punkt und aktuellem Erwartungswert mit richtigem Vorzeichen in der Rekursion berücksichtigt werden kann.

## A.5 Einbringung einer Zwangsbedingung bei der Regression

Gelegentlich ist es notwendig neben dem Kriterium zur Minimierung der Summe der Fehlerquadrate weitere Zwangsbedingungen zu berücksichtigen. Beispielsweise könnte es das Ziel der Ausgleichsrechnung sein, gleichzeitig für eine Punktemenge  $\{y_1, z_1\}$  der Größe  $n_1$  und eine zweite Punktemenge  $\{y_2, z_2\}$  der Größe  $n_2$  Ausgleichsgeraden zu bestimmen, wobei als Zwangsbedingung die Parallelität der Geraden eingebracht werden soll. Die gesuchten Ausgleichsgeraden seien beschrieben durch

$$g1 : \bar{z}_1 = a \cdot y_1 + b_1 \quad (\text{A.71})$$

$$g2 : \bar{z}_2 = a \cdot y_2 + b_2. \quad (\text{A.72})$$

Der Fehler errechnet sich aus der Differenz der Näherung zu den tatsächlichen Werten

$$e_{g1i} := z_{1i} - a \cdot y_{1i} - b_1 \quad (\text{A.73})$$

$$e_{g2j} := z_{2j} - a \cdot y_{2j} - b_2 \quad (\text{A.74})$$

für  $i = 1..n_1$  und  $j = 1..n_2$ . Das zu minimierende Kriterium ist

$$J = \sum_{i=1}^{n_1} e_{g1i}^2 + \sum_{j=1}^{n_2} e_{g2j}^2 = \sum_{i=1}^{n_1} (z_{1i} - a \cdot y_{1i} - b_1)^2 + \sum_{j=1}^{n_2} (z_{2j} - a \cdot y_{2j} - b_2)^2. \quad (\text{A.75})$$

Für ein Minimum müssen die ersten Ableitungen nach  $b_1$ ,  $b_2$  und  $a$  null werden und die zweiten Ableitungen größer null sein.

$$\begin{aligned} \frac{\partial J}{\partial b_1} &= 2 \sum_{i=1}^{n_1} (z_{1i} - a \cdot y_{1i} - b_1) \cdot (-1) \\ &\Rightarrow \sum_{i=1}^{n_1} z_{1i} - a \cdot \sum_{i=1}^{n_1} y_{1i} - n_1 \cdot b_1 \stackrel{!}{=} 0 \end{aligned} \quad (\text{A.76})$$

$$\begin{aligned}\frac{\partial J}{\partial b_2} &= 2 \sum_{j=1}^{n_2} (z_{2j} - a \cdot y_{2j} - b_2) \cdot (-1) \\ &\Rightarrow \sum_{j=1}^{n_2} z_{2j} - a \cdot \sum_{j=1}^{n_2} y_{2j} - n_2 \cdot b_2 \stackrel{!}{=} 0\end{aligned}\quad (\text{A.77})$$

$$\begin{aligned}\frac{\partial J}{\partial a} &= 2 \sum_{i=1}^{n_1} (z_{1i} - a \cdot y_{1i} - b_1) \cdot (-y_{1i}) + 2 \sum_{j=1}^{n_2} (z_{2j} - a \cdot y_{2j} - b_2) \cdot (-y_{2j}) \\ &\Rightarrow \sum_{i=1}^{n_1} (-z_{1i} \cdot y_{1i} + a \cdot y_{1i}^2 + b_1 \cdot y_{1i}) + \sum_{j=1}^{n_2} (-z_{2j} \cdot y_{2j} + a \cdot y_{2j}^2 + b_2 \cdot y_{2j}) \stackrel{!}{=} 0\end{aligned}\quad (\text{A.78})$$

$$\frac{\partial^2 J}{\partial b_1^2} = n_1 > 0 \quad (\text{A.79})$$

$$\frac{\partial^2 J}{\partial b_2^2} = n_2 > 0. \quad (\text{A.80})$$

Weiterhin gilt

$$\frac{\partial^2 J}{\partial a^2} = \sum_{i=1}^{n_1} y_{1i}^2 + \sum_{j=1}^{n_2} y_{2j}^2 > 0, \quad (\text{A.81})$$

sobald wenigstens ein  $y_{1i} \neq 0$  oder  $y_{2j} \neq 0$  ist.

Daraus ergibt sich mit

$$\begin{aligned}sy_1 &:= \sum_{i=1}^{n_1} y_{1i}, \quad sy_2 := \sum_{j=1}^{n_2} y_{2j} \\ sz_1 &:= \sum_{i=1}^{n_1} z_{1i}, \quad sz_2 := \sum_{j=1}^{n_2} z_{2j} \\ syy_1 &:= \sum_{i=1}^{n_1} y_{1i}^2, \quad syy_2 := \sum_{j=1}^{n_2} y_{2j}^2 \\ syz_1 &:= \sum_{i=1}^{n_1} (z_{1i} \cdot y_{1i}) \quad \text{und} \quad syz_2 := \sum_{j=1}^{n_2} (z_{2j} \cdot y_{2j})\end{aligned}\quad (\text{A.82})$$

das folgende zu lösende lineare Gleichungssystem

$$\begin{bmatrix} n_1 & 0 & sy_1 \\ 0 & n_2 & sy_2 \\ sy_1 & sy_2 & syy_1 + syy_2 \end{bmatrix} \cdot \begin{bmatrix} b_1 \\ b_2 \\ a \end{bmatrix} = \begin{bmatrix} sz_1 \\ sz_2 \\ syz_1 + syz_2 \end{bmatrix}. \quad (\text{A.83})$$

Dieses hat die Lösung

$$b_1 = \frac{sz_1 \cdot n_2 \cdot (syy_1 + syy_2) - sz_1 \cdot sy_2^2 + sy_2 \cdot sy_1 \cdot sz_2 - sy_1 \cdot n_2 \cdot (sy_2 z_1 + sy_2 z_2)}{\Delta} \quad (\text{A.84})$$

$$b_2 = \frac{sz_2 \cdot n_1 \cdot (syy_1 + syy_2) - sz_2 \cdot sy_1^2 + sy_2 \cdot sy_1 \cdot sz_1 - sy_2 \cdot n_1 \cdot (sy_2 z_1 + sy_2 z_2)}{\Delta} \quad (\text{A.85})$$

$$a = \frac{-sy_1 \cdot n_2 \cdot sz_1 - n_1 \cdot sy_2 \cdot sz_2 + n_1 \cdot n_2 \cdot (sy_2 z_1 + sy_2 z_2)}{\Delta}, \quad (\text{A.86})$$

wobei  $\Delta = n_1 \cdot n_2 \cdot (syy_1 + syy_2) - n_1 \cdot sy_2^2 - n_2 \cdot sy_1^2$  ist.

## A.6 Subpixelgenauigkeit

Die Bestimmung der Position eines Maximal- bzw. Minimalwertes aus einem Signalverlauf ist zunächst lediglich aufs Pixel genau. Eine Verbesserung auf Subpixel-Genauigkeit kann man durch die Bestimmung der Scheitelposition der Parabel erreichen, welche durch den Extremwert und die beiden benachbarten Signalwerte eindeutig festgelegt ist.

Nachfolgend bezeichnen  $y_0$  den Extremwert an der Stelle  $w_2 = 0$ ,  $y_{minus}$  den links davon angesiedelten Signalwert mit  $w_1 = -1$  und  $y_{plus}$  den rechts zum Extremwert gelegenen Signalwert mit Index  $w_3 = +1$ .

Das Polynom durch diese drei Punkte folgt eindeutig aus der Lösung des Gleichungssystems

$$\begin{bmatrix} y_{minus} \\ y_0 \\ y_{plus} \end{bmatrix} = \begin{bmatrix} w_1^2 & w_1 & 1 \\ w_2^2 & w_2 & 1 \\ w_3^2 & w_3 & 1 \end{bmatrix} \cdot \begin{bmatrix} a \\ b \\ c \end{bmatrix}. \quad (\text{A.87})$$

Mit  $w_1 = -1$ ,  $w_2 = 0$  und  $w_3 = +1$  gilt

$$\begin{bmatrix} y_{minus} \\ y_0 \\ y_{plus} \end{bmatrix} = \begin{bmatrix} 1 & -1 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} a \\ b \\ c \end{bmatrix} \quad (\text{A.88})$$

und daraus

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0.5 & -1 & 0.5 \\ -0.5 & 0 & 0.5 \\ 0 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} y_{minus} \\ y_0 \\ y_{plus} \end{bmatrix}. \quad (\text{A.89})$$

Der Scheitel  $w_S$  der Parabel errechnet sich aus

$$y = a \cdot w^2 + b \cdot w + c \quad (\text{A.90})$$

$$\frac{dy}{dw} = 2 \cdot a \cdot w + b \stackrel{!}{=} 0$$

$$\Rightarrow w_S = -\frac{b}{2 \cdot a}.$$

Mit Gleichung (A.89) folgt daraus

$$w_S = \frac{y_{minus} - y_{plus}}{2 \cdot (y_{minus} - 2y_0 + y_{plus})}. \quad (\text{A.91})$$

Der Wert von  $w_S$  liegt stets im Intervall  $[-0.5; 0.5]$ . Die Subpixelposition eines Extremwertes ergibt sich aus der Summe der pixelgenauen Extremwertposition mit dem Scheitelwert  $w_S$ . In Abbildung A.1 wird dies an einem Beispiel veranschaulicht. Die Abbildung zeigt die Bestimmung der Subpixelpositionen eines Minimums und eines Maximums aus dem Differenzensignal eines Intensitätssignals.

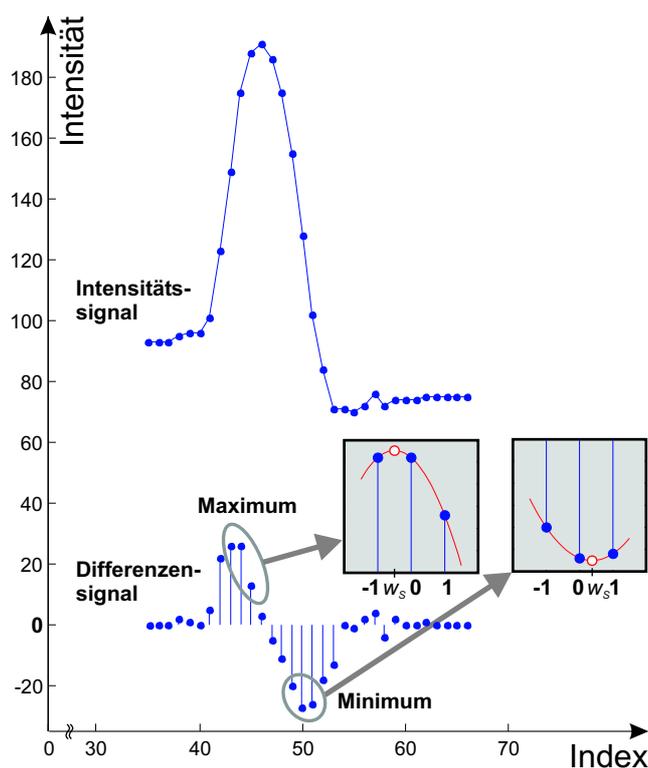


Abbildung A.1: Veranschaulichung der Subpixel-Approximation

## A.7 Bestimmung einer Ausgleichsebene

Für eine Matrix der Dimension  $n \times m$ , deren Werte z.B. die Helligkeitswerte eines rechteckigen Ausschnittes aus einem Grauwertbild sein könnten, soll eine Ausgleichsebene bestimmt werden, die den quadratischen Fehler der Abweichungen minimiert.

In diesem Abschnitt seien  $\vec{n} = [n_x, n_y, n_z]^T$  der Normalenvektor der gesuchten Ausgleichsebene,  $\vec{a} = [a_x, a_y, a_z]^T$  der Ortsvektor des Aufpunkts der Ebene und  $\vec{v} = [v_x, v_y, v_z]^T$  ein beliebiger Punkt auf der Ebene.

Diese kann dann z.B. mittels der Hesse-Normalenform eindeutig beschrieben werden

$$\vec{n}_0 \circ (\vec{v} - \vec{a}) = 0, \text{ mit } \vec{n}_0 = \vec{n} \cdot \frac{1}{|\vec{n}|} \quad (\text{A.92})$$

wobei die Orientierung von  $\vec{n}_0$  so sein soll, dass  $\text{sign}(-\vec{n}_0 \circ \vec{a}) = -1$  ist.

Daraus folgt

$$n_z \cdot v_z = -n_x \cdot v_x - n_y \cdot v_y + c \quad (\text{A.93})$$

mit  $c = n_x \cdot a_x + n_y \cdot a_y + n_z \cdot a_z$ ,

bzw.

$$v_z = \begin{bmatrix} v_x & v_y & 1 \end{bmatrix} \cdot \begin{bmatrix} p_x \\ p_y \\ p_z \end{bmatrix} \quad (\text{A.94})$$

mit  $p_x = -\frac{n_x}{n_z}$ ,  $p_y = -\frac{n_y}{n_z}$  und  $p_z = \frac{c}{n_z}$ .

Für eine Matrix  $\mathbf{Z}$  der Dimension  $n \times m$  mit dem Spaltenindex  $w_{xj}$  mit  $j = 1..m$ , dem Zeilenindex  $w_{yi}$  mit  $i = 1..n$  und dem Intensitätswert  $z_{ij}$  an der Stelle  $(i, j)$  der Matrix  $\mathbf{Z}$  errechnet sich  $\mathbf{p}$  nach Gleichung (A.6)

$$\mathbf{p} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \cdot \mathbf{z}, \quad (\text{A.95})$$

wobei

$$\mathbf{A}^T = \begin{bmatrix} w_{x1} & w_{x1} & \cdots & w_{x1} & w_{x2} & w_{x2} & \cdots & w_{x2} & \cdots & w_{xm} \\ w_{y1} & w_{y2} & \cdots & w_{yn} & w_{y1} & w_{y2} & \cdots & w_{yn} & \cdots & w_{yn} \\ 1 & 1 & \cdots & 1 & 1 & 1 & \cdots & 1 & \cdots & 1 \end{bmatrix}^T \quad (\text{A.96})$$

$$\mathbf{z}^T = \begin{bmatrix} z_{11} & z_{21} & \cdots & z_{n1} & z_{12} & z_{22} & \cdots & z_{n2} & \cdots & z_{nm} \end{bmatrix}^T.$$

Daraus folgt

$$\mathbf{A}^T \cdot \mathbf{z} = \begin{bmatrix} \sum_{j=1}^m (w_{xj} \cdot \sum_{i=1}^n z_{ij}) \\ \sum_{i=1}^n (w_{yi} \cdot \sum_{j=1}^m z_{ij}) \\ \sum_{i=1}^n (\sum_{j=1}^m z_{ij}) \end{bmatrix} \quad (\text{A.97})$$

$$\mathbf{A}^T \mathbf{A} = \begin{bmatrix} n \cdot \sum_{j=1}^m w_{xj}^2 & (\sum_{j=1}^m w_{xj}) \cdot (\sum_{i=1}^n w_{yi}) & n \cdot \sum_{j=1}^m w_{xj} \\ (\sum_{j=1}^m w_{xj}) \cdot (\sum_{i=1}^n w_{yi}) & m \cdot \sum_{i=1}^n w_{yi}^2 & m \cdot \sum_{i=1}^n w_{yi} \\ n \cdot \sum_{j=1}^m w_{xj} & m \cdot \sum_{i=1}^n w_{yi} & n \cdot m \end{bmatrix}. \quad (\text{A.98})$$

Für symmetrische Indizes  $\mathbf{w}_x$  und  $\mathbf{w}_y$  gilt

$$\begin{aligned} \sum_{j=1}^m w_{xj} &= 0 \\ \sum_{i=1}^n w_{yi} &= 0 \end{aligned} \quad (\text{A.99})$$

daraus folgt

$$\mathbf{A}^T \mathbf{A} = \begin{bmatrix} n \cdot \sum_{j=1}^m w_{xj}^2 & 0 & 0 \\ 0 & m \cdot \sum_{i=1}^n w_{yi}^2 & 0 \\ 0 & 0 & n \cdot m \end{bmatrix}. \quad (\text{A.100})$$

Mit

$$\begin{aligned} sww_m &:= \sum_{j=1}^m w_{xj}^2 \\ sww_n &:= \sum_{i=1}^n w_{yi}^2 \end{aligned} \quad (\text{A.101})$$

folgt

$$(\mathbf{A}^T \mathbf{A})^{-1} = \begin{bmatrix} \frac{1}{n \cdot sww_m} & 0 & 0 \\ 0 & \frac{1}{m \cdot sww_n} & 0 \\ 0 & 0 & \frac{1}{n \cdot m} \end{bmatrix} \quad (\text{A.102})$$

und daraus mit (A.95) und (A.97)

$$p_x = \frac{\sum_{j=1}^m (w_{xj} \cdot \overbrace{\sum_{i=1}^n z_{ij}}^{\text{Spalte } j})}{n \cdot sww_m} \quad (\text{A.103})$$

$$p_y = \frac{\sum_{i=1}^n (w_{yi} \cdot \overbrace{\sum_{j=1}^m z_{ij}}^{\text{Zeile } i})}{m \cdot sww_n} \quad (\text{A.104})$$

$$p_z = \frac{\sum_{i=1}^n (\sum_{j=1}^m z_{ij})}{n \cdot m} = b. \quad (\text{A.105})$$

Für symmetrische Indizes  $w_{xj} \rightarrow j - \frac{n+1}{2}$ ,  $j = 1..m$  und  $w_{yi} \rightarrow i - \frac{n+1}{2}$ ,  $i = 1..n$  gilt

$$sww_n = \frac{1}{12} \cdot n \cdot (n^2 - 1) \text{ und } sww_m = \frac{1}{12} \cdot m \cdot (m^2 - 1). \quad (\text{A.106})$$

Für  $n_z = 1$  ergibt sich mit Gleichung (A.94) der Normalenvektor der Ausgleichsebene zu

$$\vec{n}_0 = \vec{n} \cdot \frac{1}{|\vec{n}|} = \begin{bmatrix} -p_x \\ -p_y \\ 1 \end{bmatrix} \cdot \frac{1}{\sqrt{p_x^2 + p_y^2 + 1}} \quad (\text{A.107})$$

und der Ortsvektor des Aufpunkts aus den Gleichungen (A.93) und (A.94) zu

$$\vec{a} = \begin{bmatrix} 0 \\ 0 \\ b \end{bmatrix}. \quad (\text{A.108})$$

Als ein Spezialfall zur Bestimmung des Normalenvektors einer Ausgleichsebene mit  $n = m = 3$  ungerade kann der aus der Literatur bekannte *3x3-Prewitt-Operator* betrachtet werden.

Die Richtung einer gesuchten Kante ist orthogonal zur Projektion des Normalenvektors der Ausgleichsebene in die Bildebene. Der Richtungsvektor  $\vec{r}$  der Kante am betrachteten Bildpunkt ergibt sich demnach aus dem Kreuzprodukt des Dateneinheitsvektors  $\vec{e}_z$  mit dem Normalenvektor  $\vec{n}$  der Ausgleichsebene. Es gilt

$$\vec{r} = \vec{e}_z \times \vec{n} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \times \begin{bmatrix} -p_x \\ -p_y \\ 1 \end{bmatrix} = \begin{bmatrix} p_y \\ -p_x \\ 0 \end{bmatrix}. \quad (\text{A.109})$$

Durch die Reihenfolge der Vektoren bei der Ermittlung des Kreuzprodukts wird der Umlaufsinn festgelegt. Das heißt, der hellere Teil des Bildausschnitts liegt hier stets zur Rechten und der dunklere Teil stets zur Linken eines gerichteten Kantenpunkts.

Der Winkel zwischen dem Normalenvektor  $\vec{n}$  und dem Einheitsvektor der Datenwerte  $\vec{e}_z$  kann als Kriterium für das Vorhandensein eines signifikanten Helligkeitsübergangs verwendet werden. Er wird nachfolgend als *Normalenwinkel*  $\alpha$  bezeichnet.

Es gilt

$$\begin{aligned} \cos \alpha &= \frac{\vec{n} \circ \vec{e}_z}{|\vec{n}| \cdot |\vec{e}_z|} = \frac{1}{\sqrt{p_x^2 + p_y^2 + 1}} \\ \Rightarrow \alpha &= \arccos \left( \frac{1}{\sqrt{p_x^2 + p_y^2 + 1}} \right) \end{aligned} \quad (\text{A.110})$$

Im Quellcode wird zur Schonung der Rechenressourcen nicht der Arcuscosinus vom Kehrbuch der Wurzel der Länge des Normenvektors gebildet, sondern lediglich die Länge des Normalenvektors als Kriterium verwendet. Für eine plastische Vorstellung der Funktionsweise des Algorithmus stellt der Normalenwinkel allerdings ein wesentlich eingänglicheres Kriterium im Vergleich zur Länge des Normalenvektors dar.

Zur Vereinfachung der für die Gleichungen (A.103) und (A.104) notwendigen Summationen und um eine spätere systematische Wiederverwendbarkeit der Ergebnisse zu ermöglichen, wird an dieser Stelle noch kurz eine vereinfachte symmetrische Indizierung vorgestellt.

Es seien  $w_{xj} \rightarrow \text{sign}(j - \frac{n+1}{2})$ ,  $j = 1..m$  und  $w_{yi} \rightarrow \text{sign}(i - \frac{n+1}{2})$ ,  $i = 1..n$ . Dann gilt für  $n$  bzw.  $m$  ungerade

$$sww_n = n - 1, sww_m = m - 1 \quad (\text{A.111})$$

und für  $n$  bzw.  $m$  gerade

$$sww_n = n, sww_m = m. \quad (\text{A.112})$$

Bei vereinfachter symmetrischer Indizierung muss eine geeignete Gewichtung der  $p_x$ - und  $p_y$ -Werte vorgenommen werden, damit die Beiträge in x- und y-Richtung die gleiche Wertigkeit wie bei symmetrischer Indizierung besitzen und damit der Normalenvektor und die Richtungsinformation korrekt bestimmt werden können. Dadurch ergeben sich auch bei vereinfachter symmetrischer Indizierung sehr gute Näherungen für den Normalenvektor  $\vec{n}$  und den Richtungsvektor  $\vec{r}$ .

Die gewichteten Werte  $\tilde{p}_x$  und  $\tilde{p}_y$  errechnen sich aus  $p_x$  und  $p_y$  nach Gleichung (A.103) und (A.104) durch Multiplikation mit dem Kehrwert des x- bzw. y-Abstands der Mittelpunkte der  $\mathbf{B}_{\text{pxq}}$ -Elementarblöcke entsprechend Abbildung (4.14) bzw. Abbildung (4.17) zum Mittelpunkt des  $n \times m$  Ausschnitts. Es gilt

$$\tilde{p}_x = p_x \cdot \frac{4}{m+1} \text{ und } \tilde{p}_y = p_y \cdot \frac{4}{n+1} \quad (\text{A.113})$$

für  $n \geq 3$  bzw.  $m \geq 3$  ungerade und

$$\tilde{p}_x = p_x \cdot \frac{4}{m} \text{ und } \tilde{p}_y = p_y \cdot \frac{4}{n} \quad (\text{A.114})$$

für  $n \geq 2$  bzw.  $m \geq 2$  gerade.

Das heißt, in den Gleichungen (A.107) und (A.109) sind  $p_x$  und  $p_y$  durch die gewichteten Werte  $\tilde{p}_x$  und  $\tilde{p}_y$  zu ersetzen.

# Anhang B

## Zustandsschätzung

### B.1 Das Kalmanfilter

Das nach Rudolf Emil Kalman benannte Kalmanfilter [KALMAN 60] ist ein optimaler Zwei-Schritt-Schätzer zur rekursiven Schätzung des Zustands eines zeitdiskreten linearen dynamischen Systems nach dem Kriterium der kleinsten Summe der gewichteten quadratischen Fehler. Die statistischen Kenngrößen der Messfehler und der auf das System einwirkenden Störgrößen werden dabei als bekannt vorausgesetzt. Aufgrund seines rekursiven Filterentwurfs erlaubt es eine sehr effiziente Implementierung des Optimierungsproblems. In praktischen Anwendungen, wie z.B. bei der maschinellen Bildfolgenauswertung zur autonomen Fahrzeugführung, wird man allerdings in der Regel mit nichtlinearen Systemen und Messgleichungen konfrontiert. Deshalb wird hier aufgrund der höheren Praxisrelevanz gleich das Erweiterte Kalman Filter (EKF, engl. „*Extended Kalman Filter*“) beschrieben, das eine Erweiterung des linearen Kalmanfilters auf nichtlineare zeitdiskrete Systeme darstellt. Die nachfolgende Beschreibung ist an [WELCH 03] angelehnt.

Prinzipiell wird mittels eines Kalmanfilters versucht, an einem Zeitpunkt  $t_k$  den Zustandsvektor  $\mathbf{x}_k \in \mathbb{R}^n$  eines zeitdiskreten dynamischen Systems zu schätzen. Für die Systemgleichung eines nichtlinearen zeitdiskreten Systems wird vorausgesetzt, dass es sich durch

$$\mathbf{x}_k = f(\mathbf{x}_{k-1}, \mathbf{u}_{k-1}, \mathbf{w}_{k-1}) \quad (\text{B.1})$$

und die nichtlineare Messgleichung durch die Abbildungsvorschrift

$$\mathbf{y}_k = h(\mathbf{x}_k, \mathbf{v}_k) \quad (\text{B.2})$$

beschreiben lässt. Dabei stellt  $\mathbf{u}_{k-1} \in \mathbb{R}^l$  einen optionalen Vektor der Steuergrößen dar. Das Systemrauschen  $\mathbf{w}_{k-1}$  soll im Wesentlichen den Anteil repräsentieren, der durch Modellierungsfehler und sonstige Störeinflüsse in einem realen System stets vorhanden ist. Das Messrauschen  $\mathbf{v}_k$  beinhaltet die Anteile, die durch Modellierungsfehler der nichtlinearen Abbildungsvorschrift (dem *Messmodell*) und durch den Messfehler der Messapparatur bzw. eines verwendeten Messverfahrens bei endlicher Messgenauigkeit entstehen.

Die beiden Zufallsvariablen  $\mathbf{w}_k$  und  $\mathbf{v}_k$  seien stochastisch unabhängig voneinander, jeweils normalverteilt und weiß, und damit ihre zeitlichen Folgen mittelwertfrei und unkorreliert. Die Kovarianzmatrix des Systemrauschens sei  $\mathbf{Q}_k = E[\mathbf{w}_k \cdot \mathbf{w}_k^T]$  und die Kovarianzmatrix des Messrauschens sei durch die Matrix  $\mathbf{R}_k = E[\mathbf{v}_k \cdot \mathbf{v}_k^T]$  festgelegt.

Für den Fall, dass die den Zustandsgrößen überlagerten Störgrößen stochastisch unabhängig voneinander sind, besitzt die Matrix  $\mathbf{Q}$  Diagonalform. Ihre Einträge sind dann idealerweise so zu wählen, dass die Quadratwurzeln der Diagonalelemente den erwarteten Streuungen der modellierten Zustandsgrößen um die tatsächlichen physikalischen Werte entsprechen.

Es sei  $\mathbf{x}_k^* \in \mathbf{R}^n$  die zum Zeitpunkt  $t_{k-1}$  für den Zeitpunkt  $t_k$  prädizierte *a priori* Schätzung des Zustandsvektors  $\mathbf{x}_k$  unmittelbar vor dessen Korrektur entsprechend der durch die Messung eingebrachten Information. Mit  $\hat{\mathbf{x}}_k$  werde der *a posteriori* Schätzvektor unmittelbar nach der Korrektur bezeichnet.

Weiterhin wird der a priori Schätzfehler bzw. Prädiktionsfehler  $\mathbf{e}_k^* := \mathbf{x}_k - \mathbf{x}_k^*$  über die Differenz zwischen dem nicht zugänglichen tatsächlichen Systemzustand  $\mathbf{x}_k$  und dem prädizierten a priori Schätzvektor  $\mathbf{x}_k^*$  definiert. Für die a posteriori Größen ergibt sich analog der a posteriori Schätzfehler  $\mathbf{e}_k := \mathbf{x}_k - \hat{\mathbf{x}}_k$ . Für die a priori Schätzfehlerkovarianz folgt daraus  $\mathbf{P}_k^* = E[\mathbf{e}_k^* \cdot \mathbf{e}_k^{*T}]$  und die a posteriori Schätzfehlerkovarianz  $\mathbf{P}_k = E[\mathbf{e}_k \cdot \mathbf{e}_k^T]$ .

Mit der Systemgleichung (B.1) und dem a priori Schätzvektor

$$\mathbf{x}_k^* = f(\hat{\mathbf{x}}_{k-1}, \mathbf{u}_{k-1}, \mathbf{0}) \quad (\text{B.3})$$

folgt durch Linearisierung für den Prädiktionsfehler

$$\begin{aligned} \mathbf{e}_k^* &= \mathbf{x}_k - \mathbf{x}_k^* \\ &= f(\mathbf{x}_{k-1}, \mathbf{u}_{k-1}, \mathbf{w}_{k-1}) - f(\hat{\mathbf{x}}_{k-1}, \mathbf{u}_{k-1}, \mathbf{0}) \\ &\approx \mathbf{A}_{k-1} \cdot (\mathbf{x}_{k-1} - \hat{\mathbf{x}}_{k-1}) + \mathbf{W}_{k-1} \cdot \mathbf{w}_{k-1} \end{aligned} \quad (\text{B.4})$$

mit der Systemmatrix  $\mathbf{A}_{k-1} := \left. \frac{\partial f}{\partial \mathbf{x}} \right|_{\hat{\mathbf{x}}_{k-1}, \mathbf{u}_{k-1}, \mathbf{w}_{k-1}=\mathbf{0}}$  und der Eingangsmatrix für das Systemrauschen  $\mathbf{W}_{k-1} := \left. \frac{\partial f}{\partial \mathbf{w}} \right|_{\hat{\mathbf{x}}_{k-1}, \mathbf{u}_{k-1}, \mathbf{w}_{k-1}=\mathbf{0}}$ .

Die Linearisierung der Messgleichung (B.2) ergibt

$$\mathbf{y}_k \approx h(\mathbf{x}_k^*, \mathbf{0}) + \mathbf{C}_k \cdot (\mathbf{x}_k - \mathbf{x}_k^*) + \mathbf{V}_k \cdot \mathbf{v}_k \quad (\text{B.5})$$

mit der Matrix  $\mathbf{C}_k$  gleich der Jacobimatrix der partiellen Ableitungen von  $h$  nach  $\mathbf{x}$  und der Matrix  $\mathbf{V}_k$  als der Jacobimatrix der partiellen Ableitungen von  $h$  nach  $\mathbf{v}$ , beide an der Stelle  $\mathbf{x}_k^*$  und  $\mathbf{v}_k = \mathbf{0}$ .

Aus dem a priori Schätzwert  $\mathbf{x}_k^*$  folgt die a posteriori Zustandsschätzung  $\hat{\mathbf{x}}_k$  durch Addition eines Korrekturterms, der sich als gewichtete Differenz zwischen der aktuellen Messung  $\mathbf{y}_k$  und der erwarteten Messung  $\mathbf{y}_k^* = h(\mathbf{x}_k^*, \mathbf{0})$  ergibt:

$$\hat{\mathbf{x}}_k = \mathbf{x}_k^* + \mathbf{K}_k \cdot (\mathbf{y}_k - \mathbf{y}_k^*). \quad (\text{B.6})$$

Die Differenz zwischen der Messung und der erwarteten Messung wird *Residuum* oder *Innovation* genannt. Die  $n \times m$ -Matrix  $\mathbf{K}_k$  wird als *Kalmanverstärkung* bezeichnet und wird so gewählt, dass die a posteriori Schätzfehlerkovarianz  $\mathbf{P}_k$  minimiert wird. Eine mögliche Darstellungsform der Berechnungsvorschrift für  $\mathbf{K}_k$ , die diese Vorgabe erfüllt, lautet

$$\mathbf{K}_k = \mathbf{P}_k^* \cdot \mathbf{C}_k^T \cdot (\mathbf{C}_k \cdot \mathbf{P}_k^* \cdot \mathbf{C}_k^T + \mathbf{V}_k \cdot \mathbf{R}_k \cdot \mathbf{V}_k^T)^{-1}. \quad (\text{B.7})$$

mit

$$\mathbf{P}_k^* = \mathbf{A}_{k-1} \cdot \mathbf{P}_{k-1} \cdot \mathbf{A}_{k-1}^T + \mathbf{W}_{k-1} \cdot \mathbf{Q}_{k-1} \cdot \mathbf{W}_{k-1}^T. \quad (\text{B.8})$$

Die  $n \times n$ -Matrix  $\mathbf{A}_{k-1}$  dient der Prädiktion von Zeitpunkt  $t_{k-1}$  zum Zeitpunkt  $t_k$ . Für den mathematischen Zusammenhang, der die Korrektur der a priori Schätzfehlerkovarianz  $\mathbf{P}_k^*$  ermöglicht und dem durch die Messung  $\mathbf{y}_k$  eingebrachten Wissen durch Verringerung der a posteriori Schätzfehlerkovarianz  $\mathbf{P}_k$  Rechnung trägt, gilt:

$$\mathbf{P}_k = (\mathbf{I} - \mathbf{K}_k \cdot \mathbf{C}_k) \cdot \mathbf{P}_k^*. \quad (\text{B.9})$$

Auf die Herleitung der Gleichungen B.7 bis B.9 wurde hier verzichtet, da sie in der gängigen Standardliteratur zur Kalmanfilterung, z.B. [MAYBECK 79], [BIERMAN 77], nachzulesen ist.

Nach der Vorgabe initialer Werte für  $\hat{\mathbf{x}}_{k-1}$  und  $\mathbf{P}_{k-1}$  und der als bekannt vorausgesetzten Kovarianzmatrizen  $\mathbf{Q}_{k-1}$  und  $\mathbf{R}_k$  stellt sich ein vollständiger Filterzyklus wie folgt dar:

#### Prädiktionsschritt:

Im ersten Teil, dem *Prädiktionsschritt* (engl. „*Time Update*“), wird mittels der Gleichungen (B.3) an der Stelle  $\hat{\mathbf{x}}_{k-1}$ ,  $\mathbf{u}_{k-1}$  und  $\mathbf{w}_{k-1} = \mathbf{0}$  der a priori Schätzvektor

$$\mathbf{x}_k^* = f(\hat{\mathbf{x}}_{k-1}, \mathbf{u}_{k-1}, \mathbf{0})$$

und mit (B.8) die a priori Schätzfehlerkovarianz

$$\mathbf{P}_k^* = \mathbf{A}_{k-1} \cdot \mathbf{P}_{k-1} \cdot \mathbf{A}_{k-1}^T + \mathbf{W}_{k-1} \cdot \mathbf{Q}_{k-1} \cdot \mathbf{W}_{k-1}^T$$

für den Zeitpunkt  $t_k$  bestimmt.

#### Korrekturschritt:

Im zweiten Teil erfolgt dann der *Korrekturschritt* (engl. „*Measurement Update*“) entsprechend den Gleichungen (B.7), (B.6) und (B.9). Dabei wird der a priori Schätzvektor  $\mathbf{x}_k^*$  um das mit der Kalmanverstärkung  $\mathbf{K}_k$  gewichtete Residuum derart korrigiert, dass für den a posteriori Schätzvektor  $\hat{\mathbf{x}}_k$  die Summe der quadratischen Fehler bezüglich aller bisher eingespeisten Messwerte minimal wird. Im Anschluß wird noch die a posteriori Schätzfehlerkovarianz  $\mathbf{P}_k$  aktualisiert.

$$\mathbf{K}_k = \mathbf{P}_k^* \cdot \mathbf{C}_k^T \cdot (\mathbf{C}_k \cdot \mathbf{P}_k^* \cdot \mathbf{C}_k^T + \mathbf{V}_k \cdot \mathbf{R}_k \cdot \mathbf{V}_k^T)^{-1}$$

$$\hat{\mathbf{x}}_k = \mathbf{x}_k^* + \mathbf{K}_k \cdot (\mathbf{y}_k - h(\mathbf{x}_k^*, \mathbf{0}))$$

$$\mathbf{P}_k = (\mathbf{I} - \mathbf{K}_k \cdot \mathbf{C}_k) \cdot \mathbf{P}_k^*.$$

Nach jeder Prädiktor-Korrektor-Folge wird der Zykluszähler  $k$  um eins erhöht und der Vorgang wiederholt, indem die zuletzt berechneten a posteriori Größen dazu verwendet werden, zunächst die Transitionsmatrix  $\mathbf{A}_{k-1}$  und die Jacobimatrizen  $\mathbf{W}_{k-1}$ ,  $\mathbf{C}_{k-1}$  und  $\mathbf{V}_{k-1}$  zu aktualisieren, um dann die a priori Größen im Prädiktionsschritt zu bestimmen...

Anmerkung: Der Vorgang der Einbringung der Innovation wird nicht selten selbst als „*Innovation*“ bezeichnet. Aus dem Kontext ist dabei meist zu entnehmen, ob damit der Vorgang oder das Residuum gemeint ist. Im Englischen wird ausschließlich der Begriff „*Measurement Update*“ verwendet.

Eine anschauliche Einführung in die Thematik am Beispiel des linearen Kalmanfilters bietet das Kapitel 1.3 „The Kalman Filter: An Introduction to Concepts“ aus [MAYBECK 79].

## B.2 Anmerkungen zur verwendeten Implementierung des EKF

In [WÜNSCHE 87] werden in den Kapiteln 3.5.3.2 „Kalman Filter Rekursion“, 3.5.3.3 „Gauß-Markov Schätzung mit quasistationärer Nachfilterung“, 3.5.3.4 „Kalman Filter mit sequentieller Innovation“ und 3.5.3.5 „Kalman Filter mit  $UDU^T$  faktorisierter Kovarianzmatrix“ verschiedene Varianten der Zustandsschätzung bezüglich ihrer numerischen Stabilität und Recheneffizienz untersucht. In ähnlicher Weise werden in [MYSLIWETZ 90] in Kapitel 5.4 „Rekursive Zustandsschätzung als Interpretationsmechanismus“ die Funktionsprinzipien des 4-D Ansatzes zur Bildfolgenauswertung und die Vor- und Nachteile verschiedener Varianten des Kalmanfilters diskutiert. Als Ergebnis dieser Untersuchungen wurde am ISF für die Implementierung des EKF eine Variante nach [BIERMAN 77] ausgewählt.

Bierman führt eine UD-Zerlegung der Schätzfehlerkovarianzmatrix  $\mathbf{P} = \mathbf{UDU}^T$  durch, mit  $\mathbf{U}$  einer oberen Dreiecksmatrix und  $\mathbf{D}$  einer Diagonalmatrix, wodurch gewährleistet ist bzw. erzwungen wird, dass die Kovarianzmatrix  $\mathbf{P}$  stets symmetrisch bleibt. Dadurch wird die numerische Stabilität erhöht. Unter der Voraussetzung, dass die Messungen  $y_1, \dots, y_m$  des Messvektors  $\mathbf{y}_k$  unabhängig voneinander sind, kann die sogenannte *sequentielle Innovation* angewandt werden. Durch sequentielle Einbringung der Innovation wird erreicht, dass die numerisch aufwendige Matrizeninversion aus Gleichung (B.7) vermieden wird und die Berechnung der Kalmanverstärkung in skalarer Form durchgeführt werden kann. Weiterhin können dadurch die Algorithmen für den in der Praxis häufig auftretenden Fall ausgelegt werden, dass die Anzahl der Messungen von Zeitschritt zu Zeitschritt variiert.

Der Prädiktionsschritt findet sich in [BIERMAN 77] im Abschnitt VI.4 auf den Seiten 124ff und im Appendix VI.A auf den Seiten 131ff. Der Korrekturschritt kann in Abschnitt V.3, S.76ff und im Appendix V.A auf den Seiten 100ff nachgelesen werden.

# Literaturverzeichnis

- [AUDILEXIKON 03] Internet-Lexikon der Firma Audi: *Audi Lexikon*, <http://www.audizentrum-hannover.de/public/de/lexikonindex.htm>
- [BALLARDBROWN 82] D.H. Ballard, CH.M. Brown: *Computer vision*, Prentice-Hall, Inc. Englewood Cliffs, New Jersey, 1982.
- [BÄSSMANN 89] H. Bäsmann, Ph.-W. Besslich: *Konturorientierte Verfahren der digitalen Bildverarbeitung*, Springer Verlag, Berlin, 1989.
- [BEHRINGER 96] R. Behringer: *Visuelle Erkennung und Interpretation des Fahrspurverlaufes durch Rechnersehen für ein autonomes Straßenfahrzeug*, Dissertation, Universität der Bundeswehr München, Fakultät für Luft- und Raumfahrttechnik, Neubiberg, 1996.
- [BELLUTTA 00] P. Bellutta, R.Manduchi, L. Matthies, K.Owens, A.Rankin: *Terrain Perception for DEMO III*, IEEE Symposium on Intelligent Vehicles, USA, Dearborn, October 2000.
- [BIERMAN 77] Gerald J. Bierman: *Factorization Methods for Discrete Sequential Estimation*, Mathematics in Science and Engineering, Volume 128, Academic Press, Inc., London, 1977.
- [BISHOP 00] R. Bishop: *A Survey of Intelligent Vehicle Applications Worldwide*, IEEE Symposium on Intelligent Vehicles, USA, Dearborn, October 2000.
- [BROCKHAUS 98] *Brockhaus - Die Enzyklopädie*, F.A.Brockhaus 20.Auflage, Band 14, Leipzig/Mannheim, 1998.
- [BURT 83] P. J. Burt, E. Adelson: *The Laplacian Pyramid as Compact Image Code*, IEEE Transactions on Communications, Vol COM-31, No.4, p. 532-540, April 1983
- [BURT 95] P. J. Burt, L. Wixson, G. Salgian: *Electronically Directed "Focal" Stereo*, ICCV 95, Seite 94-101, 1995.
- [COOMBS 00] D. Coombs, K. Murphy, A. Lacaze, S. Legowik: *Driving Autonomously Offroad up to 35 km/h*, IEEE Symposium on Intelligent Vehicles, USA, Dearborn, October 2000.

- [DELLAERT 98] F. Dellaert, D. Pomerleau, C. Thorpe: *Model-Based Car Tracking Integrated with a Road-Follower*, International Conference on Robotics and Automation, May 1998.
- [DELLAERT&THORPE 97] F. Dellaert, C. Thorpe: *Robust Car tracking using Kalman filtering and Bayesian templates*, IEEE Conference on Intelligent Transportation Systems, Boston, November 1997
- [DICKMANN 94] Dirk Dickmanns: *KRONOS Benutzerhandbuch*, Universität der Bundeswehr, INF-2, Neubiberg, 1994.
- [DICKMANN 97] Dirk Dickmanns: *Rahmensystem für visuelle Wahrnehmung veränderlicher Szenen durch Computer*, Dissertation, Universität der Bundeswehr München, Fakultät für Informatik, Neubiberg, 1997.
- [DICKMANN 87] E.D. Dickmanns: *4-D Szenenanalyse mit integralen raum/zeitlichen Modellen*, In E. Paulus (Hrsg.): *Mustererkennung 1987*, Informatik Fachberichte 149, Seite 257-271, Springer-Verlag, 1987.
- [DICKMANN 88] E.D. Dickmanns: *Dynamic Computer Vision for Mobile Robot Control*, Proc. Int. Symposium on Robot Control, Sydney, Nov. 1988.
- [DICKMANN 95] E.D. Dickmanns: *Road Vehicle Eyes for High Precision Navigation*, 3rd international Workshop on HIGH PRECISION NAVIGATION, Stuttgart, April 1995.
- [DICKMANN&GRAEFE 88] E.D. Dickmanns, V. Graefe: *Dynamic Monocular Machine Vision*, Machine Vision and Applications, Springer International, Vol. 1, 1988 , pp 223-240.
- [DONGES 78] E. Donges: *Der Fahrsimulator des Forschungsinstitutes für Anthropotechnik*, Forschungsbericht Nr. 41, FGAN e.V., Forschungsinstitut für Anthropotechnik, Wachtberg-Werthhoven, Juli 1978.
- [DUDEN 01] *Duden - Das Fremdwörterbuch*, neu bearbeitete und erweiterte 7. Auflage, Herausgegeben von der Dudenredaktion, Dudenverlag, Mannheim, 2001.
- [ENKELMANN 02] M. Rüder, W. Enkelmann, R. Garnitz: *Highway Lane Change Assistant*, IEEE Symposium on Intelligent Vehicles, France, Versailles, June 2002.
- [FRANKE 97] Z. Zomotor, U. Franke: *Sensor Fusion for Improved Vision Based Lane Recognition and Object Tracking with Range-Finders*, Procs. IEEE Conference on Intelligent Transportation Systems Conference, Boston, November 1997.
- [FÜRST 98-01] S. Fürst: *CRONOS Programmbericht*, UniBW/LRT/WE13/PB95-1, letzte Version 2001.01.163, Universität der Bundeswehr, Neubiberg, 1998-2001.

- [GERN 00] A. Gern, U. Franke: *Advanced Lane Recognition - Fusing Vision and Radar*, IEEE Symposium on Intelligent Vehicles, USA, Dearborn, October 2000.
- [GERN 02] A. Gern, R. Moebus, U. Franke: *Vision-Based Lane Recognition under Adverse Weather Conditions Using Optical Flow*, IEEE Symposium on Intelligent Vehicles, France, Versailles, June 2002.
- [GREGOR 98] R. Gregor: *Off-line Learning from Perception for Improving Mission Performance of Autonomous Vehicles*, IEEE International Conference on Intelligent Vehicles, Stuttgart, 1998.
- [GREGOR 02] R. Gregor: *Fähigkeiten zur Missionsdurchführung und Landmarkennavigation*, Dissertation, Universität der Bundeswehr München, Fakultät für Luft- und Raumfahrttechnik, Neubiberg, 2002.
- [HOFMANN ET AL. 00] U. Hofmann, A. Rieder, E.D. Dickmanns: *EMS-Vision: Application to Hybrid Adaptive Cruise Control*, IEEE Symposium on Intelligent Vehicles, USA, Dearborn, October 2000.
- [HOFMANN ET AL. 01] U. Hofmann, A. Rieder, E.D. Dickmanns: *Radar and Vision Data Fusion for Hybrid Adaptive Cruise Control on Highways*, In Schiele et al. (Ed.), *Procs. of Computer Vision Systems*, Vancouver(Canada) 2001, Lecture Notes in Computer Science 2095, S. 125-138, Springer-Verlag, 2001.
- [HOFMANN ET AL. 03] U. Hofmann, A. Rieder, E.D. Dickmanns: *Radar and Vision Data Fusion for Hybrid Adaptive Cruise Control on Highways*, International Journal "Machine Vision and Applications" (2003), Volume 14(1), Seite 42-49, Springer-Verlag, 2003.
- [HOFMANN&SIEDERSBERGER 03] U. Hofmann, K.-H. Siedersberger: *Stereo and photometric image sequence interpretation for detecting negative obstacles using active gaze control and performing an autonomous jink*, SPIE-AeroSense, Proc. "Unmanned Ground Vehicle Technology", Orlando, April 2003.
- [HEIKKILÄ 97] J. Heikkilä, O. Silvén: *A four-step camera calibration procedure with implicit image correction*, IEEE Conf. on Computer Vision and Pattern Recognition, San Juan, Puerto Rico, 1997.
- [KALMAN 60] R. E. Kalman: *A New Approach to Linear Filtering and Prediction Problems*, Transaction of ASME 82 - Journal of Basic Engineering, Seiten 35-45, März 1960.
- [KASTRINAKI 03] V. Kastrinaki, M. Zervakis, K. Kalaitzakis: *A Survey of video processing techniques for traffic applications*, Image and Vision Computing, Volume 21, pp. 359-381, Elsevier Verlag, 2003

- [KUHNER 88] K.-D. Kuhnert: *Zur Echtzeit-Bildfolgenanalyse mit Vorwissen*, Dissertation, Universität der Bundeswehr München, Fakultät für Luft- und Raumfahrttechnik, Institut für Meßtechnik, Neubiberg, 1988.
- [KÜHNLE 91] A. Kuehnle: *Symmetry-based recognition for vehicle rears*, in Pattern Recognition Letter. volume 12 (1991), pp. 249-258, Number 4
- [LANGER 96] D. Langer, T. Jochem: *Fusing Radar and Vision for Detecting, Classifying and Avoiding Roadway Obstacles*, IEEE Symposium on Intelligent Vehicles, Tokyo, September 1996.
- [LN9300/90] Normenstelle Luftfahrt im DIN e.V.: *LN9300 Luftfahrtnorm*, Beuth-Verlags GmbH, Köln, 1990.
- [LÜTZELER 02] M. Lützel: *Fahrbahnerkennung zum Manövrieren auf Wegenetzen mit aktivem Sehen*, Dissertation, Universität der Bundeswehr München, Fakultät für Luft- und Raumfahrttechnik, Institut für Systemdynamik und Flugmechanik, Fortschrittsberichte VDI, Reihe 12 Verkehrstechnik/Fahrzeugtechnik, Nr.493, VDI-Verlag, Düsseldorf, 2002.
- [MATTHIES 03] L. Matthies, A. Rankin: *Negative Obstacle Detection by Thermal Signature*, Internet <http://team.caltech.edu/members/index.html>, 2003.
- [MAYBECK 79] Peter. S. Maybeck: *Stochastic Models, Estimation, and Control*, Volume 1(2), Academic Press, Inc., San Diego, 1979.
- [MEISSNER 82] H.G. Meissner: *Steuerung dynamischer Systeme aufgrund bildhafter Informationen*, Dissertation, Universität der Bundeswehr München, Fakultät für Luft- und Raumfahrttechnik, Institut für Systemdynamik und Flugmechanik, 1982.
- [MEISSNER&DICKMANN 83] H.G. Meissner, E.D. Dickmanns: *Control of an Unstable Plant by Computer Vision*, In T.S. Huang (ed): *Image Sequence Processing and Dynamic Scene Analysis*, pp 532-448, Springer Verlag, Berlin, 1983.
- [METZINGER 01] Thomas Metzinger: *Postbiotisches Bewusstsein: Wie man ein künstliches Subjekt baut und warum wir es nicht tun sollten*, In Heinz Nixdorf MuseumsForum (Hrsg.), *Computer. Gehirn. Was kann der Mensch? Was können die Computer?* Begleitpublikation zur Sonderausstellung "Computer.Gehirn" im Heinz Nixdorf MuseumsForum, Paderborn 2001.
- [MIETZEL 02] Internet Supplement zu Gerd Mietzel: *Wege in die Psychologie*, 9.Auflage, Klett-Cotta-Verlag, Oktober 2002, ISBN: 3608916601, <http://www.regiosurf.net/supplement/wahrn/wahrnh.htm>
- [MITSCHKE 72] M. Mitschke: *Dynamik der Kraftfahrzeuge*, Springer Verlag, Berlin, Heidelberg, New York, 1972.

- [MYSLIWETZ 90] B. Mysliwetz: *Parallelrechner-basierte Bildfolgen-Interpretation zur autonomen Fahrzeugführung*, Dissertation, Universität der Bundeswehr München, Fakultät für Luft- und Raumfahrttechnik, Institut für Systemdynamik und Flugmechanik, Neubiberg, 1990.
- [OTTO 88] Klaus-Dieter Otto: *Einführung in die lineare Schätztheorie*, Universität der Bundeswehr München, Interner Bericht LRT/WE 13/88-1, 1988.
- [PELLKOEFER 03] M. Pellkofer: *Verhaltensentscheidung für autonome Fahrzeuge mit Blickrichtungssteuerung*, Dissertation, Universität der Bundeswehr München, Fakultät für Luft- und Raumfahrttechnik, Neubiberg, 2003.
- [POMERLEAU 95] D. Pomerleau: RALPH: *Rapidly Adapting Lateral Position Handler*, IEEE Symposium on Intelligent Vehicles, Detroit, September 1995.
- [POMERLEAU 96] D. Pomerleau, T. Jochem: *A Rapidly Adapting Machine Vision for Automated Vehicle Steering*, IEEE Expert: Special Issue on Intelligent Systems and their Applications, pp. 19-27, April 1996.
- [RIEDER 00] André Rieder: *Fahrzeuge Sehen*, Dissertation, Universität der Bundeswehr München, Fakultät für Luft- und Raumfahrttechnik, Neubiberg, 2000.
- [RIEDER 02] A. Rieder, B. Southall, G. Salgian, R. Mandelbaum, H. Herman, P. Rander, T. Stentz: *Stereo Perception on an Off-Road Vehicle*, IEEE Symposium on Intelligent Vehicles, France, Versailles, June 2002.
- [SCHALKOFF 89] R.J. Schalkoff: *Digital image processing and computer vision*, John Wiley and Sons, Inc., New York, 1989.
- [SCHWARTZ 99] I. Schwartz: *PRIMUS: an autonomous driving robot*, Proceedings of the SPIE Vol. 3693 AeroSense Session on Unmanned Ground Vehicle Technologie, Orlando (Florida), April 1999.
- [SIEDERSBERGER 01] K.-H. Siedersberger, M. Pellkofer, M. Lützelner, E.D. Dickmanns, A. Rieder, R. Mandelbaum, L. Bogoni: *Combining EMS-Vision and Horopter Stereo for Obstacle Avoidance of Autonomous Vehicles*, In Schiele et al. (Ed.), *Procs. of Computer Vision Systems*, Vancouver (Canada) 2001, Lecture Notes in Computer Science 2095, S. 139-156, Springer-Verlag, 2001.
- [SIEDERSBERGER 01] K.-H. Siedersberger: *Koordinaten und Koordinatensysteme für EMS-Vision*, Interner Fachbericht, Universität der Bundeswehr München, UniBwM/LRT/WE13/FB 01-1, Neubiberg, 2001.
- [SIEDERSBERGER 03] K.-H. Siedersberger: *Komponenten zur automatischen Fahrzeugführung in sehenden (semi-)autonomen Fahrzeugen*, Dissertation, Universität der Bundeswehr München, Fakultät für Luft- und Raumfahrttechnik, Neubiberg, 2003.

- [SMITH 97] S.M. Smith, J.M. Brady: *SUSAN - a new approach to low level image processing*, International Journal of Computer Vision, 23(1):45-78, May 1997.
- [THOMANEK 96] F. Thomanek: *Visuelle Erkennung und Zustandsschätzung von mehreren Straßenfahrzeugen zur autonomen Fahrzeugführung*, Dissertation, Universität der Bundeswehr München, Fakultät für Luft- und Raumfahrttechnik, Neubiberg, 1996.
- [TSAI 87] R.Y. Tsai: *A versatile camera calibration technique for high-accuracy 3-D machine vision metrology using off-the-shelf TV cameras and lenses*, IEEE Journal of Robotics and Automation, 3(4):Seite 323-344, August 1987.
- [VAN DER WAL 00] G. van der Wal, M. Hansen, M. Piacentino: *The Acadia Vision Processor*, IEEE Proc. of International Workshop on Computer Architecture for Machine Perception, Padua, Italy, 2000.
- [WELCH 03] Greg Welch, Gary Bishop: *An Introduction to the Kalman Filter*, Technical Report TR 95-041, University of North Carolina at Chapel Hill, updated May 2003.
- [WÜNSCHE 87] H.-J. Wünsche: *Bewegungssteuerung durch Rechnersehen*, Dissertation, Universität der Bundeswehr München, Springer Verlag, Berlin, 1987.