

Neuronale Informationsverarbeitung für Fahrerassistenzsysteme

Uwe Handmann

aus Backnang

Dissertation

zur Erlangung des Grades

Doktor-Ingenieur

an der Fakultät

für Elektrotechnik und Informationstechnik

der Ruhr-Universität Bochum

Bochum 2000

Rückblickend auf die Jahre, die ich am Institut für Neuroinformatik verbracht habe, kann ich mit Zufriedenheit feststellen, daß es schöne Jahre waren. Die interdisziplinäre Ausrichtung des Instituts und die Arbeit in der Arbeitsgruppe *Computer Vision* bzw. *Pattern Recognition and Scene Analysis* gestalteten den Arbeitsalltag abwechslungsreich und interessant.

Herrn Prof. Dr. Werner von Seelen, dem Betreuer meiner Arbeit, ist es zu verdanken, daß neben dem fachlich hohen Niveau am Institut auch menschliche Qualitäten einen großen Stellenwert haben. Seine Art mit Menschen umzugehen wird mir immer ein Vorbild sein. Ergänzt werden diese idealen Bedingungen durch eine an Perfektion kaum zu übertreffende Infrastruktur. Im Sekretariat schaffen es Heide Berz und Katharina Weißmann, daß trotz ihrer vielfältigen Aufgaben nie Chaos aufkommt. Michael Neef und Arno Berg ist es zu verdanken, daß das bereitstehende Rechnernetz mit schier unglaublicher Stabilität läuft.

Den aktuellen und ehemaligen Mitgliedern meiner Arbeitsgruppe danke ich für die schöne Zeit, die ich hier verbringen durfte. Ich habe viele Freunde gewonnen, die ich nicht vergessen werde. Mein besonderer Dank gilt dabei dem Team der Korrekturleser Jens Gayko, Christian Goerick, Thomas Kalinke, Iris Leefken, Gesa Lorenz, Peter Stagge, Axel Steinhage, Jan Wiemer und Susanne Winter. Nur durch ihre Anmerkungen und Anregungen konnte diese Arbeit so gelingen, wie sie vorliegt.

Meinen Zimmerkollegen Thomas Kalinke und Carsten Winkel danke ich für die Unterstützung nicht nur in fachlichen Fragen. Sie standen mir immer mit Rat und Tat zur Seite. Ein besonderer Dank geht an meinen *Hiwi* Susanne. Selten habe ich ein so hohes Maß an Gewissenhaftigkeit und Zuverlässigkeit erlebt, wie es bei ihr der Fall ist. Selbst die unmöglichsten Ideen meinerseits setzte sie kurzerhand als *shared object* für IDL um.

Herrn Prof. Dr. Ermert danke ich für die unkomplizierte Übernahme des Koreferats meiner Arbeit.

Meinen Freunden Katja Trautmann, Mareike Stiller und Michael Stiller danke ich für den Zuspruch in der *heißen Phase* und die immer zur rechten Zeit geäußerten Einladungen zum Essen.

Meinen Eltern danke ich, daß sie mir immer freie Hand bei der Planung meines Lebenswegs gelassen haben und durch ihre Unterstützung mit zum Gelingen dieser Arbeit beigetragen haben.

Steffi und Finn habe ich so viel zu verdanken, daß jede Aufzählung unvollständig wäre.

Allen gilt mein herzlichster Dank !!!

*Uwe Handmann
Bochum, im Mai 2000*

Inhaltsverzeichnis

1	Einleitung	1
2	Assistenzsysteme für den Straßenverkehr	3
2.1	Fahrerassistenzsysteme	4
2.2	Erkennung von Objekten im Straßenverkehr	5
2.2.1	Sensoren	5
2.2.2	Videobildauswertung	7
2.3	Sensordatenfusion	8
2.4	Navigation im Straßenverkehr	9
2.5	Verhaltenssteuerung	10
2.6	Einordnung der Arbeit	11
2.7	Zusammenfassung	12
3	Neuronale Informationsverarbeitung	13
3.1	Neurone im Neocortex	14
3.2	Abbildungseigenschaften im Neocortex	15
3.3	Funktionelle Einheiten und hierarchische Anordnung von Neuronen im Neocortex	16
3.4	Makroskopisches Verhalten komplexer biologischer Organismen	16
3.5	Neuronale Informationsverarbeitung in technischen Systemen	17
3.6	Zusammenfassung	17
4	Analyse von Segmentierungsverfahren mittels neuronaler Karten	19
4.1	Neuronale Karten	20
4.1.1	Selbstorganisierende neuronale Karten	21
4.1.2	Lernverfahren	22
4.1.3	Kohonenkarten als Klassifikator	23
4.2	Analyse von Segmentierungsverfahren	25
4.2.1	Selbstorganisierender Prozeß	26
4.2.2	Eingangsdatenrepräsentation	27
4.2.3	Analyseergebnisse	28
4.2.4	Diskussion	32
4.3	Zusammenfassung	34

5	Neuronale Sensordatenfusion	35
5.1	Vorwärtsgerichtete neuronale Netze	36
5.1.1	Multilayer Perceptron	37
5.1.2	Lernverfahren	39
5.1.3	Datengetriebener Entwurf	40
5.1.4	Generalisierungsfähigkeit	42
5.1.5	Relevanz einzelner Eingänge	43
5.2	Fusion der Sensordaten	44
5.2.1	Fusion auf Daten- und Merkmalsebene	44
5.2.2	Strukturuntersuchung des Multilayer Perceptron	45
5.2.3	Fusion	46
5.2.4	Relevanz einzelner Komponenten	47
5.3	Ergebnisse	48
5.3.1	Segmentierungsergebnisse	48
5.3.2	Flexible Erweiterbarkeit des Verfahrens	50
5.3.3	Diskussion	53
5.4	Zusammenfassung	54
6	Neuronale Bewegungssteuerung	55
6.1	Neuronale Felder	56
6.1.1	Amarifeld	56
6.1.2	Eigenschaften	57
6.1.3	Stabilitätsbetrachtung	62
6.2	Eingangsdatenmaterial	64
6.3	Bewegungssteuerung	65
6.3.1	Vorverarbeitung	65
6.3.2	Steuerung des Lenkwinkels	68
6.3.3	Steuerung der Geschwindigkeit	71
6.4	Ergebnisse	73
6.4.1	Basisverhalten	74
6.4.2	Diskussion	83
6.5	Zusammenfassung	83
7	System zur Fahrerassistenz	85
7.1	Architektur	85
7.2	Verfahrensintegration	87
7.3	Diskussion	95
7.4	Zusammenfassung	95
8	Zusammenfassung und Ausblick	97
A	Mathematischer Anhang	99
A.1	Stetig differenzierbare Funktionen	99
A.1.1	Sigmoide $\xi(v)$	99
A.1.2	Tangens Hyperbolicus $\tanh(v)$	99

A.2 Backpropagation-Algorithmus	100
B Simulationsumgebung	103
C Notation	105
Literaturverzeichnis	109

Kapitel 1

Einleitung

Die vorliegende Arbeit liefert neue Verfahren zur Realisierung von Fahrerassistenzsystemen im Straßenverkehr. Fahrerassistenzsysteme unterstützen den Fahrer eines Kraftfahrzeugs, um eine erhöhte Sicherheit und hohen Fahrkomfort zu erreichen. Die Sicherheit soll sich hierbei für den Fahrer selbst sowie für die Umwelt wesentlich verbessern.

Die Entwicklung von robusten Fahrerassistenzsystemen stellt einen Forschungsschwerpunkt in der Automobilindustrie und den öffentlichen Einrichtungen dar. Bis zum jetzigen Zeitpunkt wurden mehrere Verfahren zur Erkennung, Klassifikation und Verfolgung anderer Verkehrsteilnehmer vorgestellt. Auf diese Verfahren baut die vorliegende Arbeit auf.

Da sich sowohl das Trägerfahrzeug des Assistenzsystems als auch andere Verkehrsteilnehmer (Objekte), die durch Sensoren erfaßt werden, in der Umwelt bewegen, ergeben sich bei Fahrerassistenzsystemen vielfältige Anforderungen.

In dieser Arbeit wird ein neues Verfahren vorgestellt, welches eine robuste und flexible Sensordatenfusion zur Segmentierung von Videobilddaten durchführt. Ein zweites Verfahren, das im Rahmen dieser Arbeit entwickelt wurde, erlaubt die Bewegungssteuerung eines autonomen Kraftfahrzeugs auf der Basis der Fusionsergebnisse.

Segmentierung

Zur Durchführung einer Objekterkennung müssen die Sensordaten verarbeitet werden. Ein wesentlicher Bestandteil bei der Objekterkennung ist die zielgerichtete Steuerung des Aufmerksamkeitsbereichs. Insbesondere in technischen Anwendungen zur Videobilddatenverarbeitung für Fahrerassistenzsysteme ist die Anforderung an eine robuste Analyse von Bildern die Grundlage für eine algorithmische Implementierung von sogenannten Segmentierungsverfahren für Videobilddaten. Ziel der Segmentierung ist es, die große Menge der Videobilddaten auf die zur Lösung der Aufgabe notwendigen Bildbereiche zu konzentrieren, um somit eine zeitlich effiziente Verarbeitung zu gewährleisten.

Die Anforderung, bei der Analyse von Straßenverkehrsszenen eine robuste Segmentierung zu erreichen und flexibel mit veränderten Randbedingungen zu operieren, ist die Grundlage für die Fusion von Segmentierungsverfahren unterschiedlicher Ausprägung.

In dieser Arbeit werden verschiedene Verfahren analysiert, die eine Segmentierung von Grauwertbilddaten erlauben, um die Eignung einzelner Komponenten für die Segmentierungsaufgabe bei Straßenverkehrsszenen zu beurteilen. Eine Fusion relevanter Komponenten wird durchge-

führt. Mit Hilfe neuronaler Informationsverarbeitung wird eine Erhöhung der Segmentierungsgüte erreicht. Die Fusion redundanter Komponenten erlaubt dabei die Realisierung eines robusten Segmentierungsverfahrens, welches beim Ausfall einzelner Sensordaten auch ein gutes Segmentierungsergebnis liefert. Weiterhin wird auf Flexibilität geachtet. Eine einfache Hinzunahme neuer Komponenten des gleichen oder eines anderen Sensortyps sowie eine Adaption der Strukturparameter bei veränderter Güte der Sensordaten wird ermöglicht. Die vorgenannten Punkte stellen wesentliche Neuerungen und Verbesserungen im Vergleich zu bisher realisierten Verfahren dar.

Bewegungssteuerung

Fahrerassistenzsysteme werden eingesetzt, um dem Fahrer eines Kraftfahrzeugs Handlungsabläufe abzunehmen. Diese Handlungsabläufe werden definiert durch eine Aufgabenstellung, die vom Fahrer an das Fahrerassistenzsystem übergeben oder systembedingt gelöst wird.

Bei komplexen Fahrerassistenzsystemen ist an eine autonome Navigation im Straßenverkehr gedacht. Es wird ein neues Verfahren vorgestellt, welches eine Bewegungssteuerung eines autonomen Fahrzeugs durchführen kann. Es werden dabei der Lenkwinkel und die Geschwindigkeit eines Kraftfahrzeugs beeinflusst. Für diese Aufgabe wird ein dynamischer Ansatz aus dem Bereich der neuronalen Informationsverarbeitung gewählt, der bei der Verhaltenssteuerung eine kontinuierliche Ausgabe der Steuergrößen des Fahrzeugs erlaubt. Für den Fahrtverlauf relevante Attribute auf unterschiedlichem Abstraktionsniveau können dabei einfach (additiv) in ein neuronales Feld eingekoppelt werden. Die Extraktion der relevanten Attribute für die Bewegungssteuerung basiert auf dem realisierten Segmentierungsprozeß.

Die Tragfähigkeit der entwickelten Verfahren wird durch die Integration in ein Gesamtsystem gezeigt.

Gliederung der Arbeit

In Kapitel 2 werden Assistenzsysteme im Straßenverkehr vorgestellt und die vorliegende Arbeit in den Bereich der Fahrerassistenzsysteme eingeordnet. Kapitel 3 gibt einen Überblick über die biologischen Grundlagen neuronaler Informationsverarbeitung und ordnet die verwendeten Ansätze in diesen Kontext ein. Kapitel 4 beschreibt ein Analyseverfahren, welches zur Auswahl von Komponenten für den Sensordatenfusionsprozeß dient. In Kapitel 5 wird das realisierte Verfahren zur robusten und flexiblen Segmentierung von Straßenverkehrsszenen vorgestellt. In Kapitel 6 wird das entwickelte Verfahren zur Bewegungssteuerung von Kraftfahrzeugen mittels neuronaler Felder diskutiert. In Kapitel 7 werden die in dieser Arbeit vorgestellten Verfahren in ein Gesamtsystem integriert und die Tragfähigkeit der Verfahren im Zusammenspiel gezeigt. Die Arbeit schließt mit einer Zusammenfassung und einem Ausblick in Kapitel 8.

Kapitel 2

Assistenzsysteme für den Straßenverkehr

Betrachtet man den Straßenverkehr in hochindustrialisierten Ländern, so läßt sich feststellen, daß durch die veränderten Lebensgewohnheiten der Menschen und die gesellschaftlichen Anforderungen eine stark ansteigende Verkehrsdichte zu verzeichnen ist. Vielfach ist eine erhöhte Mobilität der Menschen gewünscht oder gefordert. Eine starke Zunahme des öffentlichen Nah- und Fernverkehrs sowie ein Anstieg des Individualverkehrs sind die Folge. Die gleichzeitig steigenden Anforderungen an die Verkehrssicherheit und den Umweltschutz durch die Gesellschaft und die Gesetzgeber, der zunehmende Technisierungsgrad der Kraftfahrzeuge und der Verkehrssteuerungssysteme sowie ein erhöhter Komfortanspruch der Nutzer der Verkehrssysteme motivieren den Entwurf von Assistenzsystemen im Straßenverkehr.

Die Entwicklung von Assistenzsystemen für die Steuerung des Verkehrsflusses und die Führung von Kraftfahrzeugen im Straßenverkehr ist deshalb ein Forschungsschwerpunkt der Automobilindustrie und der öffentlichen Einrichtungen. Daher wurden in den letzten Jahren verschiedene Forschungsprogramme in Europa, Amerika und Asien aufgelegt. Beispiele hierfür sind *PROMETHEUS* (programme for a european traffic with highest efficiency and unprecedented safety [96, 97]), *MoTiV* (Mobilität und Transport im intermodalen Verkehr [80]), *IVI* (intelligent vehicle initiative [69]), *ITS* (intelligent transportation systems [57]), *SSVS* (super smart vehicle system [85]) und *ASV* (advanced safety vehicle [85]). Im Rahmen dieser Forschungsprogramme wurden und werden die Erhöhung der Umweltverträglichkeit und der Sicherheit sowie eine Effizienzsteigerung im Straßenverkehr angestrebt [118, 135, 79]. Projekte, in deren Rahmen Systeme entwickelt werden, die Lösungsvorschläge für einzelne Problemstellungen anbieten, werden gezielt gefördert. Dabei soll unter Beibehaltung der gesellschaftlich geforderten Mobilität, eine Verbesserung der genannten Punkte erreicht und dem gestiegenen Komfortanspruch der Menschen Rechnung getragen werden.

Die eingesetzten Systeme können sowohl stationär sein als auch mobil Anwendung finden. Stationäre Systeme sind beispielsweise Verkehrsflußanalysensysteme [65, 13, 58], die Parameter für eine intelligente Verkehrsflußsteuerung [79, 82, 83] extrahieren oder Systeme zur Optimierung von Transportsystemen [109, 118]. Mobile Assistenzsysteme sind häufig Fahrerassistenzsysteme [18, 94, 7]. Diese Systeme sind direkt in ein Fahrzeug eingebaut und unterstützen den Fahrer während der Fahrt.

Die vorliegende Arbeit liefert neue Beiträge zu Fahrerassistenzsystemen. In Abschnitt 2.1 wird der Nutzen von Assistenzsystemen in Kraftfahrzeugen motiviert. Anschließend wird in Abschnitt 2.2 das Problem der Erkennung von Objekten im Straßenverkehr beleuchtet und die

Notwendigkeit von komplexen Erkennungssystemen beschrieben. In Abschnitt 2.3 wird ein Lösungsweg zur Erkennung von Objekten in Umweltszenarien aufgezeigt. Die Anforderungen an Assistenzsysteme, die autonom in die Aktuatorik eines Kraftfahrzeugs zur Navigation eingreifen, werden in Abschnitt 2.4 diskutiert und Abschnitt 2.5 beinhaltet Vorschläge für eine mögliche Verhaltenssteuerung von Kraftfahrzeugen zur Navigation. In Abschnitt 2.6 werden die in dieser Arbeit entwickelten Verfahren in den Kontext der Objekterkennung und der Verhaltenssteuerung eingeordnet. Das Kapitel 2 schließt mit einer Zusammenfassung in Abschnitt 2.7.

2.1 Fahrerassistenzsysteme

Fahrerassistenzsysteme unterstützen den Fahrer eines Kraftfahrzeugs, um eine erhöhte Sicherheit und hohen Fahrkomfort zu erreichen. Die Sicherheit soll sich hierbei für den Fahrer selbst sowie für die Umwelt wesentlich verbessern. Beispielsweise können dem Fahrer eines Kraftfahrzeugs Handlungsabläufe durch ein Assistenzsystem abgenommen werden, um dessen Aufmerksamkeit im Straßenverkehr zu erhöhen. Dabei läßt sich dessen Reaktionszeit auf ein nicht vorhersehbares Ereignis häufig verringern und so das Unfallrisiko reduzieren. So stellte sich bei einer Untersuchung von Unfallprotokollen heraus [90], daß bei einer um eine halbe Sekunde reduzierten Reaktionszeit 60% der Auffahrunfälle, 50% der Kreuzungsunfälle und 30% der Gegenverkehrsunfälle hätten vermieden werden können. Kann die Reaktionszeit um eine Sekunde reduziert werden, könnten 90% der Auffahr- und Kreuzungsunfälle und über 60% der Gegenverkehrsunfälle verhindert werden.

Fahrerassistenzsysteme können sehr einfacher Natur sein, wie beispielsweise bei Navigationssystemen für die Wegplanung von einem Ausgangspunkt zu einem Zielort mit Hilfe von GPS¹ und elektronischen Karten. Solche Systeme sind im Handel erhältlich. Andererseits kann die Anforderung an ein System auch die Assistenz bei der Hindernisvermeidung beinhalten. Über Sensoren (zum Beispiel Videokamerasysteme), wie in Abbildung 2.1 dargestellt, werden diesem

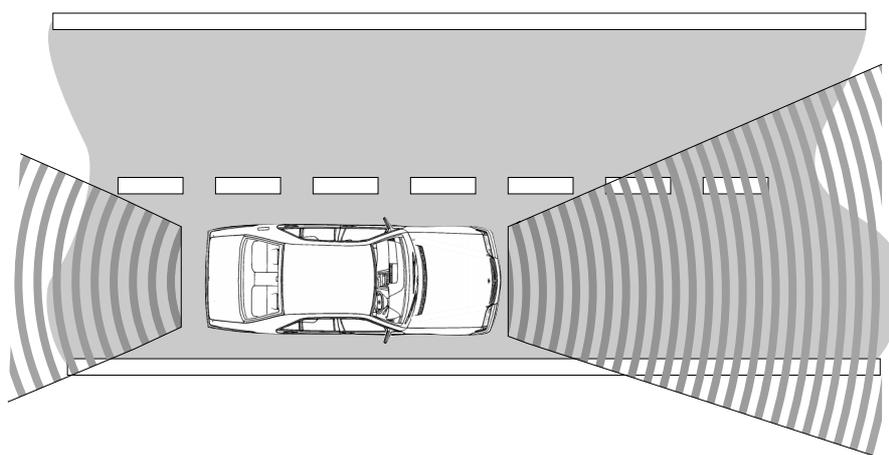


Abbildung 2.1: Trägerfahrzeug mit Sensoren, die den Fahrraum vor und hinter dem Kraftfahrzeug abtasten und dem Assistenzsystem Daten bereitstellen.

¹GPS – global positioning system

System Daten bereitgestellt, um die Umgebung zu analysieren und dem Fahrer eines Fahrzeugs eine der Aufgabenstellung entsprechende Assistenz anzubieten.

Da sich sowohl das Trägerfahrzeug der Sensoren als auch Objekte, die durch die Sensoren erfaßt werden, in der Umwelt bewegen, ergeben sich bei diesen Systemen vielfältige Anforderungen. Einerseits stellen die Sensoren oft fehlerhafte oder unvollständige Sensordaten zur Verfügung und andererseits sind unterschiedliche Sensordaten oft widersprüchlich. Um eine Verringerung der Störanfälligkeit bei der Objekterkennung zu erreichen, müssen diese Daten kombiniert werden.

In Abschnitt 2.2 wird ein Überblick der in der Literatur verwendeten Verfahren zur Erkennung von Objekten im Straßenverkehr gegeben. In Abschnitt 2.3 wird die Notwendigkeit einer Sensordatenfusion abgeleitet.

2.2 Erkennung von Objekten im Straßenverkehr

Menschen analysieren ihre Umwelt mit verschiedenen Sinnesorganen. Visuelle Information wird durch die Augen aufgenommen und über die Retina, die Ganglienzellen und die Sehnerven im visuellen Cortex weiterverarbeitet. Einen Überblick über den Aufbau des menschlichen Auges geben [50, 116, 78]. Betrachtet man die menschliche Verarbeitung visueller Reize, so kann durch psychophysikalische Untersuchungen festgestellt werden [116, 103, 71], daß einzelne Merkmale für den Prozeß der Erkennung und Verarbeitung von Objektinformationen eine besondere Relevanz haben. Beispielsweise kann stark gesättigte Farbe eine Signalwirkung haben (Verkehrsschilder), oder die Form die Detektion von Objekten erleichtern [50, 71]. Die durch die Paralaxe bedingten Unterschiede zwischen den Bildern der Augen werden als ein Merkmal für die Tiefenwahrnehmung genutzt [103, 71]. Weitere Merkmale, die von Menschen bei der visuellen Verarbeitung genutzt werden, sind der optische Fluß, Konturen und die Textur von Bildbereichen [103, 71].

Um Objekte in der Umwelt mit technischen Systemen erkennen zu können, sind Sensoren notwendig, die die Umwelt erfassen. Zur Durchführung einer Objekterkennung müssen die Sensordaten verarbeitet werden. Im folgenden wird auf die Art der möglichen Sensoren eingegangen (Abschnitt 2.2.1). Am Beispiel der Verarbeitung von Videobilddaten werden Umsetzungen von Verfahren diskutiert, die Beiträge zur Objekterkennung in der Umwelt liefern (Abschnitt 2.2.2).

2.2.1 Sensoren

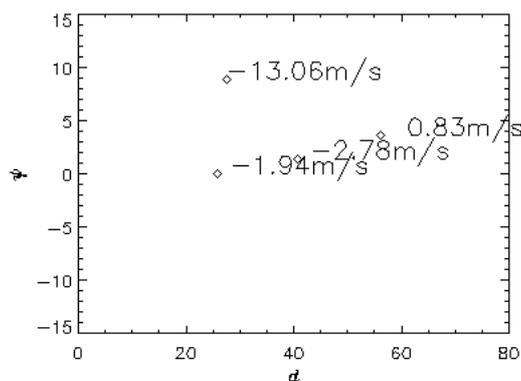
Bei Systemen zur Objekterkennung werden verschiedene Sensoren eingesetzt. Es gibt Analysensysteme, welche aktive Sensoren verwenden, wie zum Beispiel Ultraschallsensoren [131], RADAR²-Sensoren [49, 77, 12, 25, 125, 44] oder LIDAR³-Sensoren [117], um eine Objekterkennung zu realisieren. Bei diesen Sensoren werden Schallwellen oder elektromagnetische Wellen definierter Wellenlängen vom Sensor gesendet und das reflektierte Signal entsprechend seiner Eigenschaften ausgewertet. RADAR-Sensoren operieren dabei mit Mikrowellen und

²RADAR – radio detection and ranging

³LIDAR – light detection and ranging

LIDAR-Sensoren mit UV⁴-, IR⁵-, oder Wellen aus dem sichtbaren Bereich des elektromagnetischen Wellenspektrums. In Abbildung 2.2a wird dies am Beispiel eines RADAR-Sensors verdeutlicht. Dargestellt sind die Sensordaten eines fünfstrahligen Radarsensors mit einem Öffnungswinkel von 20°, der vorne zentral in einem Kraftfahrzeug (Trägerfahrzeug) angebracht ist. Die Sensordaten beschreiben die Lage von vier Objekten mit Distanz d , Winkelposition ψ und Relativgeschwindigkeit. Durch die Zeitdauer von der Aussendung eines Radarimpulses bis zum Empfang der rückgestreuten Welle lassen sich beispielsweise die Distanz und der Winkel zwischen Sensor und Objekt berechnen. Unter Ausnutzung des Doppler-Effekts [67] bei gesendeten und rückgestreuten elektromagnetischen Wellen kann die Relativgeschwindigkeit zwischen Sensor und Objekt bestimmt werden.

Weiterhin werden vielfach Ansätze verfolgt, welche Bildfolgen auswerten, die passive IR-Kameras oder Videokameras liefern. IR-Kameras [17, 133, 104, 98] verarbeiten Wärmestrahlung, die von Objekten emittiert wird. Videokameras verarbeiten emittiertes oder von Objekten reflektiertes Licht sichtbarer Wellenlänge und ordnen dieses einem Helligkeitswert (Abbildung 2.2b) oder einer Farbe zu. Der Vorteil dieser Sensoren ist, daß sie bildgebend sind. Weiterhin ist es möglich, Objekte und Szenenwissen aus einer Szene zu extrahieren, die mit aktiven Sensoren nicht oder nur sehr schwer erkannt werden können. Fahrbahnmarkierung einer Straße, Fußgänger oder Baustellen, wie in Abbildung 2.2b dargestellt, können nur mit Kameras erkannt werden [7, 132, 14]. Das motiviert den Einsatz von Kamerasensoren bei der Analyse von Straßenverkehrsszenen. Es gibt Testfahrzeuge [7, 120, 19, 24], die mit einer oder mehr-



(a)



(b)

Abbildung 2.2: Sensordaten für einen RADAR-Sensor (a) und ein Grauwertkamerabild (b) aus einer Straßenverkehrsszene. Der fünfstrahlige Radarsensor mit einem Öffnungswinkel von 20° ist vorne zentral an einem Trägerfahrzeug angebracht und erkennt vier Objekte. Aufgetragen ist für jedes Objekt die Relativgeschwindigkeit in $\frac{m}{s}$ über der Distanz in m und dem Winkel. Das Kamerabild zeigt die Voraussicht aus dem Trägerfahrzeug. Die Kamera ist vorne, hinter dem Rückspiegel, angebracht.

⁴UV – ultraviolett

⁵IR – infrarot

eren Videokameras (Farbe, Grauwert) ausgestattet sind. Diese Sensoren nehmen Videobilder der Straßenverkehrsszenen auf und geben die Bilddaten an ein System zur Verarbeitung weiter. Ein wesentlicher Bestandteil in der vorliegenden Arbeit ist die Verarbeitung von Videobildern. Die Analysemöglichkeiten von Videobildern werden deshalb in Abschnitt 2.2.2 behandelt.

2.2.2 Videobildauswertung

In der Literatur werden verschiedene Verfahren vorgestellt, die auf Basis von Videobilddaten eine Analyse durchführen und verfahrensspezifische Daten bereitstellen (Abbildung 2.3). Etabliert haben sich dabei Verfahren, die die Form und die Bewegung von Objekten analysieren oder eine Textur- und Konturanalyse auf dem Bilddatenmaterial durchführen [55, 36, 54, 46, 51]. Sind Stereokamerasysteme in das Trägerfahrzeug integriert, wird auch die Disparität zwischen verschiedenen Kamerabildern zur Analyse genutzt [71].

Bei der Videofarbbildauswertung ist aufgrund der drei vorhandenen Farbkanäle (für RGB-Kameras: rot, grün, blau) eine robustere Analyse der Videodaten möglich. Es besteht jedoch das Problem des hohen Rechenaufwands aufgrund der dreifachen Datenmenge im Verhältnis zu Grauwertbilddaten. Für die Objekterkennung kann die Farbsättigung genutzt werden [102, 45]. Bei der Bewegungsanalyse in Fahrerassistenzsystemen werden Objekte extrahiert, die eine Bewegung aufweisen, die nicht der erwarteten Bewegung entspricht [28, 30, 24, 75]. Problematisch bei diesen Verfahren ist, daß bei Eigenbewegung des Beobachters die Störanfälligkeit bei der Auswertung des sogenannten Flußfeldes sehr hoch ist, oder nur mit großem Aufwand gering gehalten werden kann [24].

Verfahren die eine Analyse des Videobilddatenmaterials unter Zuhilfenahme von Konturverläufen durchführen, werden beispielsweise in [31, 84, 115, 27, 70] vorgestellt. Diese Verfahren sind differentielle Analyseverfahren. Der Vorteil der Verfahren besteht darin, daß die anfallende Bilddatenmenge stark verringert wird. Problematisch ist jedoch, daß die Auflösung von Objekten im Fernfeld (großer Abstand zum Beobachter) ungenügend ist.

Zur Analyse der Oberflächenstruktur von Objekten oder der statistischen Verteilung von Objekt-

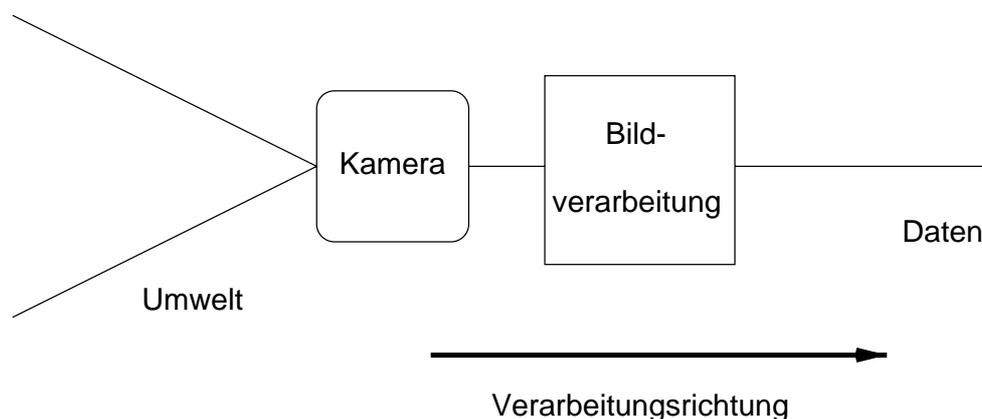


Abbildung 2.3: Prinzip der Verarbeitung von Videobilddatenmaterial. Videosignale werden durch ein Bildverarbeitungsverfahren verarbeitet. Dem System werden verfahrensspezifische Merkmale zugeführt.

merkmalen können Texturanalyseverfahren eingesetzt werden [134, 60, 38, 110, 11, 93]. Bei der Erkennung von Objekten wird mit diesen Verfahren eine lokale Analyse durchgeführt.

Stereoverfahren können zum Beispiel bei der Extraktion erhabener Objekte [71, 7] oder zur Reduzierung der Störanfälligkeit bei der Datenanalyse Anwendung finden [129, 61].

Die Formanalyse baut häufig auf Konturmerkmalen auf und gestattet im Nahfeld (geringe Distanz zum Beobachter) oder bei guter Auflösung (Wahl eines Objektivs mit großer Brennweite) verschiedene Bereiche mit einer bekannten Form zu umschließen (Modell) und diesen Bereichen eine Objektklasse (zum Beispiel PKW, LKW, Fußgänger) zuzuordnen [34, 84, 27, 107, 119]. Um erkannten Objekten eine definierte Klasse zuzuordnen oder sie über einen Klassifikationsprozeß zu verifizieren, können verschiedene in der Literatur beschriebene Klassifikatoren eingesetzt werden [34, 84, 119, 132, 91].

Sollen Objekte über der Zeit verfolgt werden, um beispielsweise deren Bewegungsrichtung analysieren zu können, werden Objektverfolger [84, 127, 74] eingesetzt.

Die durch den Verarbeitungsprozeß extrahierten und attributierten Objekte können für eine Interpretation der Videobildaten genutzt werden, um ein von der Aufgabenstellung gefordertes Systemverhalten zu erhalten. Es hat sich dabei gezeigt, daß der Einsatz einzelner Verfahren nicht ausreicht, um eine notwendige Robustheit im Systemverhalten zu erreichen [40].

Um eine robuste und flexible Analyse von Videobildern zu ermöglichen, kann eine Fusionsstrategie verfolgt werden, die die in diesem Kapitel vorgestellten Ansätze geeignet koppelt. In Abschnitt 2.3 wird das Problem der Sensordatenfusion in den Kontext der Videobildanalyse bei Straßenverkehrsszenen eingeordnet.

2.3 Sensordatenfusion

Soll mit einem gegebenen Satz von Sensoren eine Analyse einer Szene durchgeführt werden, ist eine Sensordatenverarbeitung Voraussetzung, um notwendige Informationen über die Szene zu erhalten. Unter Szene soll hier die Gesamtheit der vorhandenen Objekte inklusive deren physikalischen Eigenschaften und Beziehungen untereinander verstanden werden, die von einem Beobachter mit Hilfe von Sensoren erfaßt werden kann.

In Abbildung 2.4 ist das Prinzip eines solchen Verarbeitungsprozesses dargestellt. Mehrere Sensoren mit verschiedenen physikalischen Eigenschaften, unterschiedlicher Reichweite und Position beobachten relevante Objekte. Die resultierenden Meßgrößen werden mit Hilfe eines Datenfusionsprozesses so kombiniert, daß die Aufgabenstellung den Anforderungen entsprechend gelöst werden kann.

In der Literatur werden unterschiedliche Verfahren zur Fusion mehrerer Sensordaten vorgestellt. Dabei werden Ansätze wie Filteroperationen, heuristische Ansätze, Fuzzy-Logik, Bayesansätze und neuronale Netze zur Fusion des Sensordatenmaterials verwendet [98, 104, 53, 38]. Einen Überblick über aktuelle Arbeiten geben [10, 122, 16]. Diese Verfahren behandeln das Sensordatenfusionsproblem an unterschiedlichen Aufgabenstellungen, wie beispielsweise das Erkennen von Hindernissen beim Flug mit einem Helikopter [133] oder das Erkennen von Schußwaffen, die am Körper getragen werden [98]. Die vorgestellten Verfahren sind problemangepaßt und verringern ihre Güte bei veränderten Randbedingungen.

Die Sensordatenfusion wurde bei Fahrerassistenzsystemen noch nicht systematisch angegangen. In der Literatur finden sich dazu nur wenig Ansätze [89, 100, 120]. Die zugrundeliegenden Me-

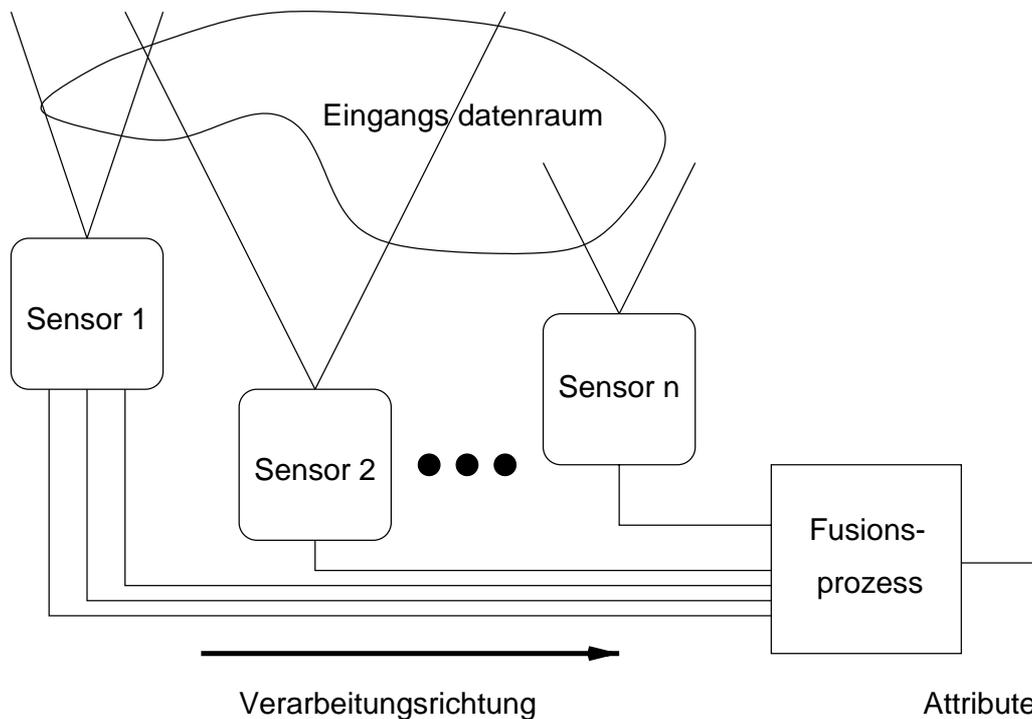


Abbildung 2.4: Prinzip einer Sensordatenfusion. Mehrere Sensoren observieren den Eingangsdatenraum und geben die Sensordaten an einen Fusionsprozeß weiter. Dieser erzeugt Attribute für den folgenden Verarbeitungsprozeß.

thoden sind dabei hochspezialisiert (heuristische Ansätze) und lassen keine oder nur im Einzelfall Erweiterungen zu. Eine gleichbleibende Güte bei veränderten Randbedingungen wird nicht erreicht.

Durch den Einsatz neuronaler Netze zur Fusion der Sensordaten ist eine robuste Struktur gegeben. Weiterhin sind flexible Erweiterungen möglich. In dieser Arbeit wird deshalb ein neuronales Netz zur Fusion von Sensordaten eingesetzt (Kapitel 5).

Werden Objekte aus der Umwelt extrahiert und ist damit die Beschreibung einer Szene möglich, können dem Fahrer eines Fahrzeugs fahrrelevante Informationen bereitgestellt [86, 126] oder ein Eingriff in die Aktuatorik des Kraftfahrzeugs durchgeführt werden. Ein Eingriff in die Aktuatorik kann zur autonomen Navigation genutzt werden. In Abschnitt 2.4 wird dieses Problem behandelt und in Abschnitt 2.5 werden Modelle für eine Verhaltenssteuerung diskutiert.

2.4 Navigation im Straßenverkehr

Fahrerassistenzsysteme werden eingesetzt, um dem Fahrer eines Kraftfahrzeugs Handlungsabläufe abzunehmen. Diese Handlungsabläufe werden definiert durch eine Aufgabenstellung, die vom Fahrer an das Fahrerassistenzsystem übergeben oder systembedingt gelöst wird. Beispielsweise kann der Fahrer einen gewünschten Zielort formulieren und vom System eine Routenplanung durchführen lassen. Greift das Fahrerassistenzsystem zur autonomen Navigation in die Aktuatorik eines Fahrzeugs ein, wie beispielsweise die Steuerung der Lenkung des Kraftfahrzeugs,

um Hindernissen auszuweichen, wird der Systemzustand des Assistenzsystems genutzt, die Zustandsvariablen des Kraftfahrzeugs zu verändern. Der Systemzustand bei Fahrerassistenzsystemen ist definiert durch Attribute vorhandener Objekte, die den Handlungsablauf beeinflussen. Als Zustandsvariablen können bei einem Fahrzeug die Bewegungsrichtung, die Beschleunigung oder die Geschwindigkeit definiert werden. Diese Variablen beschreiben den makroskopischen Systemzustand eines Fahrzeugs.

Bei einem Kraftfahrzeug ist die Bewegungsrichtung definiert durch den Lenkwinkel und die Geschwindigkeit. In Abbildung 2.5 ist eine Prinzipdarstellung dieses Problems gegeben. Um ein Verhalten zu realisieren, müssen die Aufgabe und verhaltensrelevante Attribute in einer Verhaltenssteuerung Berücksichtigung finden. In Abschnitt 2.5 wird das Problem der Verhaltenssteuerung thematisiert.

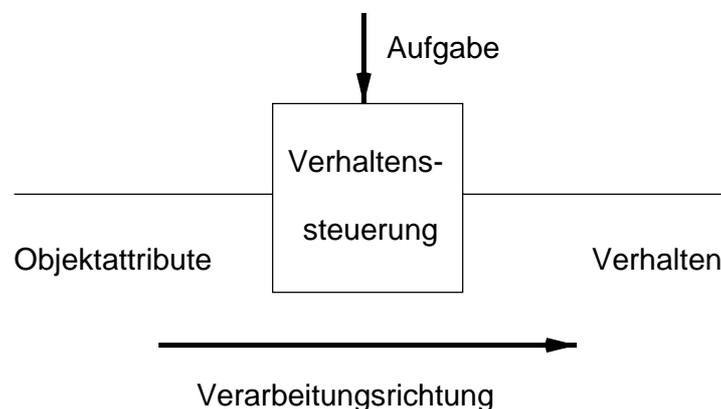


Abbildung 2.5: Prinzipdarstellung eines Moduls zur Verhaltenssteuerung. Objektattribute und eine Aufgabenstellung beeinflussen das Verhalten eines Systems. Der Prozeß der Verhaltenssteuerung muß diese Größen adäquat auf das gewünschte Verhalten abbilden.

2.5 Verhaltenssteuerung

Für die Steuerung von Verhalten gibt es unterschiedliche Ansätze [43, 137, 136, 18]. Es können beispielsweise Entscheidungsbäume, regelbasierte Ansätze oder dynamische Systeme zur Steuerung von Verhaltensvariablen [112] eingesetzt werden.

In der Literatur werden verschiedene Fahrerassistenzsysteme diskutiert, die in die Aktuatorik eines Fahrzeugs eingreifen und dadurch eine Verhaltenssteuerung ausführen. Dabei handelt es sich in jedem Fall um Prototypen, die eine geforderte Aufgabenstellung in problemangepaßten Umwelten mit unterschiedlicher Güte lösen. Beispiele werden in [18, 94, 99] vorgestellt. In [94] wird mit Hilfe verschiedener Fahrbahnmodelle ein neuronaler Klassifikator trainiert. Mit diesem Klassifikator werden Fahrbahnbilder ausgewertet und eine Steuerung des Lenkwinkels durchgeführt. Mit Hilfe dieses Systems kann die Lenkung eines Fahrzeugs automatisiert werden. Der Fahrer des Fahrzeugs muß jedoch die Längsführung übernehmen. Problematisch bei diesem System ist, daß die Fahrbahnmodelle nur eine begrenzte Auswahl an Fahrbahntypen beschreiben und Einflußgrößen auf anderem Abstraktionsniveau, wie das Vorhandensein und die Relativgeschwindigkeit eines entgegenkommenden Fahrzeugs, nicht zugelassen werden. In [18] ist ein

Versuchsaufbau in einem Fahrzeug beschrieben, in dem eine Objektdetektion und -verfolgung sowie eine Fahrbahnerkennung und -verfolgung realisiert sind. Die Längs- und Querregelung wird vom Assistenzsystem übernommen. In problemangepaßten Umwelten kann dieses Fahrzeug autonom navigieren, wobei die notwendigen Parameter für die Verhaltenssteuerung mit einem Kalman-Filter geschätzt werden [4, 35]. Bei diesem System handelt es sich um ein System, das nur schwer erweitert werden kann. In [99] wird ein dynamischer Potentialfeldansatz vorgestellt, der Gefahrenpotentiale bewertet. Dabei wird in einer Repräsentation der Umgebung ein Fahrtverlauf minimalen Risikos gesucht. Eine Kodierung der Gefahrenpotentiale erweist sich bei diesem Verfahren als schwierig. Weitere autonome Fahrzeuge werden in [120, 7] beschrieben. Eine Auflistung über den Stand verschiedenster Forschungsgruppen in Europa, Asien und Amerika gibt [24].

Im Gegensatz zu den beschriebenen Systemen wird in dieser Arbeit ein dynamischer Ansatz neuronaler Informationsverarbeitung gewählt, der bei der Verhaltenssteuerung eine kontinuierliche Ausgabe der Steuergrößen des Fahrzeugs erlaubt. Bei der Integration des Verfahrens wird weiterhin auf eine einfache Einkopplung der notwendigen Attribute geachtet, um eine gute Erweiterbarkeit des Verfahrens zu ermöglichen. Das Verfahren zeigt weiterhin ein zeitlich stabilisierendes Verhalten und verhält sich robust gegen Störungen.

2.6 Einordnung der Arbeit

In dieser Arbeit werden neue Beiträge zur Realisierung von Fahrerassistenzsystemen geliefert. Insbesondere werden Schwerpunkte in den Bereichen der Sensordatenfusion zur Analyse von Straßenverkehrsszenen und der Realisierung einer Bewegungssteuerung für autonome Kraftfahrzeuge gelegt. Bei der Realisierung der Verfahren wird eine biologisch motivierte Strategie verfolgt. Es werden drei neue Verfahren vorgestellt, die

- eine Analyse hochdimensionaler Sensordaten mehrerer Segmentierungsverfahren (Kapitel 4),
- eine flexible und robuste Sensordatenfusion zur Segmentierung von Videobilddatenmaterial (Kapitel 5) und
- eine Bewegungssteuerung eines autonomen Kraftfahrzeugs durch Beeinflussung des Lenkwinkels und der Geschwindigkeit (Kapitel 6)

zulassen. Diese Verfahren unterscheiden sich wesentlich von klassischen Ansätzen, da sie eine nichtlineare Abbildung durchführen und eine einfache Einkopplung von verschiedenen Objektattributen zulassen. Weiterhin sind sie robust gegen Störungen und flexibel erweiterbar.

Bei der Realisierung der Verfahren werden Methoden neuronaler Informationsverarbeitung eingesetzt, die Modelle für die Neurophysiologie kortikaler⁶ Prozesse bilden. Die Tragfähigkeit der Verfahren wird anhand eines in Kapitel 7 realisierten Systems gezeigt.

⁶cortikal – von Cortex, Hirnrinde; abgeleitet aus dem Lateinischen: cortex=Rinde.

2.7 Zusammenfassung

In diesem Kapitel werden Assistenzsysteme für den Straßenverkehr diskutiert. Insbesondere werden Fahrerassistenzsysteme genauer betrachtet und ein Überblick über aktuelle Arbeiten gegeben. Fahrerassistenzsysteme sind direkt in ein Fahrzeug eingebaut und unterstützen den Fahrer während der Fahrt.

Ist das betrachtete System dafür ausgelegt, in die Aktuatorik des Trägerfahrzeugs einzugreifen, ist eine Observation der Umwelt mit Sensoren notwendig. Ein wesentliches Problem ist dabei die Erkennung von Objekten im Straßenverkehr. Eine Möglichkeit die Robustheit der Erkennung zu erhöhen, bietet die Fusion von Sensordaten. Dabei ist das Ziel, relevante Objekte in der Umwelt zu erkennen und zu attributieren. Bei dem Entwurf eines Verfahrens zur Sensordatenfusion ist eine flexible zugrundeliegende Struktur notwendig, um schnell auf sich ändernde Randbedingungen der Umwelt für den Einsatz eines Fahrerassistenzsystems reagieren zu können.

Als weiteres Problem wird das autonome Verhalten von Fahrerassistenzsystemen diskutiert. Die Grundlage für einen autonomen Eingriff eines Fahrerassistenzsystems in die Aktuatorik eines Fahrzeugs bildet eine definierte Aufgabenstellung. Die Aufgabe legt die Randbedingungen für autonomes Verhalten des Trägerfahrzeugs fest. Ein wesentliches Problem ist dabei der Entwurf eines Verfahrens, das Eingriffe in die Aktuatorik eines Fahrzeugs adäquat realisiert. Zur Realisierung aufgabenbezogenen Verhaltens eines autonomen Assistenzsystems sind verschiedene Objektattribute zu berücksichtigen.

Die vorliegende Arbeit setzt bei der Sensordatenfusion in Fahrerassistenzsystemen und der Steuerung der Bewegungsrichtung eines autonomen Fahrzeugs an. Es werden mit Methoden neuronaler Informationsverarbeitung neue Verfahren entwickelt, die Lösungsvorschläge anbieten. Insbesondere wird eine Analyse hochdimensionaler Sensordaten mit einem auf neuronalen Karten basierenden Ansatz durchgeführt. Ein weiterhin entwickeltes Verfahren bietet eine flexible Struktur zur Fusion von Sensordaten an. Ein biologisch motiviertes Konzept der Verhaltenssteuerung bildet die Grundlage für ein Verfahren zur Steuerung des Lenkwinkeleinschlags und der Geschwindigkeit bei einem Kraftfahrzeug.

Da die Basis der entwickelten Verfahren die neuronale Informationsverarbeitung bildet, wird der Bereich der neuronalen Informationsverarbeitung in Kapitel 3 im biologischen Kontext beleuchtet.

Kapitel 3

Neuronale Informationsverarbeitung

Lebewesen treten mit ihrer Umwelt in vielfältiger Weise in Beziehung. Über die Sinnesorgane (z.B. Augen oder Ohren) werden Reize aus der Umwelt aufgenommen, verarbeitet und in Verhalten umgesetzt. Technisch formuliert kann dieser Zusammenhang als Übertragung von Information beschrieben werden [113]. Die Übertragung kann extern sowie intern stattfinden. Externe Informationsübertragung findet meist von einem Lebewesen zum anderen statt. Beispielsweise wird bei einem Gespräch Information von einem Menschen zu einem anderen übertragen. Interne Informationsübertragung kann die Umwandlung von Photonen in den Photorezeptoren im Auge eines Lebewesens und die darauf folgende Aktivierung einzelner Ganglienzellen [116], oder die neuronale Informationsverarbeitung, das heißt der Austausch von Transmitterstoffen und die Aktivierung von Nervenzellen sein [116, 71].

Um Eigenschaften der biologischen Informationsverarbeitung auf künstliche Systeme zu übertragen, werden künstliche neuronale Netze entworfen. Künstliche neuronale Netze haben biologische neuronale Netze als informationsverarbeitende Systeme zum Vorbild. Mit ihnen wird versucht, reale Vorgänge im Nervensystem biologischer Organismen, die komplexes Verhalten realisieren können (im folgenden komplexe biologische Organismen), nachzuvollziehen und für praktische Zwecke anzuwenden [105].

In diesem Kapitel werden Zusammenhänge der neuronalen Informationsverarbeitung komplexer biologischer Organismen veranschaulicht und die in der vorliegenden Arbeit verwendeten Verfahren neuronaler Informationsverarbeitung abgeleitet. In Abschnitt 3.1 wird als Basiselement biologischer Informationsverarbeitung die Nervenzelle diskutiert. In Abschnitt 3.2 werden die Abbildungseigenschaften im Neocortex¹ beschrieben, die die Grundlage für neuronale Karten bilden. Abschnitt 3.3 gibt Einblick in die massive und hierarchische Vernetzung von Neuronen im menschlichen Gehirn, das biologische Vorbild für vorwärtsgerichtete neuronale Netze. Das makroskopische Verhalten komplexer biologischer Organismen resultiert aus dem Zusammenspiel verschiedener mikroskopischer Einheiten. In Abschnitt 3.4 werden Modellansätze für makroskopisches Verhalten komplexer biologischer Organismen diskutiert. Die Umsetzung der vorgestellten Prinzipien in technischen Anwendungen wird in Abschnitt 3.5 motiviert. Das Kapitel 3 schließt mit einer Zusammenfassung in Abschnitt 3.6.

¹Neocortex – Cortex biologischer Organismen, die komplexes Verhalten realisieren können. Entwicklungsgeschichtlich der jüngste und größte Teil der Großhirnrinde, der mit seinem typischen sechsschichtigen Aufbau am stärksten differenziert ist [22, 5].

3.1 Neurone im Neocortex

Die Nervensysteme biologischer Organismen besitzen Gesamtarchitekturen unterschiedlicher Komplexität. Sie alle werden jedoch aus Neuronen (Nervenzellen) gebildet. Beispielsweise besteht der menschliche Neocortex aus $10^{11} - 10^{12}$ Neuronen [116, 47]. Diese Neurone lassen sich in zwei Hauptklassen von Zelltypen unterteilen: die Pyramidenzellen und die Sternzellen. In Abbildung 3.1 ist eine schematische Darstellung einer Pyramidenzelle abgebildet. Neurone empfangen Signale und geben von ihnen erzeugte Signale weiter. Die Fortsätze des Neurons, die die Eingangsinformation sammeln, werden Dendriten² genannt. Sie bilden den Haupteingabepfad einer Zelle. Dendriten übernehmen Signale aus anderen Nervenzellen in Form eines elektrischen Potentials an den Synapsen³ und führen diese dem Zellkörper zu. Überschreitet dieses Potential einen gewissen Schwellenwert, so erzeugt der Zellkörper einen elektrischen Nadelimpuls, das Ausgangssignal. Dieser Vorgang ist aufgrund der Schwellwertoperation nichtlinear. Die Ausgabesignale eines Neurons werden durch das Axon⁴ an andere Neurone weitergereicht. Die Synapsen des Axons befinden sich an den Dendriten oder dem Zellkörper der Zielneuronen. Die Synapsen sind chemische Kontakte. Der elektrische Impuls des Axons bewirkt an der Synapse die Ausschüttung eines Überträgerstoffes (Neurotransmitter) [78], der dann wiederum durch Öffnung von Ionenkanälen zu einer Potentialänderung an einem Dendrit bzw. am Zellkörper des Zielneurons führt. Je nach Art und Zustand der Synapse bewirkt ein eintreffender Impuls eine mehr oder weniger starke Potentialerhöhung (erregende Synapse) oder Potentialerniedrigung (hemmende Synapse). Während die Pyramidenzellen meist langreichweitige Axone mit exzitatorischer Signalwirkung haben, wirkt das Signal des Axons bei Sternzellen meist inhibitorisch [78]. Die massive Vernetzung dieser Neurone im Neocortex scheint die fundamentale Voraussetzung für komplexes Verhalten zu sein.

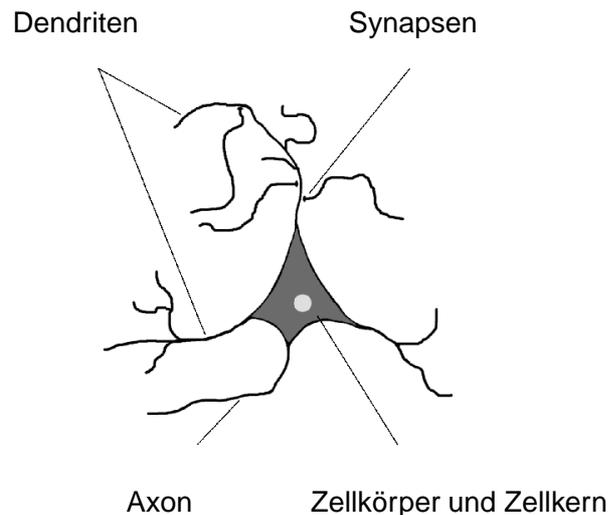


Abbildung 3.1: Schematische Darstellung eines Neurons (Pyramidenzelle) mit Zellkern, Dendriten, Synapsen und Axon.

²Dendrit – abgeleitet aus dem Griechischen: dendron = Baum.

³Synapse – abgeleitet aus dem Griechischen: synapsis = Verbindung.

⁴Axon – auch Achsenzylinder oder Nervenfasern.

In den Abschnitten 3.2, 3.3 und 3.4 werden deshalb lokale Eigenschaften des Neocortex, seine hierarchische Struktur zwischen Arealen⁵ in den einzelnen Schichten des Neocortex und Modelle für komplexes Verhalten biologischer Organismen behandelt.

3.2 Abbildungseigenschaften im Neocortex

Große Teile des Neocortex sind mit der Verarbeitung sensorischer Signale betraut. Diese Signale entstehen durch Reize, welche die Sinnesorgane aufnehmen. Diese Reize sind multidimensional. Beispielsweise sind für das Erkennen von Farbe drei Rezeptoren vorhanden, welche auf unterschiedliche Wellenlängen des Lichts reagieren [78, 116]. Es ist jedoch aus der Neurobiologie bekannt, daß viele Strukturen im Gehirn eine lineare oder planare Topologie aufweisen [21]. Das bedeutet, daß diese Strukturen ein- oder zweidimensional sind. Es ist außerdem festzustellen, daß Bereiche im Gehirn, welche auf einen Reiz besonders reagieren weitgehend in der Nachbarschaft von Bereichen liegen, welche auf ähnliche Signalausprägungen reagieren. Wird der Reiz kontinuierlich in seinen wahrnehmbaren Eigenschaften verändert, so erkennt man eine kontinuierliche Veränderung der Position der Erregung im Neocortex. Es findet also eine zweidimensionale, weitgehend topologieerhaltende Projektion,

$$f: \mathbb{R}^n \longrightarrow \mathbb{R}^2$$

der Reize auf den Cortex statt [123, 21].

Am Beispiel des visuellen Cortex kann dies verdeutlicht werden. Der visuelle Cortex bildet eine Region im Hinterlappen des menschlichen Gehirns. Viele Neurone arbeiten hier an der Decodierung und Verarbeitung der visuellen Sinneseindrücke. Es zeigt sich [116, 105], daß die visuelle Information trotz der Komplexität des Gehirns und seiner Verschaltungen, als zweidimensionale Projektion des visuellen Feldes auf den visuellen Cortex abgebildet wird. Der visuelle Cortex ist also eine Karte des visuellen Feldes. Betrachtet man andere Sinnesorgane wie beispielsweise den Tastsinn und den zugehörigen somatosensorischen Cortex, wie in [101] dargestellt, so können gleiche Eigenschaften der Projektion festgestellt werden.

Um die Abbildungseigenschaften des Neocortex auf technische Systeme zu übertragen, wurden neuronale Karten entworfen [62, 63]. Sie erlauben eine Projektion eines mehrdimensionalen Raums in eine ein- oder zweidimensionale Anordnung. Mit Hilfe neuronaler Karten kann eine Analyse hochdimensionaler Daten durchgeführt werden [64, 124].

In Kapitel 4 wird eine Analyse von Segmentierungsverfahren mit selbstorganisierenden neuronalen Karten durchgeführt. Dabei werden die Abbildungseigenschaften neuronaler Karten genutzt, um Komponenten verschiedener Segmentierungsverfahren, im Hinblick auf eine Fusion dieser, zu bewerten.

⁵Areal – aus dem Lateinischen: area = Feld, Fläche. Bereich im Neocortex mit bestimmten funktionellen Eigenschaften. Die Einteilung der Areale geht häufig auf die Einteilung nach histologischen Kriterien, d. h. die unterschiedliche Ausprägung der einzelnen Schichten, von Brodmann 1909 zurück [22, 116].

3.3 Funktionelle Einheiten und hierarchische Anordnung von Neuronen im Neocortex

Der Neocortex ist in sechs Schichten aufgebaut. Es können in jeder Schicht Gruppen von Zellkörpern und Fasern erkannt werden, wobei Schicht I an der Oberfläche und Schicht VI am Grunde des Neocortex liegt. Jede Schicht hat gesonderte funktionelle Eigenschaften [5]. Bei genauer Betrachtung des Querschnitts kann eine Ausbildung von sogenannten Säulen durch die Neurone und die Nervenfasern (Dendriten, Axone) erkannt werden, die senkrecht zu den Schichten des Cortex liegen. Die einzelnen Schichten bilden nach einer weitverbreiteten Ansicht in der Neurobiologie [116] die grundlegenden funktionellen Einheiten des Cortex. In sensorischen Cortexbereichen scheinen sie den Grundelementen der Wahrnehmungsvorgänge zu entsprechen. Der Neocortex ist weiterhin in Areale unterteilt, die einer hierarchischen Anordnung folgen und eine Verarbeitung der Reize mit

$$f: \mathbb{R}^n \longrightarrow \mathbb{R}^m$$

durchführen, um komplexes makroskopisches Verhalten zu realisieren [116]. Die hierarchische Anordnung der Areale konnte beispielsweise bei der visuellen Informationsverarbeitung bei Säugetieren nachgewiesen werden [5].

Mit sogenannten vorwärtsgekoppelten künstlichen neuronalen Netzen wurde die Eigenschaft der hierarchischen Anordnung von Neuronen im Neocortex für technische Anwendungen nutzbar gemacht. Neuronale Netze werden bei der Analyse komplexer Datensätze als Schätzer oder Klassifikator eingesetzt.

In Kapitel 5 wird ein hierarchisch angeordnetes neuronales Netz verwendet, um mehrere Komponenten einzelner Segmentierungsverfahren zu fusionieren. Dabei erweist sich der Einsatz dieser Netze als effizient und flexibel. Es wird ein neues Verfahren zur Segmentierung von Videobildmaterial vorgestellt, in dem ein neuronales Netz als Klassifikator eingesetzt wird.

3.4 Makroskopisches Verhalten komplexer biologischer Organismen

Das makroskopische Verhalten komplexer biologischer Organismen ergibt sich aus der geordneten Aktivität vieler einzelner mikroskopischer Einheiten. Die mikroskopischen Einheiten sind beispielsweise Muskeln, Neurone oder Rezeptorzellen. Zusätzlich können komplexe biologische Systeme eine Vielzahl verschiedener Verhaltensmuster organisieren. Wie diese Systeme koordiniert werden, um ein spezifisches Verhalten zu erzeugen, ist bis heute unbekannt. Das Problem ist, die relevanten Parameter dieser Systeme zu selektieren [112]. In der Literatur werden deshalb biologische Systeme nur in einem definierten Kontext untersucht. Es findet eine Beschränkung auf einzelne Basisverhalten statt. Beispielsweise wird in [72] das Jagdverhalten von Raubtieren diskutiert. Weiterhin kann die Bewegungssteuerung von Gliedmaßen biologischer Organismen untersucht werden, da das Problem der notwendigen Anzahl von Freiheitsgraden zur Durchführung einer geforderten Bewegung im Verhältnis zur Komplexität des

Gesamtsystems reduziert werden kann. Ansätze hierzu findet man beispielsweise im Bereich der Kybernetik⁶ [76] und in der Theorie dynamischer Systeme [112, 2, 111, 29].

Der kybernetische Ansatz basiert auf der Idee, das Verhalten komplexer biologischer Systeme auf Einzeltvorgänge zurückzuführen und diese durch funktionale Einheiten zu beschreiben. Das Gesamtverhalten resultiert aus der Verknüpfung aller funktioneller Einheiten. Dabei ist das beobachtbare Verhalten der biologischen Systeme und weniger die zugrundeliegende Struktur von Interesse [20]. Die Alternative hierzu bilden verhaltensorientierte Modelle aus dem Bereich der Theorie dynamischer Systeme. Bei diesen Verhaltensmodellen wird nach der zugrundeliegenden Struktur gesucht, um Verhalten zu planen und auszuführen [112]. Mit dieser Struktur wird makroskopisches Verhalten, abhängig von mikroskopischen Einheiten modelliert.

In Kapitel 6 wird ein dynamisches Verhaltensmodell eingesetzt, um eine Bewegungssteuerung in Kraftfahrzeugen durch Beeinflussung des Lenkwinkels und der Geschwindigkeit zu ermöglichen. Relevante Attribute von Objekten aus der Umgebung des Fahrzeugs (mikroskopischen Einheiten) werden zur Steuerung der Bewegung des Kraftfahrzeugs (makroskopisches Verhalten) genutzt.

3.5 Neuronale Informationsverarbeitung in technischen Systemen

Einzelne Prinzipien neuronaler Informationsverarbeitung in komplexen biologischen Organismen werden in vielfältiger Weise in technischen Systemen nutzbar gemacht. Dabei ist hauptsächlich eine algorithmische Umsetzung der beobachteten Eigenschaften biologischer Informationsverarbeitung von Interesse. Einen Überblick über diese Arbeiten gibt [52].

Die vorliegende Arbeit ordnet sich in dieses Gebiet der neuronalen Informationsverarbeitung ein, da wesentliche Aspekte dieser algorithmischen Umsetzung bei den Verfahrensentwicklungen in den Kapiteln 4, 5 und 6 berücksichtigt werden. Insbesondere die Möglichkeit der nichtlinearen Abbildung vorhandener Eingangsdatenräume auf einen Zieldatenraum erlaubt es, die Qualität der gewünschten Verfahreigenschaften zu erhöhen. Durch geeignete Wahl einer nichtlinearen Abbildungsfunktion kann beispielsweise ein Sättigungsverhalten im Verfahren kodiert werden. Damit können Redundanzen genutzt und der Einfluß von Störgrößen reduziert werden.

3.6 Zusammenfassung

In diesem Kapitel werden Grundprinzipien der neuronalen Informationsverarbeitung in komplexen biologischen Organismen vorgestellt. Dabei wird der Schwerpunkt auf Bereiche gelegt, die die Basis für algorithmische Umsetzungen künstlicher neuronaler Informationsverarbeitung bilden. Insbesondere wird die Wirkungsweise der Grundelemente biologischer Informationsverarbeitung, der Neuronen, erläutert. Die Abbildungseigenschaften im Neocortex, sowie die hierarchische Anordnung von Arealen werden beschrieben. Modelle für das makroskopische Verhalten von komplexen biologischen Organismen werden vorgestellt. Auf den Vorteil der

⁶Kybernetik – abgeleitet aus dem Griechischen: kybernetes = Lotse, Steuermann.

nichtlinearen Abbildbarkeit von Datenräumen und die daraus resultierende Störsicherheit von biologisch motivierten Verfahren wird verwiesen.

In den folgenden Kapiteln werden Verfahren vorgestellt, die unter Anwendung einzelner Prinzipien neuronaler Informationsverarbeitung, neue Lösungswege im Bereich der Fahrerassistenzsysteme aufzeigen.

Kapitel 4

Analyse von Segmentierungsverfahren mittels neuronaler Karten

Bei komplexen Fahrerassistenzsystemen ist es das Ziel, Verkehrsteilnehmer (Objekte) in einer Straßenverkehrsszene zu erkennen. Um eine möglichst genaue Beschreibung der Umgebung des Fahrzeugs zu erreichen, werden die erkannten Objekte während des Verarbeitungsprozesses attribuiert. Mögliche Attribute sind die Objektklasse, der Abstand, die Relativgeschwindigkeit und das Gefahrenpotential hinsichtlich einer beabsichtigten Eigentrajektorie. Bei der Erkennung von Objekten im Straßenverkehr für Fahrerassistenzaufgaben werden hohe Anforderungen an die verwendeten Verfahren gestellt. Nicht nur die große Vielfalt der unterschiedlichen Umweltszenarien, sondern auch das hohe Maß an Sicherheit, das die gestellte Aufgabe erfordert, bedingen ein breitbandiges und flexibles Gesamtsystem mit hoher Schätzgüte.

Ein wesentlicher Bestandteil bei der Analyse einer Szene ist die zielgerichtete Steuerung des Aufmerksamkeitsbereichs. Insbesondere ist die Anforderung an eine robuste Analyse von Bildern in technischen Anwendungen zur Videobilddatenverarbeitung für Fahrerassistenzsysteme die Grundlage für eine algorithmische Implementierung von sogenannten Segmentierungsverfahren. Mit Hilfe dieser Verfahren werden die Videobilder in Segmente unterteilt, die für den Verarbeitungsprozeß relevant sind und solche, die für den weiteren Verarbeitungsprozeß keine Verwendung finden. Mit Segmentierungsverfahren wird folglich eine Reduktion des Bilddatenmaterials erreicht.

In der Literatur werden verschiedene Verfahren zur Segmentierung von Bilddatenmaterial diskutiert [34, 59, 73, 36, 54]. Diese Verfahren ermöglichen eine Einschränkung des Betrachtungsbereichs im Bild (Datenreduktion), um einerseits eine Störunterdrückung¹ durchzuführen und andererseits komplexe und rechenintensive Algorithmen nur in Bildbereichen einzusetzen, die einer durch die Aufgabe geforderten erhöhten Aufmerksamkeit bedürfen. Die in der Literatur beschriebenen Verfahren sind häufig problemangepaßt und verringern die Güte der Segmentierung bei veränderten Randbedingungen.

Die Anforderung, bei der Analyse von Straßenverkehrsszenen eine robuste Segmentierung zu erreichen und flexibel mit veränderten Randbedingungen zu operieren, ist die Grundlage für die Fusion von Segmentierungsverfahren unterschiedlicher Ausprägung. Dabei ist es notwendig, die Eigenschaften dieser Verfahren im Hinblick auf den Einsatzbereich im Straßenverkehr bei bewegtem Beobachter zu analysieren und das Redundanzverhalten und die Güte der Segmen-

¹Störung bezeichnet hier den Einfluß von Hintergrundsegmenten auf die gestellte Aufgabe.

tierung durch Fusion zu bewerten. Redundanz verringert die Störanfälligkeit bei verrauschten Eingangsdaten und gewährleistet häufig eine Lösung der gestellten Aufgabe beim Ausfall einzelner Daten. Die Güte eines Segmentierungsergebnisses kann durch Nutzung unabhängiger, aus dem Datenmaterial extrahierter Information erhöht werden.

In diesem Kapitel werden verschiedene Verfahren analysiert, die eine Segmentierung von Grauwertbilddaten erlauben (Abbildung 4.1). Um deren Eignung für die Segmentierungsaufgabe bei Straßenverkehrsszenen zu beurteilen, werden diese mit einer zweidimensionalen neuronalen Karte fusioniert. Der Einfluß einzelner Komponenten auf das Fusionsergebnis wird untersucht. Der Einsatz neuronaler Karten wird durch eine einfache Visualisierbarkeit hochdimensionaler Datenräume motiviert [114, 48, 124]. Zweidimensionale neuronale Karten bilden durch eine nichtlineare Transformation eine Repräsentation des betrachteten Datenraums in Form einer Merkmalskarte aus. Die Merkmalskarte kann mit Expertenwissen interpretiert werden.

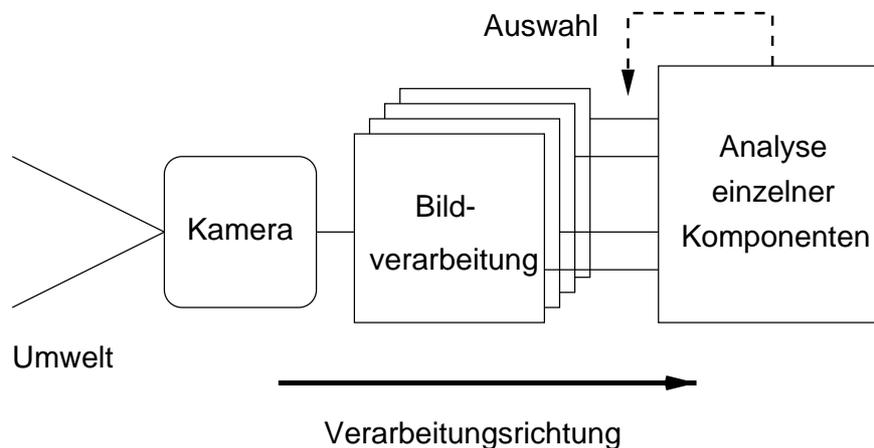


Abbildung 4.1: Prinzipbild der Analyse. Komponenten verschiedener Verfahren werden analysiert. Einzelne Komponenten werden ausgewählt.

In Abschnitt 4.1 wird der verwendete Ansatz der neuronalen Karten vorgestellt. In Abschnitt 4.2 wird eine Analyse verschiedener Verfahren durchgeführt, die eine Segmentierung von Videobilddaten ermöglichen. Dabei werden Textur-, Kontur- und Bewegungsdaten aus Grauwertbildern berücksichtigt, mit dem Ziel, die relevanten Komponenten für einen Segmentierungsprozeß, welcher in Kapitel 5 beschrieben ist, auszuwählen. Das Kapitel 4 schließt mit einer Zusammenfassung in Abschnitt 4.3.

4.1 Neuronale Karten

Die Grundlage für die Struktur neuronaler Karten basiert auf Arbeiten, in denen auf die funktionelle Organisation des visuellen Cortex bei Säugetieren eingegangen wird. Die Selbstorganisation und die Topologieerhaltung von Neuronen im visuellen Cortex wird dabei besonders untersucht [123]. Topologieerhaltung beschreibt die Beibehaltung von Stetigkeitsbeziehungen zwischen den Eingangsdaten bei der Abbildung und umgekehrt bei der Rückprojektion der Abbildung in den Eingangsdatenraum.

Neuronale Karten sind ein- oder mehrdimensionale Anordnungen von künstlichen Neuronen (im folgenden Neuronen), die über Gewichtsvektoren einen Eingangsdatenraum der Dimension n repräsentieren. Jedes Neuron der Karte bekommt den aktuellen Eingangsdatenvektor zugeführt. Es hat n Eingänge und einen Referenzvektor der Dimension n . Der Referenzvektor eines Neurons spezifiziert das Zentrum eines Bereichs des Eingangsdatenraums, für den das Neuron der Repräsentant ist.

Eine Umsetzung dieser Struktur ist beispielsweise durch selbstorganisierende neuronale Karten (SOM²) gegeben [64]. Unter Selbstorganisation wird hier die Ausbildung der Referenzvektoren der Neurone im Eingangsdatenraum in einem unüberwachten Prozeß verstanden. Selbstorganisierende neuronale Karten werden in den folgenden Abschnitten genauer diskutiert, da sie zur Analyse von Segmentierungsverfahren eingesetzt werden.

4.1.1 Selbstorganisierende neuronale Karten

Selbstorganisierende neuronale Karten wurden Anfang der achtziger Jahre von *Kohonen* eingeführt [62, 63]. Sie haben eine ein- oder zweidimensionale Anordnung. Im eindimensionalen Fall hat jedes Neuron zwei Nachbarn, man spricht von einer sogenannten Kohonenkette. Im zweidimensionalen Fall hat jedes Neuron entweder vier oder sechs Nachbarn (Rechteckgitter, bzw. hexagonales Gitter). Man spricht von einer Kohonenkarte.

Selbstorganisierende neuronale Karten bilden einen n -dimensionalen Eingangsraum \mathbf{X} im wesentlichen topologieerhaltend auf eine ein- oder zweidimensionale Struktur ab. Die Struktur der Karten ist charakterisiert durch ein Netz mit J Neuronen (Abbildung 4.2a), deren Lage in $\mathbf{X} \subset \mathbb{R}^n$ durch einen n -dimensionalen Referenzvektor \mathbf{w}_j , mit

$$\mathbf{w}_j = [w_{j1}, w_{j2}, \dots, w_{jn}]^T \in \mathbf{X}$$

und $j = \{1, 2, \dots, J\}$ beschrieben wird. Jedes Neuron hat Nachbarn, mit denen es verknüpft ist (Abbildung 4.2b). Alle Neurone empfangen den gleichen Eingangsdatenvektor \mathbf{x} , mit

$$\mathbf{x} = [x_1, x_2, \dots, x_n]^T \in \mathbf{X}.$$

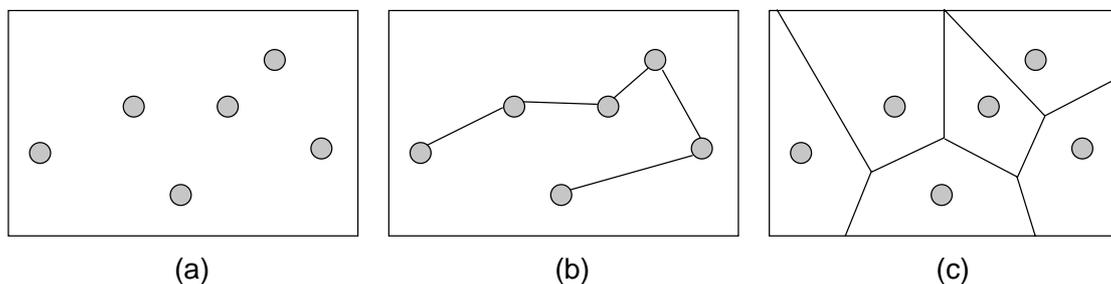


Abbildung 4.2: (a) Lage der Neurone in einem zweidimensionalen Raum, (b) Verknüpfung der Nachbarn zu einer Kohonenkette und (c) Voronoigebiete.

²SOM – self organizing map

Ein Eingangsdatenvektor \mathbf{x} wird mit den Referenzvektoren \mathbf{w}_j auf einer definierten Metrik verglichen. In technischen Anwendungen wird häufig die euklidische Distanz als Metrik verwendet. Dabei ist die Erregung eines Neurons abhängig von der Distanz d_j zwischen der Position von \mathbf{x} und der Position des Neurons, gegeben durch \mathbf{w}_j in \mathbf{X} , mit

$$d_j(\mathbf{x}) = \|\mathbf{x} - \mathbf{w}_j\|.$$

Es wird nach der kleinsten euklidischen Distanz zwischen \mathbf{x} und \mathbf{w}_j mit

$$\|\mathbf{x} - \mathbf{w}_c\| = \min_j(d_j(\mathbf{x})) \quad (4.1)$$

gesucht. Der Index c in Gleichung 4.1 kennzeichnet den Referenzvektor \mathbf{w}_c des Neurons, welcher die minimale Distanz zum Eingangsdatenvektor \mathbf{x} aufweist. Aus diesem Zusammenhang entstehen Voronoigebiete V_j , mit $V_j \subset \mathbb{R}^n$ im Eingangsdatenraum \mathbf{X} mit

$$V_j = \{\mathbf{x} \in \mathbb{R}^n \mid \arg \min_j(d_j(\mathbf{x}))\},$$

um jedes Neuron (Abbildung 4.2c). Ein Eingangsdatenvektor \mathbf{x} , welcher in V_j liegt, wird von dem zugehörigen Neuron repräsentiert.

In einem Lernvorgang ordnen sich die Neurone der Karte in \mathbf{X} an. Ist die Dimension von \mathbf{X} höher als die Dimension der Karte, können Verletzungen in der Topologieerhaltung bei der Abbildung von \mathbf{X} auf die Kohonenkarte auftreten. Es bilden jedoch benachbarte Neurone benachbarte Bereiche in \mathbf{X} ab.

4.1.2 Lernverfahren

Um mit neuronalen Karten einen gegebenen Datensatz $\tilde{\mathbf{X}}$ mit einem minimalen Fehler zu repräsentieren, werden die Referenzvektoren \mathbf{w}_j der einzelnen Neurone durch ein Lernverfahren verändert. Es ist keine explizite Fehlerfunktion für das Netz definiert. Der Lernvorgang der Netzgewichte erfolgt selbstorganisierend durch einen unüberwachten stochastischen Prozeß. Mathematisch stellt er einen sogenannten Markov-Prozeß dar [101]. Er kann mit

$$\mathbf{w}_j(t+1) = \mathbf{w}_j(t) + \vartheta_{cj}(t)[\tilde{\mathbf{x}}(t) - \mathbf{w}_j(t)]$$

beschrieben werden. Der Referenzvektor $\mathbf{w}_j(t+1)$ wird in Richtung eines zufällig aus $\tilde{\mathbf{X}}$ gezogenen Datenvektors $\tilde{\mathbf{x}}(t)$, gewichtet mit ϑ_{cj} , verändert. Durch die zufällige Auswahl wird die statistische Verteilung der $\tilde{\mathbf{x}}$ in $\tilde{\mathbf{X}}$ berücksichtigt. Die Variable t bezeichnet die diskrete Zeitachse und ϑ_{cj} eine Nachbarschaftsfunktion zwischen den Neuronen mit den Referenzvektoren \mathbf{w}_c und \mathbf{w}_j . Die Nachbarschaftsfunktion ϑ_{cj} ist eine monoton fallende Funktion mit $\vartheta_{cj}(t) = \vartheta_{cj}(\|\mathbf{r}_c - \mathbf{r}_j\|, t) \in \mathbb{R}$, wobei die Koordinaten \mathbf{r}_c und \mathbf{r}_j die Lage der zugehörigen Neurone auf der neuronalen Karte beschreiben. Um während des Lernvorgangs Konvergenz zu erreichen, muß $\vartheta_{cj}(t) \rightarrow 0$ für $t \rightarrow \infty$, sowie mit zunehmenden Abstand $\|\mathbf{r}_c - \mathbf{r}_j\|$, $\vartheta_{cj}(t) \rightarrow 0$ gelten [64].

Eine Funktion ϑ_{cj} , die diese Bedingung erfüllt, ist eine mit einer Lernrate $\alpha(t)$ gewichtete Gaußfunktion $N(0, \sigma)$ mit

$$\vartheta_{cj} = \alpha(t) \cdot e^{-\frac{\|\mathbf{r}_c - \mathbf{r}_j\|^2}{2\sigma^2(t)}}$$

und $0 < \alpha(t) < 1$. Der Parameter $\sigma(t)$ definiert die Breite der Gaußfunktion N . Über weitere Funktionen, die die geforderten Bedingungen ebenfalls erfüllen, wird ein Überblick in [64] gegeben.

In Abbildung 4.3 ist ein Beispiel dargestellt, das den Lernfortschritt nach dem beschriebenen Lernschema für eine Kohonenkette mit 100 Neuronen zeigt. Zufällig ausgewählte Datenvektoren $\tilde{\mathbf{x}}$ (x und y Koordinaten) eines gegebenen Datensatzes $\tilde{\mathbf{X}}$ werden über der Zeit der Kohonenkette präsentiert. Der Datensatz $\tilde{\mathbf{X}}$ stellt dabei eine Stichprobe eines zweidimensionalen Datenraums \mathbf{X} , der von einem Dreieck umschlossen ist (grauer Bereich), mit $\tilde{\mathbf{X}} \subset \mathbf{X}$ dar. Datenpunkte aus dem weißen Bereich werden nicht präsentiert. Die Lage der Neurone ändert sich über der Zeit, um den gegebenen Datensatz $\tilde{\mathbf{X}}$ möglichst genau abzubilden.

Aufgrund der höheren Dimension von $\tilde{\mathbf{X}}$ im Vergleich zur Dimension der Kohonenkette, finden Verletzungen der Topologieerhaltung von $\tilde{\mathbf{X}}$ auf der Kohonenkette statt (Abbildung 4.3c). Wählt man anstelle einer Kohonenkette eine zweidimensionale Kohonenkarte, ist aufgrund der gleichen Dimension eine Topologieerhaltung von $\tilde{\mathbf{X}}$ vollständig möglich.

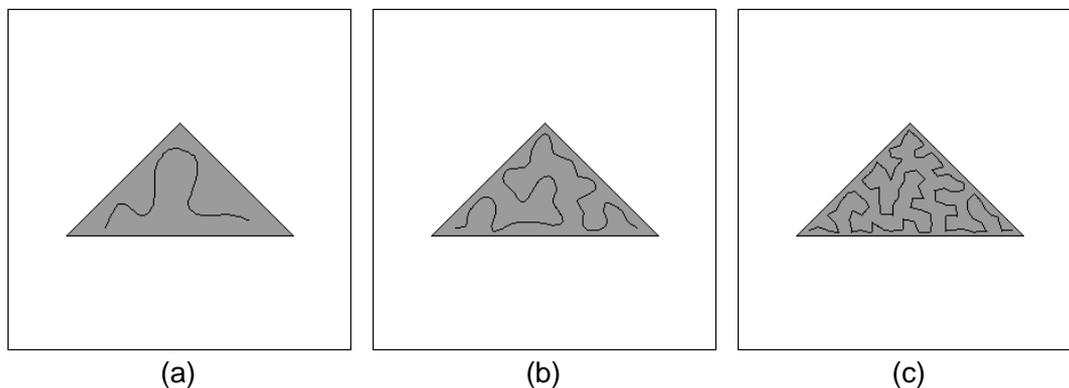


Abbildung 4.3: Abbildung eines zweidimensionalen Datenraums, welcher von einem Dreieck umschlossen ist, mit Hilfe einer Kohonenkette mit 100 Neuronen nach 12000, 24000 und 36000 Lernschritten (a,b,c).

4.1.3 Kohonenkarten als Klassifikator

Ist bei einem vorhandenen Eingangsdatenraum \mathbf{X} eine Klassifikationsaufgabe zu lösen, können Kohonenkarten eingesetzt werden [38]. Notwendig dafür sind Trainingsdaten $\tilde{\mathbf{x}}$, denen eine definierte Klasse C zugeordnet werden kann.

In Abbildung 4.4 ist ein Beispiel für eine Klassifikationsaufgabe dargestellt. In diesem Beispiel werden Datenvektoren $\tilde{\mathbf{x}}$ (x und y Koordinaten) zufällig aus einem gegebenen Datensatz $\tilde{\mathbf{X}}$ (Trainingsdatensatz mit 6000 Beispieldaten) aus drei Bereichen zweier Klassen $C_{1,2}$ ($C_1 \equiv$ „dunkelgrau“ und $C_2 \equiv$ „hellgrau“) eines zweidimensionalen Datenraums \mathbf{X} mit $\tilde{\mathbf{X}} \subset \mathbf{X}$ gezogen und die Referenzvektoren \mathbf{w}_j der Neurone über der Zeit gelernt. Datenpunkte aus dem weißen Bereich werden nicht präsentiert. Es wird eine Kartengröße von 30×20 mit 600 Neuronen gewählt. In Abbildung 4.4 ist der Selbstorganisationsprozeß über der Zeit gezeigt.

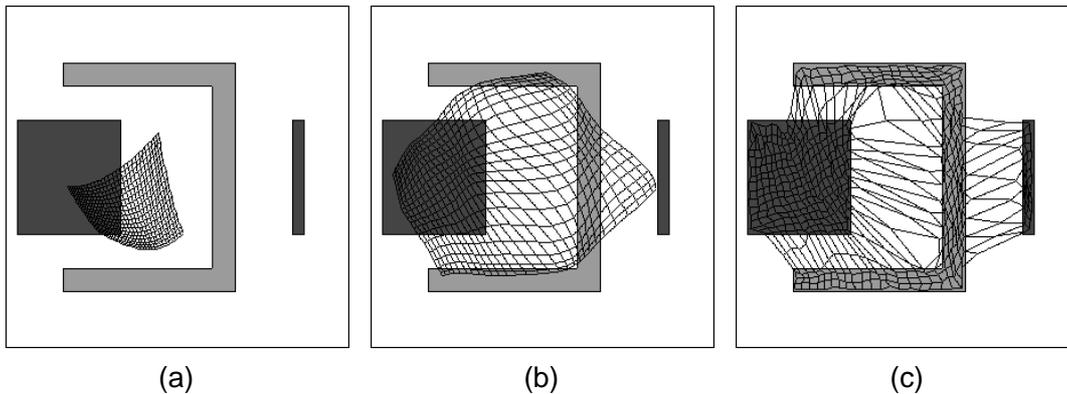


Abbildung 4.4: Abbildung eines zweidimensionalen Datenraums, welcher drei Bereiche unterschiedlicher Klassenzugehörigkeit aufweist, mit Hilfe einer Kohonenkarte mit 30×20 Elementen nach 70000, 100000 und 260000 Lernschritten (a,b,c).

Aufgrund der bekannten Klassen im Trainingsdatensatz kann jedem Neuron und damit jedem Voronoigebiet V_j der Kohonenkarte nach Abschluß der Selbstorganisation eine Klasse $C_{1,2,3}$ zugeordnet werden ($C_3 \equiv$ „unbestimmt“).

Dichtekarten visualisieren die Anzahl der von einem Voronoigebiet V_j repräsentierten Datenvektoren \tilde{x} durch einen Grauwert. In Abbildung 4.5 sind die zugehörige Dichtekarte des Netzes und die jeweiligen Karten für die repräsentierten Klassen dargestellt.

Die gleichmäßige Verteilung der Anzahl repräsentierter \tilde{x} durch die V_j läßt eine gute Repräsentation des Datensatzes \tilde{X} erkennen (Abbildung 4.5a). Eine quantitative Analyse ergibt, daß 298 Neurone der Klasse C_1 , 248 Neurone der Klasse C_2 und 54 Neurone der Klasse C_3 zugeordnet werden. Im Mittel repräsentieren die Neurone der Klassen $C_{1,2}$ jeweils 10.9890 Trainingsdaten \tilde{x} mit einer Standardabweichung $\sigma_x = 3.61998$.

Aufgrund der Topologieerhaltung der Kohonenkarte können zusammenhängende Gebiete einer Klasse $C_{1,2}$ zugeordnet werden (Abbildung 4.5b,c). Es zeigt sich, daß nur an den Grenzflächen dieser Gebiete Neurone unbestimmter Zuordnung (Klasse C_3 , weiße Bereiche in Abbildung 4.5a) vorhanden sind. Folglich kann nach Abschluß der Selbstorganisation eine Klassifikation von Eingangsdatenvektoren x aus dem Datenraum X durchgeführt werden. Eingangsdaten-

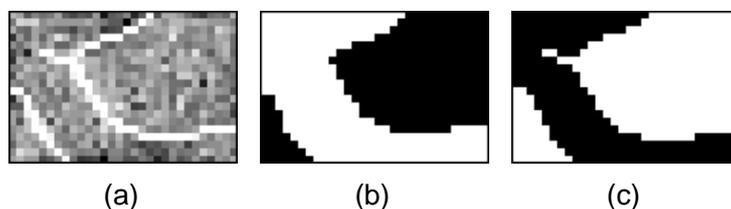


Abbildung 4.5: (a) Dichtekarte und Klassenzugehörigkeit der neuronalen Karte für (b) die Klasse C_1 und (c) die Klasse C_2 (schwarz dargestellt).

vektoren \mathbf{x} bekommen bei diesem Prozeß die Klasse des Neurons zugewiesen, in dessen Voronoi-gebiet V_j sie liegen.

Die Referenzebenen der \mathbf{w}_j beschreiben die Komponenten der Referenzvektoren und damit die Lage der Neuronen in der zugehörigen Dimension des Eingangsdatenraums \mathbf{X} . In Abbildung 4.6 sind die Referenzebenen der Kohonenkarte für die x- und die y-Koordinate dargestellt. Die Position eines Referenzvektors \mathbf{w}_j in einer Dimension ist durch die Aktivität in der Referenzebene beschrieben (hohe Aktivität entspricht einem dunklen Grauwert).

Die Darstellung der Referenzebenen kann genutzt werden, um eine Relevanzanalyse einer Dimension für die Klassifikationsaufgabe durchzuführen. Referenzebenen lassen eine Interpretation des Einflusses der Komponenten eines Eingangsdatenvektors \mathbf{x} auf das Klassifikationsergebnis zu.

Bei dem gegebenen Beispiel ist der dunkelgraue Bereich der Klasse C_1 rechts in Abbildung 4.4c (schmales Rechteck) vollständig durch die x-Koordinate beschreibbar (Abbildung 4.6a). Es ist eine lineare Trennung dieses Bereichs aus C_1 von der Klasse C_2 möglich. Die anderen Bereiche der untersuchten Klassen $C_{1,2}$ werden nichtlinear getrennt. Eine lineare Trennung dieser Bereiche durch Auswertung der Aktivität auf den Referenzebenen ist nicht möglich (Abbildung 4.6a,b).

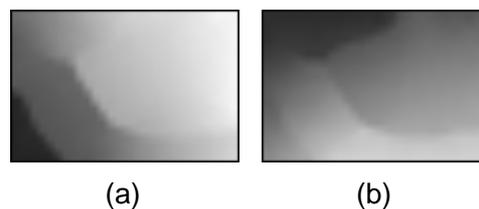


Abbildung 4.6: Abbildung der zwei Referenzebenen des gegebenen Beispiels für (a) die x- und (b) die y-Koordinate.

In Abschnitt 4.2 werden die Referenzebenen interpretiert, um eine Analyse mehrerer Komponenten einzelner Segmentierungsverfahren in einem Fusionsprozeß durchzuführen.

4.2 Analyse von Segmentierungsverfahren

Zur Segmentierung von Grauwertbildern werden in der Literatur verschiedene Ansätze umgesetzt. Soll der Eingangsdatenraum möglichst schnell reduziert werden und sind typische Konturverläufe einzelner Segmente bekannt, werden differentielle Verfahren eingesetzt, die hochfrequente Änderungen im Helligkeitsverlauf analysieren. Sind Eigenschaften der Oberflächen der Segmente bekannt, wie beispielsweise deren Homogenität, werden integrative Methoden zur Segmentierung verwendet. Ist in der zeitlichen Abfolge des Datenmaterials eine ortsvariante Dynamik von Segmenten festzustellen, kann eine Bewegungsanalyse über mehrere Bilder Hinweise auf das Vorhandensein dieser Segmente geben.

Bei der Untersuchung der Ergebnisse von Segmentierungsverfahren können nur qualitative Aussagen gemacht werden. Eine Betrachtung der Segmentierungsgüte über der Datenreduktion

kann zwar durchgeführt werden, jedoch ist eine Interpretation der Ergebnisse nur im Bezug auf den Inhalt der betrachteten Szene sinnvoll. Deshalb wird in diesem Abschnitt eine Analyse der Segmentierungsverfahren mit Expertenwissen durchgeführt.

Im folgenden wird der vorgenannte Ansatz der selbstorganisierenden neuronalen Karten genutzt, um eine Analyse der Ergebnisdaten verschiedener Segmentierungsverfahren (Komponenten), im Hinblick auf eine Fusion der Komponenten, durchzuführen. Die eingesetzte Kohonenkarte führt eine zweidimensionale Abbildung der Komponenten durch. Diese Abbildung stellt eine Fusion der Komponenten dar.

Kohonenkarten zeichnen sich außerdem durch die Eigenschaft der Topologieerhaltung aus. Benachbarte Regionen einer Kohonenkarte werden auf benachbarte Gebiete im Eingangsdatenraum \mathbf{X} abgebildet. Diese Eigenschaft erlaubt es, zusammenhängende Gebiete der Karte in Klassen aufzuteilen und eine Segmentierung der Bilddaten durchzuführen.

Es werden weiterhin die Referenzebenen der Kohonenkarte bei der Analyse der Komponenten genutzt, um eine Auswahl und eine Gruppierung einzelner Komponenten zu ermöglichen.

4.2.1 Selbstorganisierender Prozeß

Zur Datenanalyse wird eine zweidimensionale Kohonenkarte mit 40×20 Neuronen³ eingesetzt, um die untersuchten Verfahren zu fusionieren. Dafür werden verschiedene Komponenten X_i , $i = \{1, 2, \dots, n\}$, der untersuchten Verfahren zu einem n -dimensionalen Eingangsdatenvektor \mathbf{x} mit

$$\mathbf{x} = [X_1(p, q), X_2(p, q), \dots, X_n(p, q)]^T$$

zusammengefaßt. Die Variablen p und q beschreiben den betrachteten Ort im Bild.

Die Karte wird als Klassifikator gemäß Abschnitt 4.1.3 eingesetzt. Mit Hilfe einer Datenbank $\tilde{\mathbf{X}}_1$ mit $\tilde{\mathbf{X}}_1 \subset \mathbf{X}$ (Datensatz mit 30000 Beispieldaten), in der Vordergrund- und Hintergrundsegmente abgelegt sind, können Trainingsdatenvektoren $\tilde{\mathbf{x}}_1$ einer definierten Klasse C_1 oder C_2 zufällig ausgewählt und die Referenzvektoren \mathbf{w}_j der Karte gelernt werden ($C_1 \equiv$ „Vordergrundsegment“ und $C_2 \equiv$ „Hintergrundsegment“).

In einem selbstorganisierenden Prozeß, gemäß Abschnitt 4.1.2, bildet die Kohonenkarte die Referenzvektoren \mathbf{w}_j in dem durch die Komponenten aufgespannten Eingangsdatenraum \mathbf{X} aus. Bei diesem Prozeß wird die Distanz d_j , der während des Lernvorgangs dargebotenen Trainingsdaten $\tilde{\mathbf{x}}_1$ zu den Referenzvektoren \mathbf{w}_j nach Gleichung 4.1, berücksichtigt.

In Abbildung 4.7 ist der Lernvorgang an drei Komponenten $X_{1,2,3}$ verdeutlicht. Dargestellt ist die Lage einer zweidimensionalen Kohonenkarte (schwarz) im Raum während (Abbildung 4.7a,b) und nach Abschluß des Lernvorgangs (Abbildung 4.7c,d). Aufgetragen ist jeweils eine Schnittfläche parallel zu zwei Dimensionen im Raum. Außerdem eingezeichnet sind normierte Trainingsdaten $\tilde{\mathbf{x}}_1$ aus der Datenbank $\tilde{\mathbf{X}}_1$, in der Referenzkomponenten für Vordergrund- und Hintergrundsegmente abgelegt sind. Hellgrau sind dabei die Eingangsdaten der definierten Segmentklasse C_1 . Dunkelgrau eingezeichnet ist die Datenmenge der Klasse C_2 .

Aufgrund der hohen Dimensionalität ist eine Analyse der Komponenten im Eingangsdatenraum \mathbf{X} bezüglich der Klassenzugehörigkeit und dem Fusionsergebnis nicht durchführbar, wie

³Der Einsatz einer Kohonenkarte mit mehr Neuronen erbrachte keine wesentliche Verbesserung des Analyseergebnisses. Die Kartengröße von 40×20 Neuronen erbrachte ein gutes Fusionsergebnis. Bei geringerer Neuronenanzahl wurde der Eingangsdatenraum \mathbf{X} , der durch die Komponenten aufgespannt ist, nicht geeignet repräsentiert.

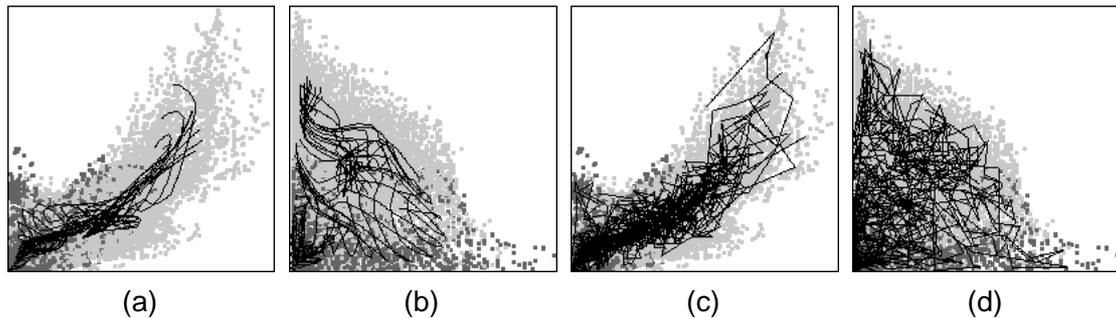


Abbildung 4.7: Ausbildung der Kohonenkarte (schwarz) im Eingangsdatenraum. Dargestellt ist jeweils eine Schnittfläche im Raum an zwei Beispielen zu zwei Zeitpunkten des Trainingsdurchlaufs. Darstellung (a) der zweiten über der ersten und (b) der dritten über der zweiten Dimension während des Trainingsvorgangs. Darstellung (c) der zweiten über der ersten und (d) der dritten über der zweiten Dimension nach Abschluß des Trainings. Die Trainingsdaten, die Vordergrund- und Hintergrundsegmente repräsentieren (hellgrau und dunkelgrau), sind ebenfalls eingezeichnet.

aus Abbildung 4.7 ersichtlich. Der Eingangsdatenraum \mathbf{X} wird jedoch mit Hilfe der Kohonenkarte geeignet repräsentiert. Es ist, wegen der geringen Dimensionalität der Karte, eine Analyse der Komponenten auf der Kohonenkarte realisierbar.

4.2.2 Eingangsdatenrepräsentation

Die Referenzvektoren der Kohonenkarte bilden sich während des selbstorganisierenden Prozesses (Lernprozeß) so aus, daß sie die Trainingsdaten $\tilde{\mathbf{x}}_1$ der Klassen $C_{1,2}$, welche zufällig aus der Trainingsdatenbank $\tilde{\mathbf{X}}_1$ gezogen wurden, geeignet repräsentieren. Anhand der Dichtekarte kann dieser Sachverhalt gemäß Abschnitt 4.1.3 überprüft werden.

In Abbildung 4.8 sind die Dichtekarten für die Menge der Trainingsdaten $\tilde{\mathbf{x}}_1$ und Testdaten $\tilde{\mathbf{x}}_2$ aus einer zweiten Datenbank $\tilde{\mathbf{X}}_2$ mit $\tilde{\mathbf{X}}_2 \subset \mathbf{X}$ (Datensatz mit 30000 Beispieldaten) dargestellt. Die Dichtekarten kodieren die Anzahl der $\tilde{\mathbf{x}}_{1,2}$ in den Voronoigebieten V_j der Kartenneurone.

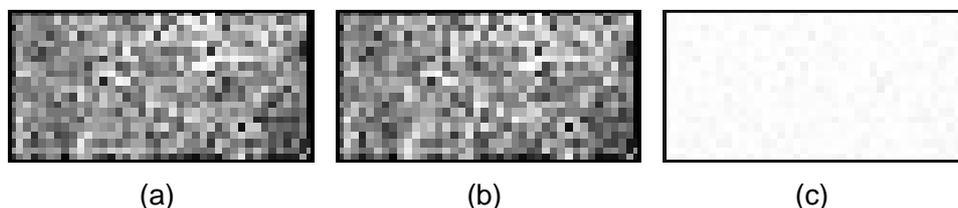


Abbildung 4.8: Dichtekarten für (a) den Trainingsdatensatz und (b) einen Testdatensatz. (c) Betrag der Differenz der Dichtekarten. Die Daten wurden zufällig aus einer Datenbank, in der Referenzkomponenten für Vordergrund- und Hintergrundsegmente abgelegt sind, gezogen.

Wie die Verteilungsdichte (kodierte durch den Grauwert) der Trainingsdaten $\tilde{\mathbf{x}}_1$ zeigt, wird der Datensatz $\tilde{\mathbf{X}}_1$ durch die Kohonenkarte gut repräsentiert. Eine quantitative Analyse ergibt, daß jedes Neuron im Mittel 37.5 Trainingsdaten $\tilde{\mathbf{x}}_1$ mit einer Standardabweichung $\sigma_x = 13.972572$ repräsentiert. Bei den weiterhin untersuchten Testdaten $\tilde{\mathbf{x}}_2$ ergeben sich keine signifikanten Unterschiede (Abbildung 4.8b,c).

Nach Beendigung des Lernvorgangs repräsentieren die Referenzebenen die statistische Verteilung der zugehörigen Komponenten. Die Referenzebenen der Karte werden genutzt, um den Einfluß einzelner Komponenten X_i auf das Fusionsergebnis zu analysieren.

4.2.3 Analyseergebnisse

Durch die Analyse der Referenzebenen, welche Repräsentanten für die Komponenten X_i sind (Abschnitt 4.1.3), kann die Segmentierungsgüte einzelner Komponenten X_i in Abhängigkeit der Klasse eines Neurons untersucht werden.

Es wurde eine Analyse verschiedener Komponenten X_i auf mehreren Grauwertbildsequenzen durchgeführt. Eine Auswahl von elf Komponenten X_1-X_{11} erweist sich dabei als geeignet. Die Auswahl ergibt sich aus verschiedenen Strategien, eine Segmentierung von Videobildmaterial aus Straßenverkehrsszenen durchzuführen. Es wurden Verfahren analysiert, die Konturverläufe und Texturverläufe im Bild sowie Bewegungsinformation beschreiben [34, 59, 24].

Mit den Ergebnissen der verarbeiteten Segmentierungsverfahren kann eine Beschreibung von relevanten Objekten vorgenommen werden. Abbildung 4.9 verdeutlicht diesen Sachverhalt anhand des Bildes eines Kraftfahrzeugs (Abbildung 4.9a) und den zugehörigen Verfahrensergebnissen (Abbildung 4.9b-f).

Die genutzten Segmentierungsverfahren und die ausgewählten Komponenten X_1-X_{11} werden im folgenden erläutert und eine Diskussion der ermittelten Ergebnisse an drei Beispielen durchgeführt.

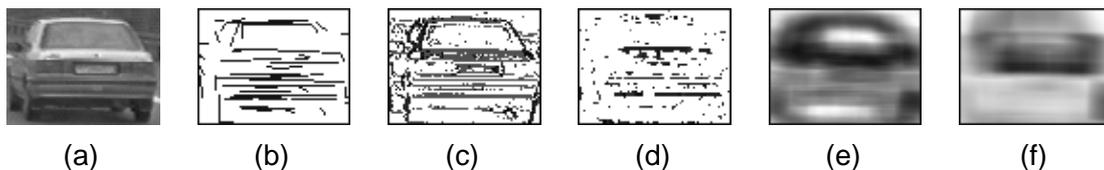


Abbildung 4.9: Beispiel für einen relevanten Bereich. Dargestellt ist ein PKW. (a) Grauwertverteilung, (b) Linienbild [23, 128], (c) LOC-Merkmale [31], (d) Flußfeld auf der Basis von LOC-Deskriptoren [54], (e) lokale Entropie [60, 59] und (f) lokale Varianz [54, 110].

Die für den Fusionsprozeß ausgewählten Komponenten X_1-X_{11} beschreiben

- die Lage von Linien im Bild, die eine Mindestlänge nicht unterschreiten (Abbildung 4.9b). Die verwendeten Linienbilder werden mit Hilfe einer *Contour Matching Hardware* [23] aus Grauwertbildern mit Hilfe eines Sobelfilteroperators [54], einer Ausdünnung und Konkatination⁴ erzeugt. Außerdem werden Linien aus einem in [128] vorgestellten Sy-

⁴Konkatination – aus dem Englischen: concatenate = verketteten, verknüpfen.

stem verwendet. Es wird eine Zuordnung von Linienpunkten zu Bildpixeln durchgeführt. Diese Zuordnung wird durch die Komponente $X_1 = K_1$ beschrieben.

- Konturverläufe im Grauwertbild (Abbildung 4.9c). Die verwendeten Konturverläufe werden unter Anwendung eines orientierungsabhängigen Nachbarschaftsoperators bestimmt. Dieser Operator (auch LOC-Operator⁵) wurde in [31] vorgestellt und erlaubt eine lokale Orientierungskodierung der Konturverläufe. Die LOC-Operatoren repräsentieren, im Gegensatz zu dem Linienbild, eine genauere Auflösung der Objektkonturen. Die Orientierungskodes werden zu vier Richtungskomponenten $K_2 - K_5$ zusammengefaßt. Die Richtungskomponenten beschreiben horizontale ($X_2 = K_2$), vertikale ($X_3 = K_3$), steigende ($X_4 = K_4$) und fallende ($X_5 = K_5$) Konturverläufe im Grauwertbild.
- die Bewegungsrichtung von Objekten durch Verschiebungsvektoren [54] (Abbildung 4.9d). Dabei wird die Verschiebungsrichtung für zwei aufeinanderfolgende Bilder einer Sequenz mit Hilfe von Deskriptoren⁶ bestimmt. Es werden vier Bewegungsgruppen $K_6 - K_9$, die den Bewegungsverlauf und Bewegungsgeschwindigkeit der Deskriptoren beschreiben, unterschieden. Die Bewegungsverläufe sind vom Fluchtpunkt aus betrachtet nach rechts oben ($X_6 = K_6$), links oben ($X_7 = K_7$), rechts unten ($X_8 = K_8$) und links unten ($X_9 = K_9$).
- den Informationsgehalt in einer Nachbarschaft durch die lokale Entropie E [60, 59] (Abbildung 4.9e). Mit diesem Maß kann die Texturiertheit von Bereichen bestimmt werden. Die lokale Entropie E ist durch die Komponente $X_{10} = E$ beschrieben.
- die lokale Varianz V der Grauwerte [54, 110] (Abbildung 4.9f). Dieses Maß erlaubt eine integrative Betrachtung der Grauwertverteilungen in einer Nachbarschaft im Bild. Die Komponente $X_{11} = V$ beschreibt die lokale Varianz im Bild.

Daraus ergibt sich ein Eingangsdatenvektor \mathbf{x} mit

$$\mathbf{x} = [K_1(p, q), \dots, K_9(p, q), E(p, q), V(p, q)]. \quad (4.2)$$

Im folgenden wird die Güte der Segmentierungsergebnisse an drei Komponenten im Zusammenhang mit dem Fusionsergebnis diskutiert. Diese Komponenten beschreiben vertikale Konturverläufe K_2 , Flußfeldvektoren K_9 , welche starke Bewegung im Bild vom Fluchtpunkt aus nach links unten beschreiben, und die lokale Entropie E in einem Bild. Die Diskussion der Ergebnisse bezieht sich auf die vorgestellte Kohonenkarte mit 40x20 Neuronen (Abschnitte 4.2.1 und 4.2.2). Die Diskussion wird jeweils an einem Grauwertbild aus einer typischen Straßenverkehrsszene durchgeführt.

In Abbildung 4.10 sind für diese drei Komponenten die Referenzebenen der Kohonenkarte dargestellt. Es ist an der unterschiedlichen Aktivität der einzelnen Ebenen zu erkennen, daß die Neurone verschiedene Regionen im Raum repräsentieren. Dabei kodiert der Grauwert die Lage eines Neurons in der jeweiligen Dimension.

⁵LOC – local orientation coding

⁶Deskriptor – aus dem Lateinischen: Deskrption = Beschreibung. Verschiebungsvektoren werden nur an markanten Punkten berechnet, die vorher durch einen *Deskriptor* ermittelt wurden.

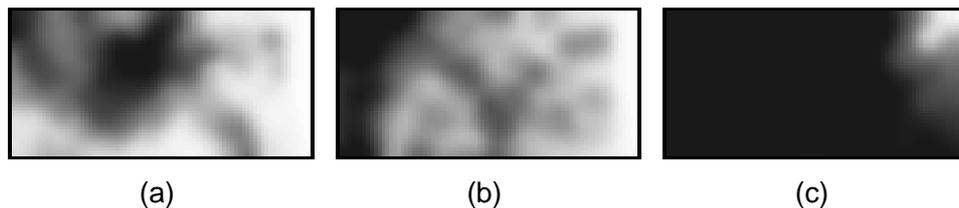


Abbildung 4.10: Referenzebenen dreier Komponenten: (a) Vertikale Konturmerkmale, (b) Flußfeldvektoren und (c) lokale Entropie. Hohe Aktivität entspricht einem dunklen Grauwert.

Aus der Aktivität der Neurone kann eine Relevanz für den Segmentierungsprozeß abgeleitet werden. Erhöhte Aktivität ist ein Indiz für das Vorhandensein eines Vordergrundsegments. Die Abbildungen 4.11, 4.12 und 4.13 verdeutlichen diesen Zusammenhang. Nach einer Schwellwertoperation werden Eingangsdatenvektoren aus Voronoigebieten mit erhöhter Aktivität in die Bildebene rückprojiziert. In den Abbildungen sind jeweils das Grauwertbild, die auf dem Grauwertbild berechnete Komponente und das Klassifikationsergebnis durch Rückprojektion der Eingangsdatenvektoren dargestellt. Grau sind dabei die Hintergrundsegmente. Im folgenden werden die Ergebnisse für die einzelnen Komponenten diskutiert.

Bei der Analyse vertikaler Konturverläufe ist zu erkennen (Abbildung 4.11b), daß bei einer Segmentierung mit der zugehörigen Komponente K_3 geometrische Strukturen im Bild segmentiert werden. Wesentliche Segmente in einer Straßenverkehrsszene sind Kraftfahrzeuge, die eine geometrische Struktur aufweisen.

Eine Segmentierung von Fahrzeugen kann folglich mit dem zugrundeliegenden Verfahren der richtungskodierenden Konturmerkmale durchgeführt werden (Abbildung 4.11c). Problematisch dabei ist, daß alle geometrischen Formen, wie Teile der Leitplanke, segmentiert werden. Durch die Aufteilung der Konturmerkmale in vier Richtungskomponenten K_2 - K_5 kann zwischen verschiedenen geometrischen Formen unterschieden werden. Die Komponente K_1 kann zur Auswahl ausgeprägter Konturverläufe (durch Festlegung einer Mindestlänge der Linien) eingesetzt werden und damit die Segmentierungsgüte von K_2 - K_5 verbessern.

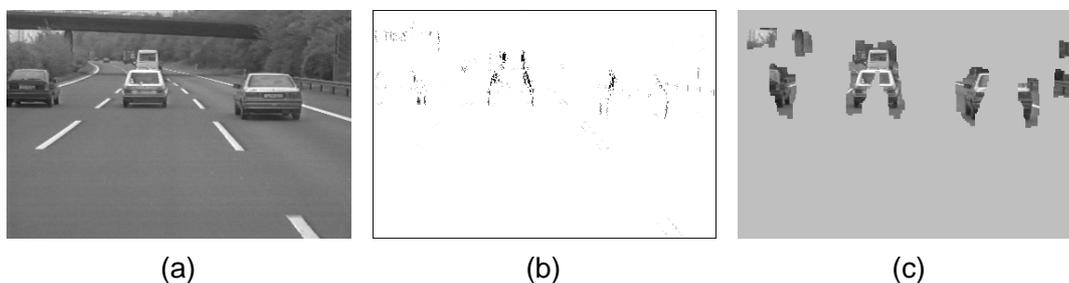


Abbildung 4.11: Segmentierung von Grauwertbildern. (a) Grauwertbild aus einer Autobahnszene, (b) Komponente vertikaler Konturmerkmale und (c) Segmentierungsergebnis durch Auswertung der Aktivitätsverteilung der zugehörigen Referenzebene.

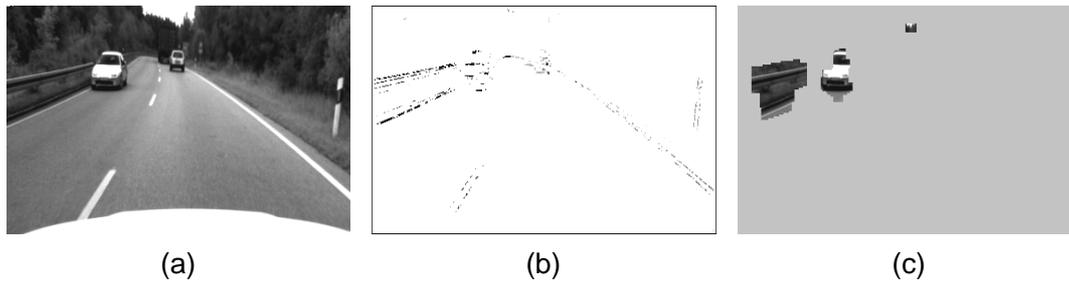


Abbildung 4.12: Segmentierung von Grauwertbildern. (a) Grauwertbild aus einer Landstraßenszene, (b) Komponente, welche hohe Relativbewegung kodiert und (c) Segmentierungsergebnis durch Auswertung der Aktivitätsverteilung der zugehörigen Referenzebene.

In Abbildung 4.12b ist das Ergebnisbild eines Flußfelddeskriptors K_9 für hohe Relativgeschwindigkeiten nach links unten dargestellt. Im analysierten Bild einer Landstraßenszene können mit dieser Komponente Bereiche hoher Relativgeschwindigkeit, vom Fluchtpunkt aus nach links unten, segmentiert werden.

Durch die Auswertung unterschiedlicher Richtungskomponenten K_6 - K_9 , kann zwischen verschiedenen Bewegungsrichtungen in der Sequenz unterschieden werden. Problematisch ist hier, daß aufgrund des bewegten Beobachters die Störanfälligkeit dieses Segmentierungsansatzes hoch ist und deshalb nur starke Relativbewegung ausgewertet werden kann. Dadurch werden Bereiche, die keine oder nur sehr geringe Relativbewegung zum Beobachter aufweisen, wie das Fahrzeug im Fernfeld, nicht segmentiert (Abbildung 4.12c).

Abbildung 4.13 beschreibt das Segmentierungsergebnis für ein Bild aus einer Innenstadtszene (Abbildung 4.13a). In diesem Bild werden anhand der Komponente der Entropie E (Abbildung 4.13b) Bereiche hoher Textur segmentiert (Abbildung 4.13c). Relevante Objekte in einer

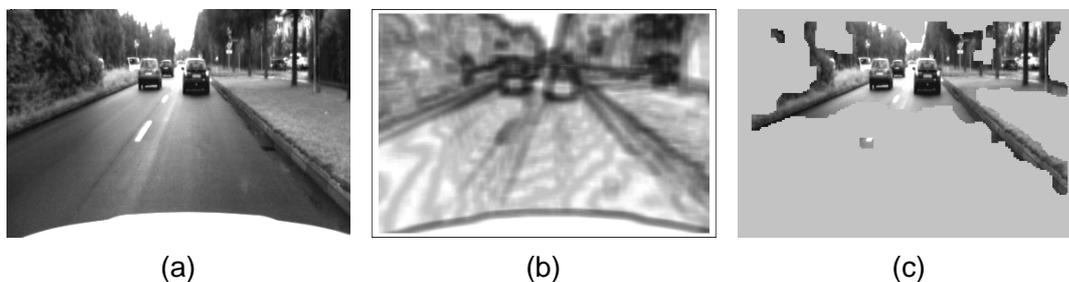


Abbildung 4.13: Segmentierung von Grauwertbildern. (a) Grauwertbild aus einer Innenstadtszene, (b) Komponente, die die lokale Entropie beschreibt und (c) Segmentierungsergebnis durch Auswertung der Aktivitätsverteilung der zugehörigen Referenzebene.

Straßenverkehrsszene sind in den segmentierten Bereichen zu finden. Problematisch ist jedoch, daß texturierte Bereiche des Hintergrunds auch segmentiert werden. Durch den Einsatz zweier integrativer Komponenten (Entropie E und Varianz V), welche differierende Eigenschaften aufweisen, kann eine Erhöhung der Segmentierungsgüte erreicht werden.

Die Segmentierungsergebnisse der Einzelkomponenten werden im folgenden im Zusammenhang mit dem Klassifikationsergebnis der Kohonenkarte betrachtet. Nach Abschluß der Selbstorganisation wird den Kartenneuronen eine definierte Klasse $C_{1,2}$ zugeordnet. Es bilden sich aufgrund der topologieerhaltenden Eigenschaften der Kohonenkarte zusammenhängende Gebiete gleicher Klasse aus.

Abbildung 4.14 zeigt die Klassenzugehörigkeit der Kartenneurone. Anhand dieser Karten kann die Zugehörigkeit der Eingangsdatenvektoren \mathbf{x} zu den Klassen C_1 und C_2 (Vordergrundsegment, Hintergrundsegment) festgelegt werden. Dabei bekommt ein Eingangsdatenvektor \mathbf{x} die Klasse $C_{1,2}$ des Neurons zugewiesen, in dessen Voronoigebiet V_j er liegt.

Die Überschneidung der Klassen C_1 und C_2 (Abbildung 4.14c) läßt den Schluß zu, daß Vordergrundsegmente und Hintergrundsegmente in Teilbereichen nicht eindeutig voneinander getrennt werden können. Aufgrund der Sicherheitsanforderung bei Fahrerassistenzsystemen werden die Neurone des Überschneidungsbereichs der Klasse C_1 zugeschlagen. Dadurch ist die Wahrscheinlichkeit höher, einen Bereich des Hintergrunds zu segmentieren, als ein relevantes Objekt zu unterdrücken. Dieser Sachverhalt ist aufgrund der Sicherheitsanforderungen gewünscht.

Wird eine Klassifikation der Eingangsdatenvektoren und eine anschließende Rückprojektion in die Bildebene durchgeführt, erhält man das in Abbildung 4.15 gezeigte Segmentierungsergebnis durch Fusion. Eine Verbesserung der Segmentierungsgüte wird erreicht. Grau dargestellt sind dabei Hintergrundsegmente. Es ist zu erkennen, daß sich die Segmentierungsergebnisse durch Fusion der beteiligten Komponenten, im Vergleich zum jeweiligen Einzelergebnis, verbessern.



Abbildung 4.14: Zugehörigkeit der Neurone (a) zu Klasse C_1 (Vordergrundsegmente), zu Klasse C_2 (Hintergrundsegmente) und (c) Überschneidungsbereiche der Klassen (schwarz dargestellt).

4.2.4 Diskussion

Es kann eine Analyse einzelner Komponenten mit einer Kohonenkarte durchgeführt und deren Einfluß auf das Fusionsergebnis untersucht werden. Das Analyseergebnis erlaubt eine konstruktive Auswahl geeigneter Komponenten für den Fusionsprozeß. Durch die nichtlineare Fusion mehrerer Komponenten verschiedener Segmentierungsverfahren wird ein Zugewinn bei der Güte des Segmentierungsergebnisses erreicht.

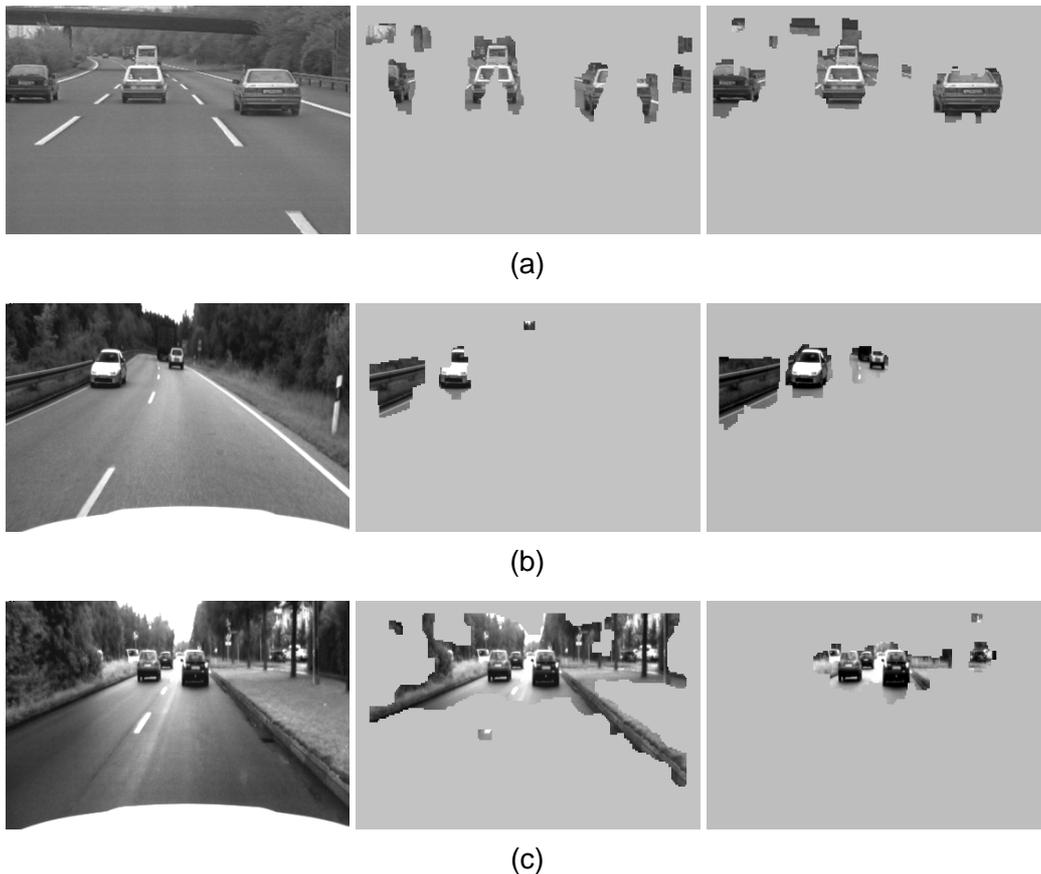


Abbildung 4.15: Segmentierung durch Fusion. Ergebnisse für ein Bild aus (a) einer Autobahnsequenz, (b) einer Landstraßensequenz und (c) einer Innenstadtsequenz. Dargestellt sind von links nach rechts das Grauwertbild, das Segmentierungsergebnis für die diskutierte Komponente (K_2 , K_8 und E) und das Segmentierungsergebnis durch Fusion.

Mit Hilfe der Segmentierung werden Datenreduktionen von über 70% erzielt. Dabei werden alle relevanten Objekte segmentiert. Die Höhe der Datenreduktion ist jedoch abhängig von den durch die Szene beschriebenen Objekten. Befinden sich mehrere relevante Objekte in der Szene, fällt die Datenreduktion geringer aus. Dies ist gewünscht, da alle relevanten Objekte aufgrund der Sicherheitsanforderungen segmentiert werden müssen. Die Ergebnisse der exemplarisch ausgewählten Bilder sind in Tabelle 4.1 dargestellt (eine ausführliche Diskussion der Segmentierungsergebnisse durch Fusion wird in Kapitel 5 durchgeführt).

Im gewählten Analyseverfahren ist eine Kartengröße von 40×20 nötig, um den Eingangsdatenraum geeignet zu repräsentieren. Die lokalen Zugehörigkeitsbereiche in Form von Voronoiflächen bedingen die große Anzahl von Kartenneuronen [8]. Während des Lernvorgangs muß demnach die Lage von 800 Kartenneuronen bestimmt werden. Bei der Segmentierungsaufgabe ist ein vollständiger Vergleich der präsentierten Eingangsdatenvektoren mit den Kartenneuronen notwendig, um die zugehörige Voronoifläche zu finden. Als Fusionsstruktur in technischen Applikationen ist der Einsatz einer Kohonenkarte daher nicht effizient.

	Autobahnbild	Landstraßenbild	Innenstadtbild
Datenreduktion auf	11.9%	8.06%	6.98%

Tabelle 4.1: Datenreduktion mit einer Segmentierung durch Fusion bei gleichzeitiger Segmentierung aller relevanter Objekte.

In Kapitel 5 wird ein Verfahren vorgestellt, welches die Fusion der in diesem Kapitel bestimmten Komponenten effizient ermöglicht. In diesem Verfahren wird eine Fusionsstruktur zugrunde gelegt, die auf einer verteilten bzw. globalen Repräsentation basiert. Es wird ein vorwärtsgekoppeltes neuronales Netz vom Typ eines Multilayer Perceptrons eingesetzt.

4.3 Zusammenfassung

In diesem Kapitel wird eine Analyse verschiedener Segmentierungsverfahren von Videobildern aus dem Straßenverkehr im Hinblick auf eine Fusion mehrerer Komponenten durchgeführt. Zentraler Bestandteil des entwickelten Verfahrens ist eine selbstorganisierende neuronale Karte (Kohonenkarte). Das zugrundeliegende Prinzip wird erläutert und der Vorteil von Kohonenkarten bei der Analyse der Segmentierungsverfahren wird abgeleitet. Insbesondere die Eigenschaften der Datenrepräsentation in Form von Merkmalskarten und die Topologieerhaltung motivieren den Einsatz einer Kohonenkarte für eine Analyse von Segmentierungsverfahren.

Für die Analyse wird das Segmentierungsproblem als Klassifikationsproblem formuliert und die Kohonenkarte zur Fusion von Komponenten eingesetzt. Eine Auswahl der verwendeten Komponenten wird mit Hilfe der Referenzebenen der Karte durchgeführt und die Fusion dieser motiviert.

Im folgenden Kapitel wird ein Verfahren vorgestellt, das die extrahierten Komponenten geeignet fusioniert.

Kapitel 5

Neuronale Sensordatenfusion

In diesem Kapitel wird ein neues Verfahren vorgestellt, welches eine Fusion verschiedener Komponenten einzelner Segmentierungsverfahren durchführt. Hierbei wird eine Erhöhung der Segmentierungsgüte durch Fusion gezeigt. Es wird in Bildkoordinaten operiert, da die Segmentierungsergebnisse im Hinblick auf nachfolgende Bildverarbeitungsalgorithmen erzeugt werden.

Ein wesentlicher Aspekt bei der Erkennung von Objekten ist der Segmentierungsprozeß. In Kapitel 4 wird eine Methode vorgestellt, welche eine Analyse verschiedener Segmentierungsverfahren zuläßt. Die dort verfolgte Methode erlaubt eine konstruktive Analyse verschiedener Komponenten im Hinblick auf das Fusionsergebnis. Jedoch ist die gewählte Struktur für die Fusion von mehreren Komponenten nicht effizient. Die lokalen Zugehörigkeitsbereiche durch Voronoiflächen bedingen eine große Anzahl von Neuronen, um den Eingangsdatenraum zu repräsentieren.

Bei der Auswahl einer geeigneten Struktur, die die extrahierten Komponenten fusioniert, steht deshalb in diesem Kapitel, neben der Robustheit und Flexibilität, die Effizienz der gewählten Fusionsstruktur im Vordergrund. Aufbauend auf die in Kapitel 4 extrahierten Komponenten, wird eine Fusion dieser Komponenten durchgeführt. Eine einfache Hinzunahme neuer Komponenten des gleichen oder eines anderen Sensortyps sowie eine Adaption der Strukturparameter bei veränderter Güte der Sensordaten wird ermöglicht.

Das in diesem Kapitel vorgestellte Verfahren verfolgt eine parallele Fusionstrategie. Diese Strategie gewährleistet eine geringere Störanfälligkeit gegenüber Ausfällen von Sensordaten. Bei paralleler Verarbeitung baut die Verarbeitung einzelner Komponenten nicht direkt aufeinander auf [15].

Wesentlicher Bestandteil des realisierten Verfahrens ist ein vorwärtsgekoppeltes neuronales Netz vom Typ eines Multilayer Perceptrons. Aufbauend auf die in Kapitel 4 durchgeführte Analyse werden integrative, differentielle und flußfeldbasierte Komponenten fusioniert.

Es wird gezeigt, daß durch die Auswahl eines vorwärtsgekoppelten neuronalen Netzes als zentrales Element des Fusionsprozesses, eine flexible Erweiterung des Eingangsdatenraums, eine Adaption der Kopplungsgewichte und eine Berücksichtigung von Redundanzen im Eingangsdatenraum ermöglicht werden. Das vorwärtsgekoppelte neuronale Netz stellt durch seine nicht-linearen Abbildungseigenschaften eine robuste Kopplungsstruktur mit geringer Störanfälligkeit dar. Der hierarchische Aufbau und der Einsatz globaler Trennungsebenen erlaubt eine effiziente Kodierung der Fusionsstruktur.

In diesem Kapitel wird in Abschnitt 5.1 der Ansatz der vorwärtsgerichteten neuronalen Netze diskutiert. In Abschnitt 5.2 wird das entwickelte Verfahren vorgestellt. Die Tragfähigkeit des Verfahrens wird an mehreren Sequenzen in Abschnitt 5.3 gezeigt. Eine flexible Erweiterbarkeit wird durch zwei Sensorfusionsprozesse demonstriert. Das Kapitel 5 schließt mit einer Zusammenfassung in Abschnitt 5.4.

5.1 Vorwärtsgerichtete neuronale Netze

Vorwärtsgerichtete neuronale Netze kommen bei technischen Anwendungen häufig zum Einsatz. Sie sind aus mehreren Schichten künstlicher Neurone aufgebaut. Bei vorwärtsgerichteten neuronalen Netzen wird eine Transformation der Eingangsdatenvektoren durchgeführt. Der Eingangsdatenvektor \mathbf{x} der Dimension n wird durch das neuronale Netz zum Ausgang mit dem Signal \mathbf{y} der Dimension m propagiert. Damit repräsentiert ein vorwärtsgerichtetes neuronales Netz eine Funktion

$$f : \mathbf{x} \rightarrow \mathbf{y}, \text{ mit } \mathbf{x} \in \mathbb{R}^n, \mathbf{y} \in \mathbb{R}^m.$$

Eine Rückkopplung findet nicht statt. Häufig werden für die Transformation der Eingangsdaten sogenannte versteckte Schichten in die Netzstruktur eingebaut. In Abbildung 5.1 ist ein solches Netz mit einer versteckten Schicht dargestellt.

Beispiele für vorwärtsgerichtete neuronale Netze sind Multilayer Perceptrons und radiale Basisfunktionsnetzwerke [47]. In radialen Basisfunktionsnetzwerken haben die Neurone lokale Zugehörigkeitsbereiche. Bei Multilayer Perceptrons beschreiben die Gewichte zu den Neuronen einer versteckten Schicht globale Trennungsebenen in einem Hyperraum, die den Datenraum aufteilen [8].

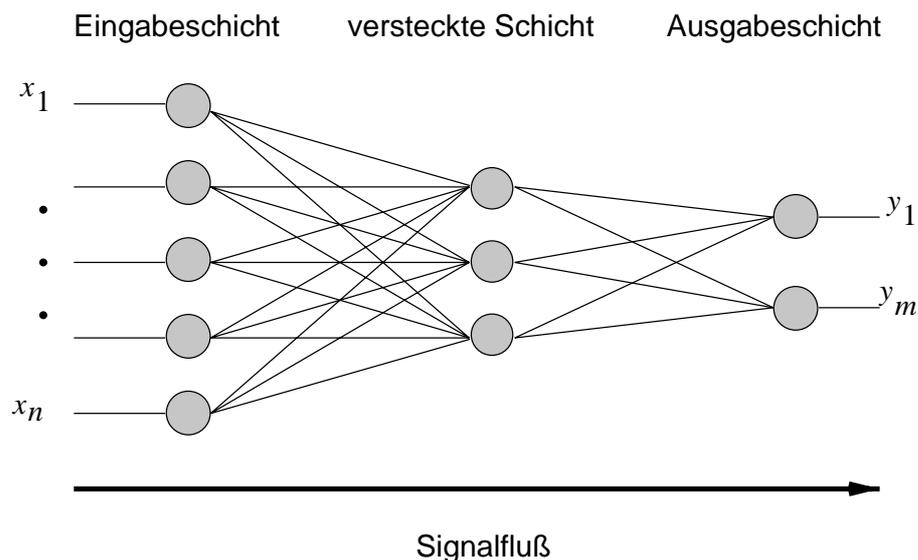


Abbildung 5.1: Vorwärtsgerichtetes, vollverbundenes neuronales Netz mit einer versteckten Schicht. Der Signalfluß ist durch den Pfeil dargestellt.

Multilayer Perceptrons werden in der Literatur als Schätzer wie auch als Klassifikatoren verwendet [47]. Hat ein Multilayer Perceptron mindestens eine versteckte Schicht, kann es als sogenannter universeller Approximator eingesetzt werden [8]. Das bedeutet, daß eine beliebige Funktion mit diesem Netztyp bei ausreichender Anzahl von Neuronen beliebig gut approximiert werden kann. Multilayer Perceptrons zeichnen sich durch gute Generalisierungsfähigkeit aus. Im folgenden wird deshalb dieser Typ genauer betrachtet.

5.1.1 Multilayer Perceptron

Multilayer Perceptrons lassen sich auf das Perceptron, welches in [106] vorgestellt wurde, zurückführen. Das Perceptron besteht aus einem Neuron, das mit einer Schwellenfunktion in der Lage ist, zwei linear separierbare Musterklassen zu trennen. Die Lage der Trennungsebene im Raum wird durch Synapsengewichte \mathbf{w} bestimmt. Ein Multilayer Perceptron hingegen besteht aus mehreren einzelnen Perceptrons, welche in Schichten organisiert sind. Jedes Neuron einer Schicht κ von J Neuronen mit der Ausgangsaktivität s_j^κ , $j = \{1, 2, \dots, J\}$, bekommt die gewichteten Ausgangsaktivitäten der Neurone $s_i^{(\kappa-1)}$, $i = \{0, 1, \dots, I\}$, der Vorgängerschicht ($\kappa - 1$) mit I Neuronen als Eingangsvektor. Dabei entspricht $(w_{j0}s_0^{(\kappa-1)})$ einem Bias. Das Neuron propagiert die Summe v_j^κ der mit w_{ji} gewichteten Einzelaktivitäten $s_i^{(\kappa-1)}$, über eine Aktivierungsfunktion φ mit

$$s_j^\kappa = \varphi(v_j^\kappa) = \varphi\left(\sum_{i=0}^I w_{ji}s_i^{(\kappa-1)}\right) \quad (5.1)$$

zur nächsten Schicht ($\kappa + 1$) weiter. In Abbildung 5.2 ist ein Modell eines Neurons veranschaulicht, welches die Ausgangsaktivitäten $s_i^{(\kappa-1)}$ einer Vorgängerschicht ($\kappa - 1$) verarbeitet. Die letzte Schicht eines Multilayer Perceptrons bildet die Ausgabeschicht mit den Ausgangsaktivitäten \mathbf{y} . Die Anzahl m der Ausgänge y_k , $k = \{1, 2, \dots, m\}$, definiert die Dimension des Ausgangsdatenvektors \mathbf{y} . In der Eingangsschicht findet keine Verarbeitung des Eingangsdatenvektors \mathbf{x} statt. Hier wird nur der Eingangsdatenvektor \mathbf{x} mit den Elementen x_i und $i = \{1, 2, \dots, n\}$ allen Neuronen der folgenden Schicht bereitgestellt.

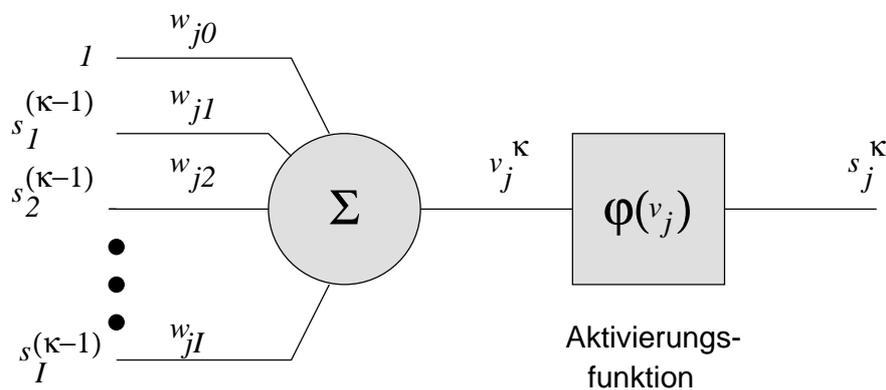


Abbildung 5.2: Neuronenmodell eines Multilayer Perceptrons: Der Datenvektor \mathbf{s} wird gewichtet summiert und über eine Aktivierungsfunktion φ zum Ausgang propagiert.

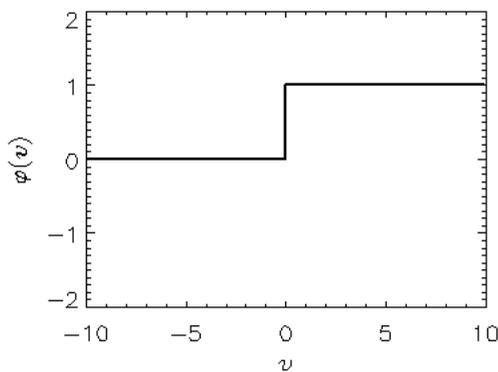
Die Aktivierungsfunktion φ eines Neurons ist normalerweise eine Funktion, die eine nichtlineare Abbildung der jeweiligen v durchführt. Als Aktivierungsfunktionen haben sich folgende Funktionen in der Literatur etabliert:

- Sprungfunktion (Abbildung 5.3a):

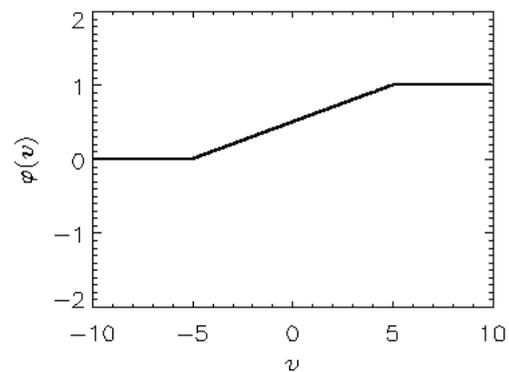
$$\varphi(v) = \begin{cases} 0 & : v \leq 0 \\ 1 & : v > 0 \end{cases} \quad (5.2)$$

- Lineare Funktion (Rampenfunktion) im Bereich $[-c, c]$ (Abbildung 5.3b):

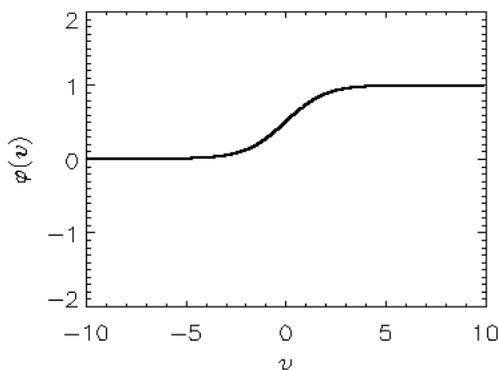
$$\varphi(v) = \begin{cases} 0 & : v < -c \\ \frac{1}{2c}v + 0.5 & : -c \leq v < c \\ 1 & : v \geq c \end{cases}$$



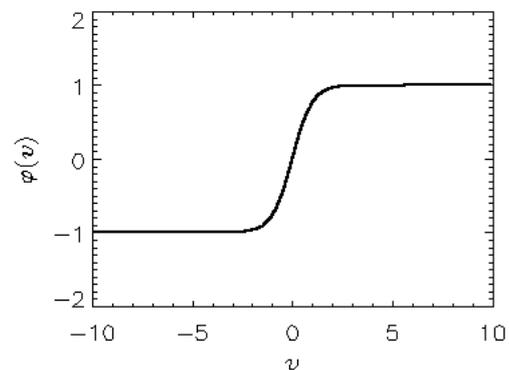
(a)



(b)



(c)



(d)

Abbildung 5.3: Aktivierungsfunktionen. (a) Sprungfunktion, (b) Rampenfunktion, (c) sigmoide Funktion und (d) Tangens Hyperbolicus.

- Sigmoid (Anhang A.1, Abbildung 5.3c):

$$\varphi(v) = \xi(v) = \frac{1}{1 + e^{-cv}}, \quad (5.3)$$

- Tangens Hyperbolicus:

$$\varphi(v) = \tanh(v), \quad (5.4)$$

wobei c ($c \in \mathbb{R}^+$) ein wählbarer Parameter ist und v den Eingangswert der Aktivierungsfunktion $\varphi(v)$ mit $v \in \mathbb{R}$ darstellt. Der Tangens Hyperbolicus (Gleichung 5.4) geht durch eine lineare Transformation aus $\xi(v)$ hervor (Anhang A.1).

In technischen Applikationen werden als Aktivierungsfunktionen häufig Sigmoid eingesetzt (Gleichung 5.3, Abbildung 5.3c). Sigmoid streben für große Werte v gegen 1 und für kleine Werte gegen 0. Sie haben einen stetig differenzierbaren Verlauf (Anhang A.1). Etablierte Lernverfahren machen sich diese Eigenschaften zunutze, wie beispielsweise der Backpropagation-Algorithmus (Anhang A.2).

5.1.2 Lernverfahren

Mit Hilfe von Lernverfahren wird eine Anpassung der Synapsengewichte \mathbf{w} eines neuronalen Netzes vorgenommen. Ziel ist das Minimieren einer definierten Fehlerfunktion zwischen dem Ausgangsdatenvektor \mathbf{y} mit den Elementen y_k des neuronalen Netzes und einem Referenzvektor \mathbf{o} mit den Elementen o_k , so daß die Abbildung des Eingangsdatenraums \mathbf{X} auf dem Zieldatenraum \mathbf{Y} optimal ist.

Als Fehlerfunktional für den Fehler an einem Ausgangsneuron mit der Aktivität y_k wird häufig der über mehrere Trainingsbeispiele L gemittelte quadratische Fehler E mit

$$E = \frac{1}{L} \sum_{l=1}^L (o_k(l) - y_k(l))^2, \quad (5.5)$$

als Gütemaß für den Fortschritt eines Lernverfahrens verwendet. Bei der Wahl eines Fehlerfunctionals nach Gleichung 5.5 wird von einer monomodalen Verteilung der Fehler ausgegangen. Liegt keine monomodale Verteilung der Fehler vor, werden andere Funktionen zur Bewertung der Qualität eines Lernverfahrens eingesetzt. Einen Überblick geben [47, 8].

Um eine gewünschte Zielfunktion mit Hilfe eines Multilayer Perceptrons approximieren zu können, werden verschiedene Lernverfahren eingesetzt. Das bekannteste ist das sogenannte Error-Backpropagation Verfahren (Backpropagation). Bei diesem Verfahren, das auf Arbeiten in [108] zurückgeht, wird abhängig von dem nach den Synapsengewichten \mathbf{w} abgeleiteten Fehler E eine Veränderung der Gewichte durchgeführt.

Wird ein Multilayer Perceptron mit einer versteckten Schicht betrachtet, kann für die Änderung der Gewichte w_{kj} aus \mathbf{w} von der versteckten Schicht zur Ausgangsschicht folgender Änderungsterm Δw_{kj} berechnet werden. Es gilt für die Änderung $\Delta w_{kj}(t+1)$ zum Zeitpunkt $(t+1)$:

$$\Delta w_{kj}(t+1) = w_{kj}(t+1) - w_{kj}(t) = -\alpha \frac{\partial E(t)}{\partial w_{kj}(t)}.$$

Dabei beschreibt α einen Lernfaktor mit $\alpha \in \mathbb{R}^+$. Wählt man als Aktivierungsfunktion $\varphi(v)$ der Neurone eine sigmoide Funktion mit $\varphi(v) = \xi(v)$ und $c = 1$ kann

$$\Delta w_{kj}(t+1) = -\alpha \cdot e_k(t) \cdot y_k(t) \cdot [1 - y_k(t)] \cdot s_j(t)$$

hergeleitet werden (Anhang A.2, Gleichung A.7). Die Variable e_k definiert den Fehler an einem Ausgangsneuron mit $e_k = o_k - y_k$. Der Änderungsterm Δw_{kj} ist weiterhin abhängig von der Aktivität s_j eines in der versteckten Schicht liegenden Neurons.

Für Neuronengewichte w_{ji} von der Eingangsschicht zu der versteckten Schicht kann zum Zeitpunkt $(t+1)$ die von dem Backpropagation-Lernverfahren geforderte Änderung

$$\begin{aligned} \Delta w_{ji}(t+1) &= w_{ji}(t+1) - w_{ji}(t) = -\alpha \frac{\partial E(t)}{\partial w_{ji}(t)} \\ &= -\alpha \sum_{k=0}^K e_k(t) \cdot y_k(t) \cdot [1 - y_k(t)] \cdot w_{kj}(t) \cdot s_j(t) \cdot [1 - s_j(t)] \cdot x_i(t) \end{aligned}$$

bei der Wahl sigmoider Aktivierungsfunktionen $\varphi(v) = \xi(v)$ mit $c = 1$ abgeleitet werden (Anhang A.2, Gleichung A.8). Dieser Änderungsterm ist zusätzlich abhängig vom Eingangswert x_i . Durch Einfügen eines Momententerms, der mit einem Gewichtungparameter ζ mit $\zeta \in \mathbb{R}^+$ multipliziert wird, kann bei schlechter Konditionierung der Eingangsdaten \mathbf{x} Oszillation während des Lernvorgangs mit

$$\Delta w(t+1) = w(t+1) - w(t) + \zeta \Delta w(t),$$

verhindert werden [8]. Weitere Lernverfahren oder andere Variationen dieses Lernansatzes können problemspezifisch genutzt werden. Einen Überblick über Lernverfahren und entsprechende Erweiterungsmöglichkeiten geben [47, 8, 68].

5.1.3 Datengetriebener Entwurf

Die Synapsengewichte \mathbf{w} neuronaler Netze können in einem modellfreien datengetriebenen Entwurf mit Hilfe von Beispielen gelernt werden. Es ist dabei möglich, lineare und nichtlineare Funktionen abzubilden.

Um eine Abbildungsfunktion unter Anwendung eines Multilayer Perceptrons lösen zu können, sind Datenvektoren $\tilde{\mathbf{x}}$, die den Eingangsdatenraum \mathbf{X} beschreiben und zugehörige Referenzvektoren $\tilde{\mathbf{o}}$ für den Zieldatenraum \mathbf{Y} in einem Datensatz $\tilde{\mathbf{X}}$ notwendig. Anhand der Datenvektoren $\tilde{\mathbf{x}}$ kann die Anpassung der Synapsengewichte \mathbf{w} mit Lernverfahren durchgeführt und der Lernfortschritt validiert werden.

Abbildung 5.4a zeigt ein Beispiel für eine Klassifikationsaufgabe. In diesem Beispiel sollen Eingangsdatenvektoren \mathbf{x} aus einem zweidimensionalen Datenraum \mathbf{X} mittels einer Abbildungsfunktion zwei Klassen $C_{1,2}$ ($C_1 \equiv$ „dunkelgrau“, $C_2 \equiv$ „hellgrau“) zugeordnet werden. Dafür werden Datenvektoren $\tilde{\mathbf{x}}_1$ (x und y Koordinaten) zufällig aus einem gegebenen Datensatz $\tilde{\mathbf{X}}_1$ (Trainingsdatensatz) aus einer Stichprobe dreier Bereiche dieser Klassen $C_{1,2}$ im Datenraum \mathbf{X} mit $\tilde{\mathbf{X}}_1 \subset \mathbf{X}$ gezogen. Die Trainingsvektoren $\tilde{\mathbf{x}}_1$ werden einem Multilayer Perceptron der Art 2-15-1 (zwei Eingangsneurone, 15 versteckte Neurone¹ und ein Ausgangsneuron) als Trainingsdaten

¹Die Anzahl der versteckten Neuronen ist abhängig von der Ordnung der abzubildenden Funktion.

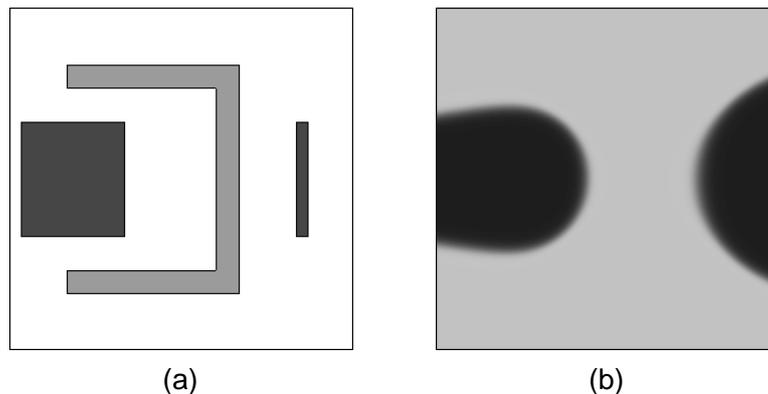


Abbildung 5.4: Zweisechichtiges Multilayer Perceptron als Klassifikator eingesetzt. (a) Trainingsdaten ordnen den Koordinaten x und y die Klassen C_1 („dunkelgrau“) und C_2 („hellgrau“) zu. Nicht verwendete Daten sind weiß markiert. (b) Die Aufteilung des Eingangsdatenraums nach Klassen C_1 und C_2 , definiert durch die Aktivität des Ausgangsneurons.

präsentiert. Unter Anwendung eines Lernverfahrens werden die Synapsengewichte \mathbf{w} des Netzes gelernt. Ein Testdatensatz $\tilde{\mathbf{X}}_2$ mit Testdaten $\tilde{\mathbf{x}}_2$ und $\tilde{\mathbf{X}}_2 \subset \mathbf{X}$ aus einer zweiten Stichprobe der drei betrachteten Bereiche, erlaubt die Validierung des Trainingsverlaufs. Bei dem dargestellten Beispiel werden 5000 Trainingsdaten und 5000 Testdaten verwendet. Weiß kodierte Koordinaten werden zum Training nicht betrachtet. Die Klassenzugehörigkeit zu C_1 und C_2 des gesamten Eingangsdatenraums \mathbf{X} nach dem Training, kodiert durch die Aktivität des Ausgangsneurons, ist in Abbildung 5.4b gegeben. Es ist eine nichtlineare Separierbarkeit von Eingangsdatenvektoren \mathbf{x} aus \mathbf{X} und eine Zuordnung dieser zu den zwei Klassen $C_{1,2}$ möglich.

Für das Training der Synapsengewichte \mathbf{w} wurde der Backpropagation-Algorithmus verwendet [47]. In Abbildung 5.5 ist der Fehler E nach Gleichung 5.5 während des Lernvorgangs über der Zeit t aufgetragen. Da keine verrauschten Daten vorliegen und keine Überschneidungen der Klassen $C_{1,2}$ existieren, erreicht der Fehler E nach $2 \cdot 10^6$ Zeitschritten den Wert $E = 0$.

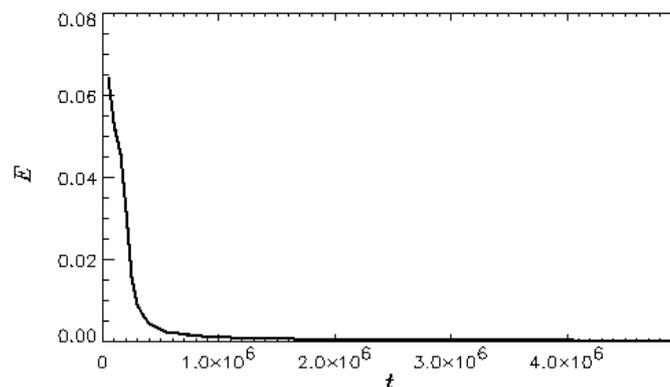


Abbildung 5.5: Zweisechichtiges Multilayer Perceptron als Klassifikator. Aufgetragen ist der über den Trainings- und den Testdatensatz gemittelte quadratische Fehler des Ausgangsneurons über der Zeit während des Lernvorgangs.

5.1.4 Generalisierungsfähigkeit

Ein Netz ist dann generalisierungsfähig, wenn es die den Daten \mathbf{x} zugrundeliegende Funktion abbilden kann. Liegen verrauschte Daten \mathbf{x} vor oder gibt es Klassenüberschneidungen bei einer Klassifikationsaufgabe, stellt sich häufig das Problem der mangelnden Generalisierungsfähigkeit der gelernten Abbildungsfunktion des neuronalen Netzes durch übertrainieren. Um hohe Generalisierungsfähigkeit zu erreichen, können verschiedene Methoden eingesetzt werden. Im folgenden wird das Problem des Übertrainierens beschrieben. Weiterhin werden zwei Methoden diskutiert, die einen Lösungsweg aufzeigen, um hohe Generalisierungsfähigkeit zu erreichen. Einen Überblick über verschiedene Vorgehensweisen geben [47, 8, 87].

Für die Durchführung eines Lernvorgangs wird ein gegebener Datensatz $\tilde{\mathbf{X}}$ in einen Trainingsdatensatz $\tilde{\mathbf{X}}_1$, einen Testdatensatz $\tilde{\mathbf{X}}_2$ und einen Generalisierungsdatensatz $\tilde{\mathbf{X}}_3$ unterteilt. Mit Hilfe der Trainingsdaten $\tilde{\mathbf{x}}_1 \in \tilde{\mathbf{X}}_1$ werden die Synapsengewichte \mathbf{w} des Netzes gelernt. Die Testdaten $\tilde{\mathbf{x}}_2 \in \tilde{\mathbf{X}}_2$ werden genutzt, um den Lernfortschritt des neuronalen Netzes zu überprüfen. Die Generalisierungsdaten $\tilde{\mathbf{x}}_3 \in \tilde{\mathbf{X}}_3$ dienen dazu, den Generalisierungsfehler festzustellen.

Werden alle Synapsengewichte \mathbf{w} des Netzes während der Initialisierungsphase zufällig ausgewählt, haben die Fehler der gegebenen Datensätze den gleichen Erwartungswert. Während der Trainingsphase wird am Anfang der Fehler für alle Datensätze reduziert, bis die globalen Minima der Fehler von Generalisierungsdatensatz $\tilde{\mathbf{X}}_3$ und Testdatensatz $\tilde{\mathbf{X}}_2$ überschritten sind. Ab diesem Zeitpunkt findet das sogenannte Übertrainieren statt. Es werden die speziellen Eigenschaften des Trainingsdatensatzes $\tilde{\mathbf{X}}_1$ gelernt. In Abbildung 5.6 ist dieser Sachverhalt an einem Klassifikationsbeispiel verdeutlicht.

In Abbildung 5.6a sind Trainingsdaten $\tilde{\mathbf{x}}_1$ eines Datensatzes $\tilde{\mathbf{X}}_1$ dargestellt. Die Klassenzugehörigkeit der $\tilde{\mathbf{x}}_1$ ergibt sich aus der Farbe. Es werden zwei Klassen (weiß und schwarz) unterschieden. Durch ein gradientenbasiertes Lernverfahren (Kapitel 5.1.2) wird der Fehler der Abbildungsfunktion des neuronalen Netzes reduziert. Erreicht der Fehler des Generalisierungs-

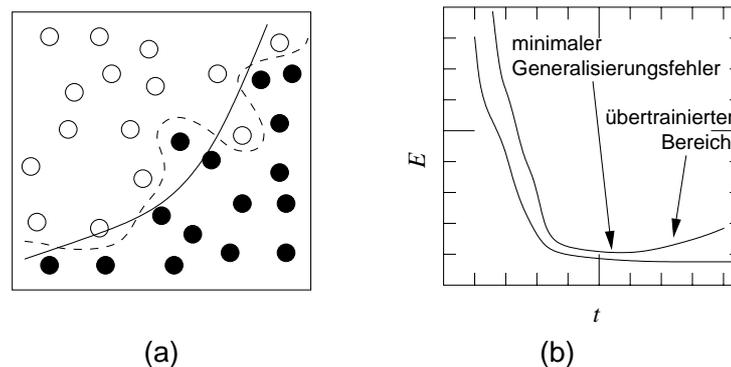


Abbildung 5.6: Prinzipbilder eines Klassifikationsproblems. (a) Datensatz mit zwei Datenklassen (weiß, schwarz). Die durchgezogene Linie beschreibt die gewünschte Klassengrenze. Gestrichelt dargestellt ist die Klassengrenze für eine Abbildungsfunktion, die übertrainiert wurde. (b) Zeitlicher Verlauf des Fehlers des Generalisierungsdatensatzes und des Trainingsfehlers für ein übertrainiertes Netz. Mit Pfeilen sind die Stellen des globalen Minimums des Generalisierungsdatensatzes und der übertrainierte Bereich markiert.

datensatzes $\tilde{\mathbf{X}}_3$ das absolute Minimum, ist die optimale Abbildungsfunktion gelernt. In Abbildung 5.6a entspricht die durchgezogene Linie der Klassengrenze und damit der optimalen Abbildungsfunktion. Abbildung 5.6b beschreibt den Fehlerverlauf des Generalisierungsfehlers (oben) und des Trainingsfehlers (unten) während des Lernvorgangs. Wird der Lernfortschritt nicht abgebrochen, findet ein Übertrainieren statt. Die Ursache liegt in Ausreißern im Datensatz $\tilde{\mathbf{X}}_1$, die den tatsächlichen Klassenzusammenhang falsch kodieren (Abbildung 5.6a, gestrichelte Linie). Um das Problem des Übertrainierens zu umgehen, gibt es die Möglichkeit, das Lernverfahren an dem Zeitpunkt abzubrechen, an dem der Generalisierungsfehler minimal ist. Diese Methode wird als *early stopping* bezeichnet. Um diesen Zeitpunkt zu finden, werden die Testdaten $\tilde{\mathbf{X}}_2$ als Schätzdaten für die Suche des minimalen Generalisierungsfehler eingesetzt. Sie beeinflussen damit den Lernvorgang der Synapsengewichte \mathbf{w} . Aus diesem Grund wird zwischen dem Testdatensatz $\tilde{\mathbf{X}}_2$ und dem Generalisierungsdatensatz $\tilde{\mathbf{X}}_3$ unterschieden. Die Schätzung des globalen Minimums ist nicht trivial, da während des Lernvorgangs lokale Minima der Fehlerfunktion aller Daten durchlaufen werden und eine Gewißheit über das Erreichen des globalen Minimums des geschätzten Generalisierungsfehlers nur im nachhinein möglich ist [95].

Eine andere Möglichkeit, das Übertrainieren eines neuronalen Netzes zu umgehen, um hohe Generalisierungsfähigkeit zu erreichen, besteht darin, die Netzstruktur geeignet zu wählen. Bei einer bekannten Ordnung des zu lösenden Problems kann die Struktur des Netzes angepaßt werden. Die Ordnung der Abbildungsfunktion wird dadurch eingeschränkt [8].

Für eine Untersuchung der Ordnung des zu lösenden Problems kann das *gain*-Spektrum² analysiert werden [33]. Dabei ist der Verstärkungsfaktor g_i eines versteckten Neurons definiert durch

$$g_j = \sqrt{\sum_{i=1}^n w_{ji}^2}, \quad (5.6)$$

für n Eingangskomponenten und für J versteckte Neurone mit $j = \{1, 2, \dots, J\}$ gemäß Gleichung 5.1. Die Variable w beschreibt ein Synapsengewicht. Mit Hilfe der Verstärkungsfaktoren g_j , die eine definierte Schwelle überschreiten, wird die benötigte Anzahl versteckter Neurone bestimmt.

Ein weiterer Ansatz, der zur Bestimmung der benötigten Anzahl versteckter Neurone genutzt werden kann, wird in [81] vorgestellt. Dieser Ansatz baut auf Verfahren aus der klassischen Signaltheorie zur Schätzung der Ordnung in linearen Modellen auf.

5.1.5 Relevanz einzelner Eingänge

Die Relevanz einzelner Eingänge eines Multilayer Perceptrons kann anhand der Verstärkungsfaktoren g_i mit

$$g_i = \sqrt{\sum_{j=1}^J w_{ji}^2}, \quad (5.7)$$

und sowie der Varianz

$$\text{VAR}_i = \frac{1}{J} \sum_{j=1}^J (w_{ji} - \bar{w}_i)^2, \quad (5.8)$$

²gain – abgeleitet aus dem Englischen: gain = Verstärkung.

der zugehörigen Einzelgewichte w_{ji} mit dem Mittelwert \bar{w}_i untersucht werden. Nach [32] werden nichtrelevante Eingänge i mit einem Verstärkungsfaktoren $g_i \rightarrow 0$ und geringer Varianz $VAR_i \rightarrow 0$ beschrieben, falls mittelwertfreie Trainingsdaten $\tilde{\mathbf{x}}_1$ ($\bar{w}_i = 0$) mit einer Standardabweichung $\sigma_x = 1$ zum Training der Synapsengewichte \mathbf{w} genutzt werden.

Mit Hilfe dieses Verfahrens kann eine Bewertung der einzelnen Komponenten eines Eingangsdatenvektors \mathbf{x} durchgeführt werden.

Der vorgestellte Ansatz der Multilayer Perceptrons wird in diesem Kapitel zur Fusion der in Kapitel 4 extrahierten Komponenten eingesetzt. In Abschnitt 5.2 wird das realisierte Verfahren vorgestellt und in Abschnitt 5.3 die Ergebnisse diskutiert.

5.2 Fusion der Sensordaten

In diesem Abschnitt wird ein vorwärtsgerichtetes neuronales Netz als Kopplungsstruktur in dem Datenfusionsprozeß eingesetzt. In einem parallelen Prozeß werden einzelne Komponenten der in Kapitel 4 untersuchten Segmentierungsverfahren fusioniert, um eine robuste Segmentierung von Videobilddaten aus dem Straßenverkehr mit bewegtem Beobachter zu realisieren. Es wird der Typ eines Multilayer Perceptrons verwendet.

5.2.1 Fusion auf Daten- und Merkmalsebene

Als Eingangsdaten werden die Ergebnisdaten verschiedener Segmentierungsverfahren (Komponenten) verwendet. Dabei werden die Analyseergebnisse aus Kapitel 4 nach Gleichung 4.2 berücksichtigt. Die Fusion erfolgt auf Daten- und Merkmalsebene [15]. Es wird eine Vorverarbeitung der konturbasierten Komponenten K_i (Linienpixel K_1 , Konturkomponenten K_2 - K_5 und Flußfelddeskriptoren K_6 - K_9) durchgeführt (Merkmalextraktion). Die resultierenden Merkmale \mathbf{x}' werden in einem parallelen Prozeß mit den Daten der integrativen Verfahren (Entropie E und Varianz V) dem Multilayer Perceptron zugeführt. In Abbildung 5.7 ist ein Prinzipbild dieser Verarbeitungskette gegeben.

In einem Vorverarbeitungsschritt wird eine Histogrammbildung über die konturbasierten Komponenten K_i mit

$$x'(i, p, q) = \sum_{\phi_1 = -\frac{\Phi}{2}}^{\frac{\Phi}{2}} \sum_{\phi_2 = -\frac{\Phi}{2}}^{\frac{\Phi}{2}} K_i(p + \phi_1, q + \phi_2)$$

für $i = \{1, 2, \dots, 9\}$ durchgeführt, wobei (p, q) die Position eines Pixels und Φ den Integrationsbereich auf einem Bild beschreiben. Die Variablen ϕ_1 und ϕ_2 beschreiben Laufvariablen. Die Histogrammbildung ist durch die Anforderung motiviert, bei der Objekterkennung zusammenhängende Bildbereiche zu segmentieren. Konturbasierte Verfahren sind differentielle Verfahren und bewerten deshalb nur Grauwertdifferenzen im Bild. Eine Histogrammbildung im Integrationsbereich Φ läßt einen Einfluß der Konturpunkte in einer Nachbarschaft, beschrieben durch $\frac{\Phi}{2}$, zu.

Für den Datenvektor \mathbf{x} an der Bildposition (p, q) folgt daraus der Zusammenhang

$$\mathbf{x} = [(\mathbf{x}'(p, q))^T, E(p, q), V(p, q)]^T. \quad (5.9)$$

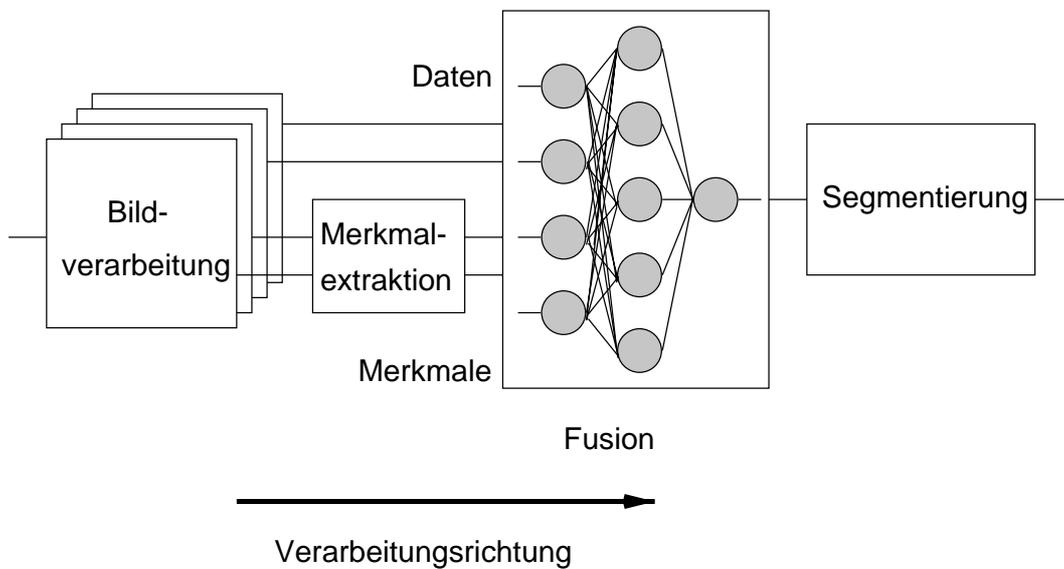


Abbildung 5.7: Prinzipbild des Verfahrens. Daten und Merkmale werden mit einem Multilayer Perceptron fusioniert, um eine Segmentierung von Videobildern zu erreichen.

Der Datenvektor \mathbf{x} wird einem Multilayer Perceptron nach Abschnitt 5.1 als Eingangsdatenvektor zugeführt. Der Eingangsdatenvektor \mathbf{x} charakterisiert dabei kontur-, bewegungs- und texturbasierte Eigenschaften des Bilddatenmaterials, die eine Relevanz für den Fusionsprozeß aufweisen.

5.2.2 Strukturuntersuchung des Multilayer Perceptron

In einem Datensatz $\tilde{\mathbf{X}}$ sind Beispieldatenvektoren $\tilde{\mathbf{x}}$ der Klassen C_1 („Vordergrundsegment“) und C_2 („Hintergrundsegment“) abgelegt. Der Datensatz $\tilde{\mathbf{X}}$ ist von einem Experten erstellt und die Datenvektoren $\tilde{\mathbf{x}}$ sind den entsprechenden Klassen $C_{1,2}$ zugeordnet. Mit einem Backpropagation Algorithmus werden die Synapsengewichte \mathbf{w} des Netzes anhand von Trainings- und Testdaten bestimmt und eine Abbildungsfunktion für eine Klassifikationsaufgabe gelernt. Nach Beendigung des Trainings wird eine Strukturuntersuchung des Multilayer Perceptron anhand des *gain*-Spektrums durchgeführt (Abschnitt 5.1.2).

Mit einer Analyse des *gain*-Spektrums wird die Fusionsstruktur überprüft und die Relevanz einzelner versteckter Neuronen untersucht. In Abbildung 5.8 ist der Verlauf des *gain*-Spektrums des Fusionsnetzwerks nach Gleichung 5.6 über der Zeit dargestellt. Bei der Wahl der Schwelle g_{min} , die die Relevanz einzelner Neurone für den Abbildungsprozeß bewertet, wird ein signifikanter Abstand der Verstärkungsfaktoren g_j zu den unberücksichtigten g_j ausgewertet. Wie aus Abbildung 5.8 ersichtlich, wird eine Schwelle von $g_{min} = 7$ festgelegt. Es läßt sich demnach eine benötigte Anzahl von sieben versteckten Neuronen mit $g_{j=\{1,2,\dots,7\}} > g_{min}$ bestimmen.

Folglich wird eine Fusion der Komponenten mit einem Multilayer Perceptron mit einer 11-7-1 Struktur (elf Eingangsneurone, sieben versteckte Neurone und ein Ausgangsneuron) durchgeführt.

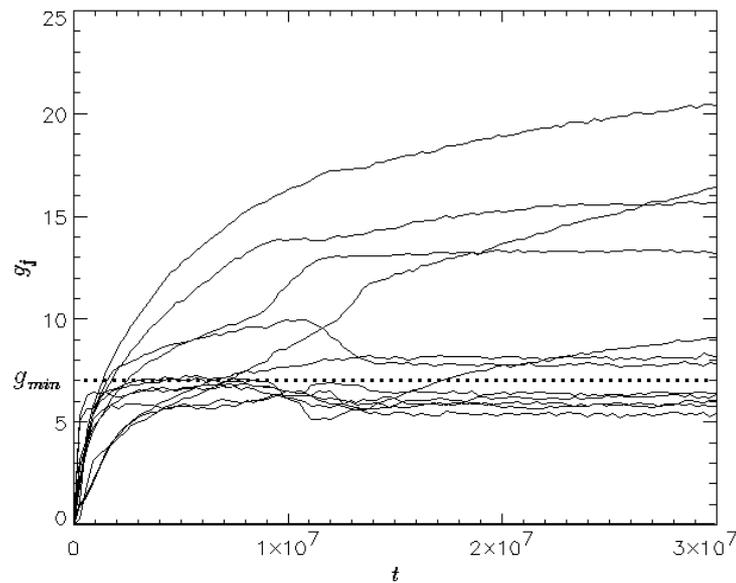


Abbildung 5.8: *Gain-Spektrum* der Neuronen Gewichte für ein Multilayer Perceptron der Struktur 11-12-1.

5.2.3 Fusion

Zur Fusion der Daten und Merkmale wird ein Multilayer Perceptron mit einer 11-7-1 Struktur verwendet. Mit einem Datensatz $\tilde{\mathbf{X}}$ von 100 000 Beispieldaten $\tilde{\mathbf{x}}$ wird das Netz trainiert. Dabei wird zwischen 40 000 Trainingsdaten $\tilde{\mathbf{x}}_1$ (Datensatz $\tilde{\mathbf{X}}_1 \subset \tilde{\mathbf{X}}$ mit $\tilde{\mathbf{x}}_1 \in \tilde{\mathbf{X}}_1$), 20 000 Testdaten $\tilde{\mathbf{x}}_2$ (Datensatz $\tilde{\mathbf{X}}_2 \subset \tilde{\mathbf{X}}$ mit $\tilde{\mathbf{x}}_2 \in \tilde{\mathbf{X}}_2$) und 40 000 Generalisierungsdaten $\tilde{\mathbf{x}}_3$ (Datensatz $\tilde{\mathbf{X}}_3 \subset \tilde{\mathbf{X}}$ mit $\tilde{\mathbf{x}}_3 \in \tilde{\mathbf{X}}_3$) unterschieden. In Abbildung 5.9 ist der Fehlerverlauf $E(t)$ des gemittelten quadratischen Fehlers E für Trainings- und Testdaten $\tilde{\mathbf{x}}_1$ und $\tilde{\mathbf{x}}_2$ über der Zeit t aufgetragen. Nach

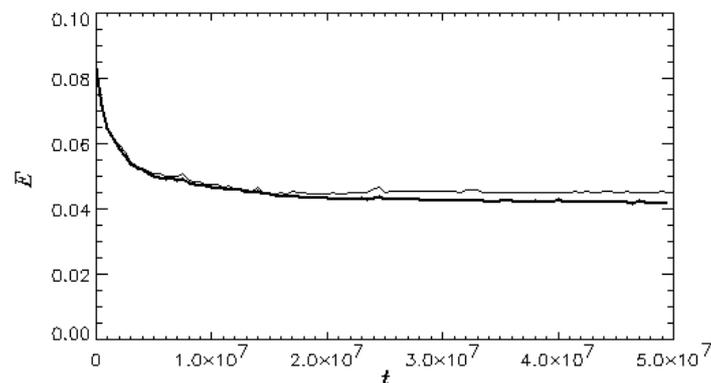


Abbildung 5.9: Zweischichtiges Multilayer Perceptron als Klassifikator für Segmentierungsaufgaben. Aufgetragen ist der über den Trainings- und den Testdatensatz gemittelte quadratische Fehler des Ausgangsneurons über der Zeit während des Lernvorgangs.

mehreren Trainingsdurchläufen wurde das beste Trainingsergebnis übernommen. Dabei wird bei einem durchschnittlichen Fehler $E = 0.04$ der Trainingsdaten $\tilde{\mathbf{x}}_1$ und einem Globalisierungsfehler $E = 0.45$ das globale Minimum der Testdaten $\tilde{\mathbf{x}}_2$ erreicht. Der Restfehler ist in der Überschneidung der Klassen begründet (siehe auch Abschnitt 4.2).

Nach Abschluß des Trainings wird die Relevanz der Komponenten untersucht und das Analyseergebnis aus Kapitel 4 quantitativ geprüft.

5.2.4 Relevanz einzelner Komponenten

Mit dem in Abschnitt 5.1.5 vorgestellten Verfahren wird eine quantitative Aussage über die Relevanz der in Kapitel 4 ausgewählten Komponenten (Gleichung 4.2) gemacht. Insbesondere werden die Synapsengewichte \mathbf{w} zwischen der Eingangsschicht und der Zwischenschicht des Multilayer Perceptron untersucht. Dabei werden die Verstärkungsfaktoren g_i (Gleichung 5.7) und die zugehörigen Varianzen VAR_i (Gleichung 5.8) betrachtet. In Abbildung 5.10 ist das Ergebnis dieser Untersuchung für mittelwertfreie Trainingsdaten $\tilde{\mathbf{x}}_1$ mit einer Standardabweichung $\sigma_x = 1$ dargestellt. Wie in Abbildung 5.10a dargestellt, ist für alle x_i der Verstärkungsfaktor $g_i > 2$. Damit liegt eine Relevanz der zugehörigen Komponenten für den Fusionsprozeß vor.

Weiterhin sind die Varianzen $VAR_i > 0.5$ für alle i gemäß Abschnitt 5.1.5 ausreichend groß (Abbildung 5.10b). Aus diesem Ergebnis wird eine Relevanz aller Komponenten nach Gleichung 5.9 für den Segmentierungsprozeß abgeleitet. Das Ergebnis bestätigt die Auswahl der Komponenten nach dem in Kapitel 4 vorgestellten Verfahren.

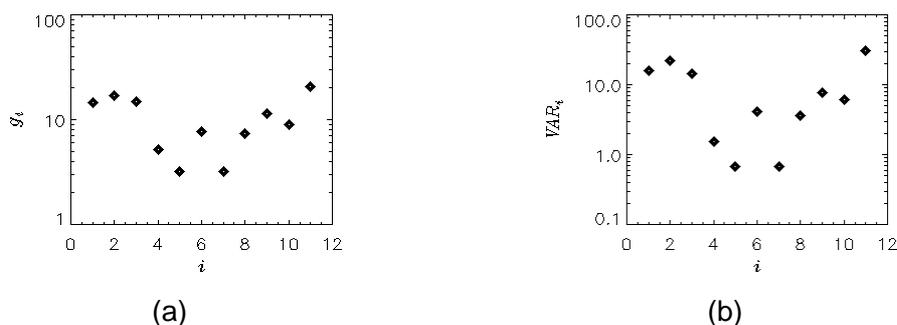


Abbildung 5.10: Relevanz der ausgewählten Komponenten auf den Fusionsprozeß. Dargestellt sind (a) die Verstärkungsfaktoren g_i der einzelnen Komponenten mit (b) zugehörigen Varianzen VAR_i .

In Abschnitt 5.3 sind die Segmentierungsergebnisse exemplarisch an drei Szenen dargestellt. Es wird das Segmentierungsergebnis für eine Autobahnscene und zwei Innenstadtszenen anhand der Datenreduktionsraten über jeweils 1000 Bilder beschrieben und an ausgewählten Bildern diskutiert. In Abschnitt 5.3.2 wird die flexible Erweiterbarkeit der gewählten Fusionsstruktur an zwei weiteren Beispielen diskutiert.

5.3 Ergebnisse

Das entwickelte Verfahren wurde auf mehreren Straßenverkehrsszenen von Autobahnen, Landstraßen und aus dem Innenstadtbereich getestet. Die Tragfähigkeit wird an drei Beispielen aus Straßenverkehrsszenen gezeigt. Das Segmentierungsergebnis wird jeweils für 1000 Bilder dargestellt. Es wird für jede Szene das zugehörige Segmentierungsergebnis anhand dreier typischer Bilder diskutiert. Weiterhin wird die flexible Erweiterbarkeit des Verfahrens an zwei Beispielen demonstriert. Eine Diskussion der Ergebnisse wird durchgeführt.

5.3.1 Segmentierungsergebnisse

Die Tragfähigkeit des entwickelten Verfahrens wurde an mehreren Sequenzen getestet. Anhand von drei Beispielsequenzen wird das Segmentierungsergebnis verdeutlicht.

In Abbildung 5.11 sind die Datenreduktionsergebnisse für diese drei Szenen für jeweils 1000 Bilder dargestellt. Es ist die Datenreduktion P über der Zeit aufgetragen. In jedem Zeitschritt t wird ein Bild präsentiert. Wie aus der Darstellung ersichtlich, wird eine Reduktion des Datenmaterials um mindestens 70% erreicht. Die Variationen der Reduktionsraten sind in der Anzahl vorhandener Objekte, sowie deren Entfernung zum Trägerfahrzeug begründet.

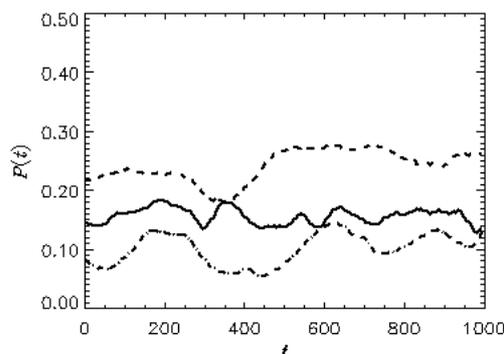


Abbildung 5.11: Datenreduktion für drei Beispielsequenzen. Dargestellt ist die Reduktion des Bilddatenmaterials für eine Autobahnsequenz (durchgezogene Linie), eine Innenstadtsequenz (Punkte und Striche) und eine Baustellendurchfahrt (gestrichelt).

Auf allen Sequenzen werden die relevanten Objekte zu 99% den Vordergrundsegmenten zugeschlagen. Wird eine Mittelung über fünf Bilder durchgeführt, ist eine vollständige Objektdetektion zu erreichen. Dies motiviert eine zeitliche Stabilisierung der Ergebnisse in Kapitel 7.

Im folgenden werden die betrachteten Szenen genauer beschrieben und jeweils wesentliche Merkmale verdeutlicht. Dies ist notwendig, da die Segmentierungsergebnisse nur im Zusammenhang mit den durch die Szene beschriebenen Objekten zu bewerten sind.

Die Autobahnsequenz zeichnet sich durch wenig texturierte Bereiche aus. In Abbildung 5.12 sind drei Bilder aus dieser Szene dargestellt. Durch den Neigungswinkel der Kamera deckt die Straße einen großen Teil des Bildes ab. Die relevanten Bereiche werden segmentiert. Alle vorhandenen Fahrzeuge befinden sich in den extrahierten Segmenten. Am rechten Fahrbahnrand werden

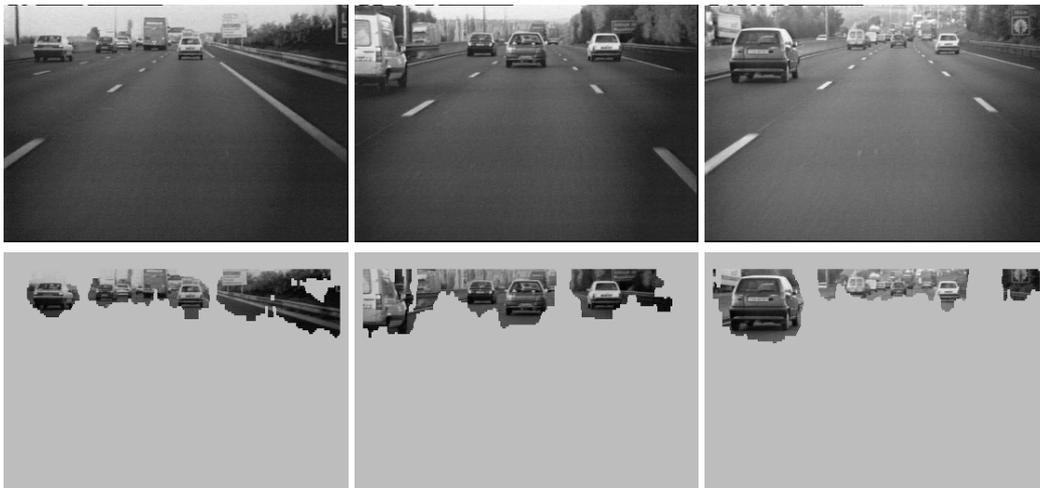


Abbildung 5.12: Segmentierung durch Sensorfusion mit einem Multilayer Perceptron. Drei Ergebnisbilder aus einer Autobahnszene bei einer Kameraposition mit großem Neigungswinkel.

zusätzlich Verkehrsschilder und die Leitplanke segmentiert. Die Segmentierung dieser Bereiche ist jedoch unkritisch für den weiteren Verarbeitungsprozeß (siehe Kapitel 7).

Abbildung 5.13 zeigt drei Bilder aus der Innenstadtszene. Mehrere Autos sind zu erkennen. Ein Fahrzeug überholt, ein weiteres kreuzt die Fahrbahn. Die relevanten Bereiche werden segmentiert. Alle vorhandenen Fahrzeuge befinden sich in den Vordergrundsegmenten. In Kreuzungsbereichen werden größere Bereiche segmentiert. Dies ist gewünscht, da eine erhöhte Aufmerksamkeit gefordert wird.

In Abbildung 5.14 sind drei Bilder der Baustellendurchfahrt bei einer starken Linkskurve dargestellt. Ein Auto fährt vor dem Trägerfahrzeug auf der Straße. Links sind mehrere Baustellenaufbauten und rechts eine betonierte Fahrbahnbegrenzung. Das führende Fahrzeug wird



Abbildung 5.13: Segmentierung durch Sensorfusion mit einem Multilayer Perceptron. Drei Ergebnisbilder aus einer Stadtszene im Kreuzungsbereich.

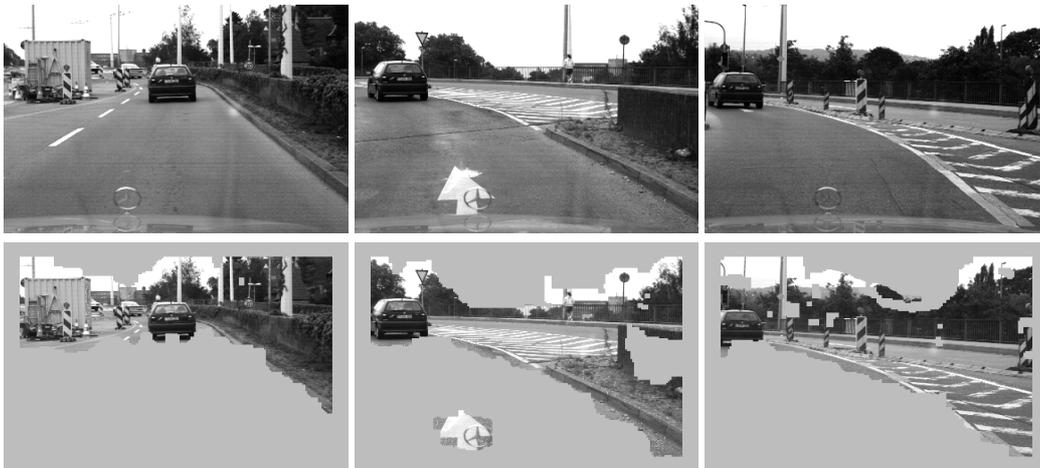


Abbildung 5.14: Segmentierung durch Sensorfusion mit einem Multilayer Perceptron. Drei Ergebnisbilder aus einer Szene an der eine Baustelle in einer starken Linkskurve passiert wird.

segmentiert. Dem Bereich der Baustelle und dem rechten Randbereich der Kurve wird ebenfalls erhöhte Aufmerksamkeit durch Segmentierung als Vordergrundsegment gewidmet. Im folgenden wird die flexible Erweiterbarkeit des realisierten Verfahrens demonstriert.

5.3.2 Flexible Erweiterbarkeit des Verfahrens

Die gewählte Fusionsstrategie eines parallelen neuronalen Prozesses auf Daten- und Merkmalsebene erlaubt eine flexible Einkopplung zusätzlicher Sensordaten. Dabei ist es möglich, sowohl neue Sensordaten des Videobildverarbeitungsprozesses zu integrieren als auch andere physikalische Sensoren auf der Pixelebene der Videobilder zu berücksichtigen.

In diesem Kapitel wird die Flexibilität anhand dreier Beispiele gezeigt. Es wird eine Erweiterung des Verfahrens durch die Integration

1. eines modellbasierten Bildverarbeitungsansatzes, der auf Basis der Analyse von Schlagschatten unter Fahrzeugen Objekthypothesen erzeugt,
2. von Radardaten eines RADAR-Sensors (3-Strahl-RADAR), welcher die Distanz und den relativen Winkel von Objekthypothesen liefert und
3. von Farbsättigungsdaten, die von einer RGB³-Farbkamera bereitgestellt werden

durchgeführt.

Der modellbasierte Ansatz der Schlagschattenanalyse wurde in [119] vorgestellt. Das dort realisierte Verfahren basiert auf der Annahme, daß Fahrzeuge auf der Straße Schlagschatten erzeugen. Die Schlagschatten weisen dabei eine Grauwertverteilung auf, die im Mittel wesentlich geringere Werte haben, als die Umgebung des Fahrzeugs. Aus dieser Annahme und der Hinzunahme der Parameter der verwendeten Kamera werden Objekthypothesen erzeugt. In dem realisierten Fusionsprozeß wird das Merkmal der Objektklasse („Fahrzeug“, „kein Fahrzeug“) als zusätzliches Eingangsmerkmal C mit $C = \{0, 1\}$ berücksichtigt (Abbildung 5.15a).

³RGB – red green blue

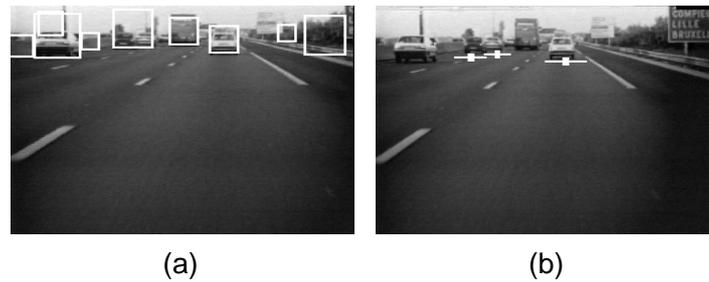


Abbildung 5.15: (a) Ergebnisbilder eines modellbasierten Ansatzes unter Ausnutzung der Kamerageometrie und (b) Objekthypothesen eines 3-Strahl-RADAR.

Zusätzlich können bei der analysierten Sequenz Radardaten eines 3-Strahl-RADARs genutzt werden. Bei diesen Daten stehen jedoch nur die Distanz und der relative Winkel der Objekthypothesen zu dem Trägerfahrzeug zur Verfügung (Abbildung 5.15b). Die Relativgeschwindigkeit der Objekthypothesen zum Trägerfahrzeug liegt nicht vor und kann deshalb für den Fusionsprozeß nicht genutzt werden. Eine Koordinatentransformation erlaubt die Erzeugung von Objekthypothesen \mathbf{R} in Bildkoordinaten. Daraus ergibt sich eine Erweiterung des Datenvektors nach Gleichung 5.9 mit

$$\mathbf{x} = [(\mathbf{x}'(p, q))^T, E(p, q), V(p, q), C(p, q), \mathbf{R}(p, q)]^T$$

und damit eine neue Netzstruktur des Multilayer Perceptrons von 13-7-1. Die Variablen p und q beschreiben die Bildkoordinaten. Die Anzahl der versteckten Neuronen kann nach einer Analyse des *gain*-Spektrums auf sieben belassen werden. Nach Durchführung eines Trainings der Synapsengewichte \mathbf{w} wird das in Abbildung 5.16 dargestellte Segmentierungsergebnis erzielt.



Abbildung 5.16: Segmentierung durch Sensorfusion mit einem Multilayer Perceptron. Zusätzliche Einkopplung von Radardaten und einem modellbasierten Ansatz auf Basis von Schlagschatten. Dargestellt sind drei Bilder und die zugehörigen segmentierten Bilder einer Sequenz.

Bedingt durch den RADAR-Sensor wird eine schärfere Trennung der Segmente erreicht. Weiterhin ist beim Ausfall eines Sensors (Videokamera, RADAR-Sensor) eine Segmentierung relevanter Bereiche noch immer möglich (Abbildung 5.12 und Abbildung 5.15b). Redundanzen im System können dabei genutzt werden. Bei einer Erweiterung des RADAR-Sensors um mehrere Radarstrahlen, läßt sich das Segmentierungsergebnis weiter verbessern.

In einer zweiten Sequenz wird das Sensordatenmaterial mit Hilfe eines RGB-Videofarbsensors um die Farbsättigung (Abbildung 5.17) erweitert. Dieses Vorgehen wird dadurch motiviert, daß in der Umwelt Farben hoher Sättigung eine Signalwirkung haben und die Aufmerksamkeit des Beobachters erhöhen.

Die Sättigung ist Teil eines an der menschlichen Farbvorstellung orientierten Farbraums, des HSV⁴-Raums, der eine Farbe durch die Attribute Farbton, Sättigung und Helligkeit beschreibt [26]. Für die Extraktion farbiger Objekte, kann die dreidimensionale Farbbeschreibung auf die Verwendung der Sättigungsordinate S , mit

$$S = \frac{\max(R, G, B) - \min(R, G, B)}{\max(R, G, B)}$$

reduziert werden. Dabei entsprechen R, G, B dem Farbwert eines Pixels im rot-, grün-, und blau-Kanal der Farbvideokamera. In dem in Abbildung 5.18 dargestellten Segmentierungsprozeß wurden Bereiche starker Sättigung durch eine Schwellwertbildung (mit der Schwelle ρ_{min}) extrahiert und als Merkmal dem Fusionsprozeß zugeführt. Das binarisierte Bild (Abbildung 5.17b) ist durch

$$S(p, q) = \begin{cases} 1 & \text{falls } S(p, q) > \rho_{min} \\ 0 & \text{sonst} \end{cases},$$

beschrieben. Daraus ergibt sich eine Erweiterung des Datenvektors nach Gleichung 5.9 mit

$$\mathbf{x} = [(\mathbf{x}'(p, q))^T, E(p, q), V(p, q), S(p, q)]^T$$

und damit eine alternative Netzstruktur des Multilayer Perceptrons vom 12-7-1. Eine Analyse des *gain*-Spektrums ergibt eine Anzahl von sieben versteckten Neuronen. Nach Durchführung

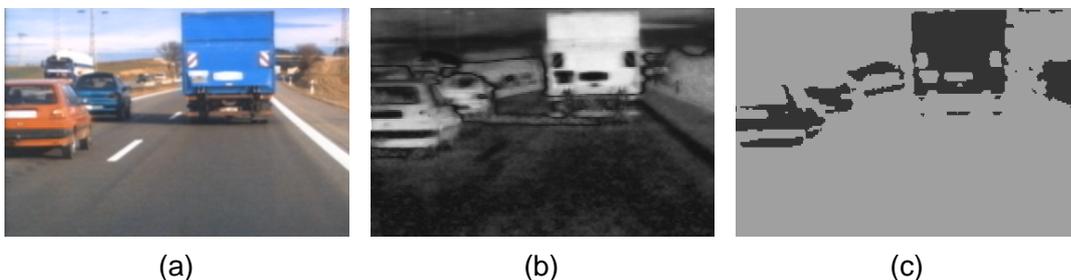


Abbildung 5.17: Farbsättigung als Merkmal für Segmentierungsaufgaben. (a) Originalbild, (b) Sättigungsbild (die Helligkeit beschreibt die Größe der Sättigung) und (c) durch eine Schwellwertoperation binarisiertes Farbsättigungsbild.

⁴HSV – hue saturation value

eines Trainings der Synapsengewichte w , wird das in Abbildung 5.18 dargestellte Segmentierungsergebnis erzielt. Es sind drei Bilder und die zugehörigen segmentierten Bilder einer Sequenz dargestellt.

Die Erweiterung um die Sättigungskomponente führt zu einer Verbesserung des Segmentierungsergebnisses in schwach strukturierten Bereichen ohne Kontur, wie dem Aufbau des LKW.

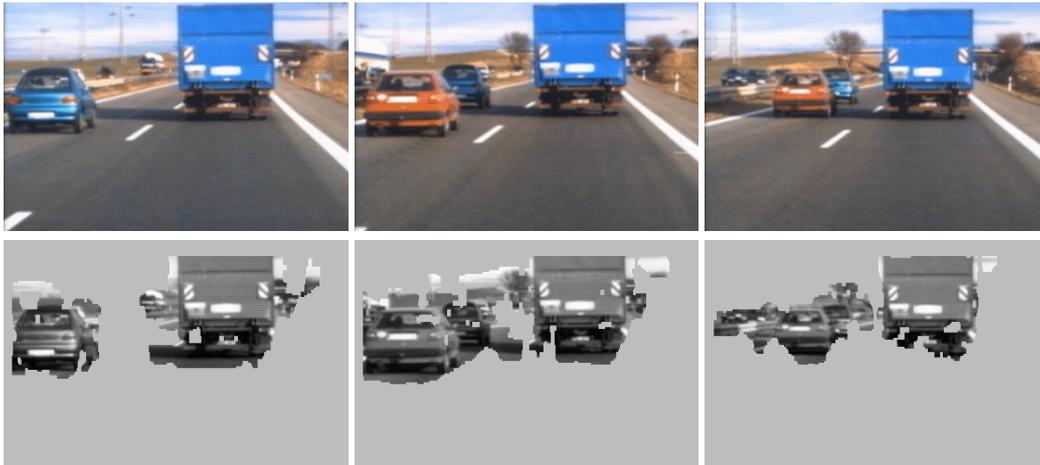


Abbildung 5.18: Segmentierung durch Sensorfusion mit einem Multilayer Perceptron. Zusätzliche Einkopplung von Farbsättigungsdaten zur Segmentierung von Farbbildern. Dargestellt sind drei Bilder und die zugehörigen segmentierten Bilder einer Sequenz.

5.3.3 Diskussion

Anhand mehrerer Szenen wurde die Tragfähigkeit des realisierten Verfahrens getestet. Dabei wird eine Reduktion des Datenmaterials von über 70% erreicht. Alle relevanten Objekte werden über der Zeit erkannt. Einzelne typische Szenen werden dargestellt. Neben einem Beispiel einer Autobahnszene, wo relevante Bereiche hauptsächlich durch Kraftfahrzeuge und Verkehrsschilder definiert sind, werden auch Ergebnisse aus dem innerstädtischen Bereich gezeigt. In diesem Bereich sind neben anderen Verkehrsteilnehmern auch besondere Markierungen und Brandungen des freien Fahrraums durch Hindernisse für den Fahrtverlauf von Interesse. Weiterhin stellen Kreuzungsbereiche erhöhte Anforderungen an die Bildverarbeitung, da hier nicht nur der eigene Fahrbahnverlauf beachtet werden muß. Alle wesentlichen Bereiche werden segmentiert.

Anhand der Ergebnisse wird die Tragfähigkeit des Verfahrens belegt. Neben der Erhöhung der Segmentierungsgüte durch Fusion, welche in Kapitel 4 diskutiert wird, wird, anhand der gezeigten Beispiele, der Einsatz eines Multilayer Perceptrons als effiziente Fusionsstruktur belegt. Durch die parallele Verarbeitung der Sensordaten wird außerdem die flexible Erweiterbarkeit gezeigt und der Einsatz verschiedener physikalischer Sensoren für den Segmentierungsprozeß motiviert. Dies stellt, neben den verbesserten Segmentierungsergebnissen, einen wesentlichen Vorteil des realisierten Verfahrens dar.

Die extrahierten Segmente können in einem weitergehenden Bildverarbeitungsprozeß genutzt werden, um Objekte aus der Umgebung des Fahrzeugs zu erkennen und zu attributieren. Diese Attribute bilden die Eingangsdaten eines Verfahrens zur Bewegungssteuerung eines autonomen Fahrzeugs, welches in Kapitel 6 vorgestellt wird.

5.4 Zusammenfassung

In diesem Kapitel wird eine robuste und flexible Sensordatenfusion der in Kapitel 4 analysierten Komponenten durchgeführt. Dabei wird eine effiziente und robuste Fusionsstruktur eingesetzt. Zentrales Element der Fusionsstruktur ist ein Multilayer Perceptron, welches als Klassifikator eingesetzt wird. Es werden textur-, kontur- und bewegungsbasierte Komponenten fusioniert. Es wird ein paralleler Fusionsprozeß auf Daten- und Merkmalsebene realisiert. Anhand von Segmentierungsergebnissen auf Videobildern unterschiedlicher Ausprägung wird die Tragfähigkeit des Verfahrens belegt. Die Flexibilität der Fusionsstruktur wird anhand zweier Erweiterungen im Eingangsdatenraum gezeigt.

Im folgenden Kapitel werden Objektattribute genutzt, die, aufbauend auf die Segmentierung, mit in der Literatur beschriebenen Verfahren extrahiert werden können [84, 127, 119, 59]. Es wird ein neues Verfahren für eine Bewegungssteuerung eines Kraftfahrzeugs vorgestellt.

Kapitel 6

Neuronale Bewegungssteuerung

Der Zustand eines Kraftfahrzeugs in der Umwelt kann durch die aktuelle Position, die Geschwindigkeit und die Bewegungsrichtung beschrieben werden. Soll ein Fahrzeug durch Einsatz eines Assistenzsystems autonom im Straßenverkehr navigieren, müssen diese Zustandsgrößen durch das System manipulierbar sein.

Die Navigation ist abhängig von einer Aufgabenstellung und von Randbedingungen, die durch die Umwelt gegeben sind. Neben sich langsam ändernden Elementen der Umwelt, wie dem Straßentyp und den Sichtverhältnissen, beeinflussen weitere Verkehrsteilnehmer und andere Hindernisse (im folgenden Objekte) das gewünschte Fahrverhalten.

In Abbildung 6.1 ist der prinzipielle Aufbau eines Assistenzsystems dargestellt, das mit einer Videokamera die Umwelt observiert. Das Bilddatenmaterial der Kamera wird von Bildverarbeitungs-Algorithmen verarbeitet und eine Extraktion von relevanten Objekten und deren Attributen in einem weiteren Verarbeitungsprozeß durchgeführt. Die Objektattribute sowie eine vorgegebene Aufgabenstellung beeinflussen das Verhalten des Fahrzeugs. Bei der Verhaltenssteuerung werden diese Größen verarbeitet und eine Veränderung der Zustandsgrößen des Fahrzeugs durchgeführt.

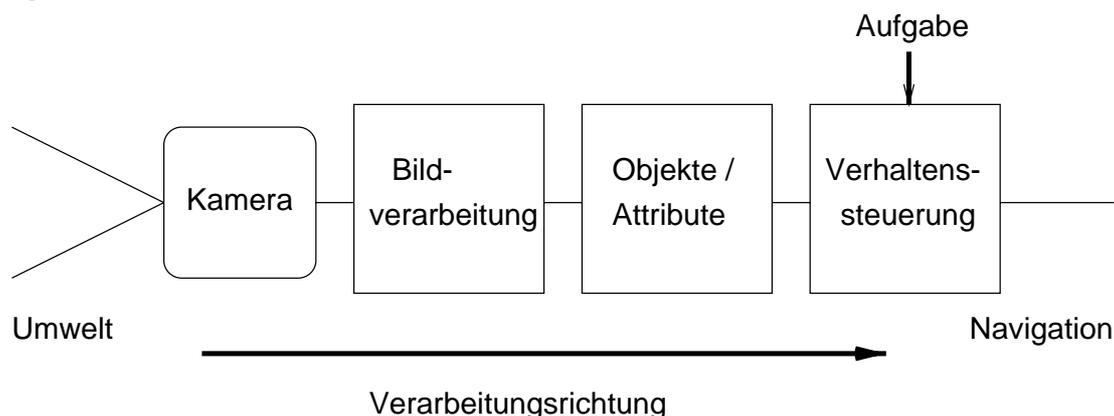


Abbildung 6.1: Blockschaltbild eines Fahrerassistenzsystems. Mit einer Videokamera wird die Umwelt observiert. Bildverarbeitungs-Algorithmen verarbeiten das von der Kamera bereitgestellte Bilddatenmaterial. In einem weiteren Verarbeitungsprozeß werden Objekte extrahiert und attributiert. Diese Objektmerkmale dienen als Basis der Verhaltenssteuerung, die die autonome Navigation des Fahrzeugs in der Umwelt realisiert.

In diesem Kapitel wird ein neues Verfahren vorgestellt, das die Bewegung eines autonomen Kraftfahrzeugs, über die Steuerung des Lenkwinkels und der Geschwindigkeit, beeinflusst. Die Basis für das realisierte Verfahren bilden zwei neuronale Felder. Der Ansatz neuronaler Felder wird in Abschnitt 6.1 diskutiert. In Abschnitt 6.2 werden die Attribute anderer Objekte beschrieben, welche die Bewegung des Fahrzeugs beeinflussen. Eine Simulationsumgebung (Anhang B) ermöglicht es, das Verhalten mehrerer Fahrzeuge zu simulieren (simulierte Objekterkennung) sowie den Lenkwinkeleinschlag und die Geschwindigkeit des Trägerfahrzeugs als Reaktion auf die Objektattribute zu testen. In Abschnitt 6.3 wird das realisierte Verfahren beschrieben und dessen Verhalten an einer Beispielszene diskutiert. Abschnitt 6.4 verdeutlicht die Robustheit des Verfahrens an verschiedenen typischen Straßenverkehrsszenen im innerstädtischen Bereich. Das Kapitel 6 schließt mit einer Zusammenfassung in Abschnitt 6.5.

6.1 Neuronale Felder

Neuronale Felder sind nichtlineare dynamische Systeme. Sie wurden als Verhaltensmodelle für die Neurophysiologie kortikaler Prozesse eingeführt [130, 1]. Neuronale Felder können als makroskopische Beschreibung der Dynamik der Aktivität von Neuronen in Ort und Zeit verstanden werden [112, 6, 29]. Sie gehören zu der Klasse der Integro-Differentialgleichungen. In Abschnitt 6.1.1 wird der Ansatz des Amarifelds diskutiert, der ein zentrales Element in dem realisierten Verfahren zur Bewegungssteuerung darstellt.

6.1.1 Amarifeld

Eine Realisierung eines neuronalen Feldes ist durch einen Ansatz nach *Amari* [1, 3] gegeben. Dieser Feldansatz hat den Vorteil, daß er mathematisch einfach zu beschreiben ist und eine gute Analyse seiner dynamischen Eigenschaften im Betrachtungsraum Γ über der Zeit und dem Ort zuläßt.

Die Dynamik des Feldes (auch Amarifeld) ist gegeben durch die nichtlineare dynamische Integro-Differentialgleichung

$$\begin{aligned} \tau \dot{u}(\mathbf{z}, t) &= -u(\mathbf{z}, t) + h + S(\mathbf{z}, t) \\ &+ \int_{\Gamma} w(\mathbf{z}, \mathbf{z}') \varphi(u(\mathbf{z}', t)) d\mathbf{z}'. \end{aligned} \quad (6.1)$$

Die erste Zeile in Gleichung 6.1 beschreibt die lokale Dynamik der Aktivität des Feldes $u(\mathbf{z}, t)$ mit $u(\mathbf{z}, t) \in \mathbb{R}$ am Ort \mathbf{z} mit $\mathbf{z} \in \Gamma$ zum Zeitpunkt t mit $t \geq 0$. Die zeitliche Ableitung von $u(\mathbf{z}, t)$ ist gegeben durch

$$\dot{u}(\mathbf{z}, t) = \frac{\partial u(\mathbf{z}, t)}{\partial t}.$$

Die Aktivität des Feldes ändert sich abhängig von einer Zeitkonstanten τ mit $\tau \in \mathbb{R}^+$. Die Variable h entspricht einer konstanten Voraktivierung mit $h = \text{const}$ und $h \in \mathbb{R}$. Das Eingangssignal $S(\mathbf{z}, t)$ (der Stimulus) ist ebenfalls abhängig von \mathbf{z} und t mit $S(\mathbf{z}, t) \in \mathbb{R}$. Es wird additiv in das Feld eingekoppelt. Ohne Stimulus entspricht die Aktivität u aller Feldneurone im eingeschwungenen Zustand einem globalen stationären Feld.

Eine biologisch motivierte Wechselwirkung zwischen den Feldneuronen wird durch das Integral in der zweiten Zeile der Gleichung 6.1 beschrieben. Dabei führt die Aktivierungsfunktion $\varphi(u)$ eine nichtlineare Transformation der Aktivität u der einzelnen Feldneurone durch. Sie ist eine stetig differenzierbare Funktion (siehe Anhang A.1), mit

$$\begin{aligned}\lim_{u \rightarrow -\infty} \varphi(u) &= 0 \text{ und} \\ \lim_{u \rightarrow \infty} \varphi(u) &= 1.\end{aligned}$$

Das Amarifeld ist homogen [29]. Folglich kann das Integral als Faltung mit

$$w(\mathbf{z}) * \varphi(u(\mathbf{z}, t)) = \int_{\Gamma} w(\mathbf{z}, \mathbf{z}') \varphi(u(\mathbf{z}', t)) d\mathbf{z}'$$

zwischen dem Wechselwirkungskern $w(\mathbf{z}, \mathbf{z}')$ mit $w(\mathbf{z}, \mathbf{z}') = w(\mathbf{z} - \mathbf{z}')$, $(\mathbf{z} - \mathbf{z}') \in \Gamma$, $w \in \mathbb{R}$ und $\varphi(u(\mathbf{z}', t))$ angesehen werden [9]. Durch die Faltung findet eine Interaktion zwischen einzelnen Neuronen statt. Der Wechselwirkungskern w mit $w(\mathbf{z} - \mathbf{z}') = w(\Delta\mathbf{z})$ ist in diesem Ansatz nicht abhängig von t .

Häufig wird für w eine Funktion mit der symmetrischen Eigenschaft $w(\Delta\mathbf{z}) = w(-\Delta\mathbf{z})$ verwendet. Als eindimensionales Beispiel ist in Abbildung 6.2 ein sogenannter *mexican hat*-Verlauf eines Wechselwirkungskerns $w(\Delta\mathbf{z})$ dargestellt.

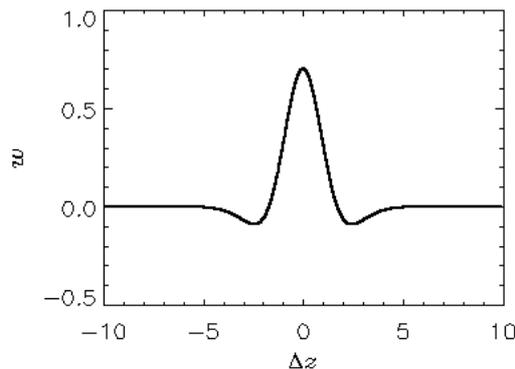


Abbildung 6.2: Wechselwirkungskern: *mexican hat*.

Ein Wechselwirkungskern dieses Typs hat Bereiche mit exzitatorischer, d. h. $w(\Delta\mathbf{z}) \geq 0$, sowie Bereiche mit inhibitorischer Wirkung, d. h. $w(\Delta\mathbf{z}) < 0$. Er konvergiert für $|\Delta\mathbf{z}| \rightarrow \infty$ vom Negativen kommend gegen Null. Mathematisch kann die zugrundeliegende Funktion durch eine Differenz zweier Gaußfunktionen $N_1(0, \sigma_1)$ und $N_2(0, \sigma_2)$ mit $w = c_1 N_1 - c_2 N_2$, $\sigma_2 > \sigma_1$, $c_1 > 0$ und $c_2 > 0$, beschrieben werden. Die Variablen c_1 , c_2 bezeichnen dabei Gewichtungsfaktoren und σ_1 bzw. σ_2 definieren die Breiten der Gaußfunktionen $N_{1,2}$.

6.1.2 Eigenschaften

Um das prinzipielle Verhalten von Amarifeldern zu untersuchen, werden wesentliche Zustände des Feldes qualitativ abgeleitet und dabei das dynamische Verhalten des Feldes analysiert. Hierbei werden Teile der Feldgleichung betrachtet und daraus das Gesamtverhalten ermittelt.

Wird nur die lokale Dynamik der Aktivität u der Feldneurone mit einer konstanten Voraktivierung ($h = -1$) unter Einwirkung eines zeitlich und örtlich begrenzten Stimulus S betrachtet, so erhält man für ein eindimensionales Feld $u(z, t)$ mit der Feldvariablen z die Differentialgleichung

$$\tau \dot{u}(z, t) = -u(z, t) + h + S(z, t). \quad (6.2)$$

Der exponentielle Anstieg der Felderregung ist bestimmt durch die Höhe des Stimulus S und den Wert der Zeitkonstanten τ . Wird der Stimulus $S(z, t)$ abgeschaltet, klingt die Aktivität $u(z, t)$ des Feldes exponentiell ab. Wie in Gleichung 6.2 beschrieben, findet keine Wechselwirkung statt. In Abbildung 6.3 ist der Verlauf des Eingangssignals und die Dynamik des beschriebenen Feldes dargestellt.

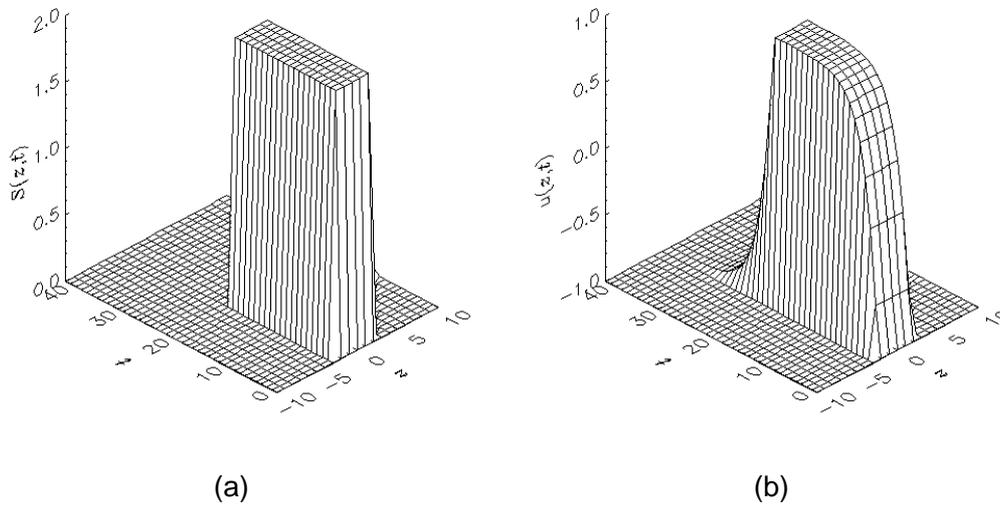


Abbildung 6.3: Dynamik eines eindimensionalen Feldes nach Gleichung 6.2 dargestellt über der Zeit mit $h = -1$. (a) Eingangssignal $S(z, t)$ und (b) zugehörige Feldaktivität $u(z, t)$.

Betrachtet man ein eindimensionales dynamisches Feld mit einer linearen Aktivierungsfunktion, kann die Feldgleichung mit

$$\tau \dot{u}(z, t) = -u(z, t) + h + \int_{\Gamma} w(z - z') u(z', t) dz' + S(z, t) \quad (6.3)$$

beschrieben werden. In Abbildung 6.4 ist der Verlauf dieser Feldaktivität unter Einkopplung eines Eingangssignals nach Abbildung 6.3a gezeigt. Dabei wird ein Wechselwirkungskern vom Typ eines *mexican hat* (Abbildung 6.2) gewählt.

Die Energie W des Wechselwirkungskerns, mit

$$W(z) = \int_0^z w(\Delta z) d(\Delta z),$$

charakterisiert dabei mit $W_{max} = \max_{z > 0} (W(z))$ die maximale Stärke der Wechselwirkung zwischen den Feldneuronen.

Wird der Wechselwirkungskern so dimensioniert, daß eine sehr schwache Wechselwirkung mit $W_{max} \ll -h$ stattfindet (Abbildung 6.4a), bildet sich an der Stelle der Einkopplung in Abhängigkeit von τ ein Maximum aus, das nach Abschalten des Eingangssignals abklingt. Der inhibitorische Teil des Wechselwirkungskerns bedingt hierbei eine sogenannte *peak-Lösung*¹. Findet eine starke Wechselwirkung mit $W_{max} \gg -h$ zwischen den Feldneuronen statt, ist die Stabilität nicht mehr sichergestellt. Die Feldaktivität kann mit $t \rightarrow \infty$ ins Unendliche wachsen (Abbildung 6.4b).

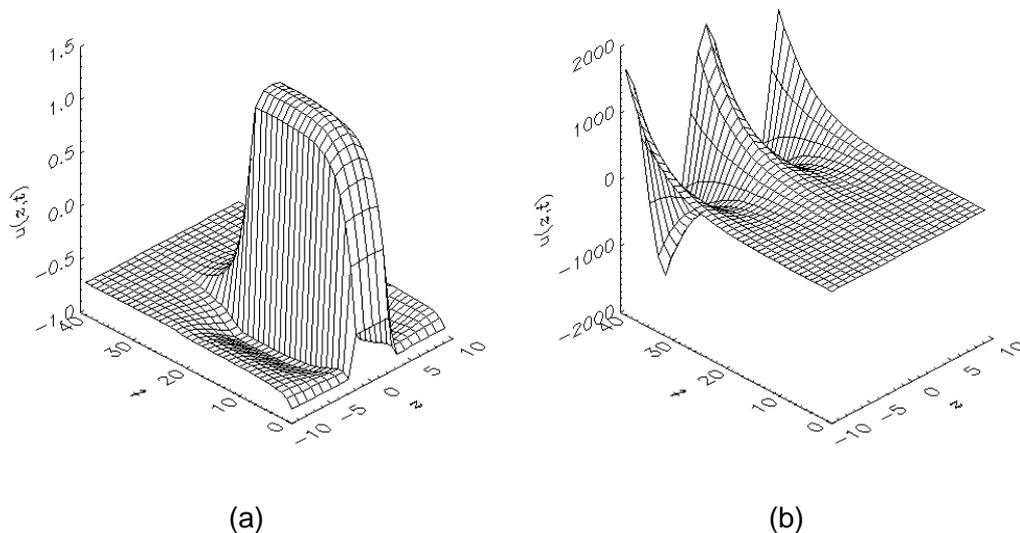


Abbildung 6.4: Dynamik eines eindimensionalen Feldes nach Gleichung 6.3 mit einem Eingangssignal wie in Abbildung 6.3a dargestellt. (a) sehr schwache Wechselwirkung ($W_{max} \ll -h$) und (b) starke Wechselwirkung ($W_{max} \gg -h$).

Betrachtet man schließlich die Felderregung eines eindimensionalen Amarifeldes nach Gleichung 6.1, wird durch die Hinzunahme einer sigmoiden Nichtlinearität $\varphi = \xi$ Stabilität erreicht. In Abbildung 6.5 ist die Felderregung bei Wechselwirkungen unterschiedlicher Stärke gezeigt. Bei geeigneter Wahl von Wechselwirkungskern w und Voraktivierung h zeigt das neuronale Feld nach dem Wegfall des Stimulus S mit $S = 0$ ein zeitlich stabilisierendes Verhalten der *peak-Lösung* (Abbildung 6.5c,d). Es sind monostabile sowie multistabile Zustände möglich. Im folgenden werden mögliche Lösungen betrachtet.

Bei der Betrachtung von Amarifeldern wird zwischen vier möglichen Lösungen im eingeschwungenen Zustand unterschieden [1].

1. Ist die Aktivität $u(z)$ des Feldes im gesamten Bereich $u(z) \leq 0$, so liegt eine sogenannte \emptyset -Lösung vor.
2. Ist die Aktivität $u(z)$ des Feldes im gesamten Bereich $u(z) > 0$, so liegt eine sogenannte ∞ -Lösung vor.

¹peak – abgeleitet aus dem Englischen: peak = Spitze, Höchststand.

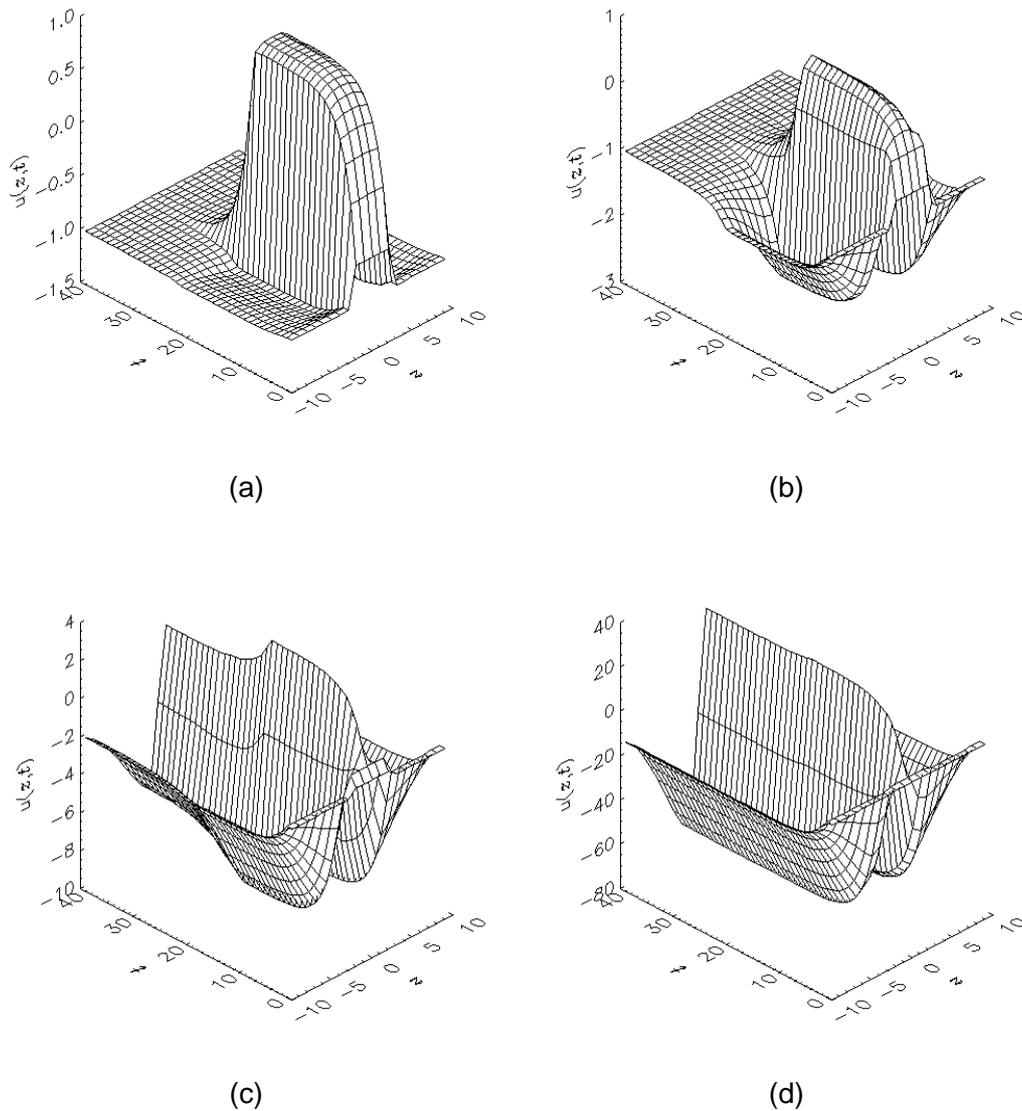


Abbildung 6.5: Dynamik eines eindimensionalen Feldes nach Gleichung 6.1 mit einem Eingangssignal wie in Abbildung 6.3a dargestellt. (a) sehr schwache Wechselwirkung ($W_{max} \ll -h$), (b) schwache Wechselwirkung ($W_{max} < -h$), (c) typische Wechselwirkung ($W_{max} \geq -h$) und (d) starke Wechselwirkung ($W_{max} \gg -h$).

3. Eine lokal begrenzte Aktivität $R(u) = (z_1, z_2)$, mit den Grenzen z_1 und z_2 , sowie der Länge $a = z_2 - z_1$ für $z_1 < z_2$, wird als a -Lösung bezeichnet. Dabei wird die Aktivität $R(u)$ mit $R(u) = \{z | u(z) > 0\}$, $u(z_1) = 0$ und $u(z_2) = 0$ definiert.
4. Ist die Aktivität eines Feldes durch eine Vielzahl von a -Lösungen gegeben, so liegen *multi-peak*-Lösungen vor.

In Abbildung 6.6a-d sind Beispiele für mögliche Lösungen eines eindimensionalen Amarifeldes bei einem kurzzeitig anliegenden Eingangssignal S über der Zeit aufgetragen.

Das Stabilitätsverhalten von Amarifeldern wird ausführlich in [1, 2] diskutiert, so daß in Kapitel 6.1.3 nur die notwendigen Stabilitätsbedingungen für eindimensionale Amarifelder wiedergegeben sind.

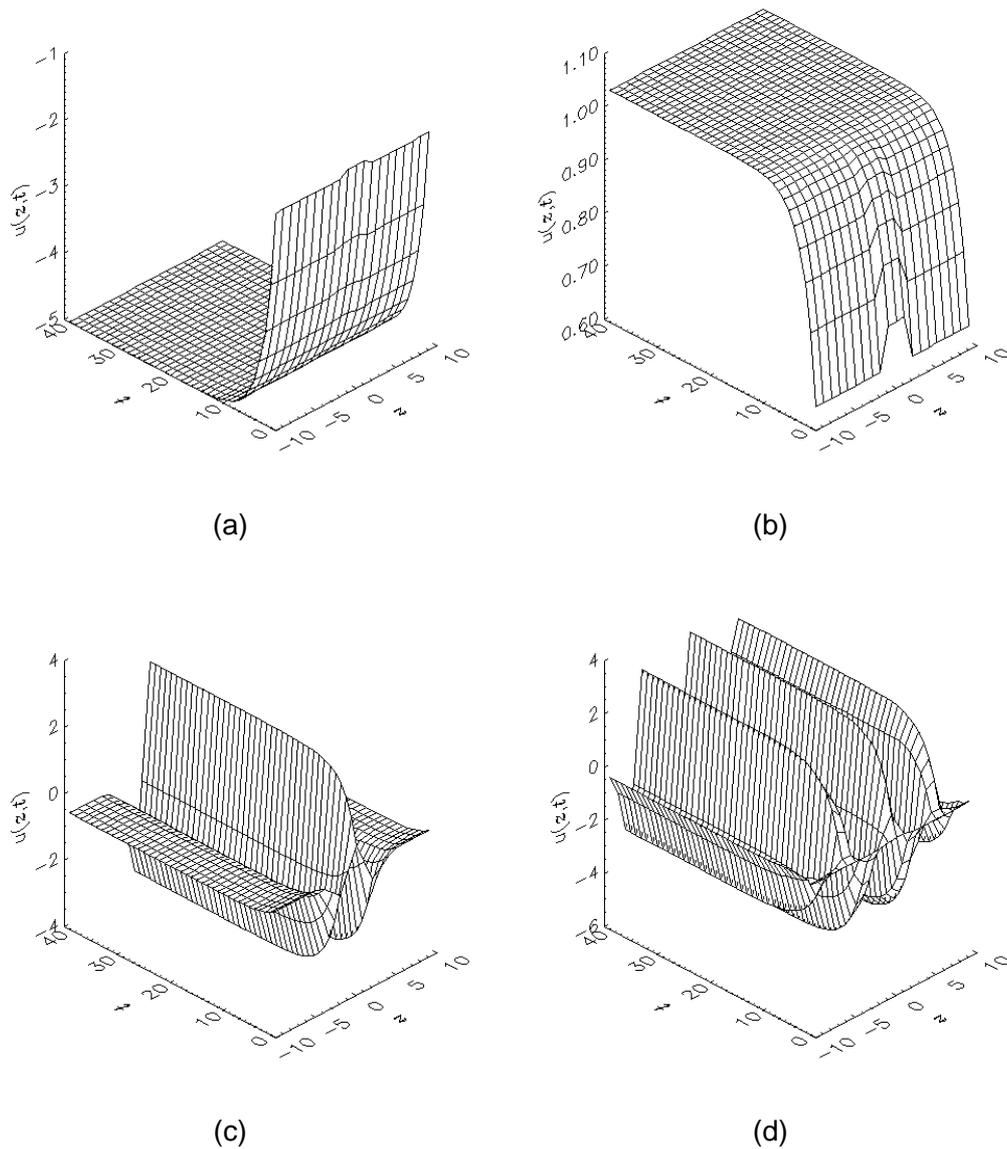


Abbildung 6.6: Mögliche Zustände und Lösungen eines eindimensionalen Amarifeldes über der Zeit betrachtet. (a) 0 -Lösung, (b) ∞ -Lösung, (c) lokale a -Lösung und (d) *multi-peak*-Lösung.

6.1.3 Stabilitätsbetrachtung

Das dynamische Verhalten von Amarifeldern wird in [1] anhand eines eindimensionalen Amarifeldes untersucht. Dabei wird als Nichtlinearität $\varphi(u)$ eine Sprungfunktion (Gleichung 5.2) eingesetzt. Diese kann durch eine Sigmoidfunktion $\xi(u)$ (Gleichung 5.3) beliebig genau approximiert werden mit

$$\lim_{c \rightarrow \infty} \xi(u) = \lim_{c \rightarrow \infty} \left(\frac{1}{1 + e^{-cu}} \right) = \begin{cases} 0 & : u < 0 \\ 0.5 & : u = 0 \\ 1 & : u > 0 \end{cases}$$

Im folgenden wird das Stabilitätsverhalten eindimensionaler Amarifelder diskutiert. Ist die Stabilität höherdimensionaler Amarifelder zu untersuchen, wird auf [2] verwiesen. Dort wird mit Hilfe eines Lyapunov Funktional [92, 66] die Stabilität für höherdimensionale Amarifelder analysiert.

In dem in dieser Arbeit realisierten Verfahren werden nur eindimensionale Felder eingesetzt, weshalb das Stabilitätsverhalten dieser Felder näher betrachtet wird.

Am Beispiel einer a -Lösung mit den Grenzen z_1, z_2 und $R(u) = (z_1, z_2)$ wird eine Stabilitätsanalyse unter Einwirkung eines zeitlich und örtlich begrenzten Stimulus $S(z, t)$ für ein eindimensionales neuronales Feld $u(z, t)$ durchgeführt. Dieser Fall ist für das realisierte Verfahren von besonderem Interesse. Aufgrund der homogenen Eigenschaften des Feldes wird ohne Beschränkung der Allgemeinheit $u'(z) = u(z - z_1)$ betrachtet. Es wird davon ausgegangen, daß sich das System noch nicht im eingeschwungenen Zustand befindet und noch keine stationäre Lösung existiert. Daraus folgt $R(u') = (0, a)$. Somit gilt für eine nichtstationäre a -Lösung

$$R(u'(z, t + dt)) = (0, a(t + dt)).$$

Da eine a -Lösung begrenzt ist, gilt weiterhin

$$\begin{aligned} u'(0, t) = 0 \quad \wedge \quad u'(a, t) = 0 & \quad \text{zum Zeitpunkt } t \quad \text{und} \\ u'(0, t + dt) = 0 \quad \wedge \quad u'(a + da, t + dt) = 0 & \quad \text{zum Zeitpunkt } t + dt \quad , \end{aligned}$$

für $a(t + dt) = a + da$. Mit Hilfe einer Taylorentwicklung [9] kann eine dynamische Gleichung für die Änderung der Begrenzung $a(t)$ Erregungspeaks $R(u')$ in Abhängigkeit der Energie W formuliert werden [1]. Daraus folgt:

$$\begin{aligned} \frac{da}{dt} &= \frac{1}{\tau} \left(\frac{1}{c_1} + \frac{1}{c_2} \right) (W(a) + h), \text{ mit} \\ c_1 &= \frac{\partial u(z_1, t)}{\partial z}, \\ c_2 &= -\frac{\partial u(z_2, t)}{\partial z}, \text{ und} \\ W(a) &= \int_0^a w(z) dz. \end{aligned} \tag{6.4}$$

Im eingeschwungenen Zustand mit $\frac{da}{dt} = 0$ ist eine stationäre Lösung gegeben. Es ergibt sich Stabilität [56, 92] für

$$\frac{dW(a)}{da} < 0 \quad (6.5)$$

und Instabilität für

$$\frac{dW(a)}{da} > 0. \quad (6.6)$$

In Abbildung 6.7 ist das Integral über den Wechselwirkungskern (Stammfunktion) nach Gleichung 6.4 aufgetragen. Werden mögliche Lösungen a_i betrachtet, ergibt sich für $h < 0$ in a_1 eine instabile und a_2 eine stabile (Gleichungen 6.6 und 6.5) *peak*-Lösung. Dabei gilt $W_{max} > -h$, mit $W_{max}(a) = \max_{a>0}(W(a))$. Für $W_{max} < -h$ ist die stabile Lösung die \emptyset -Lösung.

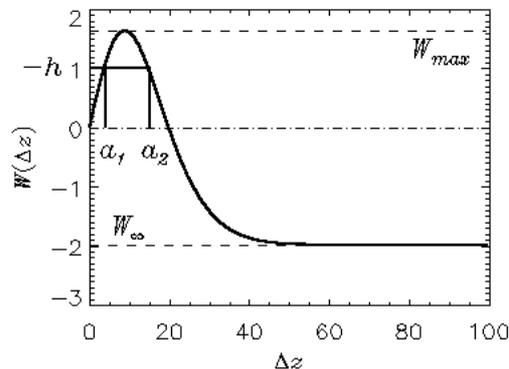


Abbildung 6.7: Verlauf der Stammfunktion W eines Wechselwirkungskerns w der Form eines *mexican hat*.

Aufgrund der vorgenannten Eigenschaften kann ein eindimensionales Amarifeld $u(z, t)$ durch geeignete Wahl der Voraktivierung h und des Wechselwirkungskerns w als zeitlich stabiler Ansatz für die Realisierung eines Verhaltensmodells genutzt werden, in dem eine a -Lösung den Zustand einer Verhaltensvariable beschreibt.

Der Ansatz der neuronalen Felder erlaubt bei der Verhaltenssteuerung eine kontinuierliche Ausgabe der Verhaltensvariable unter Berücksichtigung eines additiv einkoppelnden Stimulus. Dieser Sachverhalt wird in dieser Arbeit genutzt, um eine Bewegungssteuerung zu realisieren. Bei einfacher (additiver) Einkopplung von Objektattributen, beeinflusst der Feldzustand eines eindimensionalen neuronalen Feldes den Lenkwinkel bzw. die Geschwindigkeit eines Trägerfahrzeugs.

In den folgenden Abschnitten wird ein neues Verfahren vorgestellt, welches die genannten Eigenschaften aufweist.

6.2 Eingangsdatenmaterial

Das autonome Navigieren eines Kraftfahrzeugs im Straßenverkehr bedingt, daß Position, Bewegungsrichtung und Relativgeschwindigkeit anderer Objekte beurteilt werden können, um eine unfallfreie Steuerung des Fahrzeugs zu garantieren. Die Funktionstüchtigkeit des implementierten Verfahrens muß anhand dieser Größen an verschiedenen Verkehrsszenarien getestet werden. Dabei sind kritische Szenen, welche hohe Anforderungen an das Verfahren stellen, von besonderem Interesse. Solche Szenen beinhalten jedoch ein großes Gefahrenpotential für die Verkehrsteilnehmer.

Durch die Entwicklung einer Simulationsumgebung (Anhang B), die die Umsetzung kritischer Szenen erlaubt, ist es möglich, die Position, Bewegungsrichtung und Relativgeschwindigkeit von Objekten zu beeinflussen und das realisierte Verfahren ohne Gefahr für andere zu testen. In Abbildung 6.8 sind Ergebnisbilder der Simulation gegeben.

Abbildung 6.8a stellt ein Bild einer definierten Szene aus der Vogelperspektive dar. In dieser Szene fährt das autonome Trägerfahrzeug (schwarz, mit angedeutetem Kamerasensor) auf einer zweispurigen Straße in einer Rechtskurve an parkenden Fahrzeugen vorbei und folgt einem weiteren Fahrzeug. Abbildung 6.8b zeigt ein Ergebnisbild des simulierten Kameradatenmaterials für eine Videokamera (technische Daten: Sensorchipfläche: 36mm^2 , Höhe: 1.2m , Öffnungswinkel von 90° und Auslenkung der Kamera aus der Fahrbahnparallelen Ebene: 4° in Richtung Fahrbahn).

Die aus den simulierten Sensordaten berechneten Größen, die dem entwickelten Verfahren der Bewegungssteuerung als Eingangsgrößen zur Verfügung stehen, sind die relative Distanz $d_O(\psi)$ zu den Objekten O , sowie deren Relativgeschwindigkeiten $v_{xO}(\psi)$ und $v_{yO}(\psi)$ in x - und y -Richtung. Dabei bezeichnet die y -Koordinate die Fahrtrichtung und die x -Koordinate die Senkrechte auf y (Abbildung 6.9). In Abbildung 6.10 sind diese Größen der in Abbildung 6.8 dargestellten Szene in Abhängigkeit von der Winkelposition ψ aufgetragen. Dabei bemißt der Winkel ψ den von dem Sensor abgedeckten Winkelbereich.

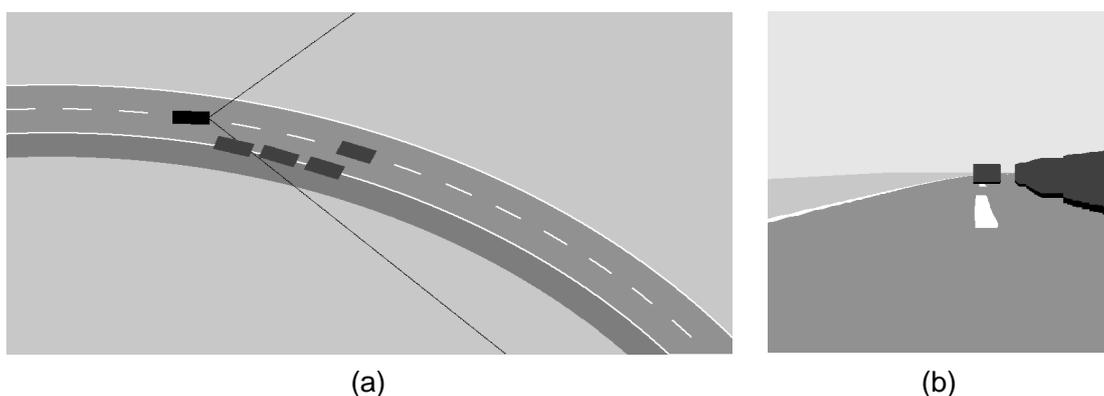


Abbildung 6.8: Simuliertes Sensordatenmaterial für eine Straßenverkehrsszene bei einer leichten Rechtskurve mit parkenden Autos am rechten Fahrbahnrand. Die Fahrtrichtung des Trägerfahrzeugs (schwarz) ist von links nach rechts. (a) Vogelperspektive und (b) Kamerabild für einen Öffnungswinkel von 90° .

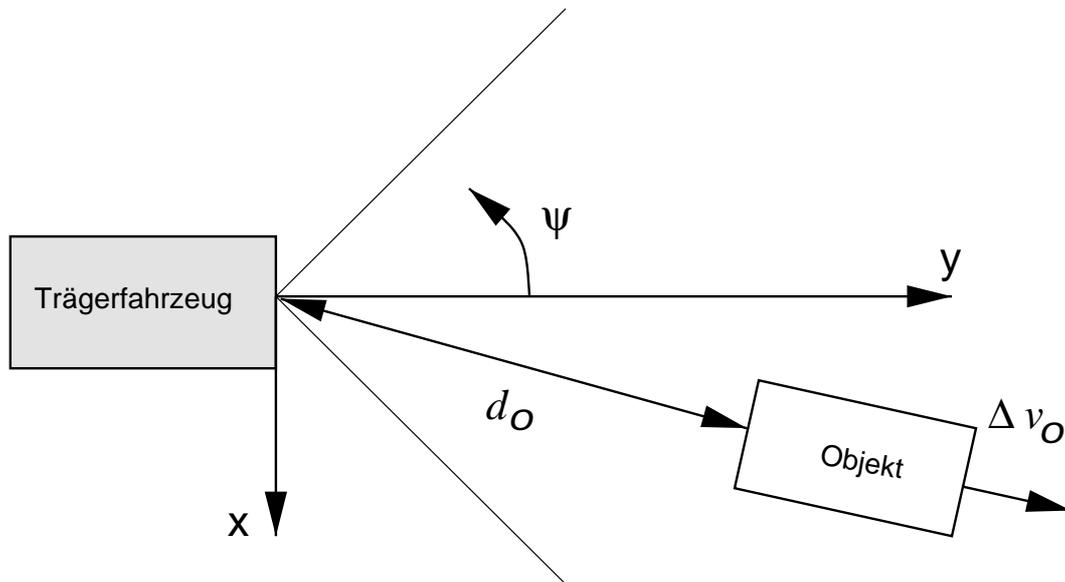


Abbildung 6.9: Koordinatensystem der betrachteten Zustandsgrößen. Die y -Koordinate bezeichnet die Fahrtrichtung und die x -Koordinate die Senkrechte auf y . Der Winkel ψ bemisst den von dem Sensor abgedeckten Winkelbereich.

Es werden vier Fahrzeuge im Winkelbereich ψ von $\psi \simeq -9^\circ$ bis $\psi \simeq -45^\circ$ bei unterschiedlicher Distanz (Abbildung 6.10b) erkannt. Abbildung 6.10a beschreibt die Belegung $b_O(\psi) = 1$ im observierten Bereich in Abhängigkeit des Winkels ψ . Drei am rechten Fahrbahnrand parkende Autos werden erkannt. Das vorherfahrende Fahrzeug fährt mit nahezu gleicher Geschwindigkeit wie das Trägerfahrzeug (Abbildung 6.10c,d).

Die von der Simulation bereitgestellten Daten werden für eine neuronale Bewegungssteuerung als Eingangsgrößen genutzt.

6.3 Bewegungssteuerung

Für die Realisierung eines Verfahrens zur Bewegungssteuerung mit einem Amarrifeldansatz ist es notwendig, die vorhandenen Attribute der Objekte vorzuverarbeiten, so daß die Ergebnisse additiv über den Stimulus S in die neuronalen Felder für die Lenkwinkel- und Geschwindigkeitssteuerung eingekoppelt werden können. In Abschnitt 6.3.1 wird diese Vorverarbeitung beschrieben und in den Abschnitten 6.3.2 und 6.3.3 die Feldzustände und deren Einfluß auf das Verhalten des Fahrzeugs diskutiert.

6.3.1 Vorverarbeitung

Die für die Navigation eines Fahrzeugs im Straßenverkehr notwendigen Zustandsgrößen vorhandener Objekte sowie der freie Fahrraum werden von der simulierten Objekterkennung bereitgestellt. Es stehen dabei die objektbezogenen Attribute der relativen Distanz $d_O(\psi)$ sowie der Relativgeschwindigkeit $v_{xO}(\psi)$ und $v_{yO}(\psi)$ in x - und y -Richtung zur Verfügung. Aus den

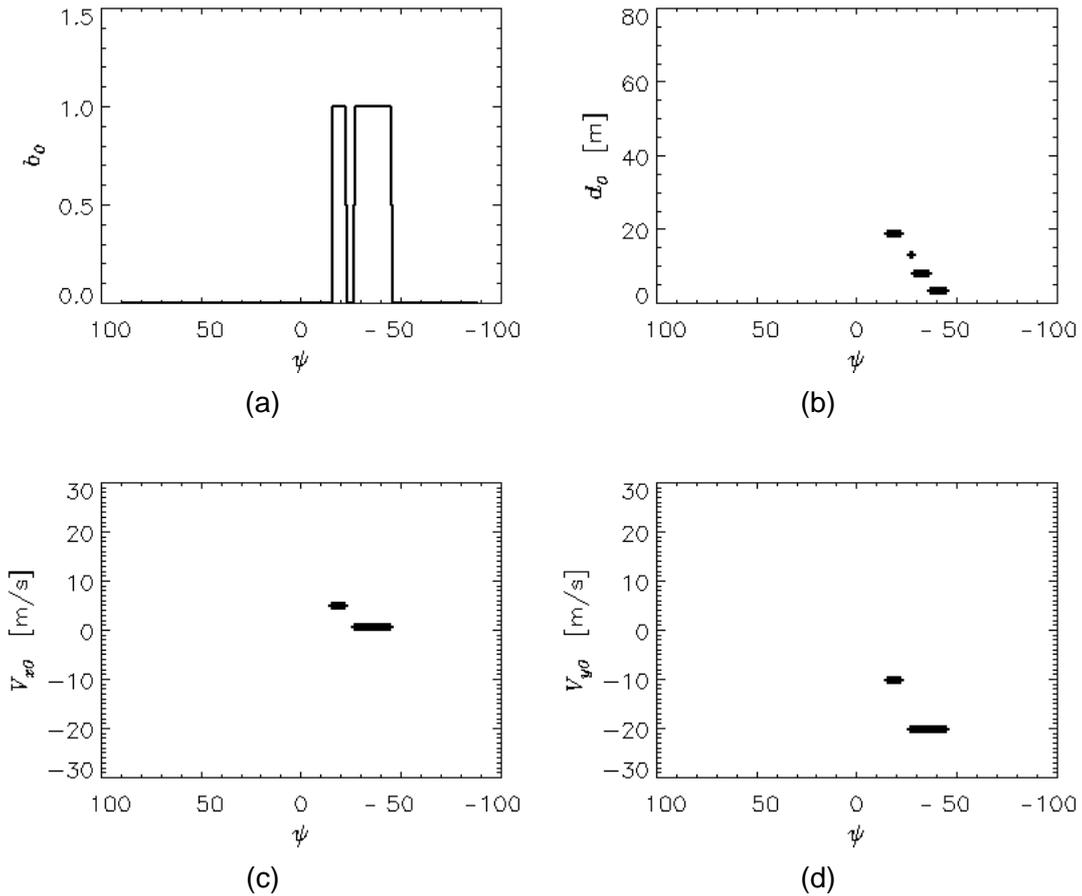


Abbildung 6.10: Sensordatenmaterial für eine Straßenverkehrsszene bei einer leichten Rechtskurve mit Gegenverkehr, einem führenden Fahrzeug und parkenden Autos am rechten Fahrbahnrand. (a) Objekte über den Winkel aufgetragen, (b) Distanz der Objekte, (c) Relativgeschwindigkeit in x und (d) Relativgeschwindigkeit in y .

Objekten kann ein führendes Fahrzeug mit $v_{xF}(\psi)$, $v_{yF}(\psi)$ und $d_F(\psi)$ extrahiert werden. Weiterhin ist aus dem simulierten Kamerabild der Fahrspurverlauf bekannt.

Vorverarbeitung in Abhängigkeit des Lenkwinkels

Die vorliegenden winkelabhängigen Größen der Objekte werden in einer Vorverarbeitung für die Steuerung des Lenkwinkels in ein Gefahrenpotential $O(\psi)$ mit

$$O(\psi) = \zeta_o \cdot \frac{(\min(v_{xo}(\psi), 0))^2 + (\min(v_{yo}(\psi), 0))^2}{d_o(\psi)}$$

umgerechnet und der Einflußbereich auf die Nachbarschaft im Stimulus $S_o(\psi)$ durch eine Faltung mit einem Faltungskern der Art eines *mexican hat* mit

$$S_o(\psi) = - \int_{-\gamma_1}^{\gamma_1} \left(\alpha_o \cdot e^{-\nu_o(\psi-\psi')^2} - \beta_o \cdot e^{-\delta_o(\psi-\psi')^2} \right) \cdot O(\psi') d\psi' + \eta_o$$

berechnet. Dabei beschreibt der Bereich von $-\gamma_1$ bis γ_1 die Grenzen des betrachteten Winkelbereichs. Die geschwindigkeitsabhängigen Gewichtungsfaktoren ζ_O ($\zeta_O \in \mathbb{R}^+$), α_O ($\alpha_O \in \mathbb{R}^+$), β_O ($\beta_O \in \mathbb{R}^+$), ν_O ($\nu_O \in \mathbb{R}^+$) und δ_O ($\delta_O \in \mathbb{R}^+$) sowie die Schwelle η_O ($\eta_O \in \mathbb{R}^+$) können aufgrund der Komplexität des Lösungsraums nicht berechnet werden. Sie werden heuristisch festgelegt. Die Notwendigkeit einer Schwelle η_O ergibt sich aus dem inhibitorischen Anteil des Faltungskerns, der nur bei einem sehr hohem Gefahrenpotential $O(\psi)$ als positiver Stimulus $S_O(\psi)$ im Lenkwinkelfeld wirken soll.

Die Wahl eines Faltungskerns vom Typ eines *mexican hat* liegt in den exzitatorischen und inhibitorischen Eigenschaften begründet, wodurch das Verhalten des neuronalen Feldes bei additiver, bzw. subtraktiver Einkopplung der Stimuli einfach beeinflusst werden kann (Gleichung 6.1).

Der Fahrspurverlauf $L(\psi)$ wird kodiert durch den Verlauf der Fahrspurmitte in der Umwelt in einer definierten Distanz d_L . Er ist abhängig von ψ . Ziel der autonomen Steuerung des Lenkwinkels ist eine möglichst genaue Einhaltung der Fahrspur unter der Randbedingung, den Fahrkomfort nicht zu beeinträchtigen. Daraus ergibt sich die Forderung, daß Trägerfahrzeug bei freiem Fahrraum entlang der Straße zu führen, ohne Kurven zu schneiden oder den Randbereich der Fahrbahn zu übertreten. Diese Anforderungen an den Stimulus $S_L(\psi)$ lassen sich mathematisch durch eine Faltung von der Art

$$S_L(\psi) = \int_{-\gamma_1}^{\gamma_1} \left(\alpha_L \cdot e^{-\nu_L(\psi-\psi')^2} - \beta_L \cdot e^{-\delta_L(\psi-\psi')^2} \right) \cdot L(\psi') d\psi' \quad (6.7)$$

beschreiben. Die komfortabhängigen Gewichtungsfaktoren α_L ($\alpha_L \in \mathbb{R}^+$), β_L ($\beta_L \in \mathbb{R}^+$), ν_L ($\nu_L \in \mathbb{R}^+$) und δ_L ($\delta_L \in \mathbb{R}^+$) sowie die Distanz d_L ($d_L > 0$) der berücksichtigten Spurlage werden heuristisch festgelegt.

Ist durch die Aufgabenstellung eine definierte Fahrtrichtung $F_\psi(\psi)$ vorgegeben, wie beispielsweise bei Folgefahrten, kann dieser Sachverhalt durch den Stimulus

$$S_{\psi F}(\psi) = \int_{-\gamma_1}^{\gamma_1} \left(\alpha_{1F} \cdot e^{-\nu_{1F}(\psi-\psi')^2} - \beta_{1F} \cdot e^{-\delta_{1F}(\psi-\psi')^2} \right) \cdot F_\psi(\psi') d\psi'$$

beschrieben werden. Auch hier sind die Gewichtungsfaktoren α_{1F} ($\alpha_{1F} \in \mathbb{R}^+$), β_{1F} ($\beta_{1F} \in \mathbb{R}^+$), ν_{1F} ($\nu_{1F} \in \mathbb{R}^+$) und δ_{1F} ($\delta_{1F} \in \mathbb{R}^+$) komfortabhängig und werden heuristisch festgelegt.

Die berechneten Größen $S_O(\psi)$, $S_L(\psi)$ und $S_{\psi F}(\psi)$ werden additiv als Stimuli in ein neuronales Feld zur Steuerung des Lenkwinkels eingekoppelt (Abbildung 6.11). In diesem Feld ist die mögliche Lenkwinkeländerung ψ als Feldgröße definiert.

Die Formulierung des jeweiligen Verhaltenssteuerungssystems als neuronales Feld erlaubt es, einen kontinuierlichen Stimulus S als gewichtete Summe der genannten Anforderungen darzustellen. Das Feld ermittelt dabei in einem dynamischen Prozeß eine Lösung, die sich in der Position des *peaks* einer *a*-Lösung ablesen läßt.

Die Festlegung der in diesem Abschnitt betrachteten Gewichtungsfaktoren ist in weiten Bereichen unkritisch. Sie werden anhand der Aufgabenstellung, gekoppelt an die konkrete Geometrie und Physik des Systems, festgelegt.

Vorverarbeitung in Abhängigkeit der Geschwindigkeit

Zur Steuerung der Geschwindigkeit des Fahrzeugs wird die Differenz $D_v(\Delta v)$ zwischen aktueller und gewünschter Geschwindigkeit, vorgegeben beispielsweise durch Verkehrsregeln, betrachtet

und ein Stimulus $S_D(\Delta v)$ mit

$$S_D(\Delta v) = \int_{-\gamma_2}^{\gamma_2} \left(\alpha_D \cdot e^{-\nu_D(\Delta v - \Delta v')^2} - \beta_D \cdot e^{-\delta_D(\Delta v - \Delta v')^2} \right) \cdot D_v(\Delta v') d\Delta v'$$

und den Gewichtungsfaktoren α_D ($\alpha_D \in \mathbb{R}^+$), β_D ($\beta_D \in \mathbb{R}^+$), ν_D ($\nu_D \in \mathbb{R}^+$) sowie δ_D ($\delta_D \in \mathbb{R}^+$) bestimmt. Dabei beschreibt der Bereich von $-\gamma_2$ bis γ_2 die Grenzen des betrachteten Geschwindigkeitsbereichs. Die Variable Δv beschreibt die Feldvariable des Geschwindigkeitsfelds.

Die Relativgeschwindigkeit eines vorherfahrenden Kraftfahrzeugs (bzw. eines Hindernisses auf der Fahrbahn), gegeben durch v_{yF} und v_{xF} und dessen Abstand d_F werden genutzt, um die Geschwindigkeit des Trägerfahrzeugs an die Verkehrsverhältnisse anzupassen. Es lässt sich ein Gefahrenpotential $F_v(\Delta v)$ an der Stelle $\Delta v = v_F$ mit

$$v_F = \text{sign}(v_{yF}) \sqrt{v_{xF}^2 + v_{yF}^2} \text{ und}$$

$$F_v(v_F) = \zeta_F \frac{|v_F|}{d_F^2}$$

berechnen. Daraus wird ein Stimulus $S_{vF}(\Delta v)$ für das neuronale Feld

$$S_{vF}(\Delta v) = \int_{-\gamma_2}^{\gamma_2} \left(\alpha_{2F} \cdot e^{-\nu_{2F}(\Delta v - \Delta v')^2} - \beta_{2F} \cdot e^{-\delta_{2F}(\Delta v - \Delta v')^2} \right) \cdot F_v(\Delta v') d\Delta v'$$

mit den Gewichtungsfaktoren ζ_F ($\zeta_F \in \mathbb{R}^+$), α_{2F} ($\alpha_{2F} \in \mathbb{R}^+$), β_{2F} ($\beta_{2F} \in \mathbb{R}^+$), ν_{2F} ($\nu_{2F} \in \mathbb{R}^+$) und δ_{2F} ($\delta_{2F} \in \mathbb{R}^+$) abgeleitet. Die Gewichtungsfaktoren werden heuristisch festgelegt.

Die in diesem Abschnitt beschriebenen Gewichtungsfaktoren werden anhand der Aufgabenstellung, gekoppelt an die konkrete Geometrie und Physik des Systems, festgelegt. Ihre Festlegung ist in weiten Bereichen unkritisch.

Die Stimuli $S_D(\Delta v)$ und $S_{vF}(\Delta v)$ werden in ein neuronales Feld zur Steuerung der Geschwindigkeit eingekoppelt (Abbildung 6.13). In diesem Feld ist die mögliche Geschwindigkeitsänderung Δv als Feldgröße definiert.

In den Abschnitten 6.3.2 und 6.3.3 werden die Feldgleichungen abgeleitet und die Ergebnisse diskutiert.

6.3.2 Steuerung des Lenkwinkels

Bei der Lenkwinkelsteuerung ist die Aufgabenstellung, alle Objekte zu umfahren, die makroskopische Beschreibung des gewünschten Verhaltens, das durch den Lenkwinkel als Verhaltensvariable kodiert wird (Kapitel 2 und 3). Die Zustandsgrößen des vorverarbeiteten Datenmaterials werden als mikroskopische Größen verwendet, die additiv mit

$$S(\psi, t) = S_O(\psi, t) + S_L(\psi, t) + S_{\psi F}(\psi, t)$$

in Abhängigkeit von ψ und t als Stimulus $S(\psi, t)$ in das neuronale Feld

$$\tau \dot{u}(\psi, t) = -u(\psi, t) + S(\psi, t) + h + \int_{-\gamma_1}^{\gamma_1} w(\psi - \psi') \varphi(u(\psi', t)) d\psi' \quad (6.8)$$

eingekoppelt werden. Dabei beschreibt die Variable ψ mögliche Lenkwinkeländerungen als Feldvariable des Amarrifeldes im Bereich von $-\gamma_1$ bis γ_1 . Nach dem in Abschnitt 6.1 beschriebenen Verhalten neuronaler Felder bildet das eindimensionale Feld bei Anliegen eines positiven Eingangssignals $S(\psi, t)$ und bei geeigneter Wahl des Wechselwirkungskerns w sowie der Voraktivierung h nach einer Einschwingzeit in Abhängigkeit von τ eine sogenannte a -Lösung aus.

In der in Abbildung 6.11 dargestellten Aktivitätsverteilung des neuronalen Feldes (durchgezogene Linie) werden die Eingangsdaten aus der in Abschnitt 6.2 beschriebenen Szene zugrunde gelegt (Abbildung 6.8). Es wird die Voraktivierung auf $h = -1$ gesetzt. Der Wechselwirkungskern vom Typ eines *mexican hat* wird so dimensioniert, daß ein stabiles Verhalten des Feldes gewährleistet ist und eine a -Lösung existiert (siehe Abschnitt 6.1). Die Neuronen des Feldes kodieren mögliche Winkel ψ in einem Wertebereich von $-90^\circ \leq \Delta\psi \leq 90^\circ$.

Das Argument $N(t)$ des Maximums der a -Lösung des Feldes mit

$$N(t) = \arg \max_{\psi} (u(\psi, t))$$

kodiert die gewünschte Lenkwinkeländerung $N(t)$ in Grad zum aktuellen Lenkwinkel in Abhängigkeit von t . In Abbildung 6.11 entspricht der Feldzustand $u(\psi, t_1)$ zum dargestellten Zeitpunkt t_1 einer gewünschten Winkeländerung von $N(t_1) = -2^\circ$.

In Abbildung 6.11 sind zur Veranschaulichung die mikroskopischen Zustandsgrößen $S_O(\psi)$, $S_L(\psi)$ und $S_{\psi F}(\psi)$ mit einer verschobenen Nulllinie ($\tilde{S}_O(\psi) = S_O(\psi) + 4$, $\tilde{S}_L(\psi) = S_L(\psi) + 8$ und $\tilde{S}_{\psi F}(\psi) = S_{\psi F}(\psi) + 12$) eingetragen.

Um einen komfortablen Fahrtverlauf zu gewährleisten, wird eine Unterdrückung abrupter Lenk-

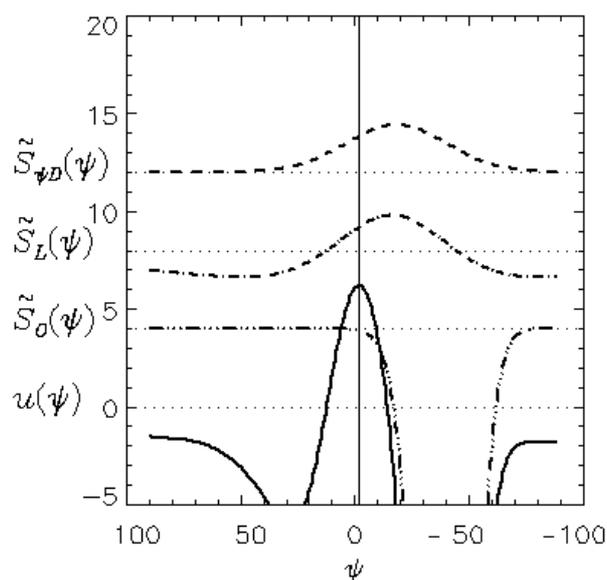


Abbildung 6.11: Aktivität des Neuronalen Feldes sowie die Stärke der Einkopplung von Hindernissen, Straßenverlauf und gewünschter Fahrtrichtung zum Zeitpunkt $t_1 = 65$. Sensordatenmaterial für eine Straßenverkehrsszene bei einer leichten Rechtskurve mit vorherfahrendem Fahrzeug und parkenden Autos am rechten Fahrbahnrand.

winkeländerungen durch eine Minimumoperation mit

$$\Delta W = \text{sign}(N) \cdot \alpha_N \min(|N|, N_{max}), \text{ und}$$

$$W = W + \Delta W$$

durchgeführt. Der Gewichtungsfaktor α_N ($\alpha_N \in \mathbb{R}^+$) ist geschwindigkeitsabhängig und die Variable N_{max} beschreibt die maximal zulässige Lenkwinkeländerung des aktuellen Lenkwinkels W . Wie in Abbildung 6.11 dargestellt, werden im negativen Winkelbereich Objekte erkannt. Die verschobenen Maxima von $S_L(\psi)$ und $S_F(\psi)$ weisen auf einen Ausschervorgang des Trägerfahrzeugs nach links hin.

In Abbildung 6.12a,b ist der Verlauf des Stimulus S und die Feldaktivität u über der Zeit aufgetragen. Dabei ist der negative Anteil des Stimulus S im Bereich der Winkelposition der parkenden Fahrzeuge zu beachten. Das Maximum der Feldaktivität u wird bei Auftreten dieser

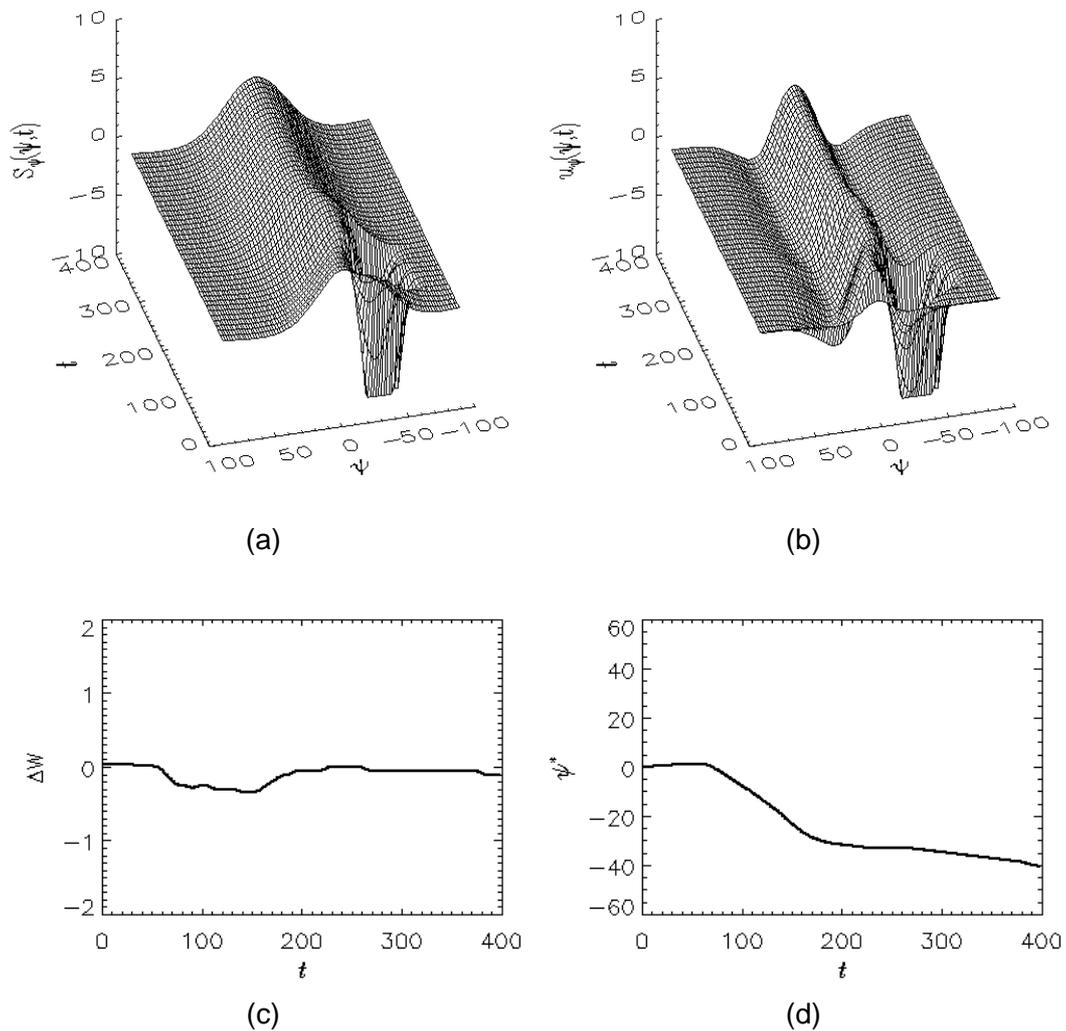


Abbildung 6.12: Darstellung (a) des Stimulus, (b) der Aktivität der Feldes, (c) der Lenkwinkeländerung und (d) der Winkelposition des Trägerfahrzeugs zu einem stationären Beobachter über der Zeit.

Hindernisse nach links verschoben. Betrachtet man den Verlauf des Lenkwinkels über der Zeit, wird das Durchfahren einer leichten Rechtskurve erkannt. In Abbildung 6.12c,d ist der Verlauf der Lenkwinkeländerung, sowie der Winkel ψ^* zu einem stationären Beobachter, der sich an der Ausgangsposition des Trägerfahrzeugs befindet, dargestellt.

Der Verlauf der Lenkwinkeländerung liegt über die gesamte Sequenz in der Nähe von Null. Es werden keine starken Lenkbewegungen unternommen. Wird die Position des Trägerfahrzeugs von dem stationären Beobachter beurteilt, kann die Durchfahrt einer Rechtskurve, sowie das Ausweichmanöver bei den parkenden Fahrzeugen abgelesen werden. Die annähernd konstante Veränderung der Winkelposition ψ^* zum Beobachter beim Durchfahren der Rechtskurve ist während des Manövers in Richtung des positiven Winkelbereichs verschoben (Abbildung 6.12d).

6.3.3 Steuerung der Geschwindigkeit

In diesem Abschnitt wird die prinzipielle Anwendbarkeit des gewählten Ansatzes für die Steuerung der Geschwindigkeit gezeigt. Es wird eine Kopplung beider Felder über die Umwelt durchgeführt, da eine Geschwindigkeitsabhängigkeit der Stimuli vorliegt.

Bei der Geschwindigkeitssteuerung wird entsprechend Abschnitt 6.3.2 vorgegangen. Der Stimulus S ist dabei mit

$$S(\Delta v, t) = S_D(\Delta v, t) + S_{vF}(\Delta v, t)$$

in Abhängigkeit von Δv und t definiert. Das neuronale Feld wird beschrieben durch

$$\tau \dot{u}(\Delta v, t) = -u(\Delta v, t) + S(\Delta v, t) + h + \int_{-\gamma_2}^{\gamma_2} w(\Delta v - \Delta v') \varphi(u(\Delta v', t)) d\Delta v'.$$

Die Variable Δv kodiert die möglichen Geschwindigkeitsänderungen als Feldvariable des Amari-feldes im Bereich von $-\gamma_2$ bis γ_2 . Der Wechselwirkungskern w vom Typ eines *mexican hat* und die Voraktivierung h ($h = -1$) werden gemäß Abschnitt 6.1 so gewählt, daß sich eine stabile a -Lösung ausbildet.

In Abbildung 6.13 ist die Aktivitätsverteilung des neuronalen Feldes (durchgezogene Linie) dargestellt. Die Neuronen des Feldes kodieren mögliche Relativgeschwindigkeiten Δv in einem Wertebereich von $-50 \frac{m}{s} \leq \Delta v \leq 50 \frac{m}{s}$.

Das Argument $M(t)$ des Maximums der a -Lösung des Feldes mit

$$M(t) = \arg \max_{\Delta v} (u(\Delta v, t))$$

kodiert die gewünschte Geschwindigkeitsänderung $M(t)$ in $\frac{m}{s}$ zur aktuellen Geschwindigkeit in Abhängigkeit von t . In Abbildung 6.13 entspricht der Feldzustand $u(\Delta v, t_1)$ zum dargestellten Zeitpunkt t_1 einer gewünschten Geschwindigkeitsänderung von $M(t_1) = -9 \frac{m}{s}$.

In Abbildung 6.13 sind zur Veranschaulichung die Stimuli $S_D(\Delta v)$, und $S_{vF}(\Delta v)$ mit einer verschobenen Nulllinie ($\tilde{S}_D(\Delta v) = S_D(\Delta v) + 8$ und $\tilde{S}_{vF}(\Delta v) = S_{vF}(\Delta v) + 12$) eingetragen.

Eine Unterdrückung abrupter Geschwindigkeitsänderungen wird durch eine Minimumoperation mit

$$\begin{aligned} \Delta V &= \text{sign}(M) \cdot \alpha_M \min(|M|, M_{max}), \text{ und} \\ V &= V + \Delta V \end{aligned}$$

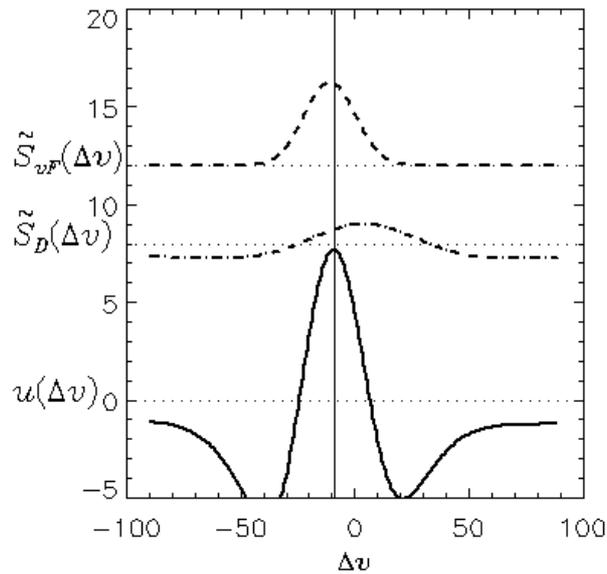


Abbildung 6.13: Aktivität des Neuronalen Feldes, sowie die Stärke der Einkopplung der Stimuli zum Zeitpunkt $t_1 = 65$.

erreicht und ermöglicht einen Fahrtverlauf unter Berücksichtigung physikalischer Randbedingungen. Der Gewichtfaktor α_M ($\alpha_M \in \mathbb{R}^+$) ist komfortabhängig und die Variable M_{max} beschreibt die maximal zulässige Geschwindigkeitsänderung der aktuellen Geschwindigkeit W . Wie in Abbildung 6.13 dargestellt, weist das führende Fahrzeug, beschrieben durch $\tilde{S}_{vF}(\Delta v)$, eine geringere Geschwindigkeit auf. Es wird eine Anpassung der Geschwindigkeit des Trägerfahrzeugs durchgeführt.

In Abbildung 6.14 ist der Verlauf des Stimulus und die Feldaktivität über der Zeit aufgetra-

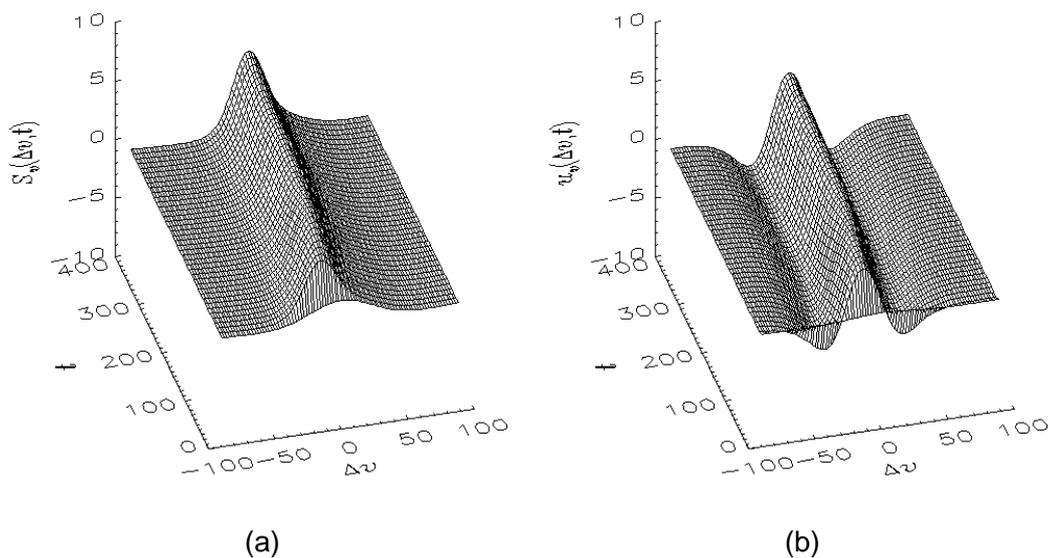


Abbildung 6.14: Darstellung (a) des Stimulus und (b) der Feldaktivität über der Zeit.

gen. Es bildet sich eine a -Lösung aus, die eine starke Verringerung der Geschwindigkeit V des Trägerfahrzeugs bedingt.

In Abbildung 6.15 ist der Geschwindigkeitsverlauf über der Zeit aufgetragen. Von einer anfänglichen Geschwindigkeit von $V \approx 90\text{km/h}$ ($25\frac{\text{m}}{\text{s}}$) wird das Auto auf bis $V \approx 50\text{km/h}$ ($\sim 14\frac{\text{m}}{\text{s}}$) abgebremst.

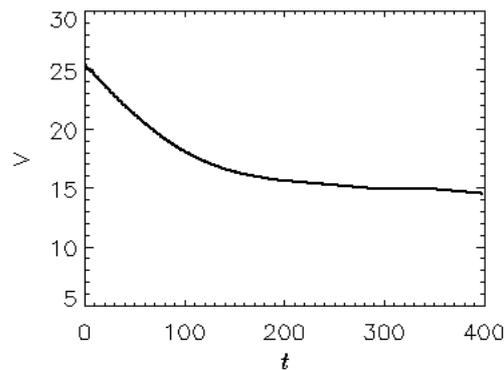


Abbildung 6.15: Darstellung der Geschwindigkeit des Trägerfahrzeugs über der Zeit.

Das Fahrverhalten des Trägerfahrzeugs wird mit Hilfe der Simulationsumgebung an komplexen Szenen getestet. In Abschnitt 6.4 werden mehrere Szenen dargestellt und das Fahrverhalten des Trägerfahrzeugs anhand der neuronalen Lenkwinkelsteuerung diskutiert.

6.4 Ergebnisse

In diesem Abschnitt wird das Verhalten des entwickelten Verfahrens zur Bewegungssteuerung an sechs Beispielszenarien für den innerstädtischen Bereich verdeutlicht. Die ausgewählten Szenarien beinhalten Anforderungen an typische Basisverhalten (z. B. Hindernisvermeidung, Folgefahren und Spurhalten) von Kraftfahrzeugen im Stadtverkehr. In diesem Abschnitt wird im speziellen die Lenkwinkelsteuerung betrachtet, welche den wesentlichen Aspekt bei der Navigation eines Fahrzeugs im Stadtverkehr darstellt.

Bei den Szenarien wird eine simulierte Videokamera mit den technischen Daten nach Abschnitt 6.2 eingesetzt. Es sind jeweils die Vogelperspektive an einem ausgewählten Zeitschritt, sowie das zugehörige Kamerabild dargestellt. Das Verhalten des Stimulus S sowie der Feldzustand u über der Zeit wird verdeutlicht. Die Lenkwinkeländerungen ΔW und die Winkeländerung ψ^* zu einem stationären Beobachter, der sich am Ausgangspunkt des Trägerfahrzeugs befindet, werden diskutiert.

Die ausgewählten Szenarien beschreiben folgende Situationen:

- Folgefahrt des Trägerfahrzeugs in einer Rechtskurve an einem parkendem Fahrzeug vorbei ohne Fahrspurinformation.
- Folgefahrt wie oben, es ist jedoch der Fahrspurverlauf bekannt.
- Folgefahrt an Hindernissen vorbei bei einem entgegenkommenden Fahrzeug.

- Fahrt auf einer Straße bei links abbiegenden führendem Fahrzeug und Gegenverkehr.
- Fahrt durch eine Straßenschlucht ohne führendes Fahrzeug.
- Folgefahrt bei einem rückwärts ausparkenden Fahrzeug und Gegenverkehr.

Eine Diskussion der Ergebnisse wird abschließend durchgeführt.

6.4.1 Basisverhalten

Das Basisverhalten von Kraftfahrzeugen wird anhand von sechs Beispielen durch Beeinflussung des Lenkwinkels untersucht. Es werden dabei die wesentlichen Basisverhalten im Stadtverkehr berücksichtigt. Bei den vorgestellten Beispielen werden auch Situationen beim Ausfall einzelner Stimuli und bei widersprüchlichen Stimuli betrachtet. Insbesondere werden

- Spurhalten,
- Folgefahren,
- Hindernisvermeidung,
- Verhalten bei Gegenverkehr,
- das Einhaltung von Sicherheitsabständen und
- das Verhalten des Verfahrens in Gefahrensituationen

untersucht.

Die aufgezählten Basisverhalten werden in verschiedenen Zusammensetzungen demonstriert, um die Tragfähigkeit des realisierten Verfahrens zu untermauern. Insbesondere werden Szenarien dargestellt, die das Fahrverhalten des Trägerfahrzeugs wesentlich beeinflussen. Beispielsweise stellen Hindernisse auf der Fahrbahn ein großes Problem für autonome Kraftfahrzeuge im Stadtverkehr dar. Mit Hilfe des vorgestellten Verfahrens wird dieses Problem aufgegriffen und eine Lösung durch eine adäquate Lenkwinkelsteuerung beschrieben.

Es wird jeweils eine Rechtskurve durchfahren. Dies stellt keine Einschränkung der Allgemeinheit dar. Aufgrund des Kurvenverlaufs wird eine Anpassung des Lenkwinkels durch das Trägerfahrzeug erzwungen. Dies ist erwünscht, um die Tragfähigkeit des Verfahrens zu zeigen.

Folgefahrt ohne Fahrspurinformation

Das dargestellte Szenario beschreibt die Fahrt des Trägerfahrzeugs auf einer Straße an einem auf der Straße parkenden Fahrzeug vorbei (Abbildung 6.16). Der Fahrspurverlauf ist unbekannt. Der fehlende Fahrspurverlauf beschreibt einen Ausfall des Fahrspurstimulus S_L . Es wird ein anderes Fahrzeug verfolgt. Die Basisverhalten sind Spurhaltung, Folgefahren und Hindernisvermeidung.

In Abbildung 6.17 sind die Ergebnisse über der Zeit dargestellt. Das parkende Fahrzeug wird erkannt und ein Ausweichmanöver durchgeführt. Der Stimulus S wird während der Vorbeifahrt an dem parkenden Fahrzeug im entsprechenden Winkelbereich stark negativ (Abbildung 6.17a), was zu einer Verschiebung der a -Lösung des Feldes führt (Abbildung 6.17b). Es resultiert eine

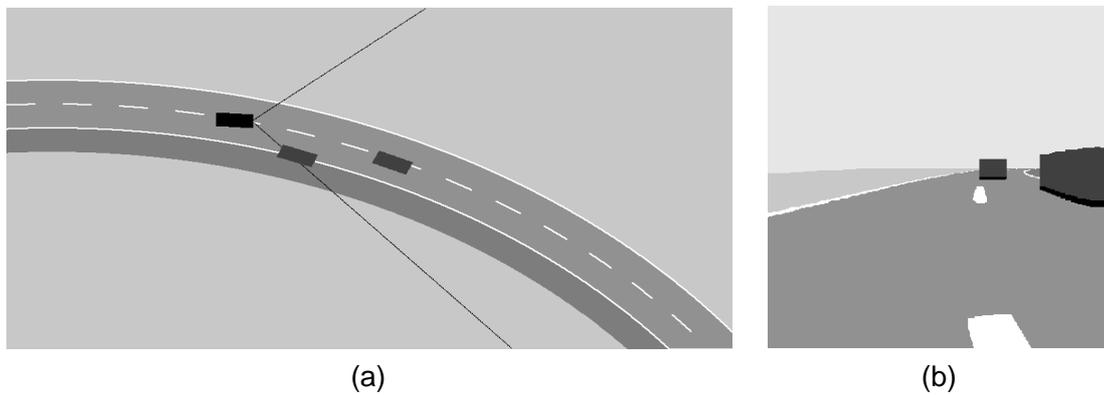


Abbildung 6.16: Folgefahrt ohne Fahrspurinformation. Simuliertes Sensordatenmaterial für eine Straßenverkehrsszene bei einer leichten Rechtskurve mit einem parkenden Auto am rechten Fahrbahnrand zum Zeitpunkt $t_1 = 175$. Die Fahrtrichtung des Trägerfahrzeugs (schwarz) ist von links nach rechts einem führenden Fahrzeug folgend. (a) Vogelperspektive und (b) Kamerabild für einen Öffnungswinkel von 90° .

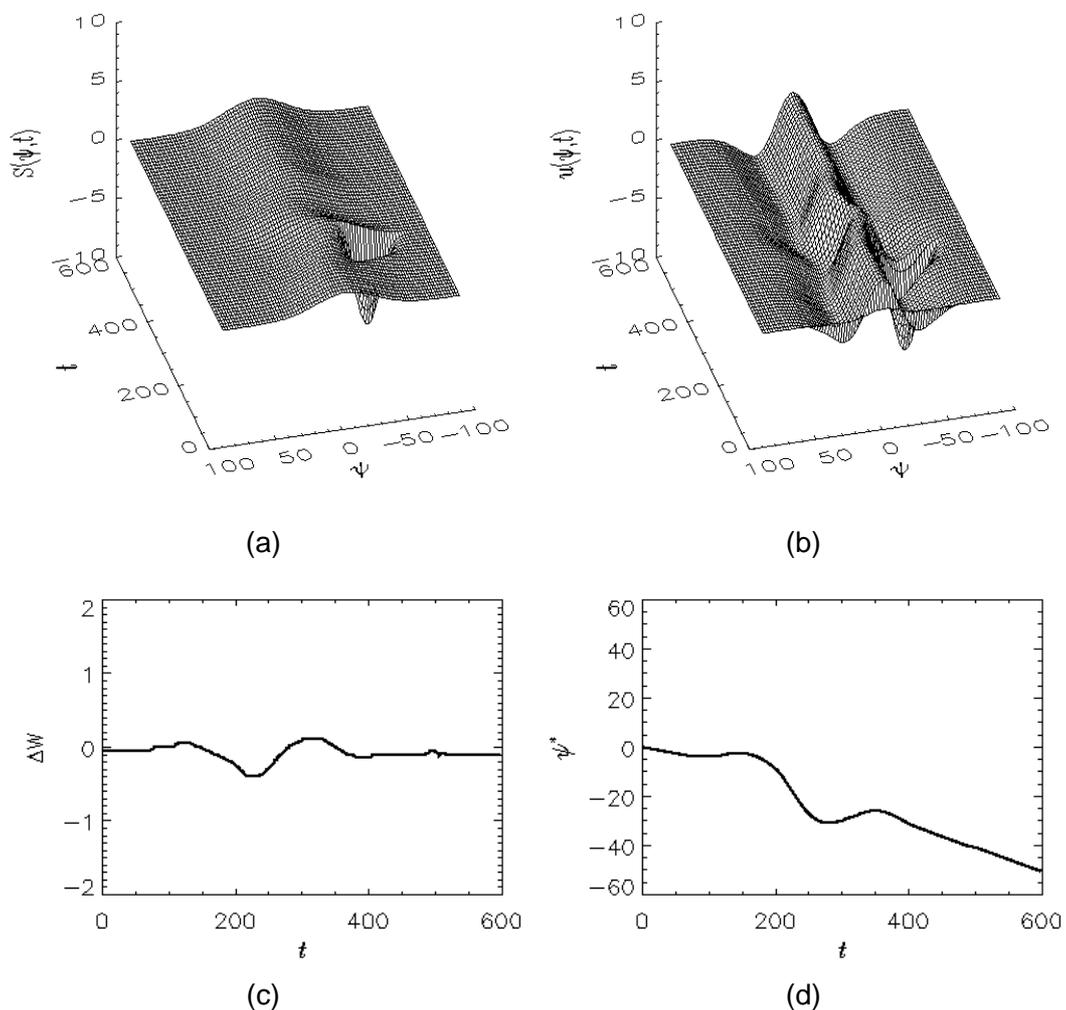


Abbildung 6.17: Folgefahrt ohne Fahrspurinformation. Darstellung (a) des Stimulus, (b) der Aktivität der Feldes, (c) der Lenkwinkeländerung und (d) der Winkelposition des Trägerfahrzeugs zu einem stationären Beobachter über der Zeit.

Änderung des Lenkwinkelschlags (Abbildung 6.17c). Diese Änderung führt zu einer lateralen Positionsverschiebung auf der Fahrbahn nach links (Abbildung 6.17d). Nach Beendigung des Ausweichmanövers wird die Folgefahrt auf der Fahrbahn fortgesetzt. Aufgrund des verschobenen Maximums beim Stimulus S (Kurvenfahrt) und des fehlenden Fahrspurstimulus S_L wird die Kurve leicht geschnitten.

Folgefahrt mit Fahrspurinformation

Das dargestellte Szenario beschreibt die Fahrt des Trägerfahrzeugs auf einer Straße an einem auf der Straße parkenden Fahrzeug vorbei (Abbildung 6.18). Es wird ein weiteres Fahrzeug verfolgt. Die Basisverhalten entsprechen denen im vorhergehenden Abschnitt. Dabei ist der Straßenverlauf jedoch bekannt.

In Abbildung 6.19 sind die Ergebnisse für diese Szene über der Zeit dargestellt. Der Fahrtverlauf des Trägerfahrzeugs ist annähernd der gleiche, wie im vorhergehenden Beispiel. Aufgrund der vorhandenen Fahrspurinformation wird das Ausweichmanöver später eingeleitet und früher beendet (Abbildung 6.19c,d). Die bekannte Fahrspur bedingt ein ausgeprägtes Maximum beim Stimulus, das den Fahrspurverlauf genauer beschreibt. Im Gegensatz zum vorhergehenden Beispiel wird die Kurve nicht geschnitten. Dies stellt eine Verbesserung des durchgeführten Fahrmanövers dar.

Folgefahrt bei Gegenverkehr an parkenden Fahrzeugen vorbei

In diesem Szenario ist die Durchfahrt einer Kurve an parkenden Fahrzeugen vorbei beschrieben. Diese Fahrzeuge blockieren Teile der Fahrbahn, so daß ein Ausweichmanöver notwendig ist. Ein entgegenkommendes Fahrzeug auf der linken Spur muß bei dem Ausweichmanöver berücksichtigt werden (Abbildung 6.20). Die Basisverhalten sind Spurhaltung, Folgefahren, Hindernisvermeidung, Einhaltung von Sicherheitsabständen und Beachten des Gegenverkehrs. In Abbildung 6.21 sind der Stimulus S sowie die Ergebnisse der Lenkwinkelsteuerung über der Zeit dargestellt. Während der Fahrt entlang einer Rechtskurve werden die parkenden Fahrzeuge erkannt und ein Ausweichmanöver durchgeführt. Der darauf folgende Gegenverkehr verursacht den starken Abfall des Stimulus S (Abbildung 6.21a,b) im positiven Winkelbereich. Bedingt durch diesen Einbruch kommt es zu einer starken Verschiebung der a -Lösung von u . Das Trägerfahrzeug wird schnell (starker Lenkwinkelschlag) in die eigene Spur zurückgeführt. Die Folgefahrt wird fortgesetzt.

Parkende Fahrzeuge und Gegenverkehr bei links abbiegendem führenden Fahrzeug

Das dargestellte Szenario beschreibt die Durchfahrt einer Kurve an parkenden Fahrzeugen vorbei. Das führende Fahrzeug verläßt während der Fahrt die Fahrbahn, um links abzubiegen. Ein weiteres Fahrzeug kommt entgegen. In diesem Szenario werden die Basisverhalten Spurhaltung, Hindernisvermeidung, Einhaltung von Sicherheitsabständen und Beachten des Gegenverkehrs bei widersprüchlichen Stimuli $S_{\psi F}$ und S_L gezeigt. Das führende Fahrzeug verläßt die Fahrbahn bei Beibehaltung der Fahrtrichtung des Trägerfahrzeugs.

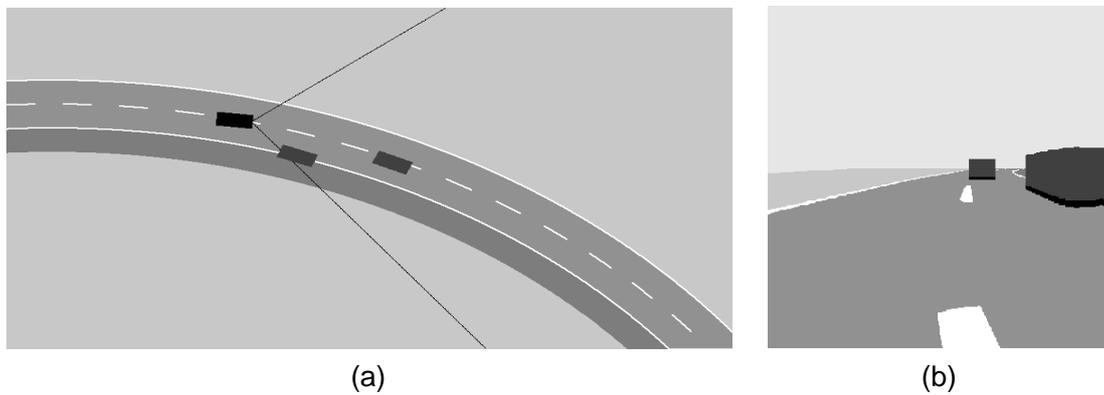


Abbildung 6.18: Folgefahrt mit Fahrspurinformation. Simuliertes Sensordatenmaterial für eine Straßenverkehrsszene bei einer leichten Rechtskurve mit einem parkenden Auto am rechten Fahrbahnrand zum Zeitpunkt $t_1 = 175$. Die Fahrrichtung des Trägerfahrzeugs (schwarz) ist von links nach rechts einem führenden Fahrzeug folgend. (a) Vogelperspektive und (b) Kamerabild für einen Öffnungswinkel von 90° .

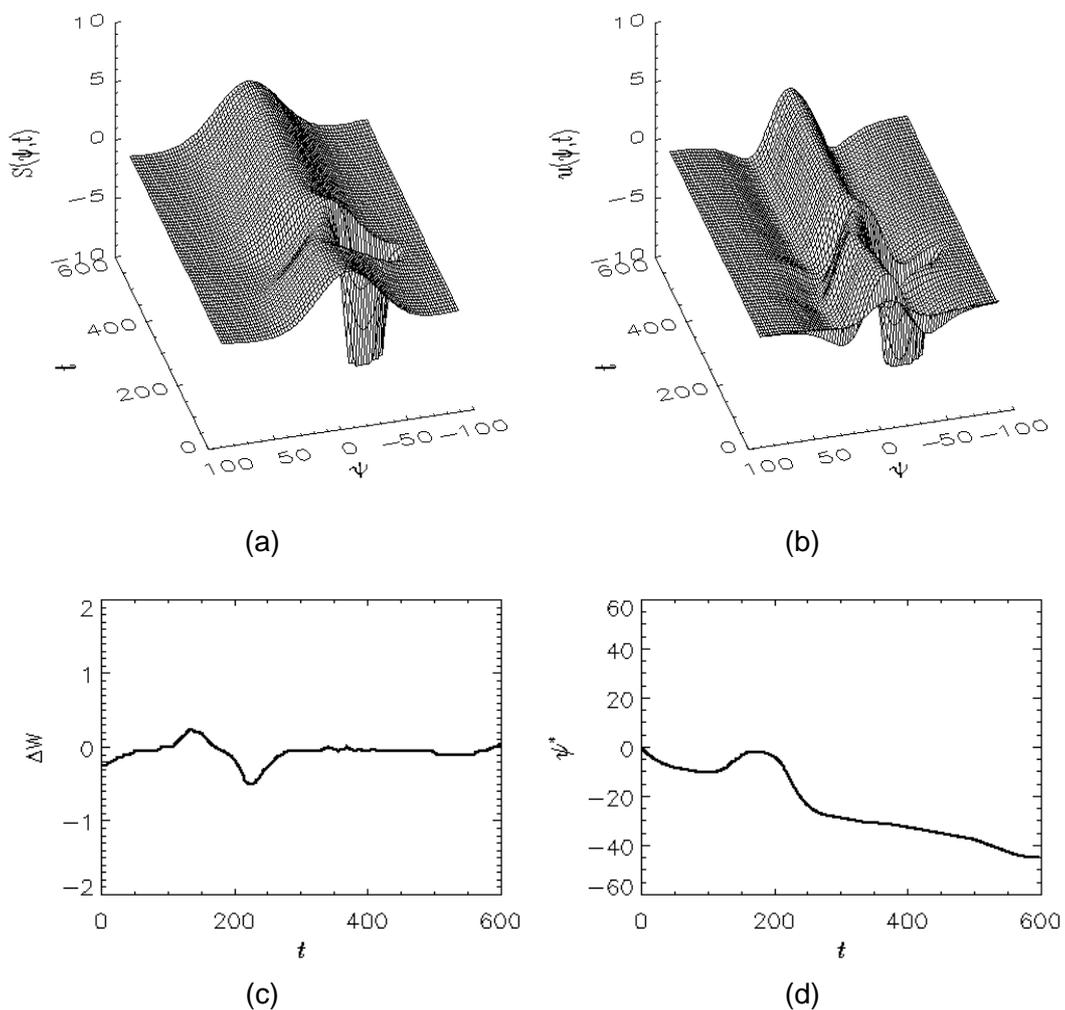


Abbildung 6.19: Folgefahrt mit Fahrspurinformation. Darstellung (a) der Höhe des Stimulus, (b) der Aktivität der Feldes, (c) der Lenkwinkeländerung und (d) der Winkelposition des Trägerfahrzeugs zu einem stationären Beobachter über die Zeit.

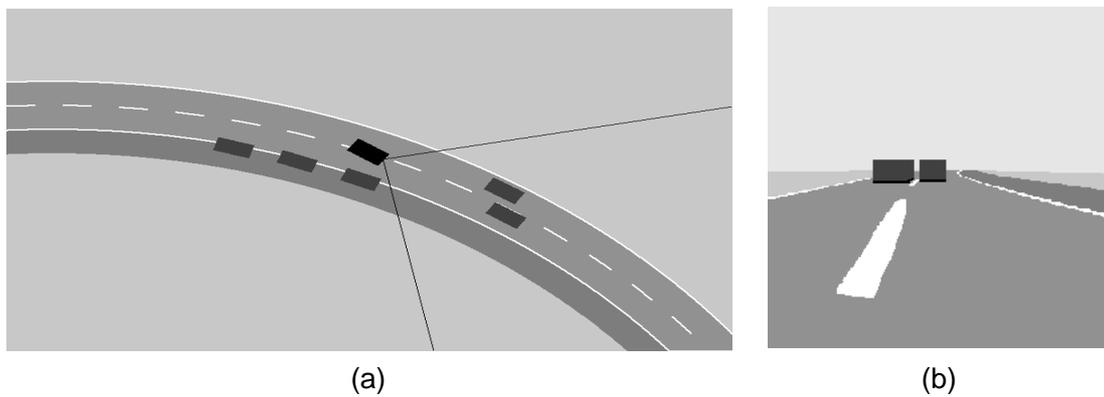


Abbildung 6.20: Folgefahrt bei Gegenverkehr an parkenden Fahrzeugen vorbei. Simuliertes Sensordatenmaterial für eine Straßenverkehrsszene bei einer leichten Rechtskurve mit parkenden Autos am rechten Fahrbahnrand zum Zeitpunkt $t_1 = 300$. Die Fahrtrichtung des Trägerfahrzeugs (schwarz) ist von links nach rechts einem führenden Fahrzeug folgend. Ein weiteres Fahrzeug kommt auf der linken Spur entgegen. (a) Vogelperspektive und (b) Kamerabild für einen Öffnungswinkel von 90° .

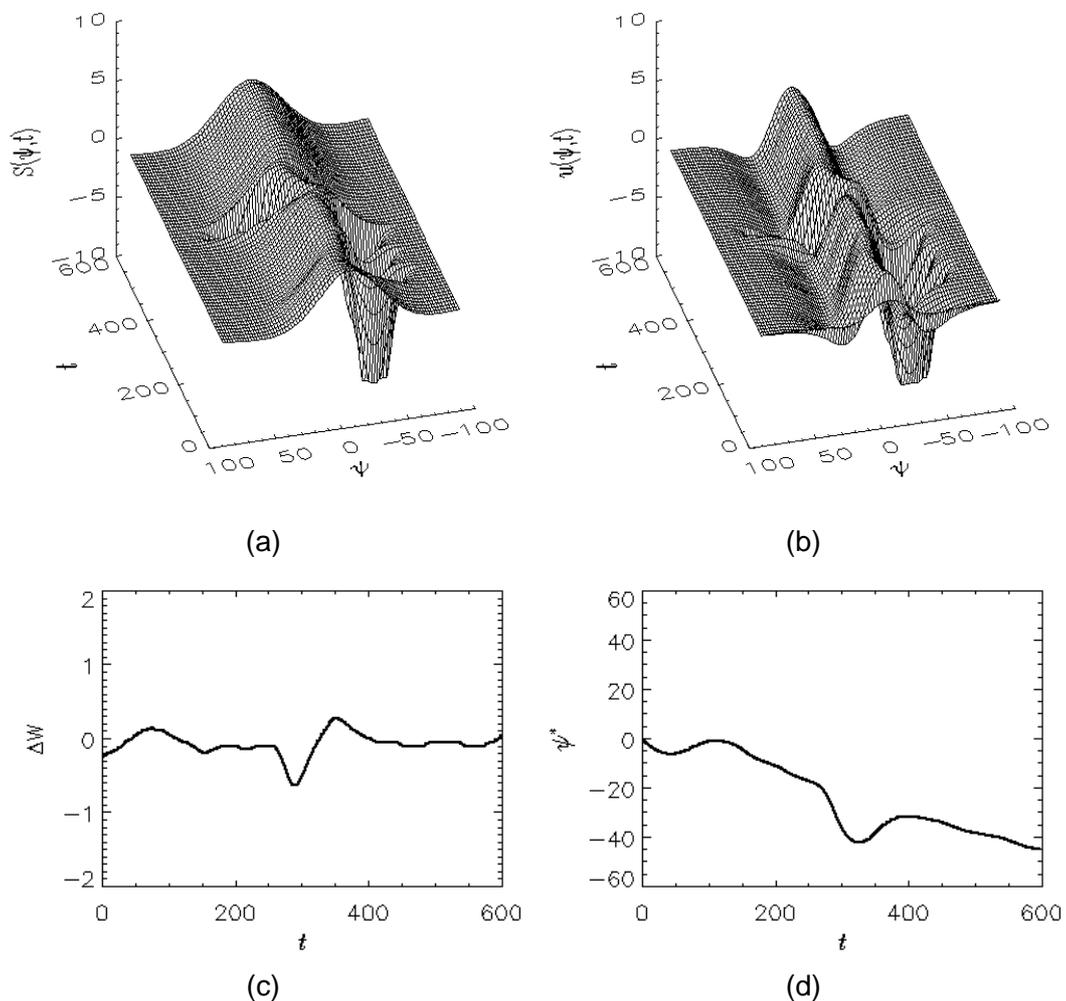


Abbildung 6.21: Folgefahrt bei Gegenverkehr an parkenden Fahrzeugen vorbei. Darstellung (a) des Stimulus, (b) der Aktivität der Feldes, (c) der Lenkwinkeländerung und (d) der Winkelposition des Trägerfahrzeugs zu einem stationären Beobachter über der Zeit.

In Abbildung 6.23 sind die Ergebnisse über der Zeit dargestellt. Das Verhalten des führenden Fahrzeugs bedingt eine Verbreiterung des Stimulus S , der jedoch aufgrund der Felddynamik ausgeglichen wird (Abbildung 6.23a,b). Wegen des inhibitorischen Anteils des winkelabhängigen Fahrspurstimulus S_L (Gleichung 6.7), wird die Ausbildung eines zweiten Maximums im Feld unterdrückt. Zuerst wird das führende Fahrzeug verfolgt, dann aber die Spur gehalten. Der Gegenverkehr wird erkannt.

Straßenschlucht

In diesem Szenario ist die Fahrt des Trägerfahrzeugs auf einer Straße mit parkenden Fahrzeugen am linken und rechten Fahrbahnrand beschrieben. Es wird kein Fahrzeug verfolgt (Abbildung 6.24). Als Basisverhalten werden die Spurhaltung, die Hindernisvermeidung und die Einhaltung von Sicherheitsabständen ohne führendes Fahrzeug untersucht.

In Abbildung 6.25 sind die Ergebnisse über der Zeit dargestellt. Die parkenden Fahrzeuge bedingen den starkenden inhibitorischen Anteil des Stimulus (Abbildung 6.25a,b). Der fehlende Stimulus S_{VF} des führenden Fahrzeugs verursacht einen stärkeren Einfluß der parkenden Fahrzeuge auf den Fahrtverlauf. Die Fahrt des Fahrzeugs wird dennoch auf der Straße, unter Einhaltung des notwendigen Sicherheitsabstands zu den parkenden Fahrzeugen, auf der eigenen Spur durchgeführt (Abbildung 6.25c,d).

Rückwärts ausparkendes Fahrzeug

Das dargestellte Szenario, beschreibt die Fahrt des Trägerfahrzeugs auf einer Straße mit drei parkenden Fahrzeugen (Abbildung 6.24). Das Trägerfahrzeug folgt einem führenden Fahrzeug. Während der Fahrt fährt ein Fahrzeug rückwärts auf die Straße. Ein Fahrzeug kommt auf der anderen Fahrspur entgegen.

In diesem Szenario wird das Zusammenspiel aller Basisverhalten untersucht. Betrachtet werden die Spurhaltung, das Folgefahren, die Hindernisvermeidung, die Einhaltung von Sicherheitsabständen, das Beachten des Gegenverkehrs und das Verhalten des Verfahrens in Gefahrensituationen.

In Abbildung 6.27 sind die Ergebnisse über der Zeit dargestellt. Das rückwärts ausparkende Fahrzeug bedingt einen starken Abfall im Stimulus (Abbildung 6.23a). Aufgrund der hohen Relativgeschwindigkeit ist erhöhte Aufmerksamkeit gefordert. Der Lenkwinkel wird für ein schnelles Ausweichmanöver korrigiert (Abbildung 6.23c). Nachdem das ausparkende Fahrzeug zum Stillstand gekommen ist, reduziert sich der inhibitorische Anteil im Stimulus. Ein normales Ausweichmanöver wird eingeleitet und durchgeführt. Ein entgegenkommendes Fahrzeug bedingt die starke Lenkwinkeländerung zurück auf die eigene Fahrspur.

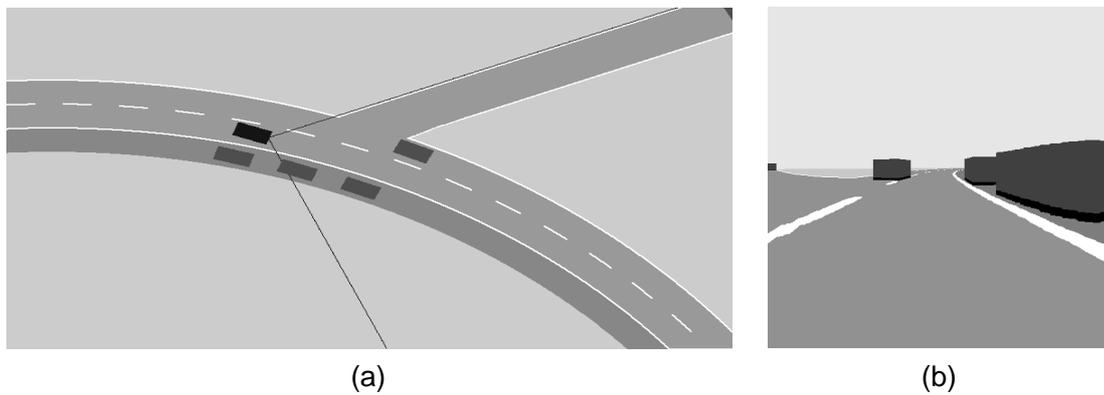


Abbildung 6.22: Parkende Fahrzeuge und Gegenverkehr bei links abbiegendem führenden Fahrzeug. Simuliertes Sensordatenmaterial für eine Straßenverkehrsszene bei einer leichten Rechtskurve mit parkenden Autos am rechten Fahrbahnrand bei Gegenverkehr zum Zeitpunkt $t_1 = 500$. Die Fahrtrichtung des Trägerfahrzeugs (schwarz) ist von links nach rechts. (a) Vogelperspektive und (b) Kamerabild für einen Öffnungswinkel von 90° .

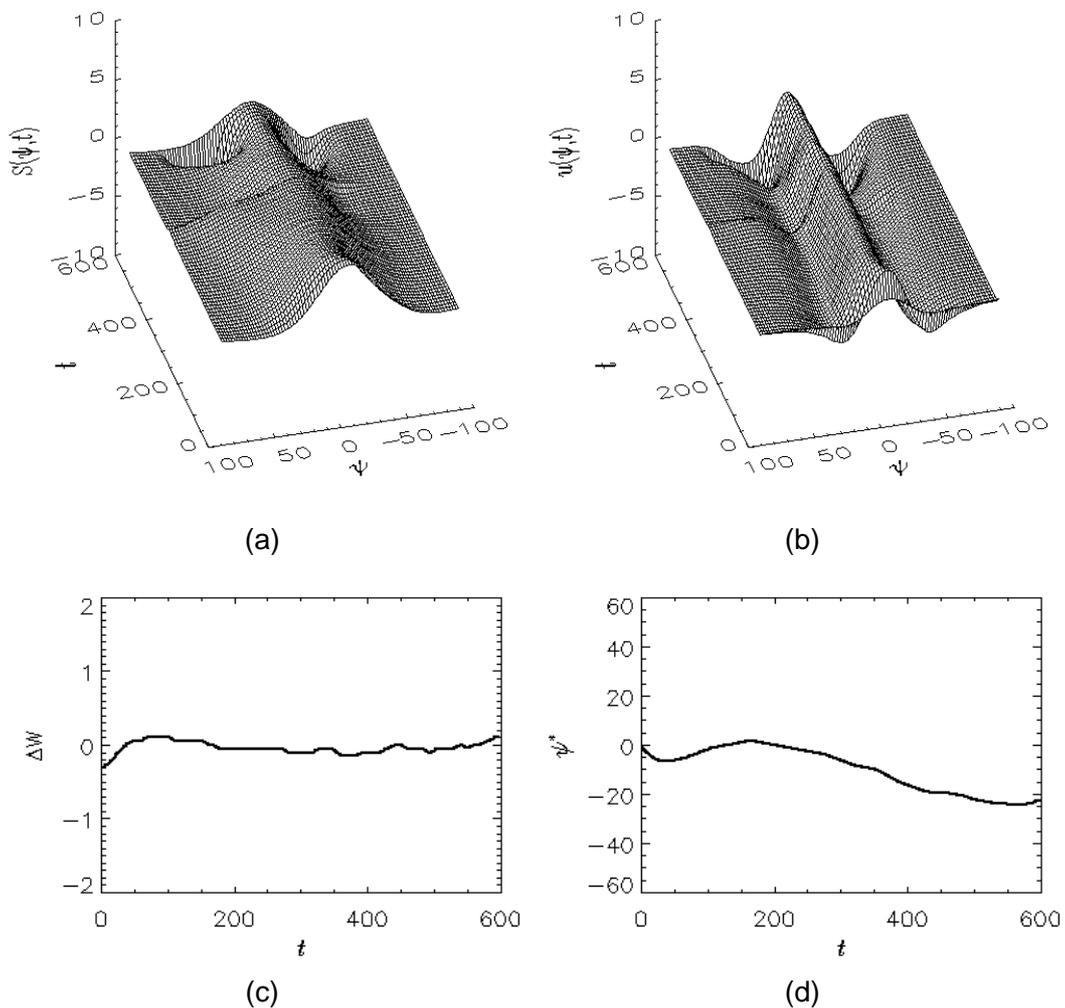


Abbildung 6.23: Parkende Fahrzeuge und Gegenverkehr bei links abbiegendem führenden Fahrzeug. Darstellung (a) des Stimulus, (b) der Aktivität der Feldes, (c) der Lenkwinkeländerung und (d) der Winkelposition des Trägerfahrzeugs zu einem stationären Beobachter über der Zeit.

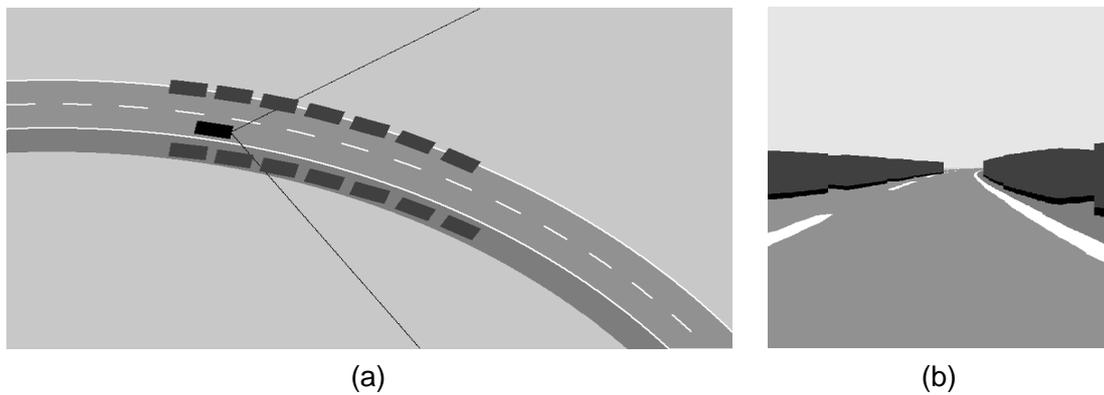


Abbildung 6.24: Straßenschlucht. Simuliertes Sensordatenmaterial für eine Straßenverkehrsszene bei einer leichten Rechtskurve mit parkenden Autos an beiden Fahrbahnrandern zum Zeitpunkt $t_1 = 180$. Die Fahrtrichtung des Trägerfahrzeugs (schwarz) ist von links nach rechts. (a) Vogelperspektive und (b) Kamerabild für einen Öffnungswinkel von 90° .

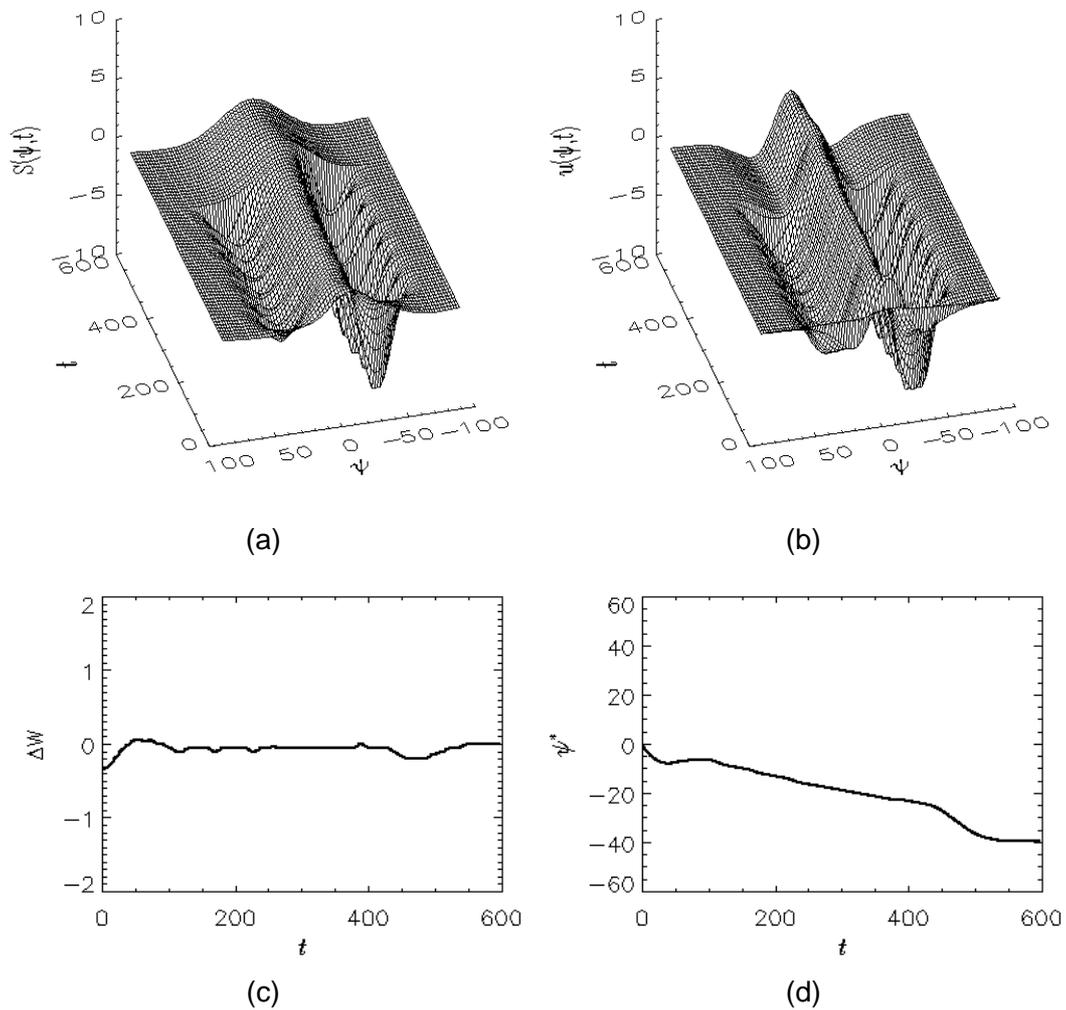


Abbildung 6.25: Straßenschlucht. Darstellung (a) des Stimulus, (b) der Aktivität der Feldes, (c) der Lenkwinkeländerung und (d) der Winkelposition des Trägerfahrzeugs zu einem stationären Beobachter über der Zeit.

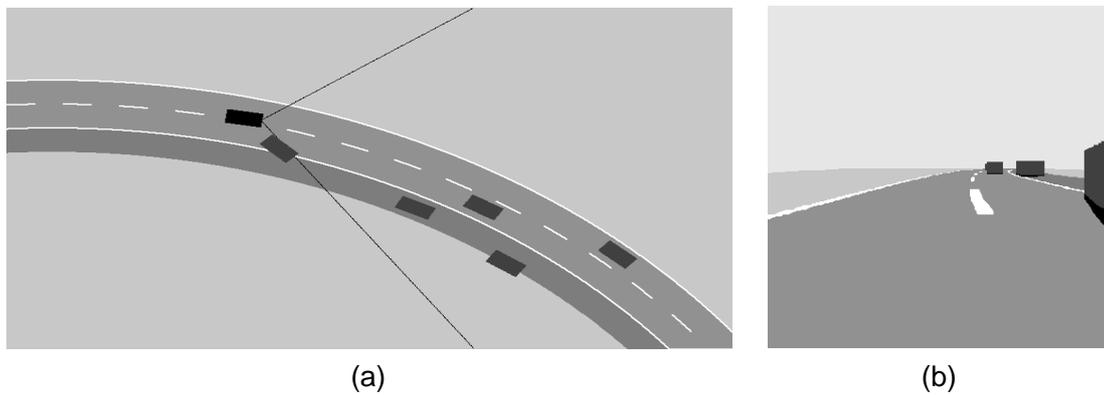


Abbildung 6.26: Rückwärts ausparkendes Fahrzeug. Simuliertes Sensordatenmaterial für eine Straßenverkehrsszene bei einer leichten Rechtskurve mit parkenden Autos am rechten Fahrbahnrand bei Gegenverkehr und ausparkendem Fahrzeug zum Zeitpunkt $t_1 = 175$. Die Fahrtrichtung des Trägerfahrzeugs (schwarz) ist von links nach rechts. (a) Vogelperspektive und (b) Kamerabild für einen Öffnungswinkel von 90° .

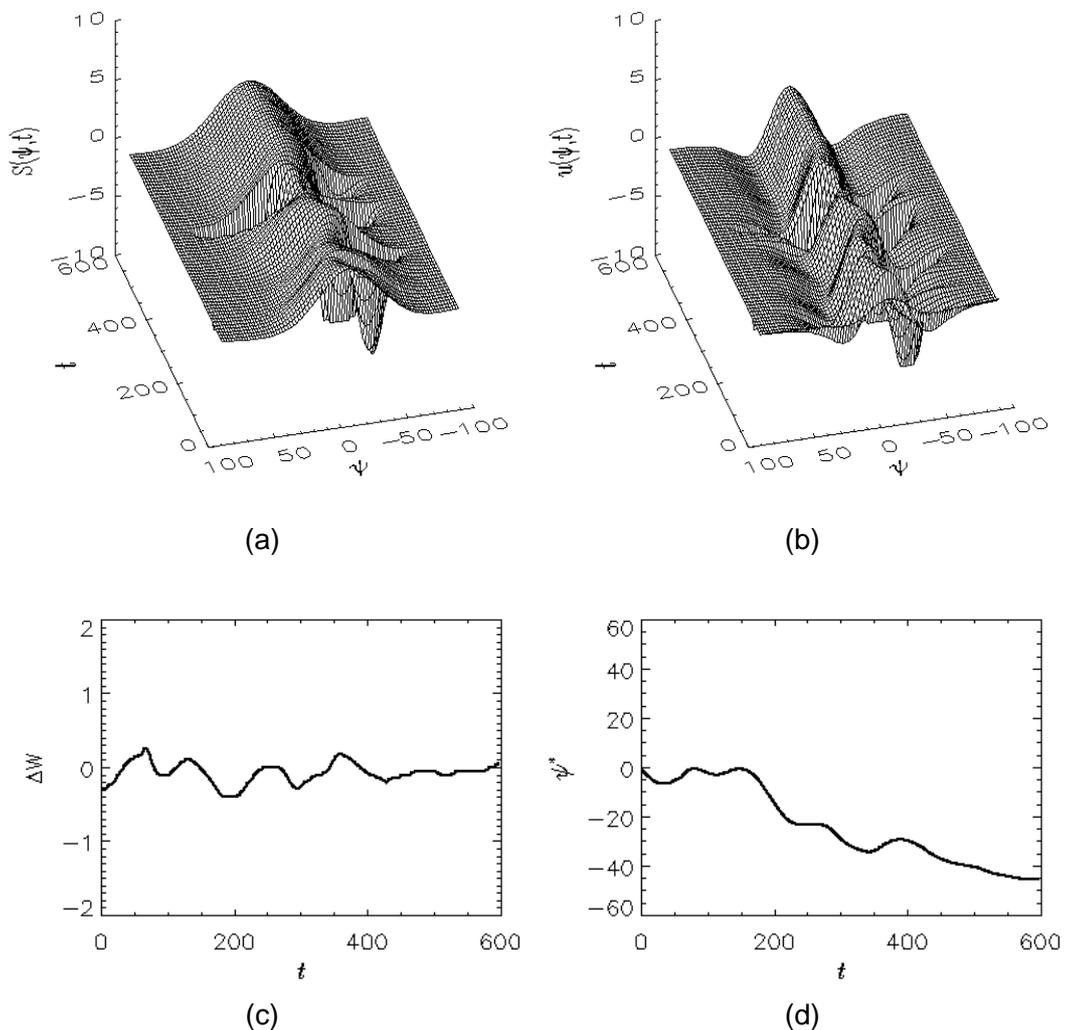


Abbildung 6.27: Rückwärts ausparkendes Fahrzeug. Darstellung (a) der Höhe des Stimulus, (b) der Aktivität der Feldes, (c) der Lenkwinkeländerung und (d) der Winkelposition des Trägerfahrzeugs zu einem stationären Beobachter über der Zeit.

6.4.2 Diskussion

Die Tragfähigkeit des realisierten Verfahrens zur Bewegungssteuerung eines Kraftfahrzeugs wird anhand von relevanten Basisverhalten belegt. Dabei ist insbesondere das Verhalten im innerstädtischen Bereich von Interesse.

Bei der Geschwindigkeitssteuerung wird die Anpassung dieser bei auftretenden Hindernissen in der Fahrspur beschrieben (Anschnitt 6.3.3). Es wird eine Anpassung anhand des Zustands des neuronalen Feldes bei einem langsam vorausfahrenden Fahrzeug durchgeführt.

Die Steuerung des Lenkwinkels wird besonders untersucht. Im innerstädtischen Bereich hat diese einen wesentlichen Anteil bei der Navigation eines Fahrzeugs. Die Steuerung des Lenkwinkels muß dabei hohen Anforderungen genügen. Insbesondere müssen verschiedene Basisverhalten Einfluß auf das Lenkverhalten haben.

Die Beispiele werden gezielt ausgewählt, um das Verhalten des entwickelten Verfahrens an unterschiedlichen Basisverhalten zu testen. Dabei wird die Spurhaltung, das Folgefahren, die Hindernisvermeidung, die Einhaltung von Sicherheitsabständen, das Beachten des Gegenverkehrs und das Wahrnehmen von Gefahrensituationen bei der Steuerung des Fahrzeugs berücksichtigt. Bei den Szenarien wird weiterhin der Ausfall von Stimuli und das widersprüchliche Verhalten von Stimuli untersucht. Das richtige Verhalten des Trägerfahrzeugs wird belegt.

Es kann anhand der beschriebenen Beispiele sinnvolles Verhalten der entwickelten Bewegungssteuerung abgeleitet werden. Die gewählten Szenarien stellen zwar nur eine Untermenge möglicher Szenarien im innerstädtischen Bereich dar, sie beschreiben aber die wesentlichen Basisverhalten für eine Navigation im innerstädtischen Bereich.

6.5 Zusammenfassung

In diesem Kapitel wird ein neues Verfahren vorgestellt, das eine Bewegungssteuerung für ein Kraftfahrzeug im Straßenverkehr umsetzt. Der Realisierung liegt ein nichtlinearer Ansatz zweier eindimensionaler neuronaler Felder mit Integro-Differentialgleichungen für die Steuerung des Lenkwinkels und der Geschwindigkeit zugrunde. Die Eigenschaften der Felder werden diskutiert und eine Stabilitätsuntersuchung durchgeführt. Das von einer simulierten Objekterkennung gelieferte Sensordatenmaterial wird zur Steuerung der Feldaktivitäten genutzt. Eine Vorverarbeitung berücksichtigt als Randbedingung Sicherheit und komfortablen Fahrtverlauf. Die resultierenden Stimuli für Objekte, Fahrspur und führendes Fahrzeug werden in die eindimensionalen Felder eingekoppelt und ein Vorschlag für eine Lenkwinkeländerung und eine Geschwindigkeitsänderung über den Zustand der neuronalen Felder abgeleitet. Die Tragfähigkeit des Verfahrens wird an typischen Basisverhalten bei geschlossenem Regelkreis gezeigt.

In Kapitel 7 wird das in diesem Kapitel vorgestellte Verfahren im Zusammenspiel mit realen Daten getestet. Aufbauend auf die in Kapitel 5 beschriebene Fusion zur Segmentierung, werden Objektattribute erzeugt und das Verhalten der Bewegungssteuerung analysiert.

Kapitel 7

System zur Fahrerassistenz

In dem vorliegenden Kapitel werden die entwickelten Verfahren der Segmentierung und der Lenkwinkelsteuerung in ein Gesamtsystem integriert und die Tragfähigkeit an Videobilddaten gezeigt. Die Grundlage dieses Systems bildet eine in [42] vorgestellte Architektur für Fahrerassistenzsysteme.

In Abschnitt 7.1 werden verschiedene Architekturansätze für Fahrerassistenzsysteme diskutiert und die verwendete Architektur motiviert. In Abschnitt 7.2 wird das System beschrieben und die Tragfähigkeit demonstriert. Dabei wird das in Kapitel 5 vorgestellte Verfahren zur Segmentierung und das in Kapitel 6 vorgestellte Verfahren der Bewegungssteuerung genutzt. Eine Diskussion der Ergebnisse wird in Abschnitt 7.3 durchgeführt. Das Kapitel 7 schließt mit einer Zusammenfassung in Abschnitt 7.4.

7.1 Architektur

Bei der Realisierung von Fahrerassistenzsystemen ist die Wahl einer geeigneten Architektur wesentlich, um verschiedene Aufgabenstellungen, wie das Folgefahren oder das Erkennen von Hindernissen, entsprechend der Vorgaben zu lösen. Das bedeutet, daß das realisierte System eine flexible Architektur für implementierte Verfahren bereitstellen muß. Ansätze zur Realisierung einer Architektur, die systematisch die Problematik von Fahrerassistenzsystemen angehen sind beispielsweise [19, 121] und [42].

Eine vierdimensionale Repräsentation mit den drei Weltkoordinaten und einer Zeitkoordinate ist die Basis in [19], um eine Analyse der Umwelt durchzuführen. Ein Kontroll- und Datenflußplan für ein Bildverarbeitungssystem in autonomen Fahrzeugen, welcher in einer Hierarchie mit drei Schichten aufgebaut ist, wird in [121] vorgestellt. In den drei Schichten werden Merkmale, Objekte und Situationen kodiert. Die Kommunikation zwischen den Schichten ist in Form eines Entscheidungsbaums implementiert.

Ein Ansatz, welcher eine Architektur für ein modulares Fahrerassistenzsystem realisiert, wird in [42] vorgestellt. Neue Anforderungen und Informationen können durch den modularen Aufbau einfach hinzugefügt werden. Die hierarchische Anordnung der einzelnen Module erlaubt eine Verarbeitung von Information auf verschiedenen Abstraktionsniveaus, wie die Verarbeitung von Sensordaten und das Lösen einer an das System übergebenen abstrakten Aufgabenstellung. Das im Rahmen dieser Arbeit entwickelte System nutzt diesen Architekturansatz. Im folgenden wird der Architekturansatz diskutiert.

Die Architektur (Abbildung 7.1) unterstützt verschiedene Verhalten, die zur Erfüllung einer abstrakten Aufgabe dienen. Die zugrundeliegenden Module *Wissensdatenbank*, *objektbezogene Analyse*, *Szeneninterpretation* und *Verhaltenssteuerung* bilden die Basis für diese Architektur.

In der Wissensdatenbank sind statisches und dynamisches Wissen repräsentiert. Sie beinhaltet die Beschreibung physikalischer Zusammenhänge (statisch), Verhaltensregeln für den Straßenverkehr (statisch, dynamisch), sowie Objekt- und Szenenwissen (dynamisch). Externes Wissen, wie zum Beispiel GPS, kann in die Wissensdatenbank eingebunden werden.

Im Modul der objektbezogenen Analyse werden die Sensordaten verarbeitet. Diese Verarbeitung wird in einem sensorbezogenen Koordinatensystem durchgeführt. Objekthypothesen werden generiert. In einer internen Repräsentation werden die verarbeiteten Sensordaten über der Zeit stabilisiert. Die Auswertung kann sowohl für jeden Sensor einzeln, als auch für eine Gruppe von Sensoren über Fusion erfolgen. Eine Realisierung einer objektbezogenen Analyse auf Basis von visuellen Sensordaten wird beispielsweise in [39] und [40] vorgestellt. Objekthypothesen werden in der sensorbezogenen Verarbeitung durch eine initiale Segmentierung, eine Klassifikation und eine Verfolgung von ROIs¹ erzeugt und attribuiert.

Im Modul der Szeneninterpretation werden die Hypothesen der objektbezogenen Analyse interpretiert und auf einen konsistenten Stand gebracht. Verhaltensrelevante Informationen, wie beispielsweise die Distanz oder die Geschwindigkeit von Objekten, werden extrahiert. Ansätze bieten hierzu [41].

Das Modul der Verhaltenssteuerung erlaubt den Abgleich von Vorgaben, die in Form von Aufgaben an das System übergeben werden, mit dem internen Systemzustand, um eine Steuerung der Verhaltensvariablen durchzuführen.

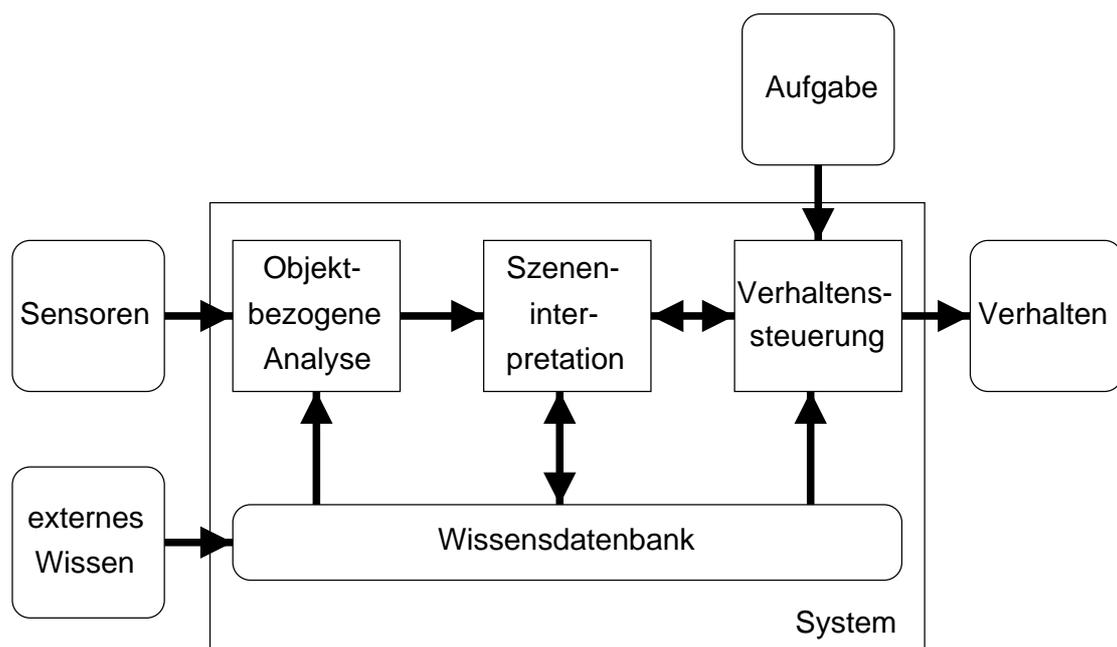


Abbildung 7.1: Modulare Architektur für Fahrerassistenzsysteme: *objektbezogene Analyse*, *Wissensdatenbank*, *Szeneninterpretation* und *Verhaltenssteuerung*.

¹ROI – region of interest

Die vorgestellte Architektur läßt eine Strukturierung des Aufbaus von Fahrerassistenzsystemen zu. Es ist möglich, einzelne Problemstellungen zu kapseln und modular zu lösen. Insbesondere ist das Hinzufügen neuer Verfahren durch die Modularität einfach realisierbar.

Die in dieser Arbeit realisierten Verfahren bieten neue Beiträge zur objektbezogenen Analyse und der Verhaltenssteuerung. Die Simulationsumgebung bildet eine Repräsentation, die die Basis für eine Szeneninterpretation ist. Diese ist jedoch nicht dynamisch ausgelegt, so daß die Planung von Verhalten nur eingeschränkt möglich ist.

In Abschnitt 7.2 werden die realisierten Verfahren integriert und die Tragfähigkeit des Gesamtsystems demonstriert.

7.2 Verfahrenintegration

In diesem Abschnitt wird die Integration der realisierten Verfahren in ein Gesamtsystem beschrieben, das der in Abschnitt 7.1 vorgestellten Architektur folgt. Nach der Integration kann das Verhalten der entwickelten Verfahren im Zusammenspiel mit anderen Verfahren, wie Objektverfolgern [84, 74, 127] oder Klassifikatoren [34, 119] getestet werden. An zwei Beispielen, die reale Videoaufnahmen von Innenstadtszenen beinhalten, werden die Ergebnisse des Gesamtsystems diskutiert. In den Abbildungen 7.2- 7.4 ist der Aufbau des Gesamtsystems verdeutlicht. Grau unterlegt sind die Verfahren, die im Rahmen dieser Arbeit realisiert wurden. Die schraffierten Module wurden vom Autor erstellt und in [43, 37] vorgestellt. Abbildung 7.2 zeigt ein Blockschaltbild der objektbezogenen Analyse. In Abbildung 7.3 ist das Blockschaltbild des Moduls der Szeneninterpretation abgebildet und Abbildung 7.4 beschreibt das Modul der Verhaltenssteuerung. Die Module bauen gemäß Abschnitt 7.1 aufeinander auf. Die Übergänge sind mit Nummern markiert.

Der Aufbau der objektbezogenen Analyse (Abbildung 7.2) folgt dabei dem im folgenden beschriebenen Schema. Die Videobilddaten werden vorverarbeitet („Seg.-Verfahren“) und analysiert („Komp. Analyse“, Kapitel 4). Relevante Komponenten werden fusioniert und eine Segmentierung durchgeführt („Fusion“, Kapitel 5). Ein einfaches Verfahren [37] wertet in den segmentierten Gebieten horizontale Linien aus („ROIs“). Die Linienbilder werden mit Hilfe zweier in [23, 128] vorgestellter Ansätze erzeugt. Die resultierenden ROIs werden bewegungssensitiv stabilisiert [43] („Bew.sens. Rep.“) und mit einem neuronalen Merkmalklassifikator [34] („Klass.“) bewertet. Regionen hoher Güte werden dem Objektverfolger [84] („Verfolger“) übergeben, der diese über der Zeit verfolgt. Die verfolgten ROIs werden zur Stabilisierung in die sensorbezogene Repräsentation zurückgekoppelt.

Der Aufbau einer sensorbezogenen Repräsentation basiert auf der Annahme, daß Objekte einen stetigen Bewegungsverlauf haben. Das bedeutet, daß Objekte, welche für die Aufgabe relevant sind, bei der genutzten Bildwiederholffrequenz f für $f > 25Hz$ in aufeinanderfolgenden Bildern nur einen geringen Verschiebungsvektor aufweisen. Daher wird die verwendete bewegungssensitive Repräsentation als pixelbasiertes Akkumulatorfeld in Form bewegungssensitiver Karten nach [43] organisiert. Aus der Wissensbasis werden physikalische Regeln und gespeicherte Objektattribute abgerufen.

In Abbildung 7.3 ist das Modul der Szeneninterpretation dargestellt. Die verfolgten Objekte werden in die Vogelperspektive („Vogelperspekt.“) transformiert [42] und in der Simulation (Anhang B) verarbeitet. Der Fahrspurverlauf („Fahrspur“) wird nach [88] bestimmt und ebenfalls an

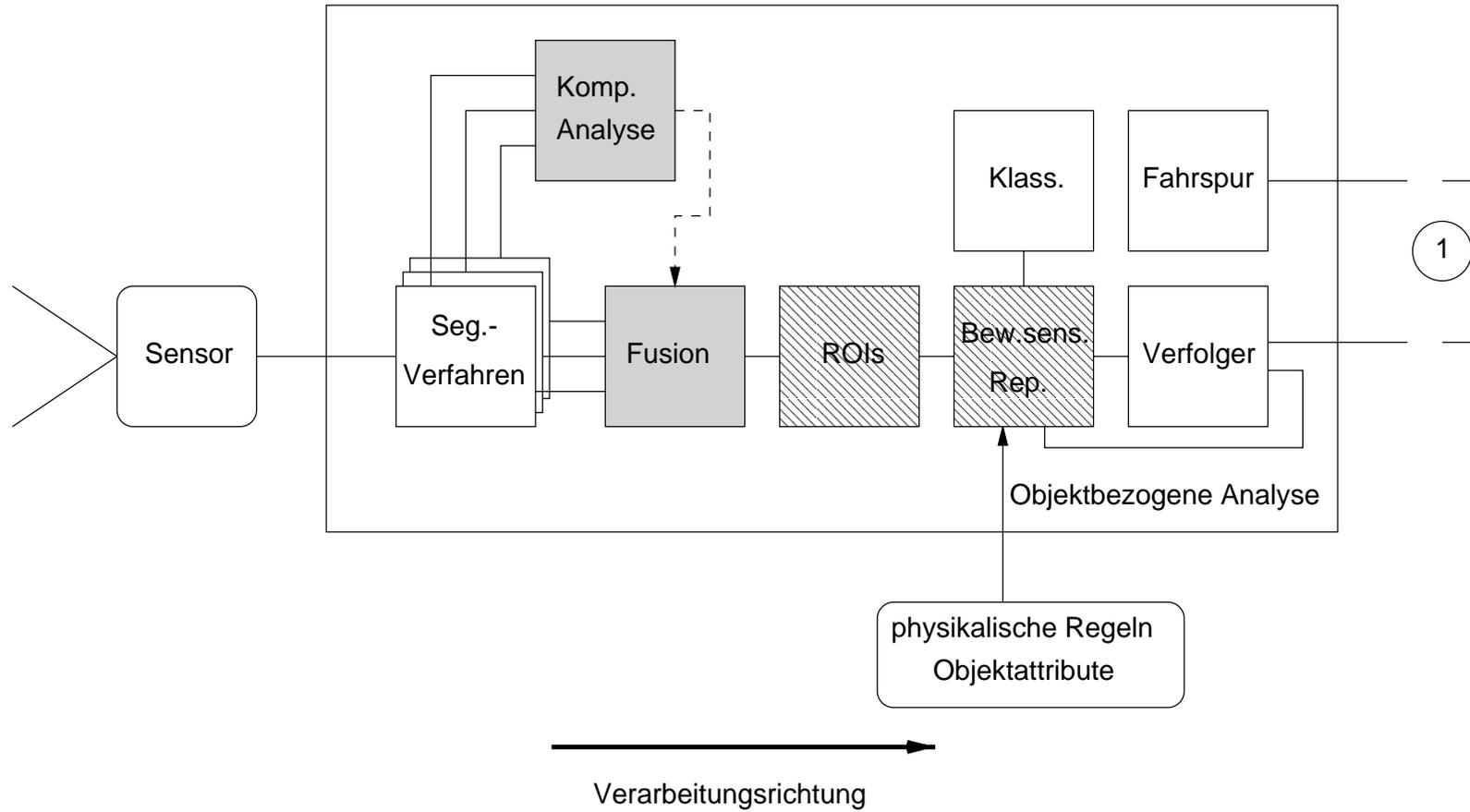


Abbildung 7.2: Objektbezogene Analyse. Verarbeitung der Sensordaten und Generierung von attribuierten Objekthypothesen (ROIs).

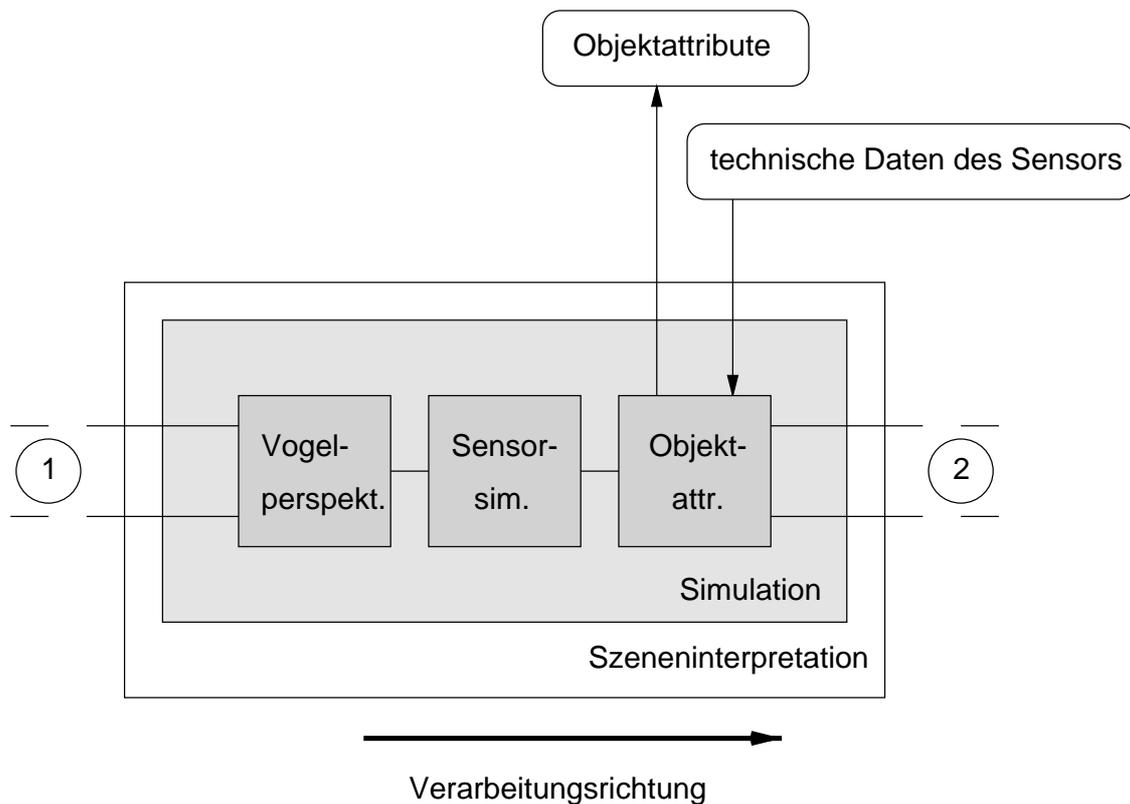


Abbildung 7.3: Szeneninterpretation. Aufarbeitung der attribuierten Objektlisten (ROIs) aus der objektbezogenen Analyse und Erzeugung relevanter Attribute für das Modul der Verhaltenssteuerung.

die Simulation übergeben. Die technischen Daten der Videokamera werden bei der Umsetzung der simulierten Sensoren („Sensorsim.“) berücksichtigt (Sensorchipfläche: $6\text{mm} \cdot 5\text{mm}$, Höhe: 1.3m , Öffnungswinkel von 30° und Auslenkung der Kamera aus der Fahrbahnparallelen Ebene: 3.5° in Richtung Fahrbahn) und relevante Objektattribute („Objektattr.“) an die Wissensdatenbank übergeben.

Die Simulation liefert die notwendigen Eingangsdaten der Verhaltenssteuerung. Abbildung 7.4 zeigt die genutzten Eingangsdaten der neuronalen Lenkwinkelsteuerung („Lenkw.-steuerung“) und der neuronalen Geschwindigkeitssteuerung („Geschw.-steuerung“), welche in Kapitel 6 beschrieben sind. Es werden neben den aufbereiteten Objektattributen aus der Szeneninterpretation die Fahrhinweise bzw. die Sollgeschwindigkeit des Fahrzeugs berücksichtigt und physikalische Regeln und Sicherheitsanforderungen beachtet.

Im folgenden wird die Umsetzung des Systems an zwei Beispielen diskutiert und die Tragfähigkeit der Verfahren gezeigt. Die Simulation dient dabei der Verifikation der berechneten Objektattribute.

Fahrt entlang einer geraden Strecke

In dem ersten Beispiel ist eine Fahrt des Trägerfahrzeugs auf einer geraden Strecke dargestellt. In Abbildung 7.5a ist das Segmentierungsergebnis zu erkennen. Abbildung 7.5b beschreibt die

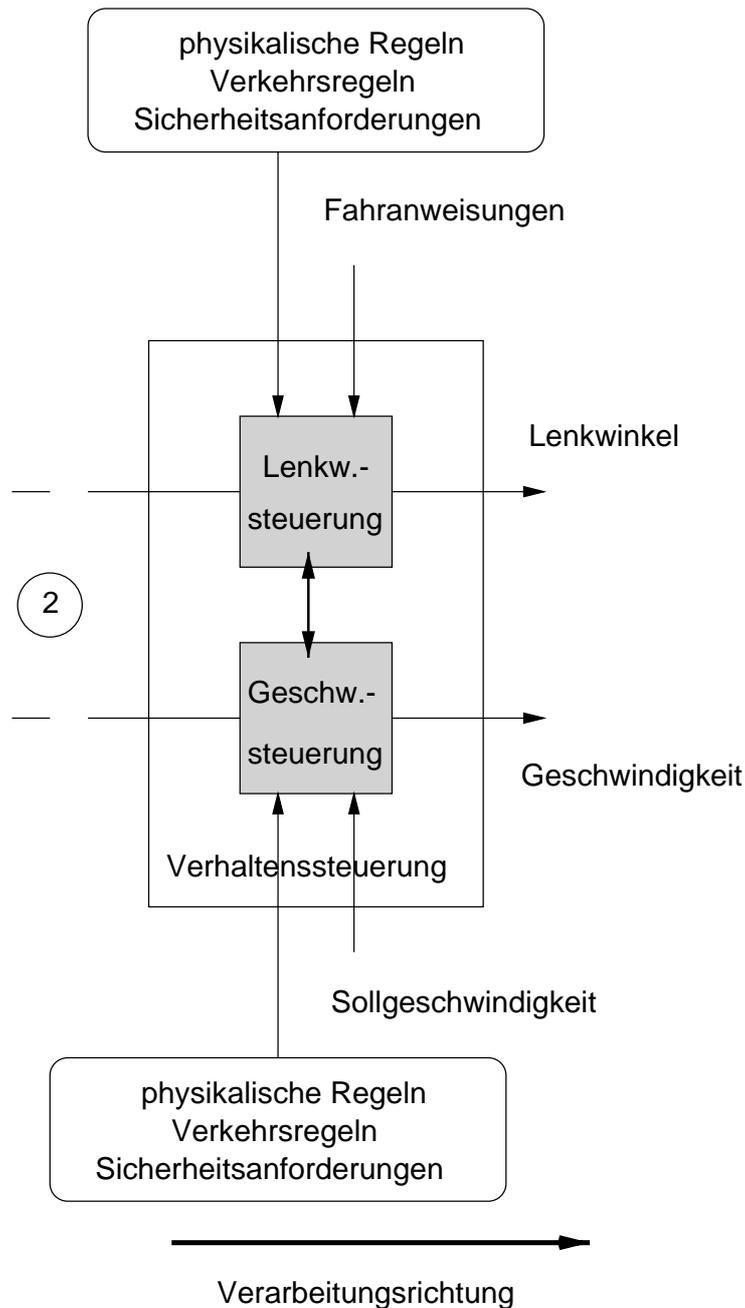


Abbildung 7.4: Verhaltenssteuerung. Steuerung des Lenkwinkels und der Geschwindigkeit unter Berücksichtigung von aufbereiteten Objektattributen aus der Szeneninterpretation, der Fahrerweisungen, der Sollgeschwindigkeit des Fahrzeugs, von physikalischen Regeln und den Sicherheitsanforderungen.

extrahierten Objekthypothesen. Es wird der vorherfahrende LKW erkannt und verfolgt (Abbildung 7.5c).

Die bewegungssensitive Repräsentation (Abbildung 7.6a-c, Aktivitätsverlauf von hell nach dunkel aufsteigend) erkennt ein vorausfahrendes Objekt (Abbildung 7.6b). Zwei weitere Objekte werden am linken Fahrbahnrand erkannt und verfolgt. Eine dritte Objekthypothese wird ver-

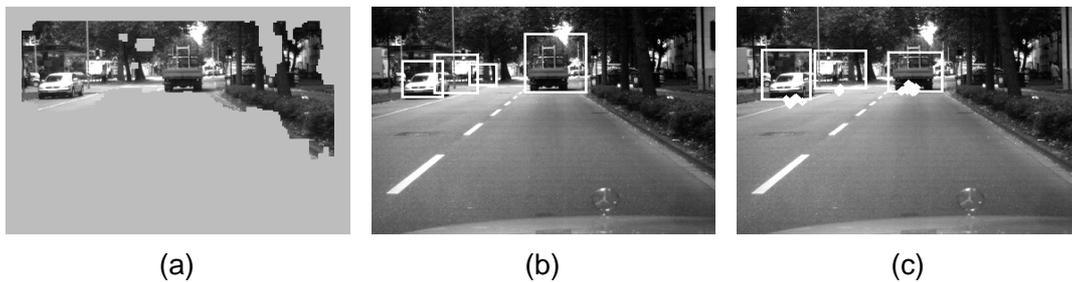


Abbildung 7.5: Fahrt entlang einer geraden Strecke, Objekthypothesenerzeugung. (a) Neuronale Segmentierung, (b) Objekthypothesen und (c) verfolgte Objekte über der Zeit (eingezeichnet sind umschreibende Rechtecke und der Verlauf der Verfolgung über der Zeit).

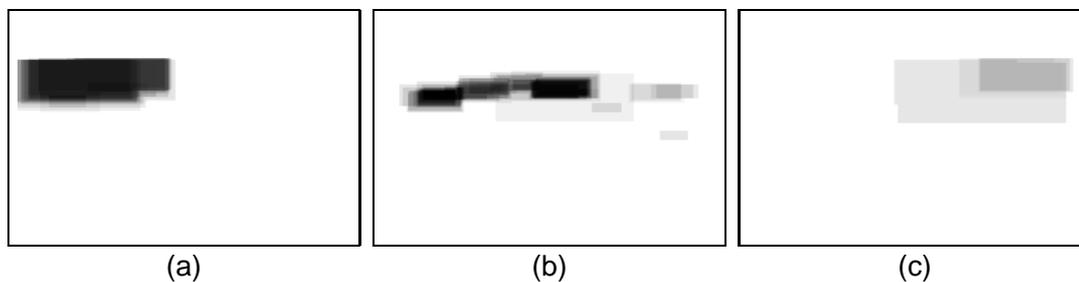


Abbildung 7.6: Fahrt entlang einer geraden Strecke, bewegungssensitive Karten. (a) Auswertung der Bewegungsrichtung für entgegenkommenden Verkehr auf der linken Seite, (b) Auswertung sich konstant bewegender Objekte vor dem Trägerfahrzeug und (c) Auswertung der Bewegungsrichtung für entgegenkommenden Verkehr auf der rechten Seite.

worfen. Die bewegungssensitive Karte weist dabei eine starke Aktivität für Objekte auf der linken Seite auf (Abbildung 7.6a). Am rechten Fahrbahnrand kommen keine Objekte entgegen (Abbildung 7.6c).

In Abbildung 7.7a ist die Vogelperspektive der Simulation dargestellt. Das Trägerfahrzeug befindet sich dabei konstant in der Mitte. Ein externer Beobachter ist nicht vorhanden. Der Winkelbereich der Kamera sowie der geschätzte Spurverlauf sind eingezeichnet. Abbildung 7.7b zeigt das simulierte Sensorbild. Ein Vergleich mit Abbildung 7.5 bestätigt die Richtigkeit der berechneten Größen. Mit Hilfe dieser Darstellung werden die Eingangsdaten für die neuronale Bewegungssteuerung erzeugt.

Abbildung 7.8b beschreibt das Verhalten des neuronalen Feldes für die Lenkwinkeländerung über der Zeit. Der Lenkwinkel wurde nahezu konstant gehalten (Abbildung 7.8c). Auftretende falsche Objekthypothesen, dargestellt durch negative Anteile im Stimulus (Abbildung 7.8a) verursachen dabei Korrekturen in der Position der α -Lösung des Feldes. Die durchgeführte Fahrt entlang einer geraden Straße wird demnach richtig interpretiert.

Bei kurzzeitigen Fehlhypothesen zeigt das neuronale Feld ein zeitlich stabilisierendes Verhalten. Treten durch den Bildverarbeitungsteil der objektbezogenen Analyse länger anhaltende Fehler auf, führt eine Fehlhypothese zur Veränderung des Lenkwinkeleinschlags. Einen Ausweg bietet der Einsatz verschiedener Sensoren mit unabhängigen physikalischen Eigenschaften, um die

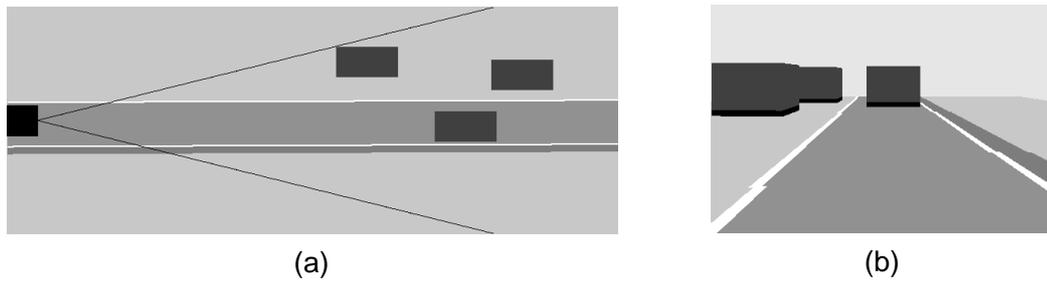


Abbildung 7.7: Fahrt entlang einer geraden Strecke, simuliertes Sensordatenmaterial. Die Fahrtrichtung des Trägerfahrzeugs (schwarz) ist von links nach rechts. (a) Vogelperspektive und (b) Kamerabild.

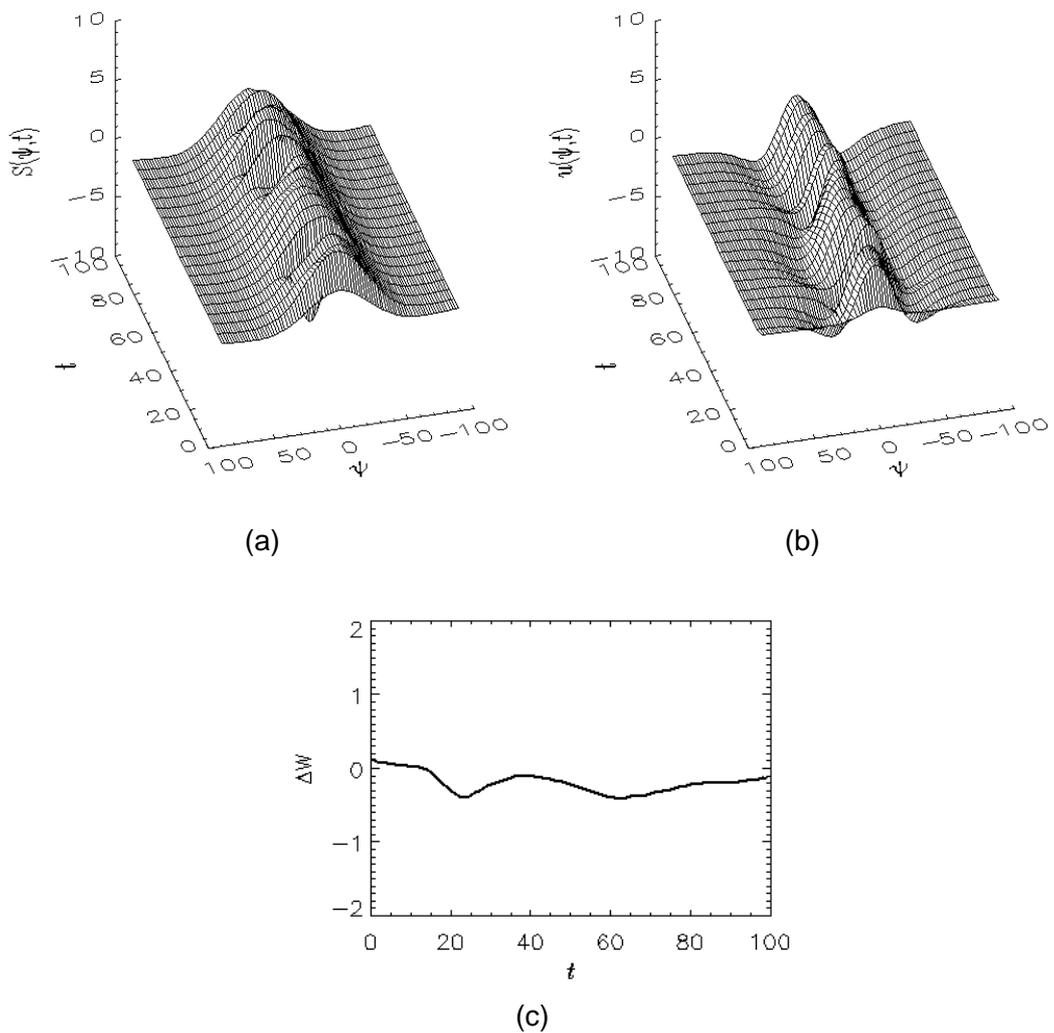


Abbildung 7.8: Fahrt entlang einer geraden Strecke. Darstellung (a) des Stimulus, (b) der Aktivität der Feldes, (c) der Lenkwinkeländerung über der Zeit.

Störanfälligkeit bei der Objekterkennung zu reduzieren. Ein Lösungsvorschlag für eine Sensordatenfusion verschiedener physikalischer Sensoren wird in Kapitel 5 am Beispiel einer Fusion von RADAR-Daten und Videobilddaten gegeben.

Fahrt entlang einer Baustelle durch eine Linkskurve

Im zweiten Beispiel wird die Fahrt des Trägerfahrzeugs entlang einer Baustelle durch eine Linkskurve diskutiert. In Abbildung 7.9 sind die Ergebnisse der objektbezogenen Analyse dargestellt. Es werden mehrere Objekthypothesen generiert und klassifiziert. Zwei Fahrzeuge werden erkannt (Abbildung 7.9b).

Die zeitliche Stabilisierung in der sensorbezogenen Repräsentation liefert eine ausgeprägte Objekthypothese für das vorausfahrende Fahrzeug (Abbildung 7.10b). Dieses wird verfolgt (Abbildung 7.9c). Eine erhöhte Aktivität wird auch im Bereich der Baustelle erzeugt (Abbildung 7.10a). Am rechten Fahrbahnrand werden Teile der Berandungsmauer als Hindernisse erkannt und in die Repräsentation eingetragen (Abbildung 7.10c).

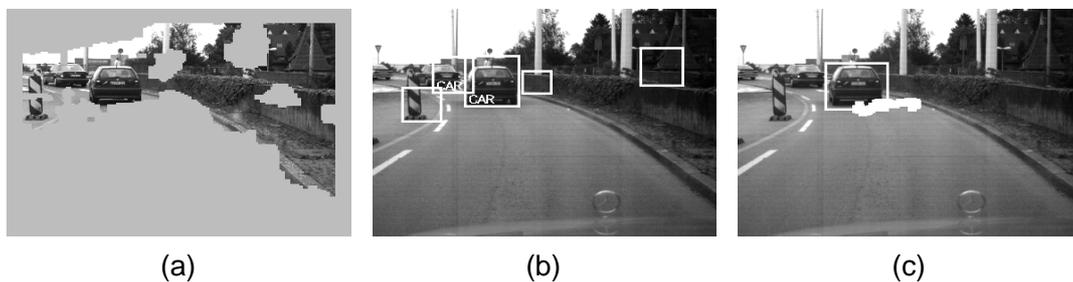


Abbildung 7.9: Fahrt entlang einer Baustelle durch eine Linkskurve, Objekthypothesenerzeugung. (a) Neuronale Segmentierung, (b) Objekthypothesen mit Klassenzugehörigkeit und (c) verfolgte Objekte über der Zeit (eingezeichnet sind umschreibende Rechtecke und der Verlauf der Verfolgung über der Zeit).

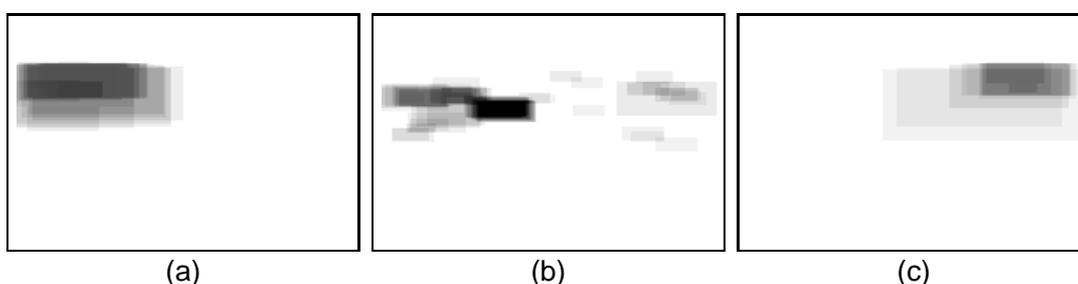


Abbildung 7.10: Fahrt entlang einer Baustelle durch eine Linkskurve, bewegungssensitive Karten. (a) Auswertung der Bewegungsrichtung für entgegenkommenden Verkehr auf der linken Seite, (b) Auswertung sich konstant bewegnender Objekte vor dem Trägerfahrzeug und (c) Auswertung der Bewegungsrichtung für entgegenkommenden Verkehr auf der rechten Seite.

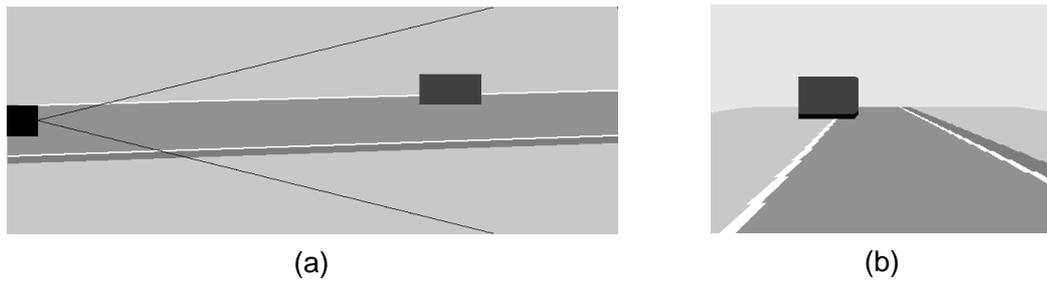


Abbildung 7.11: Fahrt entlang einer Baustelle durch eine Linkskurve, simuliertes Sensordatenmaterial. Die Fahrtrichtung des Trägerfahrzeugs (schwarz) ist von links nach rechts. (a) Vogelperspektive und (b) Kamerabild.

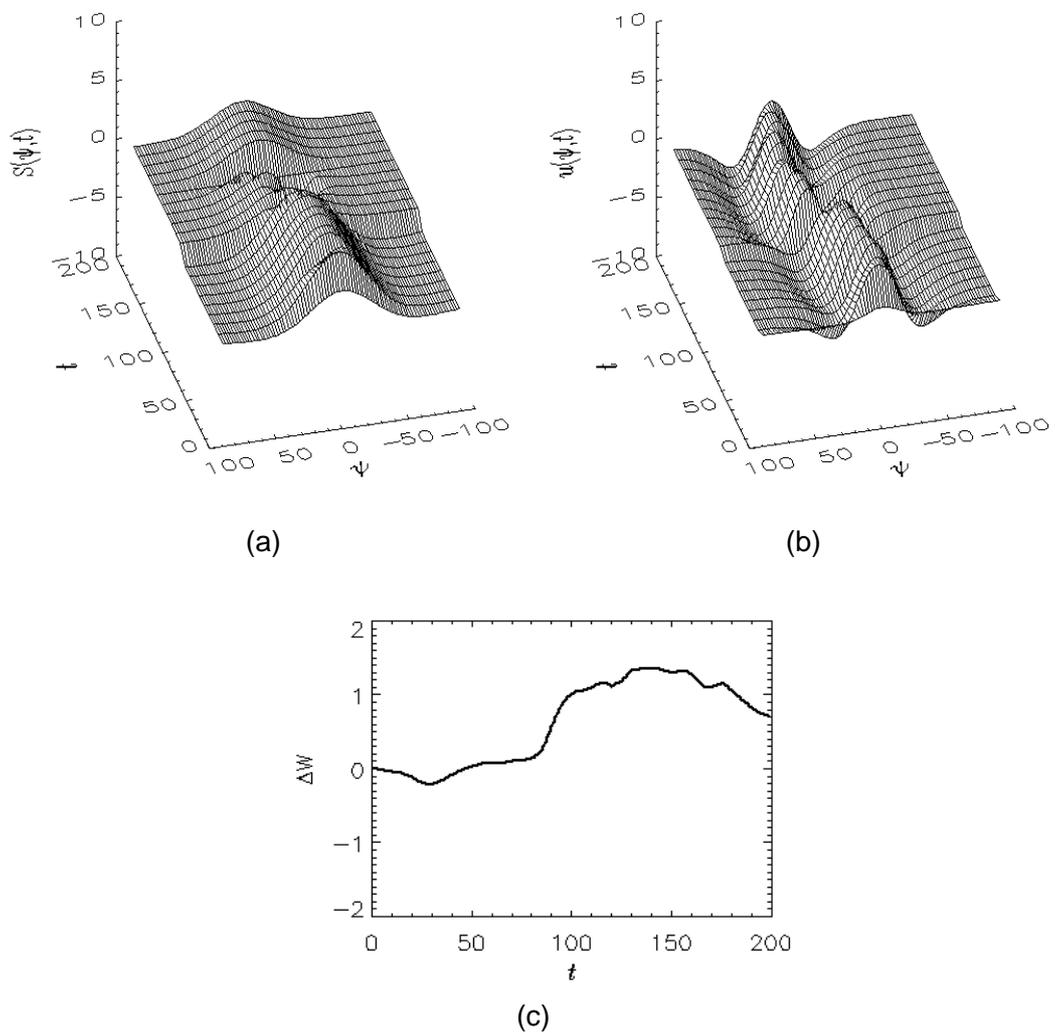


Abbildung 7.12: Fahrt entlang einer Baustelle durch eine Linkskurve. Darstellung (a) des Stimulus, (b) der Aktivität der Feldes, (c) der Lenkwinkeländerung über der Zeit.

In Abbildung 7.11a ist die Vogelperspektive der Simulation dargestellt. Das Trägerfahrzeug befindet sich dabei am linken Rand. Der Winkelbereich der Kamera sowie der geschätzte Spurverlauf sind eingezeichnet. Der Fahrspurverlauf wird aufgrund der starken Linkskurve nicht richtig interpretiert. Das führt dazu, daß das führende Fahrzeug außerhalb der Fahrspur detektiert wird. Im Laufe der Szene geht der Spurverlauf ganz verloren. Das simulierte Kamerabild ist in Abbildung 7.11b gezeigt.

Abbildung 7.12b beschreibt das Verhalten des neuronalen Feldes über der Zeit. Es wird eine Lenkbewegung nach links gefordert (Abbildung 7.8c). Die durchgeführte Fahrt entlang einer Baustelle durch eine Linkskurve wird demnach richtig interpretiert. Einem anfänglich konstanten Lenkwinkelschlag (Einfahrt in die Kurve) folgt ein starker Einschlag des Lenkwinkels nach links (Kurvendurchfahrt). Bei der Ausfahrt der Kurve wird der Lenkwinkelschlag nach rechts korrigiert.

7.3 Diskussion

Die Tragfähigkeit des realisierten Systems wird anhand zweier Videobildsequenzen im Stadtverkehr gezeigt. Das Zusammenspiel mehrerer Module und das daraus resultierende Verhalten des entwickelten Verfahrens zur Bewegungssteuerung wird demonstriert. Am Beispiel des gewünschten Lenkwinkelschlags im Bezug zum tatsächlich durchgeführten Fahrmanöver kann die Tragfähigkeit des Verfahrens belegt werden.

Es wird gezeigt, daß in realen Systemen mit Ausfällen einzelner Module bzw. mit widersprüchlichen Informationen zu rechnen ist. Der Einsatz unabhängiger Sensordaten ist notwendig, um Widersprüche mit Hilfe von Redundanzen aufzudecken und eine Steigerung der Robustheit des Systems zu erreichen. Das in dieser Arbeit entwickelte Verfahren zur Sensordatenfusion (Kapitel 5) zeigt einen Lösungsweg auf, die Robustheit eines Fahrerassistenzsystems durch Sensordatenfusion zu erhöhen.

7.4 Zusammenfassung

In diesem Kapitel wird die Tragfähigkeit der in den Kapiteln 4, 5 und 6 vorgestellten Verfahren an einem realisierten Fahrerassistenzsystem gezeigt. Das System folgt einer in [42] vorgestellten Architektur. Eine Motivation für den Einsatz dieser Architektur wird abgeleitet. Die Verfahrensintegration wird anhand zweier Beispiele gezeigt. Es wird dabei das Zusammenspiel mit in der Literatur beschriebenen Verfahren getestet.

Die aufbauend auf die Sensordatenfusion extrahierten objektspezifischen Merkmale werden für das Verfahren der Bewegungssteuerung genutzt und die Tragfähigkeit des Gesamtsystems anhand von Lenkwinkeländerungsvorschlägen gezeigt.

Kapitel 8

Zusammenfassung und Ausblick

In dieser Arbeit werden Methoden der neuronalen Informationsverarbeitung zum Entwurf von Verfahren für Fahrerassistenzsysteme im Straßenverkehr eingesetzt. Insbesondere werden neue Verfahren zur Fusion anfallender Sensordaten sowie zur Bewegungssteuerung eines Kraftfahrzeugs realisiert.

Zusammenfassung

Aufgrund der hohen Störanfälligkeit bei der Erkennung von Objekten in der natürlichen Umwelt mit Videosensoren, muß die Robustheit der Segmentierungsverfahren bei Fahrerassistenzsystemen verbessert werden. Von der Güte der Segmentierung hängt das Gesamtverhalten des Assistenzsystems entscheidend ab. Eine Verbesserung der Güte wird durch eine geeignete Auswahl und anschließende Fusion der Sensordaten erreicht. Somit wird eine Erhöhung der Segmentierungsgüte durch Systemredundanz erzielt. In dem ersten entwickelten Verfahren werden verschiedene Bildverarbeitungsalgorithmen mit selbstorganisierenden neuronalen Karten analysiert und damit relevante Komponenten für den Fusionsprozeß ausgewählt (Kapitel 4). Integraler Bestandteil des realisierten Fusionsprozesses ist ein vorwärtsgekoppeltes künstliches neuronales Netz vom Typ eines Multilayer Perceptrons, welches als nichtlinearer Klassifikator eingesetzt wird (Kapitel 5). In einem parallelen Prozeß wird die Fusion der ausgewählten Sensordaten, sowie die flexible Hinzunahme neuer Sensordaten ermöglicht. Es wird eine robuste Segmentierung des Videobildmaterials durchgeführt, so daß nachfolgend die Extraktion und Attributierung relevanter Objekte bei variierender Umgebung möglich ist.

Mit der autonomen Steuerung des Lenkwinkels und der Geschwindigkeit in Kraftfahrzeugen muß das Fahrverhalten des Fahrers modelliert werden. Deshalb sind wesentliche Bestandteile des zweiten realisierten Verfahrens der Bewegungssteuerung zwei neuronale Felder (Kapitel 6). Neuronale Felder beschreiben Verhaltensmodelle komplexer biologischer Organismen durch eine Integro-Differentialgleichung. Bei diesem Verfahren kodieren zwei eindimensionale neuronale Felder als Systemzustand den Lenkwinkel und die Geschwindigkeit. Diese werden durch Eingangsdaten auf unterschiedlichen Abstraktionsniveaus gesteuert. Berücksichtigt werden die gewünschte Fahrtrichtung, der Fahrbahnverlauf, die Distanz und die Relativgeschwindigkeit von Objekten in der Umgebung des Fahrzeugs. In einer entwickelten Simulationsumgebung wird die Steuerung des Lenkwinkels und der Geschwindigkeit bei geschlossenem Regelkreis evaluiert.

Die Tragfähigkeit beider Verfahren wird anhand eines realisierten Fahrerassistenzsystems an unterschiedlichen Sequenzen gezeigt (Kapitel 7).

Ausblick

Der Einsatz der Verfahren auf realen Daten zeigt, daß die entwickelten Verfahren in der realen Anwendung tragfähig sind.

Eine Erweiterung des Sensordatenmaterials kann zu einer weiteren Erhöhung der Güte bei der Objekterkennung führen. Insbesondere die Kombination eines RADAR-Sensors mit einem Videosensor erscheint vielversprechend, um eine weiterführende Robustheitssteigerung bei der Objekterkennung zu erreichen. Dies wird in Kapitel 5 verdeutlicht. Dabei ist die Kombination eines passiven bildgebenden Sensors mit einem aktiven Sensor, im Hinblick auf den Einsatz eines Fahrerassistenzsystems bei unterschiedlichen Wetterverhältnissen, interessant.

Eine Integration einer dynamischen Repräsentation in der Szenenanalyse läßt eine Fahrtplanung realistisch erscheinen. Insbesondere ist das „Durchspielen“ von möglichen Verhalten zur Verringerung des Unfallrisikos mit einem Fahrerassistenzsystem sinnvoll. Eine Erweiterung der vorhandenen Repräsentation um eine verhaltenskodierende Funktion, wie sie durch neuronale Felder gegeben ist, ermöglicht die Lösung einer Aufgabe dieser Art. Die Fahrtplanung führt dabei weit über die normale Prädiktion von Trajektorien hinaus und basiert auf einer „Prüfung“ relevanter Basisverhalten.

Die Steuerung der relevanten Verhaltensvariablen (Lenkwinkel und Geschwindigkeit) wird anhand mehrerer Basisverhalten demonstriert. Eine Erweiterung der Attribute um andere Variablen und das Einbringen weiterer Randbedingungen erscheint sinnvoll. Von Vorteil wäre dabei eine lernende Anpassung der Strukturparameter der verhaltensbeschreibenden Integro-Differentialgleichung.

Anhang A

Mathematischer Anhang

A.1 Stetig differenzierbare Funktionen

Neuronen haben in der Regel im gesamten Definitionsbereich monoton steigende und asymptotisch gegen einen Sättigungswert strebende Aktivierungsfunktionen $\varphi(v)$. Für die Umsetzung von Lernverfahren wird häufig stetige Differenzierbarkeit gefordert. Im folgenden werden deshalb zwei typische Funktionen vorgestellt und die geforderten Eigenschaften gezeigt.

A.1.1 Sigmoides $\xi(v)$

$$\begin{aligned}\varphi(v) &= \xi(v) = \frac{1}{1 + e^{-cv}}, & (A.1) \\ \varphi'(v) &= \frac{d\xi(v)}{dv} = \frac{ce^{-cv}}{(1 + e^{-cv})^2} = \xi(v)(1 - \xi(v)) \\ \lim_{v \rightarrow \infty} \xi(v) &= 1, \text{ mit } \lim_{v \rightarrow \infty} e^{-cv} = 0 \\ \lim_{v \rightarrow -\infty} \xi(v) &= 0, \text{ mit } \lim_{v \rightarrow -\infty} e^{-cv} = \infty.\end{aligned}$$

A.1.2 Tangens Hyperbolicus $\tanh(v)$

$$\begin{aligned}\varphi(v) &= \tanh(v) \\ &= \frac{e^v - e^{-v}}{e^v + e^{-v}} \\ &= \frac{1 - e^{-2v}}{1 + e^{-2v}} = \frac{2}{1 + e^{-2v}} - \frac{1 + e^{-2v}}{1 + e^{-2v}} \\ &= \frac{2}{1 + e^{-2v}} - 1 = 2 \cdot \xi\left(\frac{v}{2}\right) - 1 \\ &\rightarrow \text{lineare Transformation von } \xi(v).\end{aligned}$$

Daraus folgt

$$\begin{aligned}
 \varphi'(v) &= 2 \frac{d \left[\xi \left(\frac{v}{2} \right) - 1 \right]}{dv} \\
 &= 2 \frac{e^{-\frac{v}{2}}}{\left(1 + e^{-\frac{v}{2}} \right)^2} \\
 &= 2 \xi \left(\frac{v}{2} \right) \left(1 - \xi \left(\frac{v}{2} \right) \right) \\
 \lim_{\frac{v}{2} \rightarrow \infty} \xi \left(\frac{v}{2} \right) &= 1, \text{ mit } \lim_{\frac{v}{2} \rightarrow \infty} e^{-\frac{v}{2}} = 0 \\
 \lim_{\frac{v}{2} \rightarrow -\infty} \xi \left(\frac{v}{2} \right) &= 0, \text{ mit } \lim_{\frac{v}{2} \rightarrow -\infty} e^{-\frac{v}{2}} = \infty.
 \end{aligned}$$

A.2 Backpropagation-Algorithmus

Eine häufig verwendete Lernregel zur Adaption von Gewichten in Multilayer-Perceptron (MLP) Netzen ist der Backpropagation-Algorithmus [47], der im folgenden am Beispiel eines MLP mit einer Eingabeschicht $\mathbf{x}_i, i = \{1, 2, \dots, n\}$, mit n Neuronen und einem Bias x_0 , einer verdeckten Schicht $s_j, j = \{1, 2, \dots, J\}$, mit J Neuronen und einem Bias s_0 und einer Ausgabeschicht $y_k, k = \{1, 2, \dots, m\}$ mit m Neuronen hergeleitet wird. Die Aktivierungsfunktion $\varphi(v)$ eines Neurons v entspricht der Sigmoiden nach Gleichung A.1,

$$\varphi(v) = \xi(v)$$

mit $c = 1$, Die Eigenschaften eines solchen Netzwerkes lassen sich wie folgt mathematisch beschreiben:

$$\begin{aligned}
 y_k &= \varphi(v_k), \\
 v_k &= \sum_{j=0}^J w_{kj} s_j, \tag{A.2}
 \end{aligned}$$

$$\begin{aligned}
 s_j &= \varphi(v_j), \\
 v_j &= \sum_{i=0}^n w_{ji} x_i. \tag{A.3}
 \end{aligned}$$

Daraus ergibt sich für den Fehler e_k am k -ten Ausgang die Differenz aus Sollwert o_k und Ausgangswert y_k für den l -ten Trainingsvektor, mit $l \in L$ bei L Trainingsbeispielen,

$$e_k(l) = o_k(l) - y_k(l)$$

und für den quadratischen Fehler $E(l)$ des l -ten Trainingsvektors (zur Vereinfachung mit Proportionalitätsfaktor $c = \frac{1}{2}$)

$$E(l) = \frac{1}{2} \sum_{k=0}^m e_k^2(l) \tag{A.4}$$

und für den durchschnittlichen Fehler E über alle Trainingsbeispiele L gemittelt

$$E = \frac{1}{L} \sum_{l=0}^L E(l).$$

Als Gradientenabstieg wird eine Bewegung der Gewichte im Raum \mathbb{R}^m in Richtung des negativen Gradienten bezeichnet mit

$$\Delta w = -\alpha \nabla_w E(w). \quad (\text{A.5})$$

Die Variable $\alpha \in \mathbb{R}^+$ beschreibt dabei einen Lernparameter.

Der quadratische Fehler wird durch schrittweises Verändern der Neuronengewichte w durch einen Gradientenabstieg (Gleichung A.5) minimiert. Der Gradient von E (die Richtung des stärksten Anstiegs der Fehlerfunktion) ergibt sich zu

$$\nabla_w E(w) = \left(\frac{\partial E}{\partial w_1}, \frac{\partial E}{\partial w_2}, \dots, \frac{\partial E}{\partial w_m} \right)^T.$$

Für ein hinreichend kleines $\alpha > 0$ bewegt man sich in Richtung des steilsten Abfalls von E (Gleichung A.6). Für die Fehleränderung ΔE gilt bei einem solchen Schritt näherungsweise:

$$\Delta E \approx \sum_w \frac{\partial E}{\partial \Delta w} \Delta w = -\alpha \sum_w \left(\frac{\partial E}{\partial w} \right)^2 \leq 0. \quad (\text{A.6})$$

Es ergibt sich dadurch die Änderung eines einzelnen Gewichtes w_{ik} (Gewicht zu einem Ausgangsneuron nach Gleichung A.2) mit der Lernrate α unter Anwendung der Kettenregel zu

$$\begin{aligned} \Delta w_{kj} &= -\alpha \frac{\partial E}{\partial w_{kj}} \\ \text{mit } \frac{\partial E}{\partial w_{kj}} &= \frac{\partial E}{\partial e_k} \cdot \frac{\partial e_k}{\partial y_k} \cdot \frac{\partial y_k}{\partial v_k} \cdot \frac{\partial v_k}{\partial w_{kj}} \\ &= e_k \cdot (-1) \cdot \phi'(v_k) \cdot s_j \\ &= -e_k \cdot y_k \cdot (1 - y_k) \cdot s_j. \end{aligned} \quad (\text{A.7})$$

Für Neuronengewichte w_{ji} (Gewichte zu einem Neuron in der verdeckten Schicht nach Gleichung A.3) folgt daraus

$$\begin{aligned} \Delta w_{ji} &= -\alpha \frac{\partial E}{\partial w_{ji}} \\ \text{mit } \frac{\partial E}{\partial w_{ji}} &= \frac{\partial E}{\partial e_k} \cdot \frac{\partial e_k}{\partial y_k} \cdot \frac{\partial y_k}{\partial v_k} \cdot \frac{\partial v_k}{\partial s_j} \cdot \frac{\partial s_j}{\partial v_j} \cdot \frac{\partial v_j}{\partial w_{ji}} \\ &= -\sum_{k=0}^m e_k \cdot y_k \cdot (1 - y_k) \cdot w_{kj} \cdot \phi'(v_j) \cdot x_i \\ &= -\sum_{k=0}^m e_k \cdot y_k \cdot (1 - y_k) \cdot w_{kj} \cdot s_j \cdot (1 - s_j) \cdot x_i. \end{aligned} \quad (\text{A.8})$$

Anhang B

Simulationsumgebung

Basis der realisierten Simulationsumgebung ist eine Repräsentation in der Vogelperspektive (Abbildung B.1a). Diese wird zu jedem Zeitschritt aktualisiert. Die simulierten Sensoren sind an einem Trägerfahrzeug angebracht, welches sich in einem definierten Szenario bewegt. Dabei wird die Position des Fahrzeugs von den ausgewerteten Sensordaten gesteuert. In Kapitel 6 wird eine Bewegungssteuerung vorgestellt, die einen Videosensor auswertet (Abbildung B.1b).

Für komplexere Aufgabenstellungen, wie die Realisierung eines Überholvorgangs, unter Berücksichtigung von Fahrzeugen im Rückraum des Trägerfahrzeugs, reicht jedoch ein Videosensor nicht aus. Aus diesem Grund können mit Hilfe der Simulationsumgebung beliebige Positionen verschiedener Sensoren genutzt werden, um eine Rundumsicht beim Trägerfahrzeug zu realisieren.

In Abbildung B.1b-d ist das am Beispiel dreier Kameras demonstriert. Zwei Kameras observieren in der dargestellten Sensoranordnung den seitlichen Rückraum des Fahrzeugs. Für die Realisierung komplexeren Verhaltens können weiterhin andere Sensortypen, wie RADAR-Sensoren, für verschiedene Aufgabenstellungen (Observation des Nah- und Fernfeldes, Abbildung B.1e-h) simuliert werden.

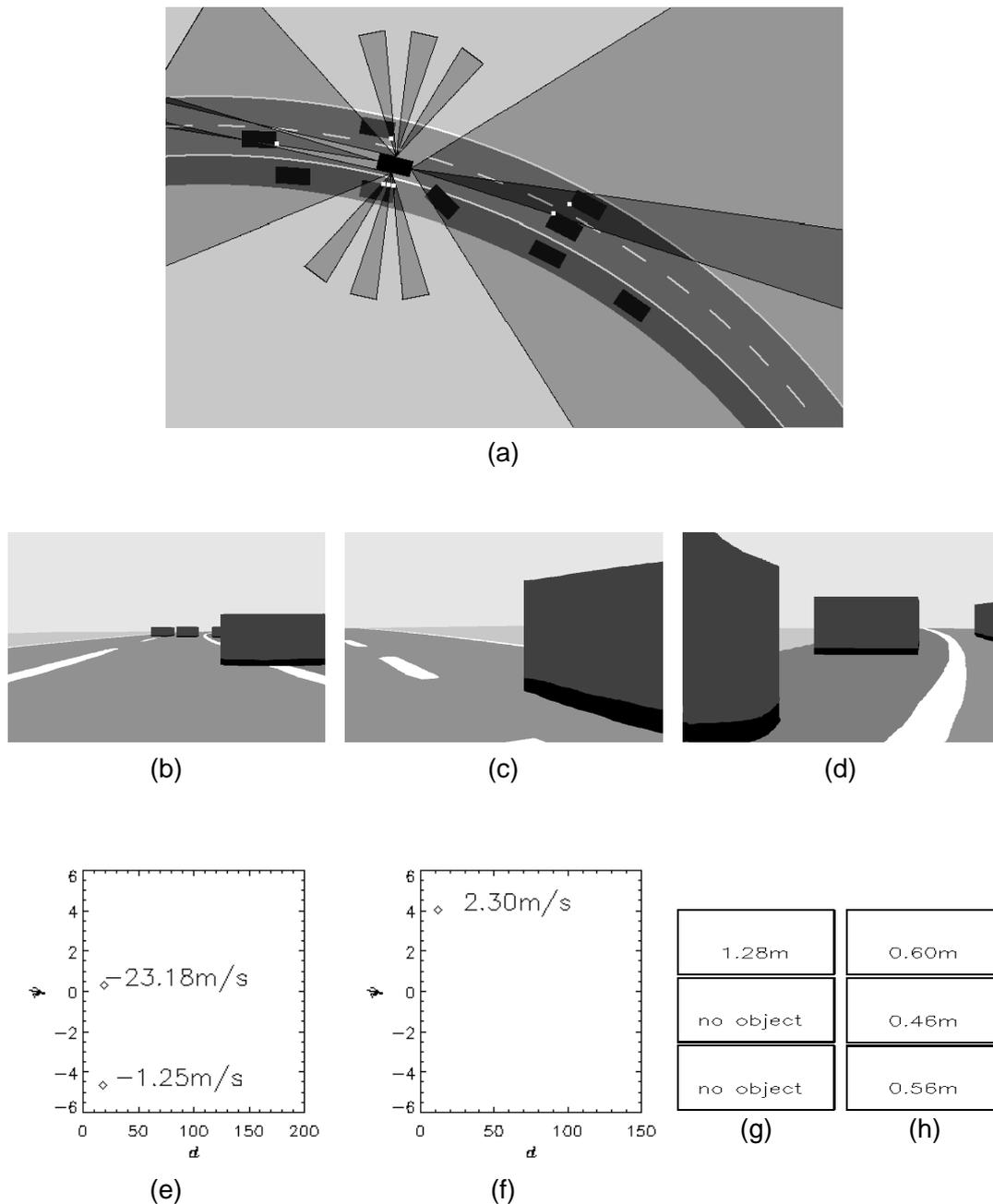


Abbildung B.1: Ergebnisbilder der Simulation. Ein Trägerfahrzeug (schwarz) ist mit mehreren Sensoren ausgestattet. (a) Der observierte Bereich ist in der Vogelperspektive dunkel unterlegt. (b) Eine Weitwinkelkamera observiert in Kombination mit (e) einem RADAR-Sensor (Distanz d , Winkel ψ und Relativgeschwindigkeit) den Bereich vor dem Fahrzeug. Der Seitenbereich wird durch (c,d) zwei rückwärtsgerichtete Kameras beobachtet. (f) Ein zweiter RADAR-Sensor observiert den Rückraum (Distanz d , Winkel ψ und Relativgeschwindigkeit). (g,h) Sechs Nahbereich-RADAR-Sensoren geringerer Wellenlänge überprüfen die Bereiche neben dem Fahrzeug (Distanz).

Anhang C

Notation

Allgemeine Notation

a	Skalar
\mathbf{a}	Spaltenvektor
\mathbf{a}^T	transponierter Vektor zu \mathbf{a}
a_i	i -tes Element des Vektors \mathbf{a}
$f(\mathbf{a})$	Funktion von \mathbf{a}
\bar{a}	Mittelwert von \mathbf{a}
σ_a	Standardabweichung über \mathbf{a}
\mathbb{R}	Menge der reellen Zahlen
t	diskrete Zeitkoordinate
\mathbf{a}'	Ableitung von \mathbf{a}
$\dot{\mathbf{a}} = \frac{\partial \mathbf{a}}{\partial t}$	partielle Ableitung von \mathbf{a} nach t
c	Parameter
i, j, k, l	Indizes
I, J, K, L	Gesamtzahl der Elemente eines Vektors
(p, q)	Pixelkoordinaten

Notation für Kapitel 4

\mathbf{X}	Eingangsdatenraum
n	Dimension von \mathbf{X}
\mathbf{x}	Eingangsdatenvektor aus \mathbf{X}
$\tilde{\mathbf{X}}$	Datensatz aus \mathbf{X}
$\tilde{\mathbf{x}}$	Datenvektor aus $\tilde{\mathbf{X}}$
\mathbf{w}	Referenzvektor eines Kartenneurons
d	euklidische Distanz
V	Voronoigebiet
ϑ	Nachbarschaftsfunktion
N	Gaußfunktion
α, σ	Lernparameter
C	Klasse
X	Komponenten einzelner Segmentierungsverfahren
K, E, V	ausgewählte Komponenten

Notation für Kapitel 5

\mathbf{X}	Eingangsdatenraum
n	Dimension von \mathbf{X}
\mathbf{x}	Eingangsdatenvektor aus \mathbf{X}
$\tilde{\mathbf{X}}$	Datensatz aus \mathbf{X}
$\tilde{\mathbf{x}}$	Datenvektor aus $\tilde{\mathbf{X}}$
$\tilde{\mathbf{o}}$	Sollwertvektor aus $\tilde{\mathbf{X}}$
\mathbf{Y}	Zieldatenraum
m	Dimension von \mathbf{Y}
\mathbf{y}	Ausgangsdatenvektor aus \mathbf{Y}
\mathbf{o}	Sollwertvektor zu \mathbf{y}
\mathbf{w}	Synapsengewichte eines Multilayer Perceptrons
v	Summe der Einzelaktivitäten einer Vorgängerschicht
s	Ausgangsaktivität eines Neurons
κ	Schicht eines Multilayer Perceptrons
φ	Aktivierungsfunktion
$\xi(x)$	sigmoide Funktion
$\tanh(x)$	Tangens Hyperbolicus
e	Differenz zwischen Soll- und Istwert
E	quadratischer Fehler
α	Lernparameter
ζ	Gewichtsparameter
C	Klasse
g	Verstärkungsfaktor
VAR	Varianz
Φ	Integrationsbereich
ϕ	Laufvariable in Φ
K, E, V, C, R, S	ausgewählte Komponenten einzelner Segmentierungsverfahren
P	prozentuale Reduktion des Datenmaterials
R, G, B	Farbwerte
S	Sättigung

Notation für Kapitel 6

u	Aktivität
\mathbf{z}	Ort von u
Γ	Betrachtungsbereich
h	Voraktivierung
S	Stimulus
φ	Aktivierungsfunktion, Nichtlinearität
w	Wechselwirkung
W	Energie von w
N	Gaußfunktion
σ	Parameter von N
R	a -Lösung von u
a	Definitionsbereich von R
$\xi(u)$	sigmoide Funktion
d	Distanz
v_x	Differenzgeschwindigkeit in x -Richtung
v_y	Differenzgeschwindigkeit in y -Richtung
b	Belegung
Ψ	Winkelbereich
v	Geschwindigkeit
Δ	Differenz zweier Variablen gleichen Typs
$-\gamma, \gamma$	Faltungsgrenzen
$\zeta, \alpha, \beta, \nu, \delta, \eta$	Parameter mikroskopischer Größen
O	Objekt
L	Fahrspurverlauf
F	führendes Fahrzeug
D	Differenz zur Sollgeschwindigkeit
N, M	Argumente der Maxima von u
W	Lenkwinkel des Trägerfahrzeugs
V	Geschwindigkeit des Trägerfahrzeugs

Literaturverzeichnis

- [1] S. Amari. Dynamics of pattern formation in lateral-inhibition type neural fields. *Biological Cybernetics*, 27:77–87, 1977.
- [2] S. Amari. Dynamical stability of formation of cortical maps. In M.A. Arbib und S. Amari (Hrsg.), *Dynamic Interactions in Neural Networks: Models and Data*, Seite 15–34. Springer-Verlag, New York, USA, 1989.
- [3] S.-I. Amari. Field Theory of Self-Organizing Neural Nets. *IEEE Transactions on Systems, Man, and Cybernetics*, 13:741–748, 1983.
- [4] S. Baten und N. Müller. Navigation with an Autonomous Vehicle on Arbitrary Roads. In *3rd International Workshop High Precision Navigation*, 1995.
- [5] A. Benninghoff. *Anatomie, Band 3*. Urban & Schwarzenberg, München, Germany, 1985.
- [6] T. Bergener, C. Bruckhoff, P. Dahm, H. Janßen, F. Joublin, R. Menzner, A. Steinhage und W. von Seelen. Complex Behavior by means of Dynamical Systems for an Anthropomorphic Robot. *special issue of Neural Networks*, 1999.
- [7] M. Bertozzi und A. Broggi. GOLD: a Parallel Real-Time Stereo Vision System for Generic Obstacle and Lane Detection. *IEEE Transactions on Image Processing*, 7(1):62–81, 1998.
- [8] C.M. Bishop. *Neural Networks for Pattern Recognition*. Oxford University Press Inc., New York, 1995.
- [9] N. Bronstein und K.A. Semendjajew. *Taschenbuch der Mathematik*. Verlag Harri Deutsch, Frankfurt/Main, Germany, 23 Ausgabe, 1987.
- [10] R.R. Brooks und S.S. Ivengar. *Multi-Sensor Fusion*. Prentice Hall PTR, Upper Saddle River, New Jersey, 1998.
- [11] H. Burkhardt (Hrsg.). *WTA 98, Workshop on Texture Analysis 1998*, Freiburg, Germany, 1998. Albert-Ludwigs-Universität, Institut für Informatik.
- [12] M. Camiade, D. Domnesque, P.F. Alleaume, A. Mallet und H. Daembkes. GAAS MMICS for automotive radar applications. In *IV'98, IEEE International Conference on Intelligent Vehicles 1998*, Stuttgart, Germany, 1998. IEEE.
- [13] R. Cucchiara, M. Piccardi und P. Mello. Image Analysis and Rule-based Reasoning for a Traffic Monitoring System. In *ITSC'99, IEEE Conference on Intelligent Transportation Systems 1999*, Seite 758 – 763, Tokyo, Japan, 1999. IEEE.
- [14] C. Curio, J. Edelbrunner, T. Kalinke, C. Tzomakas und W. von Seelen. Walking Pedestrian Recognition. In *ITSC'99, IEEE Conference on Intelligent Transportation Systems 1999*, Seite 292 – 297, Tokyo, Japan, 1999. IEEE.
- [15] B. V. Dasarathy. *Decision Fusion*. IEEE Computer Society Press, Los Alamitos, 1994.
- [16] B.V. Dasarathy (Hrsg.). *Sensor Fusion: Architectures, Algorithms, and Applications*, Band 3376 von *Proceedings of SPIE*, Bellingham, USA, 1998. SPIE.
- [17] D. de Ridder, K. Schutte und P. Schwering. Detection of vehicles in infrared imagery using shared weight neural network feature detectors. In *Proceedings of SPIE Vol. 3374*, Seite 247–258, Orlando, USA, 1998. SPIE. Session Sensor Fusion 1998.

- [18] E.D. Dickmanns, R. Behringer, D. Dickmanns, T. Hildebrandt, M. Maurer, F. Thomanek und J. Schiehlen. The seeing passenger car VaMoRs-P. In *Proceedings of the Intelligent Vehicles '94 Symposium, Paris, France*, Seite 260–265, 1994.
- [19] E.D. Dickmanns et al. Vehicles capable of dynamic vision. In *15th International Joint Conference on Artificial Intelligence (IJCAI)*, Seite 1–16, Nagoya, Japan, 1997.
- [20] M. Dose. *Wegeplanung autonomer mobiler Roboter mittels dynamischer Systeme*. Dissertation, Institut für Neuroinformatik, Ruhr-Universität Bochum, Germany, 1994.
- [21] R. Durbin und G. Mitchison. A dimension reducing framework for understanding cortical maps. *Nature*, 343:644–647, 1990.
- [22] P. Duus. *Neurologisch-topische Diagnostik*. Thieme, Stuttgart, Germany, 1990.
- [23] ELTEC Elektronik GmbH, Mainz. THINEDGE-Processor for Contour Matching. Hardware Manual, Rev. 1A, 1991.
- [24] W. Enkelmann. *Entwicklung von Systemen zur Interpretation von Straßenverkehrsszenen durch Bildfolgenauswertung*. Infix, Sankt Augustin, 1997.
- [25] D. Farkas, J. Young B. Baertlein und Özgüner. forward-looking radar navigation system for 1997 AHS demonstration. In *ITSC'97, IEEE Conference on Intelligent Transportation Systems 1997*, Boston, USA, 1997. IEEE.
- [26] J.D. Foley, A. van Dam, S.K. Feiner und J.F. Hughes. *Computer Graphics: Principles and Practice*. Addison-Wesley Publishing Company, Inc., 1990.
- [27] D.M. Gavrilu. Real-Time Object Detection Using Distance Transforms. In *IV'98, IEEE International Conference on Intelligent Vehicles 1998*, Stuttgart, Germany, 1998. IEEE.
- [28] A. Giachetti, M. Campani, R. Sanni und A. Succi. The Recovery of Optical Flow for Intelligent Cruise Control. In *Proceedings of the Intelligent Vehicles '94 Symposium, Paris, France*, Seite 91–96, 1994.
- [29] M.A. Giese. *Dynamic Neural Field Theory for Motion Perception*. Kluwer, Norwell, USA, 1999.
- [30] W. Gillner. Motion based vehicle detection on motorways. In *Proceedings of the Intelligent Vehicles '95 Symposium, Detroit, USA*, Seite 483–487, 1995.
- [31] C. Goerick. Local Orientation Coding and Adaptive Thresholding for Real Time Early Vision. Internal Report IRINI 94-05, Institut für Neuroinformatik, Ruhr-Universität Bochum, D-44780 Bochum, Germany, Juni 1994.
- [32] C. Goerick. *Beiträge zur Theorie der Lerndynamik künstlicher neuronaler Netze*. Nummer 557 in Fortschrittberichte, Reihe 10. VDI-Verlag, Düsseldorf, 1998. Dissertation, Ruhr-Universität Bochum.
- [33] C. Goerick. Considerations of the Gain Spectrum. Internal Report IRINI 98-02, Institut für Neuroinformatik, Ruhr-Universität Bochum, D-44780 Bochum, Germany, April 1998.
- [34] C. Goerick, D. Noll und M. Werner. Artificial Neural Networks in Real Time Car Detection and Tracking Applications. *Pattern Recognition Letters*, 1995.
- [35] M.S. Grewal und A.P. Andrews. *Kalman Filtering*. Prentice-Hall, Inc., Englewood Cliffs, New Jersey, USA, 1993.
- [36] P. Haberäcker. *Digitale Bildverarbeitung, Grundlagen und Anwendung*. Carl Hanser Verlag, München, Wien, 3 Ausgabe, 1989.
- [37] U. Handmann et al. The Pattern Recognition and Scene Analysis Group, Driver Assistance Systems. URL: http://www.neuroinformatik.ruhr-uni-bochum.de/ini/PROJECTS/PATREC/pr_top.html, Institut für Neuroinformatik, Ruhr-Universität Bochum, 1999.
- [38] U. Handmann und T. Kalinke. Fusion of texture and contour based methods for object recognition. In *ITSC'97, IEEE Conference on Intelligent Transportation Systems 1997*, Boston, 1997. IEEE. Session 35: Intelligent Vehicles: Vision(3).

- [39] U. Handmann, T. Kalinke, C. Tzomakas, M. Werner und W. von Seelen. An Image Processing System for Driver Assistance. In *IV'98, IEEE International Conference on Intelligent Vehicles 1998*, Seite 481 – 486, Stuttgart, Germany, 1998. IEEE.
- [40] U. Handmann, T. Kalinke, C. Tzomakas, M. Werner und W. von Seelen. An Image Processing System for Driver Assistance. *Image and Vision Computing (Elsevier)*, 18(5):367 – 376, 2000.
- [41] U. Handmann und I. Leefken. Scene Interpretation and Behavior Planning for Driver Assistance. In *Proceedings of SPIE*, Orlando, 2000. SPIE. Session Enhanced and Synthetic Vision 2000.
- [42] U. Handmann, I. Leefken, C. Tzomakas und W. von Seelen. A Flexible Architecture for Driver Assistance. In *SPIE's International Symposium on Intelligent Systems and Advanced Manufacturing 1999 (Mobile Robots and Autonomous Systems)*, *Proceedings of SPIE Vol. 3838*, Seite 2 – 11, Boston, 1999. SPIE.
- [43] U. Handmann, I. Leefken, C. Tzomakas und W. von Seelen. A Flexible Architecture for Intelligent Cruise Control. In *ITSC'99, IEEE Conference on Intelligent Transportation Systems 1999*, Seite 958 – 963, Tokyo, Japan, 1999. IEEE.
- [44] U. Handmann, G. Lorenz, T. Schnitger und W. von Seelen. Fusion of Different Sensors and Algorithms for Segmentation. In *IV'98, IEEE International Conference on Intelligent Vehicles 1998*, Seite 499 – 504, Stuttgart, Germany, 1998. IEEE.
- [45] U. Handmann, G. Lorenz und W. von Seelen. Fusion von Basisalgorithmen zur Segmentierung von Straßenverkehrsszenen. In *Mustererkennung 1998*, Seite 101– 108, Heidelberg, 1998. Springer-Verlag.
- [46] R.M. Haralick und L.G. Shapiro. *Computer and Robot Vision*, Band I. Addison-Wesley, Reading, Massachusetts, 1992.
- [47] S. Haykin. *Neural Networks*. Macmillan College Publishing Company, New York, 1994.
- [48] J. Himberg. Enhancing the som based data visualization by linking different data projections. In *Proceedings of the International Symposium on Intelligent Data Engineering and Learning (IDEAL'98)*, Seite 427–434, Hong Kong, October 1998.
- [49] M. Hischke. Collision Warning Radar Inference. In *Proceedings of the Intelligent Vehicles '95 Symposium, Detroit, USA*, Seite 13–18, 1995.
- [50] D.H. Hubel. *Auge und Gehirn*. Spektrum der Wissenschaft Verlagsgesellschaft mbG & Co, Heidelberg, Germany, 1989.
- [51] D.P. Huttenlocher. Comparing Images Using the Hausdorff Distance. *IEEE Transactions on PAMI*, 15(9), September 1993.
- [52] IEEE (Hrsg.). *Special Issue on Neural Networks*, Band 78 (9) von *Proceedings of the IEEE*. IEEE, 1990.
- [53] R. Inguva. Fusion of LWIR sensor data by Bayesian methods. In *Proceedings of SPIE Vol. 3376*, Seite 161–174, Orlando, USA, 1998. SPIE. Session Sensor Fusion 1998.
- [54] B. Jähne. *Digitale Bildverarbeitung*. Springer Verlag, Berlin, Heidelberg, New York, 3 Ausgabe, 1993.
- [55] A.J. Jain. *Fundamentals of digital image processing*. Prentice-Hall, Inc., Englewood Cliffs, New Jersey, USA, 1986.
- [56] G. Jetschke. *Mathematik der Selbstorganisation*. F. Vieweg & Sohn Verlagsgesellschaft mbH, Braunschweig, Germany, 1989.
- [57] C.M. Johnson. The National ITS Program: Where We've Been and Where We're Going. URL: www.tfhr.gov/pubrds/pr97-10/p6.htm, U.S. Department of Transportation, Federal Highway Administration Research, Development, and Technology, Turner-Fairbank Highway Research Center, 6300 Georgetown Pike McLean, Virginia 22101-2296, 1999.
- [58] Y.-K. Jung und Y.-S. Ho. Traffic Parameter Extraction using Video-based Vehicle Tracking. In *ITSC'99, IEEE Conference on Intelligent Transportation Systems 1999*, Seite 764 – 769, Tokyo, Japan, 1999. IEEE.

- [59] T. Kalinke. *Texturbasierte dynamische Erkennung veränderlicher Objekte*. Dissertation, Institut für Neuroinformatik, Ruhr-Universität Bochum, Germany, 1999.
- [60] T. Kalinke und W. von Seelen. Entropie als Maß des lokalen Informationsgehalts in Bildern zur Realisierung einer Aufmerksamkeitssteuerung. In *Mustererkennung 1996*, Seite 627–634, 1996.
- [61] C. Knöppel, U. Regensburger und B. Michaelis. Detektion und Bestimmung des Abstandes von Straßenfahrzeugen in großer Entfernung. In *Mustererkennung 1999*, Seite 258–265, Heidelberg, 1999. Springer-Verlag.
- [62] T. Kohonen. Analysis of a simple self-organizing process. *Biological Cybernetics*, 44:135–140, 1982.
- [63] T. Kohonen. Self-organized formation of topologically correct feature maps. *Biological Cybernetics*, 43:59–69, 1982.
- [64] T. Kohonen. *Self-Organizing Maps*. Springer-Verlag, Berlin, Heidelberg, Germany, 1995.
- [65] D. Koller, J. Weber und J. Malik. Robust multiple car tracking with occlusion reasoning. In *Third European Conference on Computer Vision, Stockholm, Sweden*, Seite 189 – 196, Berlin, Heidelberg, New York, 1994. Springer Verlag.
- [66] B. Kosko. *Neural networks and fuzzy systems*. Prentice-Hall, Inc., Englewood Cliffs, New Jersey, USA, 1992.
- [67] L.D. Landau und E.M. Lifschitz. *Klassische Feldtheorie*. Akademie-Verlag, Berlin, DDR, 1989.
- [68] Y. LeCun, L. Bottou, G.B. Orr und K.-R. Müller. *Neural Networks: Tricks of the Trade (editor: G.B. Orr and K.-R. Müller)*, Kapitel Efficient BackProp. Springer-Verlag, Berlin, Heidelberg, Germany, 1998.
- [69] C. Little. The Intelligent Vehicle Initiative: Advancing "Human-Centered" Smart Vehicles. *Public Roads*, 61(2), 1997.
- [70] L. Lohmann, C. Nowack und J. Ruiz del Solar. A robust architecture for edge detection in real-world scenes. In *WTA 98, Workshop on Texture Analysis 1998*, Seite 125–135, Freiburg, Germany, 1998. Albert-Ludwigs-Universität, Institut für Informatik.
- [71] H.A. Mallot. *Sehen und Verarbeitung visueller Informationen*. Vieweg, 1998.
- [72] D. McFarland (Hrsg.). *Animal Behaviour*. Oxford University Press, Oxford, Great Britain, 1987.
- [73] R. Mech und M. Wollhorn. A noise robust method for segmentation of moving objects in video sequences. In *ICASSP'97 Proceedings*, Seite 2657 – 2660, Munich, Germany, 1997.
- [74] E.B. Meier und F. Ade. Tracking Cars in Range Images Using the Condensation Algorithm. In *ITSC'99, IEEE Conference on Intelligent Transportation Systems 1999*, Seite 129 – 134, Tokyo, Japan, 1999. IEEE.
- [75] E. Mémin und P. Pérez. Dense estimation and object-based segmentation of the optical flow with robust techniques. *Transactions on Image Processing*, 7(5):703–719, may 1998.
- [76] F.W. Merkel. *Orientierung im Tierreich*. Gustav Fischer Verlag, Stuttgart, Germany, 1980.
- [77] N.P. Morenc und G.G. Lajiness. 76 GHz collision warning transceiver component requirements and current status. In *ITSC'97, IEEE Conference on Intelligent Transportation Systems 1997*, Boston, USA, 1997. IEEE.
- [78] K.D. Mörike, E. Betz und W. Mergenthaler. *Biologie des Menschen*. Quelle und Meyer Verlag, Wiesbaden, Heidelberg, 12 Ausgabe, 1989.
- [79] M. Morita und T. Yokota. Realization of Offset Optimization Systems which takes Crossroad Traffic Demand into Account. In *ITSC'99, IEEE Conference on Intelligent Transportation Systems 1999*, Seite 392 – 397, Tokyo, Japan, 1999. IEEE.
- [80] MoTiV-Büro. MoTiV – Mobilität und Transport im intermodalem Verkehr. URL: www.motiv.de, MOTIV-Büro, c/o Walter Scholl, Paradiesstr. 34, 73230 Kirchheim/Teck, 2000.
- [81] N. Murata, S. Yoshizawa und S. Amari. Network Information Criterion – Determining the Number of Hidden Units for an Artificial Neural Network Model. *IEEE Transactions on Neural Networks*, 5:865–872, 1994.

- [82] M. Nagai, K. Yamazaki und S. Miyata. The Saturated Traffic flow Rate Analysis for Improvement of 'STREAM'. In *ITSC'99, IEEE Conference on Intelligent Transportation Systems 1999*, Seite 382 – 387, Tokyo, Japan, 1999. IEEE.
- [83] S. Niikura, Y. Mashiyama, M. Kurwahara, T. Yoshii und T. Oda. An Analysis of Estimating Turning Distribution Rate at Intersections. In *ITSC'99, IEEE Conference on Intelligent Transportation Systems 1999*, Seite 388 – 391, Tokyo, Japan, 1999. IEEE.
- [84] D. Noll. *Ein Optimierungsansatz zur Objekterkennung*. Nummer 454 in Fortschrittberichte, Reihe 10. VDI-Verlag, Düsseldorf, 1996. Dissertation, Ruhr-Universität Bochum.
- [85] Japanese Ministry of Construction. Basic Government Guidelines of advanced Information and Communications in the Fields of Roads, Traffic and Vehicles. URL: www.nihon.net/ITS/ITSinJapan/mri2.html, Japanese Ministry of Construction, ITS-Homepage, 1998.
- [86] R. Onken. DAISY, an Adaptive, Knowledge-based Driver Monitoring and Warning System. In *Proceedings of the Intelligent Vehicles '94 Symposium, Paris, France*, Seite 544–549, 1994.
- [87] G.B. Orr und K.-R. Müller (Hrsg.). *Neural Networks: Tricks of the Trade*. Springer-Verlag, Berlin, Heidelberg, Germany, 1998.
- [88] F. Paetzold und U. Franke. Road recognition in urban environment. In *IV'98, IEEE International Conference on Intelligent Vehicles 1998*, Seite 87 – 91, Stuttgart, Germany, 1998. IEEE.
- [89] P. Palacharla, P.C. Nelson und V.P. Sisoioiku. Data Fusion Using Fuzzy Valued Logic. In *Proceedings of the Intelligent Vehicles '94 Symposium, Paris, France*, Seite 115–6, 1994.
- [90] F. Panik. Automobiltechnik als Korrektiv menschlichen Unvermögens? In *FhG-Berichte*, Band 4. Fraunhofer-Gesellschaft, München, Germany, 1987.
- [91] C. Papageorgiou, T. Evgeniou und T. Poggio. A Trainable Pedestrian Detection System. In *IV'98, IEEE International Conference on Intelligent Vehicles 1998*, Stuttgart, Germany, 1998. IEEE.
- [92] L. Perko. *Differential Equations and Dynamical Systems*. Springer, New York, USA, 1991.
- [93] O. Pichler, A. Teubner und B.J. Hosticka. An unsupervised texture segmentation algorithm with feature space reduction and knowledge feedback. *Transactions on Image Processing*, 7(1):53–61, january 1998.
- [94] D. Pomerleau. RALPH: Rapidly Adapting Lateral Position Handler. In *Proceedings of the Intelligent Vehicles '95 Symposium, Detroit, USA*, Seite 506–511, 1995.
- [95] L. Prechelt. *Neural Networks: Tricks of the Trade (editor: G.B. Orr and K.-R. Müller)*, Kapitel Early Stopping – But When? Springer-Verlag, Berlin, Heidelberg, Germany, 1998.
- [96] Prometheus Office. *Proceedings of the workshop PROMETHEUS*, Wolfsburg, Germany, 1989. PROMETHEUS OFFICE c/o Daimler Benz AG, Postfach 60 02 02, D-7000 Stuttgart 60.
- [97] Prometheus Office. *Proceedings of the 3rd workshop PROMETHEUS*, Torino, Italy, 1990. PROMETHEUS OFFICE c/o Daimler Benz AG, Postfach 60 02 02, D-7000 Stuttgart 60.
- [98] L. C. Ramac, M. K. Uner, P. K. Varshney, M. Alford und D. Ferris. Morphological filters and wavelet based image fusion for concealed weapons detection. In *Proceedings of SPIE Vol. 3376*, Seite 110–119, Orlando, USA, 1998. SPIE. Session Sensor Fusion 1998.
- [99] D. Reichardt und J. Schick. Collision Avoidance in dynamic environments applied to autonomous vehicle guidance on the motorway. In *Proceedings of the Intelligent Vehicles '94 Symposium, Paris, France*, Seite 74–78, 1994.
- [100] R. Reynaud und T. Maurin. On Board Data Fusion and Detection System used for Obstacle Detection: 2D vision versus 1D sensor Fusion. In *Proceedings of the Intelligent Vehicles '94 Symposium, Paris, France*, Seite 243–248, 1994.
- [101] H. Ritter, T. Martinetz und K. Schulten. *Neuronale Netze: Eine Einführung in die Neuroinformatik selbstorganisierter Netzwerke*. Addison-Wesley, Bonn, München, Reading, Massachusetts (u.a.), 2 Ausgabe, 1991.

- [102] W. Ritter, F. Stein und R. Janssen. Traffic Sign Recognition Using Colour Information. *Mathl. Comput. Modelling*, 22(4–7):149–161, 1995.
- [103] I. Rock. *Perception*. Scientific American Library, New York, USA, 1984.
- [104] O. Rockinger und T. Fechner. Pixel-Level Image Fusion: The Case of Image Sequences. In *Proceedings of SPIE Vol. 3374*, Seite 378–388, Orlando, USA, 1998. SPIE. Session Sensor Fusion 1998.
- [105] R. Rojas. *Theorie der neuronalen Netze*. Springer Verlag, Berlin, Heidelberg, New York, 4 Ausgabe, 1996.
- [106] F. Rosenblatt. The perceptron: A probabilistic model for information storage and organisation in the brain. *Psychological Review*, 65:386–408, 1958.
- [107] C. Rother und H.-H. Nagel. Analyse initialer Positionsschätzungen bei Bildfolgenauswertung. In *Mustererkennung 1999*, Seite 189–196, 1999.
- [108] D.E. Rumelhart, G.E. Hinton und R.J. Williams. *Parallel Distributed Processing (editor: Rumelhart and McClelland)*, Kapitel Learning internal representations by error propagation. MIT-Press, Cambridge, USA, 1986.
- [109] W.T. Scherer, B.L. Smith und C.C. McGhee. The development of integrated transportation system management (itsm) in North Virginia. In *ITSC'97, IEEE Conference on Intelligent Transportation Systems 1997*, Boston, 1997. IEEE. Session 8: Integrated Transportation System Management.
- [110] T. Schnitger und U. Handmann. Fusion von Bildanalyseverfahren mittels einer neuronalen Kopplungsstruktur. Internal Report IRINI 98-01, Institut für Neuroinformatik, Ruhr-Universität Bochum, D-44780 Bochum, Germany, April 1998.
- [111] G. Schöner, M. Dose und C. Engels. Dynamic of behavior: theory and applications for autonomous robot architectures. *Robotics and Autonomous Systems*, 16:213–245, 1995.
- [112] G. Schöner und J.A.S. Kelso. Dynamic pattern generation in behavioral and neural systems. *Science*, 239:1513–1520, 1988.
- [113] C.E. Shannon. A mathematical theory of communication. *Bell Systems Technical Journal*, 27:379–423, 623–656, 1948.
- [114] O. Simula, E. Alhoniemi, J. Hollmén und J. Vesanto. Analysis of complex systems using the self-organizing map. In *Proceedings of the International Conference on Neural Information Processing and Intelligent Information Systems (ICONIP'97)*, Seite 1313–1317, 1997.
- [115] M. Tabb und N. Ahuja. Multiscale image segmentation by integrated edge and region detection. *Transactions on Image Processing*, 6(5):642–655, may 1997.
- [116] R.F. Thompson. *Das Gehirn*. Spektrum Akademischer Verlag, Heidelberg, Berlin, Oxford, 2 Ausgabe, 1994.
- [117] L. Trassoudaine, D. Hutbert, P. Checchin, J. Alizon, J. Gallice und M. Thonnat. Building an Environment Map around the Prolab2 Vehicle Using a Controllable Range Sensor. In *Proceedings of the Intelligent Vehicles '94 Symposium, Paris, France*, Seite 562–567, 1994.
- [118] K. Tsubaki. Environment Protection Management Systems (EPMS). In *ITSC'99, IEEE Conference on Intelligent Transportation Systems 1999*, Seite 998 – 1002, Tokyo, Japan, 1999. IEEE.
- [119] C. Tzomakas. *Contributions to the Visual Object Detection and Classification*. Dissertation, Institut für Neuroinformatik, Ruhr-Universität Bochum, Germany, 1999.
- [120] B. Ulmer. VITA II -Active Collision Avoidance in Real Traffic. In *Proceedings of the Intelligent Vehicles '94 Symposium, Paris, France*, Seite 1–6, 1994.
- [121] V. v. Holt und S. Baten. Perceptual architecture for a vision system of autonomous vehicles. In *IV'98, IEEE International Conference on Intelligent Vehicles 1998*, Seite 539 – 544, Stuttgart, Germany, 1998. IEEE.
- [122] P.K. Varshney (Hrsg.). *Special Issue on Data Fusion*, Band 85 (1) von *Proceedings of the IEEE*. IEEE, 1997.

- [123] C. v.d. Malsburg. Self-organizing of orientation sensitive cells in the striata cortex. *Kybernetik*, 14:84–100, 1973.
- [124] J. Vesanto. Som-based data visualization methods. *Intelligent Data Analysis*, 3:111–126, 1999.
- [125] M. Wahl, D. Georges und M. Dang. data pre-processing for obstacle detection in automotive applications. In *ITSC'97, IEEE Conference on Intelligent Transportation Systems 1997*, Boston, USA, 1997. IEEE.
- [126] A. Walsdorf, H. Putzer und R. Onken. The cognitive process and its application within cockpit assistant systems. In *ITSC'99, IEEE Conference on Intelligent Transportation Systems 1999*, Seite 874 – 879, Tokyo, Japan, 1999. IEEE.
- [127] M. Werner. *Objektverfolgung und Objekterkennung mittels der partiellen Hausdorff-Distanz*. Nummer 574 in Fortschrittberichte, Reihe 10. VDI-Verlag, Düsseldorf, 1999. Dissertation, Ruhr-Universität Bochum.
- [128] R. Wertheimer et al. Ein hochsprachenprogrammierbares System zur Vollbilddauswertung im Videotakt: Anwendungen zur Interpretation monokularer, semi-strukturierter Bildfolgen bei natürlicher Beleuchtung und schnell bewegter Kamera. In *Mustererkennung 1998*, Seite 37– 65, Heidelberg, 1998. Springer-Verlag.
- [129] T. Williamson und C. Thorpe. Detection of Small Obstacles at Long Range Using Multibaseline Stereo. In *IV'98, IEEE International Conference on Intelligent Vehicles 1998*, Stuttgart, Germany, 1998. IEEE.
- [130] H.R. Willson und J.D. Cowan. A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. *Kybernetik*, 13:55–80, 1973.
- [131] B. Wirtzner, W.M. Grimm, H. Schmidt und R. Klinnert. Inference Cancellation in Ultrasonic Sensor Arrays by Stochastic Coding and Adaptive Filtering. In *IV'98, IEEE International Conference on Intelligent Vehicles 1998*, Stuttgart, Germany, 1998. IEEE.
- [132] C. Woehler, J.K. Anlauf, T. Portner und U. Franke. A Time Delay Neural Network Algorithm for Real-Time Pedestrian Recognition. In *IV'98, IEEE International Conference on Intelligent Vehicles 1998*, Stuttgart, Germany, 1998. IEEE.
- [133] K. Yamamoto und K. Yamada. Image Processing and Fusion to detect Navigation Obstacles. In *Proceedings of SPIE Vol. 3374*, Seite 337–346, Orlando, USA, 1998. SPIE. Session Sensor Fusion 1998.
- [134] J. Zhang und H. Nagel. Texture-Based Segmentation of Road Images. In *Proceedings of the Intelligent Vehicles '94 Symposium, Paris, France*, Seite 260–265, 1994.
- [135] L. Zhao und C. Thorpe. Stereo- and Neural Network-Based Pedestrian Detection. In *ITSC'99, IEEE Conference on Intelligent Transportation Systems 1999*, Seite 298 – 303, Tokyo, Japan, 1999. IEEE.
- [136] Q. Zhuang. *Optimierung eines Fuzzy-Fahrreglers mit Hilfe der Evolutionsstrategie*. Nummer 804 in Fortschrittberichte, Reihe 8. VDI-Verlag, Düsseldorf, 1999. Dissertation, Ruhr-Universität Bochum.
- [137] Q. Zhuang, J. Gayko und M. Kreutz. Optimization of a Fuzzy Controller for a Driver Assistant System. In *Proceedings of the Fuzzy-Neuro Systems 98*, Seite 376 – 382, München, Germany, 1998.