



Internal Report 97-13

Technische Bildverarbeitung dynamischer Szenen

by

Werner von Seelen, Walter Gillner, Christian Goerick, Uwe Handmann,
Thomas Kalinke, David Kastrup, Iris Leefken, Gesa Lorenz, Detlev Noll,
Christos Tzomakas, Bernd Völpel, Martin Werner, Carsten Winkel

Ruhr-Universität Bochum
Institut für Neuroinformatik
44780 Bochum



IR-INI 97-13
July 1997
ISSN 0943-2752

Inhaltsverzeichnis

1	Computer Vision für Fahrerassistenzsysteme	3
1.1	Lokale Orientierungskodierung	6
1.2	Symmetrie	6
1.3	Flußfeldanalyse zur Segmentierung bewegter Objekte	7
1.4	Orts-Zeit-Analyse	9
1.5	Farbquantisierung zur Erweiterung der Merkmalbasis	10
1.6	Textur	12
1.6.1	Lokale Bildentropie	12
1.6.2	Kookkurrenz Matrizen	13
1.6.3	Fraktale Dimension	14
1.6.4	Texturbasierte Objektverfolgung - Kreuzentropie	14
1.7	Schätzung von Orts- und Skalierungsparametern	15
1.8	Die Hausdorff Distanz zur Objektverfolgung und -klassifikation	16
1.9	Parametrische Optimierung	18
1.10	Fusion zur initialen Segmentierung	19
1.10.1	Neuronale Kopplung	20
1.10.2	Segmentierung durch Fusion	20
1.11	LOC-Klassifikation	21
1.12	Ein modulares neuronales Netzwerk zur Klassifikation von Fahrzeugen	22
2	Erkennung von Sprengstoffzündern in Fluggepäck	23
3	Meeresumweltüberwachung	24
	Literaturverzeichnis	25

Technische Bildverarbeitung dynamischer Szenen

Werner von Seelen, Walter Gillner, Christian Goerick, Uwe Handmann,
Thomas Kalinke, David Kastrup, Iris Leefken, Gesa Lorenz, Detlev Noll,
Christos Tzomakas, Bernd Völpel, Martin Werner, Carsten Winkel

Zusammenfassung

Systeme zur automatisierten Bildanalyse sind vielfältig einsetzbar und gewinnen aufgrund technologischer Weiterentwicklungen und gesellschaftlicher Akzeptanz zunehmend an Bedeutung. Schwerpunkt im Bereich der „Technischen Bildverarbeitung dynamischer Szenen“ ist die Entwicklung von Methoden, die bei der Interpretation von Bildern aus verschiedenen Sensordaten Verwendung finden. Dies sind neben den herkömmlichen Kamerabildern im wesentlichen Röntgen- und Radarbilder. Unter geeigneter Berücksichtigung der durch die jeweiligen Anwendungen vorgegebenen Randbedingungen werden daraus entsprechende Verfahren abgeleitet.

Derzeitige Projekte beschäftigen sich mit der Analyse von Straßenverkehrsszenen (Abschnitt 1), der Detektion von Sprengstoffzündern bei der Durchleuchtung von Fluggepäck (Abschnitt 2) sowie mit der Bestimmung von Art und Ausdehnung von Ölverschmutzungen bei der Meeresüberwachung (Abschnitt 3).

1 Computer Vision für Fahrerassistenzsysteme

Autonom oder teilautonom geführte Fahrzeuge, insbesondere für den Einsatz im Straßenverkehr, implizieren wegen der durch die natürliche Umwelt aufgestellten Randbedingungen hohe Anforderungen an die zu entwickelnden Algorithmen. Am Institut für Neuroinformatik werden in Kooperation mit mehreren Partnern aus der Automobilindustrie Methoden entwickelt, die Beiträge zur Fahrzeugführung mittels Computer Vision zum Gegenstand haben.

Die grundsätzliche Problematik besteht darin, daß ein bewegter Beobachter in einer weitgehend natürlichen Umwelt und mit einer aufgabenabhängigen Wiederholfrequenz bewegte Objekte klassifiziert sowie ihre Positionen schätzt. Im speziellen wird versucht, Verkehrsteilnehmer aus dem Videobild zu extrahieren und die so gewonnenen Objekthypothesen weiter zu attribuieren (z.B. Objektklasse, Abstand, Geschwindigkeit, Gefahrenpotential hinsichtlich der beabsichtigten Eigentrajektorie etc.). Die laterale Querführung des Fahrzeugs und die damit verbundene Schätzung von Straßenbegrenzungen ist hingegen Schwerpunkt der Arbeit unserer Projektpartner.

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. Wegen der Anforderung an die Sicherheit der zu erreichenden Lösung, der Varianz der geometrischen Erscheinungsformen der involvierten Objekte und der Variation der Umweltbedingungen, die sowohl deterministischer als auch statistischer Art sind, werden zunächst eine Vielzahl von Teillösungen benötigt, die auf unterschiedlichen Repräsentationen der Umwelt beruhen. Das hat zur Folge, daß Komplexität und Struktur des Gesamtsystems offen sein müssen, um mit zunehmender Systemkomplexität neue Module aufnehmen zu können, ohne daß die erreichte Teillösung degeneriert. Somit sind bereits „einfache“ Anwendungen notwendigerweise mit Überlegungen bezüglich des Gesamtsystems belastet.

Im wesentlichen läßt sich das Gesamtsystem in Basis-Algorithmen und Algorithmen zur Fusion bzw. Integration teilen. Als Basis dienen Methoden, die spezielle Teillösungen bei gegebenen Randbedingungen (z.B. Umweltverhältnissen) liefern. Die Ergebnisse und Anwendungen der einzelnen Algorithmen sind nicht unabhängig, was eine Erhöhung der Redundanz zur Folge hat, die in geeigneter Weise gekoppelt ein Gesamtsystem sicher bzw. zuverlässig macht. Die hierbei einzusetzenden Methoden der Fusion und Integration sorgen sowohl für ein flexibles Zusammenspiel als auch für eine integrative Ergebnisbildung der Basiskomponenten. Ebenso ist in sequentieller Art und Weise ein Datentransfer und damit auch eine Systemdynamik nötig, um rückgekoppelt ein Gesamtsystem aufzubauen, das zur Lösung einer vielseitigen Aufgabe imstande ist.

Auf der Seite der Vorverarbeitung sind vor dem Hintergrund des aktuellen Problems eine Vielzahl unterschiedlicher Verfahren zur initialen Segmentierung, zur Objektverfolgung und zur Objektklassifikation entstanden. Hierbei sind einige inhärente Tendenzen bemerkenswert.

- Bisherige Arbeiten stützen sich oft auf die Verwendung höherer Merkmale im Sinne der Gewinnung einer Kette von Merkmalen beginnend auf der ikonischen (Bildpunkt bezogenen) Seite hin zur symbolischen Seite. Dies hat im wesentlichen zwei Gründe:
 1. Aufgrund der historisch zu sehenden starken Verwurzelung der Bildverarbeitung in der Inspektion von Werkstoffen und Oberflächen in der Qualitätskontrolle existierten hier bereits theoretisch fundierte und praktisch erprobte Algorithmen.
 2. Symbolische Merkmale dienen in der Regel zur kompakten Kodierung, so daß die Datenmenge stark reduziert werden kann und so den begrenzten Rechnerressourcen Rechnung getragen werden kann.

Durch die rasante Weiterentwicklung der Prozessoren hat sich gerade der letzte Punkt stark entzerrt. Zudem konnte festgestellt werden, daß insbesondere bei begrenzter Sensorauflösung (z.B. im Fernfeld) Algorithmen zum Einsatz kommen, die auf statistischen Eigenschaften größerer „früher“ (im Sinne der Verarbeitungskette) Merkmalmengen beruhen. Zudem ergänzen diese Algorithmen das Methodenspektrum der selbstverständlich ebenfalls weiter verwendeten, eher im klassischen Bereich anzusetzenden Algorithmen.

- Oft lassen sich durch eine Formulierung als Optimierungsproblem implizit robuste Problemlösungen finden, die die Nachteile expliziter Verfahren (z.B. Zuordnungen von Modell- zu Bildmerkmalen, Korrespondenzproblem) vermeiden. Auch hierbei hat der Anstieg an verfügbarer Rechenleistung die Entwicklungen begünstigt.
- Gerade beim Einsatz in natürlichen Umwelten werden bevorzugt flexible Algorithmen eingesetzt, die ein gewisses Maß an durch Eingangsdaten getriebener Lernfähigkeit besitzen.

Am Institut für Neuroinformatik sind eine Reihe von Algorithmen verfügbar, die einerseits die angesprochenen Teillösungen abdecken als auch andererseits auf die Architektur eines Gesamtsystems hinzielen:

- Fourier Mellin und Cepstrum Analyse,
- Auf binokularem Stereo basierende Verfahren,
- Invers perspektive Abbildung,
- Lokale Orientierungskodierung
- Verwendung lokaler Symmetrie zur initialen Segmentierung,
- Flußfeldanalyse zur Segmentierung bewegter Objekte,
- Orts-Zeit Analyse zur Segmentierung bewegter Objekte,
- Farbquantisierung zur Erweiterung der Merkmalbasis,
- Verwendung Nichtlinearer Dynamik zur lateralen und longitudinalen Regelung,
- Texturanalyse, wie die Berechnung lokaler Entropie oder Kookkurrenz zur initialen Segmentierung,
- Verwendung der Kreuzentropie zur Objektverfolgung,
- Verwendung von Ordnungsstatistiken zur Objektverfolgung,
- Parametrische Optimierung und lokale Adaption,
- Hausdorff Distanz zur Objektverfolgung und Klassifikation,
- Auf Kookkurrenzmatrizen basierende Maße zur Klassifikation,
- Fusion unterschiedlicher Repräsentationen und Sensoren,
- Neuronale Netzwerke für unterschiedliche Klassifikations- und Fusionsaufgaben.

Im folgenden werden einige ausgewählte am Institut für Neuroinformatik entwickelte Verfahren detaillierter vorgestellt.

1.1 Lokale Orientierungskodierung

Christian Goerick, Thomas Kalinke

Die lokale Orientierungskodierung (LOC) ist eine differenzielle Vorverarbeitungsmethode. Als Bildmerkmale werden binäre Codes erstellt, die eine Richtungskodierung der Grauwertvariation in einer Pixelumgebung repräsentieren. Der LOC-Operator ist definiert als

$$b'(n, m) = \sum_{i,j} k(i, j) \cdot u(b(n, m) - b(n + i, m + j) - t(i, j)), \quad (i, j) \in \text{Nachbarschaft},$$

wobei $b(n, m)$ die Grauwerte des Eingangsbildes, $b'(n, m)$ die Codes des Ausgangsbildes, $\mathbf{k}$ eine Koeffizientenmatrix, $\mathbf{t}$ eine Schwellwertmatrix und $u(z)$ die Einheitssprungfunktion darstellen. Das Ausgangsbild kodiert die Nachbarschaftsstruktur des Grauwertbildes. Für eine N_4 und eine N_8 Nachbarschaftsumgebung eines regelmäßigen karthesischen Gitters werden von uns folgende Koeffizientenmatrizen definiert:

$$\begin{bmatrix} 0 & 1 & 0 \\ 2 & R & 4 \\ 0 & 8 & 0 \end{bmatrix} \quad \begin{bmatrix} 1 & 2 & 4 \\ 8 & R & 16 \\ 32 & 64 & 128 \end{bmatrix} \quad \begin{matrix} n \\ \uparrow \searrow \\ m \end{matrix},$$

wobei R die Referenzposition ist. Für eine N_4 Umgebung erzeugt der LOC-Operator ein Ausgangsbild mit einer 4-Bit Kodierung. Folgende Eigenschaften implizieren die Wahl der Koeffizienten und die Formulierung des Operators:

- Eine Erhaltung der Information der lokalen Orientierung im Bild ist durch die eindeutige Separierbarkeit der Summe in ihre einzelnen Komponenten gewährleistet.
- Eine absolute Grauwertinvarianz ist dem Operator inhärent.
- Bei der Suche nach bestimmten Bildstrukturen müssen nur bestimmte Codes ausgewertet werden. Zur Repräsentation horizontaler Strukturen sind die Codes 1, 8 und 9 zu wählen.

Ein Adaptionsmechanismus für die Schwellwertmatrix $\mathbf{t}$ sorgt für ein hohes Maß an Flexibilität unter variierenden Beleuchtungsverhältnissen [Goe94] [GNW96].

1.2 Symmetrie

Thomas Kalinke

Ein signifikantes Merkmal des Prozesses der Objekterkennung ist die bilaterale Symmetrie [BG85]. Nicht nur in technischen Anwendungen, sondern auch in der Natur hat sich dieses Merkmal als ein Indikator für Nahrung und Feinde ausgebildet [SK92]. Pflanzen und Tiere zeigen oft eine symmetrische Struktur, die meist mit der Bewegungsrichtung übereinstimmt. Die Detektion von Symmetrie-Eigenschaften ist besonders bei der Jagd (Objektverfolgung) von großen Interesse. Diese Fähigkeit resultiert aus komplexen biologischen

Verarbeitungsschritten höherer Ordnung [BT90]. Hierbei ist die Erkennung der vertikalen Symmetrien die schnellste und am besten ausgebildetesteste [Roc84] [Roy81]. In der Computer Vision lassen sich die meisten Algorithmen in intensitäts- und kantenbasierte Methoden teilen [Ros86]. Ein Algorithmus, der die geraden und ungeraden Anteile der horizontalen Intensitätsverteilung der Grauwerte nutzt, wird in [Mar89] vorgestellt. Eine Kombination beider Methoden zeigt eine Arbeit von [ZBvS93]. Auch neuronale Netzwerke sind in der Lage eine Schätzung über die Art und Position der Symmetrieachsen zu geben [BS89] [Lis92] [JLS94].

Unser Interesse an der Symmetriedetektion bezieht sich auf Objekterkennung in Verkehrsszenen. Insbesondere Fahrzeugrück- und -frontansichten weisen eine starke Symmetrie auf. Unter dem Gesichtspunkt der Objektdetektion im Fernfeld hat man zusätzlich mit niedriger Ortsauflösung, Rauschen und Teilverdeckungen zu kämpfen. Wir setzen ein neuronales Netz ein, das hypothetische vertikale Achsen hinsichtlich ihrer Symmetrieintensität bewertet. Das Netz ist weiterhin in der Lage die laterale Genauigkeit der Achsenhypothese zu verbessern. Ein Verlauf der Asymmetrie und die in den Minima detektierten Symmetrieachsen sind in Abbildung 1 gezeigt. Hierbei erreicht der Objekt- zu Hintergrundanteil (*signal-to-noise ratio*) in den Objekthypothesen ungefähr eins.

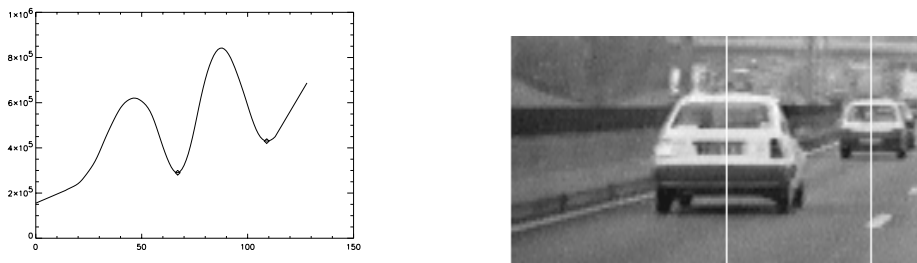


Abbildung 1: Verlauf der Größe der Asymmetrie, aufgetragen über die horizontalen Achsenhypothesen. Detektierte Minima ergeben maximale Symmetrie im Bild.

1.3 Flußfeldanalyse zur Segmentierung bewegter Objekte

Gesa Lorenz, Martin Werner, Walter Gillner

Überholende bzw. einscherende Fahrzeuge auf Autobahnen oder Landstraßen bergen ein erhöhtes Gefahrenpotential für die beteiligten Verkehrsteilnehmer. Ihre Detektion bildet daher einen wesentlichen Bestandteil bei der Realisierung einer computergestützten teilautonomen Fahrzeugführung.

Das dazu entwickelte Verfahren nutzt einen Algorithmus zur Bewegungssegmentierung für eine schnelle Generierung von Objekthypothesen. Diese werden wiederum durch weitere im Gesamtsystem integrierte Klassifikations- und Trackingverfahren gestützt oder auch verworfen. Die Suche nach initialen Regionen wird auf den rechten und linken Bildrandbereich beschränkt, in dem die interessierenden Objekte zu erwarten sind. Dadurch kann die Eigen- und Objektbewegung in guter Näherung als rein translatorisch angenommen werden, was

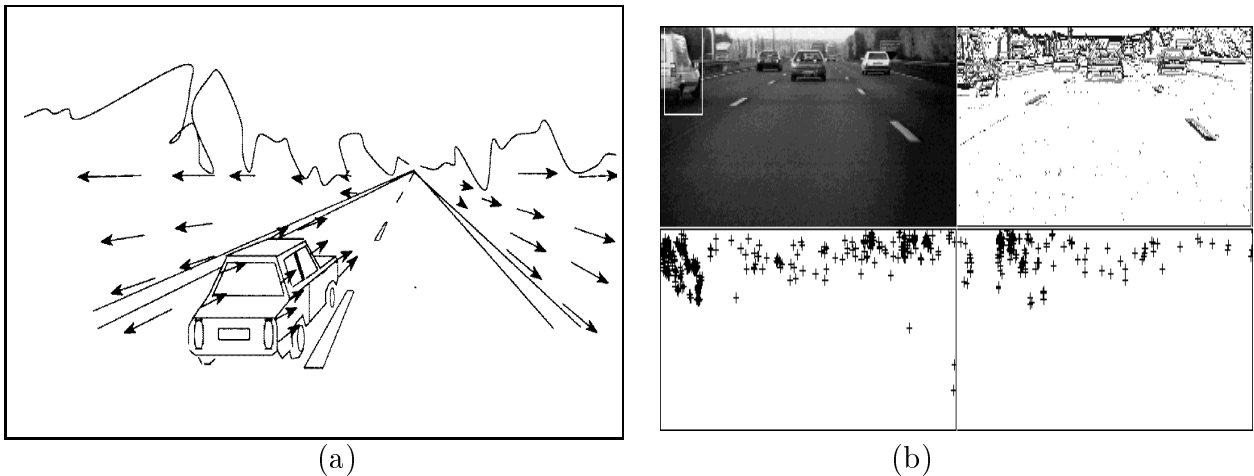


Abbildung 2: (a): Flußfeld. (b): von links oben nach rechts unten: Grauwertbild, LOC-Bild, Flußfeldanteile mit Komponenten nach rechts, Flußfeldanteile mit Komponenten nach links

sich später bei der Analyse der Bewegungskomponenten als vorteilhaft erweist. Die Größe der Suchregionen wird an die Kamerageometrie und den Fahrspurverlauf angepaßt.

Die Bewegungssegmentierung wird über die Berechnung und anschließende Analyse des optischen Flusses durchgeführt. Um kurze Rechenzeiten zu erhalten, haben wir einen diskreten merkmalsbasierten Ansatz gewählt. Zum Aufbau des Flußfeldes werden dabei geeignete Bildmerkmale extrahiert und Korrespondenzen in aufeinanderfolgenden Bildern bestimmt.

In Abbildung 2 (a) ist das Flußfeld, wie es bei einem Überholvorgang auftreten würde, skizziert. Aufgrund der Eigenbewegung des Beobachters entstehen Flußfeldkomponenten, die radial vom Fluchtpunkt ausgehen und daher in den Bildrandbereichen unterschiedliche horizontale Richtungen aufweisen. Eine unabhängige Objektbewegung erzeugt dagegen ein abweichendes Flußfeld, welches zur Objektdetektion verwendet werden kann. Im hier dargestellten Fall eines überholenden Fahrzeugs entstehen Flußfeldkomponenten mit einem dem eigenbewegungsinduzierten Flußfeld in diesem Bildbereich genau entgegengerichteten Anteil.

Als Vorverarbeitung wird die am Institut entwickelte *lokale Orientierungskodierung* (LOC) verwendet (Abschnitt 1.1, [Goe94] [GNW96]). Als Resultat dieser Vorverarbeitung liegen Merkmale in der Form von Bit-Strings vor, welche die Richtung von Grauwertdifferenzen in einer 4er Nachbarschaft kodieren. Zur Berechnung des Flußfeldes wird versucht, in zwei aufeinanderfolgenden Bildern Korrespondenzen für die derartig extrahierten Merkmale zu finden. Dabei wird in der 8er Nachbarschaft des Referenzpixels nach einem eindeutigen perfekten Match gesucht. Zur Steigerung der Auflösung arbeiten wir auf drei Pyramidenstufen, so daß auch größere Verschiebungen und damit größere Fahrzeuggeschwindigkeiten erfaßt werden.

In Abbildung 2 (b) sind das Grauwertbild, das LOC-Bild sowie das daraus berechnete Flußfeld dargestellt, wobei Flußfeldanteile entsprechend ihrer horizontalen Ausrichtung in zwei unterschiedliche Bilder eingetragen wurden. Das linke Bild enthält darin die Flußfeldanteile, die eine Komponente nach rechts aufweisen, das rechte dagegen diejenigen, die eine

Komponente nach links aufweisen. In dem linken Bild ist deutlich zu erkennen, daß das überholende Fahrzeug ein Flußfeld erzeugt, dessen Richtung sich von der eigenbewegungs-induzierten Richtung absetzt. Dieses Fahrzeug kann mit Hilfe von Schwellwertbildungen auf Flußfeldhistogrammen detektiert werden, die entlang der Ortskoordinaten gebildet werden. Das Ergebnis ist durch den im Originalbild eingezeichneten Rahmen angedeutet. Eine zeitliche Prädiktion führt zu einer weiteren Stabilisierung des Verfahrens.

1.4 Orts-Zeit-Analyse

Iris Leefken, Walter Gillner

Zur Segmentierung von bewegten Objekten in Bildern, bei deren Aufnahme sich der Beobachter bewegt, ist es notwendig, zwischen den Effekten der Eigen- und der Objektbewegung zu unterscheiden.

Unter Ausnutzung der zeitlichen Information von Bildsequenzen wurde nach der Entwicklung eines Verfahrens zur Bestimmung des optischen Flusses (Abschnitt 1.3) ein weiteres Verfahren zur Analyse von Bildsequenzen ausgearbeitet. Dieses Verfahren basiert auf einer Orts-Zeit-Betrachtung [GS97].

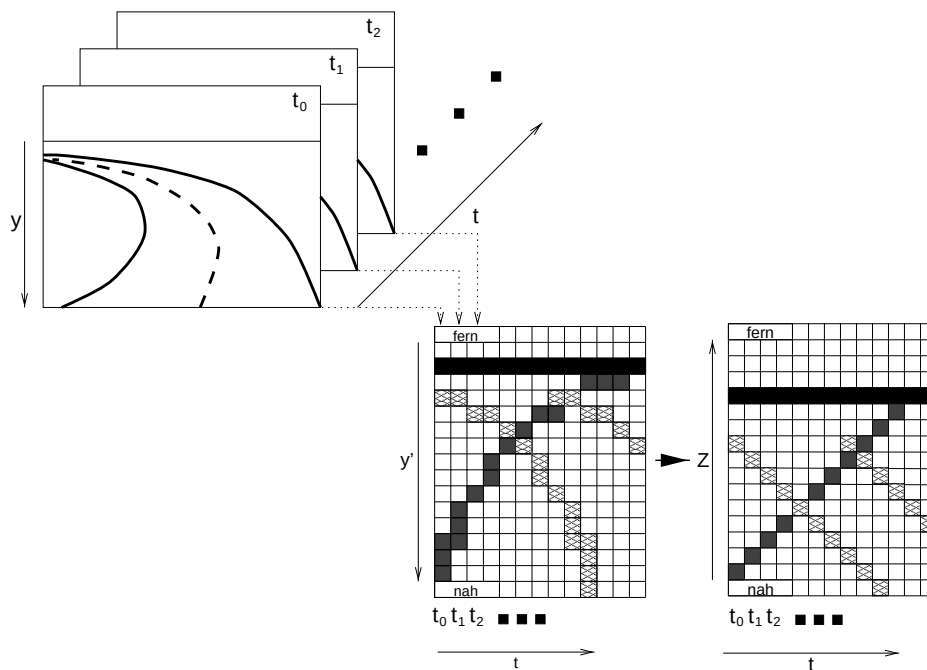


Abbildung 3: Abtastung von Grauwerten unterhalb einer Trajektorie (gestrichelte Linie). Die Ortskoordinate entlang der Trajektorie wird auf die vertikale Achse des Orts-Zeit-Bildes projiziert. Die horizontale Achse entspricht der Zeitachse der Bildfolge. Durch Rückprojektion kann die Bemaßung der Achse für die Höhe im Bild (y -Achse) auf die Entfernung in der Fahrebene (z -Koordinate) abgebildet werden.

Hierbei wird eine Betrachtung des zeitlichen Verlaufs von aus den Bildern abgetasteten Grauwerten vorgenommen. Die Abtastung erfolgt über Trajektorien, die aufgrund der prädiagnostizierten Eigenbewegungsrichtung ermittelt werden. Durch die Auswertung der entstehenden Ort-Zeit-Bilder kann die Geschwindigkeit bestimmt werden, mit der sich Objekte relativ zum Beobachter entlang der Trajektorien bewegen (Abbildung 3). Anhand der ermittelten Daten können Regionen mit unterschiedlichen Objektgeschwindigkeiten bestimmt werden. Ist die Eigengeschwindigkeit bekannt, so kann die statische Umwelt aus der Darstellung entfernt werden. Die verbleibenden Regionen können als bewegliche Objekte identifiziert und mit einer Relativgeschwindigkeit attribuiert werden (Abbildung 4).

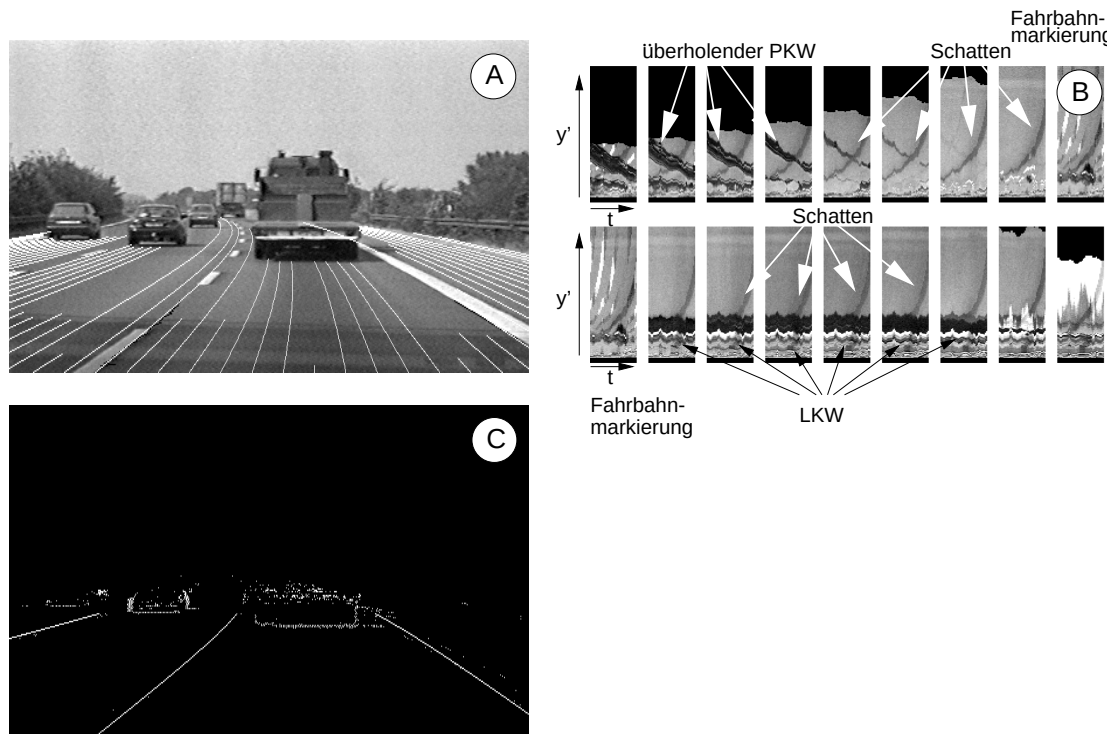


Abbildung 4: Segmentierungsvorgang einer Autobahnszene. Das Originalbild mit den betrachteten Trajektorien ist in Bild A dargestellt. Bild B zeigt die Orts-Zeit-Diagramme für die Trajektorien und Bild C das korrespondierende Binärbild nach der Elimination des Eigengeschwindigkeitsanteils. Zur Orientierung sind die Fahrspuren eingezeichnet.

1.5 Farbquantisierung zur Erweiterung der Merkmalbasis

Gesa Lorenz

Die Verwendung von Farben stellt für Bildanalyseaufgaben eine wertvolle Ergänzung zur Grauwertinformation dar. Dabei ist insbesondere der Farbkontrast die wesentliche informationstragende Größe, die auch vom Menschen beim Sehprozeß vornehmlich ausgewertet wird. Durch die Nutzung von Farbinformationen kann ein deutlicher Gewinn an Stabilität

und Robustheit erzielt werden.

Die Integration von Farbinformation in das unserem Gesamtsystem zugrundeliegende Merkmalspektrum bietet verschiedene Vorteile. Zum einen ermöglicht Farbe als Oberflächeneigenschaft von Objekten den Übergang von kantenbasierten zu flächenbasierten Objektbeschreibungen. Zum anderen beeinflusst die Farbinformation selbst die Aufmerksamkeit des Betrachters, insbesondere in den hier analysierten Straßenverkehrsszenen, und kann daher auch für eine maschinelle Aufmerksamkeitssteuerung genutzt werden.

Zur Generierung der Farbflächen führen wir eine Farbquantisierung mit dem *neural gas* Verfahren [MBS93] durch. Dabei handelt es sich um ein selbstorganisierendes neuronales Netz, das eine wählbare Zahl an Farbrepräsentanten datengetrieben derart im Farbraum anordnet, daß sie die Verteilungsdichte der Farbwerte aller Bildpixel approximieren. Durch Zuordnung der den ursprünglichen Farbwerten am nächsten liegenden Farbrepräsentanten entstehen Bildregionen mit homogenen Farbwerten, d.h. Farbflächen. Das hier verwendete *neural gas* Verfahren zeichnet sich gegenüber anderen neuronalen Netzen zur Vektorquantisierung durch eine hohe Konvergenzgeschwindigkeit und einen geringen Rekonstruktionsfehler aus. Er ist zudem nach einmaliger Initialisierung der Farbrepräsentanten mit relativ geringem Aufwand an sich ändernde Bildstatistiken adaptierbar, so daß die Verarbeitung von Bildsequenzen möglich ist. Das Ergebnis einer Farbquantisierung auf 16 Farben ist in Abbildung 5 dargestellt. Man erkennt insbesondere an den Fahrzeugen und im Himmel die Ausbildung von Flächen homogener Farbe.

Die hier beschriebene Flächenkodierung wird derzeit von zwei am Institut entwickelten Verfahren zur Objektverfolgung genutzt, die die Kreuzentropie bzw. Hausdorff-Distanz zum Objektmatching nutzen (Abschnitte 1.6.4 u. 1.8). Diese Maße arbeiten auf der Farbverteilungsdichte bzw. auf aus den Farbflächen generierten Kanten. Es zeigt sich, daß ein Stabilitätsgewinn gegenüber grauwertbasierten Verfahren erzielt werden kann. Gegenstand aktueller Arbeiten ist eine weitere Verbesserung des Farbquantisierungsergebnisses bezogen auf die Homogenität der Farbflächen. Dazu muß neben der reinen Farbinformation auch Ortsinformation in das Quantisierungsverfahren eingebracht werden.



Abbildung 5: Links: Originalbild, Rechts: Farbquantisiertes Bild

1.6 Textur

Thomas Kalinke

In der Bildanalyse werden neben den Methoden, die auf differenziellen Operatoren (wie Grauwertdifferentiation und der LOC aus 1.1) basieren, integrierende Verfahren wie z.B. Texturanalysen eingesetzt. Für den Begriff der Textur wird in der Literatur keine genaue Definition gegeben. Allgemeingültig kann man sagen, daß Textur als eine Beschreibung von Pixeln eines Bildes bezeichnet wird, die, wenn sie gruppiert oder auch für sich selbst in einem bestimmten räumlichen und farblichen Verhältnis zu anderen Gruppen oder Pixeln stehen, einer bestimmten Texturklasse angehören. Textur ist also per Definition skalierungsvariant. Das räumliche und farbliche Verhältnis der Elemente zueinander gibt an, von welcher Art die Textur ist. Starke Farbschwankungen auf engem Raum ergeben feine Texturen und eine größere Anzahl größerer Flächen gleicher Farbe ergeben grobe Texturen. Desweiteren kann man leichte und starke Texturen unterscheiden. Leichte Texturen, bei welchen sich keine genaue Struktur definieren läßt, lassen sich durch statistische Maße beschreiben. Starke Texturen hingegen werden durch die genaue Angabe (Grammatik) der Texturelemente und deren räumlichen Abstand zueinander bestimmt. Zum heutigen Zeitpunkt werden in der Wissenschaft vornehmlich zwei Arten der Texturbeschreibung unterschieden: statistische und syntaktische. In unserem Anwendungsfeld der Objekterkennung in Verkehrsszenen beschäftigen wir uns im Schwerpunkt mit den statistischen Verfahren. Hierbei wird die Textur so beschrieben, daß eine statistische Mustererkennung erfolgen kann. Ziel ist es, jede Textur durch einen Merkmalvektor in einem mehrdimensionalen Vektorraum zu repräsentieren, um durch geeignete Entscheidungsregeln die Textur einer bestimmten Klasse zuzuordnen. Die Erstellung eines Merkmalvektors kann auf unterschiedlichen Methoden basieren.

1.6.1 Lokale Bildentropie

Zur Texturanalyse kann die lokale Bildentropie [KvS96] genutzt werden. Hierbei wird der Bildinformationsgehalt des Grauwertsignals in einer lokalen Nachbarschaft jedes Pixels bestimmt. Die Realisierung einer Aufmerksamkeitskarte auf Basis dieses Maßes läßt eine Teilung des Bildes in Objekt- und Hintergrundbereich zu. Abbildung 6 zeigt Ergebnisse einer initialen Segmentierung durch die lokale Bildentropie. Hierbei sind zum einen die Objektbereiche und zum anderen der Hintergrundbereich durch eine Detektionsschwelle ausgeblendet.

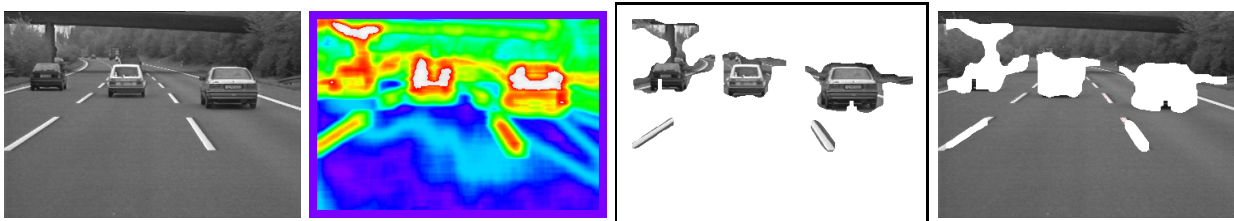


Abbildung 6: Originalbild, lokale Bildentropie, segmentierte Objekt- und Hintergrundbereiche

1.6.2 Kookkurrenz Matrizen

Eines der elementaren Werkzeuge der Texturbeschreibung steht mit den von Haralick [HSD73] vorgeschlagenen Kookkurrenz-Matrizen zur Verfügung. Hierbei wird die Häufigkeit des Wiederauftretens von Pixelpaaren unter gegebenen Randbedingungen bewertet. Diese Randbedingungen werden durch das Farbwertverhältnis und den räumlichen Abstand in Winkel und Betrag parametrisiert. Die Kookkurrenz-Matrix ist wie folgt definiert: In einem Bildausschnitt $\mathbf{I}$ der Größe $M \times N$ und einer maximalen Anzahl verschiedener Grauwerte Q wird die Kookkurrenz-Matrix $\mathbf{P}$ für eine vorgegebene Richtung α in einem vorgegebenen Abstand d wie folgt bestimmt:

$$P_{d,\alpha}(i, j) = \frac{\text{Anzahl der Paare } (x, y), (x', y'), \text{ die } (d, \alpha) \text{ genügen und } I(x, y) = i \text{ und } I(x', y') = j}{\text{Anzahl aller Paare des Bildausschnitts } (x, y), (x', y')}$$

Für die sinnvolle Berechnung von Texturmerkmalen werden die Kookkurrenz-Matrizen in den meisten Anwendungen für alle vier möglichen Richtungen ($\alpha = 0, 45, 90$ und 135) und verschiedene Distanzen $d = 1, 2, \dots$ pro Bild bestimmt. Hierbei sorgt eine Summierung aller Matrizen einer Distanz $d = \text{const}$ für eine Rotationsinvarianz. Eine Berechnung über verschiedene Abstände d schränkt den Raum der Skalierungsvarianz ein.

Julesz erbrachte den Beweis, daß das menschliche visuelle Wahrnehmungssystem Texturen auf der Basis von Kookkurrenz-Statistiken auswertet. Haralick, Shanmugan und Dinstein schlugen in [HSD73] insgesamt 14 verschiedene statistische Merkmale vor, die sich aus den

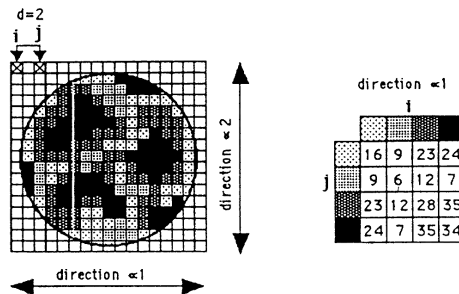


Abbildung 7: Bildausschnitt mit 4 verschiedenen Grauwerten und die dazugehörige Kookkurrenz-Matrix für $\alpha = 0$ (horizontale Richtung) und $d = 2$.

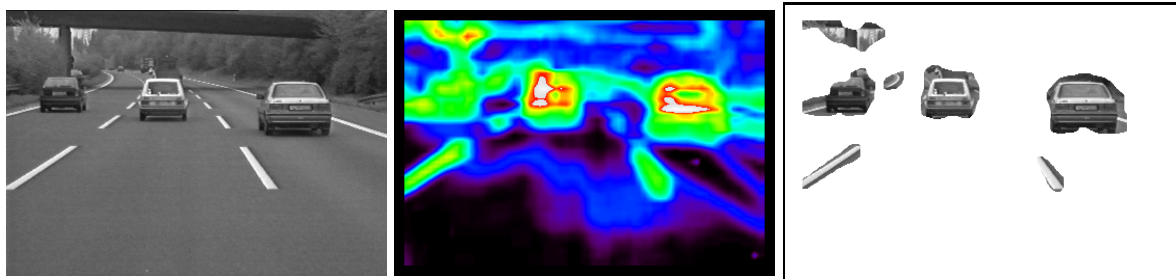


Abbildung 8: Grauwertbild, Kookkurrenz basierte Aufmerksamkeitskarte, segmentierte Objektbereiche

Matrizelementen der Kookkurrenz-Matrizen berechnen lassen.

In unserem Anwendungsfeld werden die Kookkurrenz-Matrizen vornehmlich zur initialen Segmentierung erfolgreich eingesetzt (Abbildung 8). In einander überlappenden Teilbildbereichen werden die Kookkurrenz-Matrizen berechnet und objektoptimierte Merkmale wie Energie, Entropie, Kontrast, Korrelation und größte Kookkurrenz aus [HSD73] in Kombination zur Segmentierung genutzt.

1.6.3 Fraktale Dimension

In einem anderen Ansatz läßt sich eine Texturbeschreibung mit Hilfe der fraktalen Dimension verwirklichen. Pentland [Pen84] hat eine Beziehung zwischen der fraktalen Dimension und der Texturstruktur hergestellt und Keller [KCC89] hat eine Klasse von Texturmaßen basierend auf der fraktalen Dimension und Lakunen eingeführt. Yahiaoui [YdSB93] hat eine Fahrbahnerkennung auf der Basis der „Hausdorff-Besicovitch“-Dimension [PNHA84] und einem neuronalen Netz vorgestellt. Unser Ziel ist es, eine Segmentierung des Bildes auf Basis dieses Maßes durchzuführen. Abbildung 9 zeigt ein Segmentierungsergebnis für eine typische Verkehrsszene.

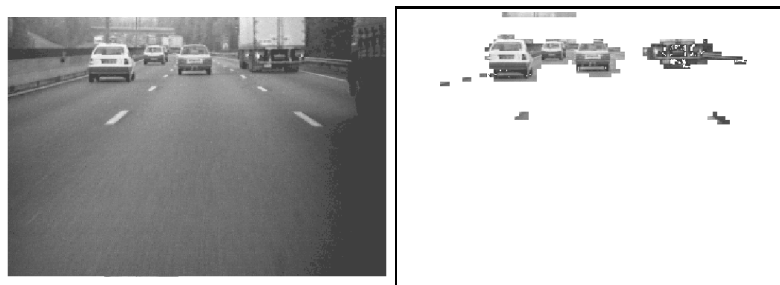


Abbildung 9: Grauwertbild, Segmentierung auf Basis der lokalen fraktalen Dimensionen

1.6.4 Texturbasierte Objektverfolgung - Kreuzentropie

Eine der einfachsten Beschreibungen der Textur ist die der Grauerthistogramme (Statistiken erster Ordnung). Diese lassen sich erfolgreich zur Objektverfolgung einsetzen. Besonders nicht-rigide Objekte wie Fußgänger und Zweiradfahrer, die über weitere rotatorische Freiheitsgrade gegenüber anderen Verkehrsteilnehmern verfügen, lassen sich mit Hilfe der Kreuzentropie verfolgen.

Wie in [KvS97] beschrieben, bestimmt die Kreuzentropie (Kullback-Leibler Divergenz) ein Ähnlichkeitsmaß zwischen zwei Verteilungsdichten. In unserem Anwendungsbereich werden Vergleiche eines Modells zum Zeitpunkt $(t - 1)$ mit Objekthypothesen zum Zeitpunkt t durchgeführt. In Abbildungen 10 und 11 ist eine Objektverfolgung über eine Bildsequenz basierend auf Grauwert- und Gradienten-Verteilungsdichten gezeigt.



Abbildung 10: Tracking eines Fußgängers im Mittelfeld basierend auf Grauwertverteilungen.



Abbildung 11: Tracking von Fußgängern im Nahfeld auf Basis des LOC.

1.7 Schätzung von Orts- und Skalierungsparametern

Martin Werner

Oft werden Algorithmen benötigt, die eine Objektverfolgung noch vor der eigentlichen Klassifikation ermöglichen, so daß modellbasierte Verfahren nicht anwendbar sind. Daher verwenden wir einen modellfreien Ansatz zur Objektverfolgung, der auf der Schätzung der Lage- und Skalierungsparameter einer Merkmalverteilung durch Ausnutzung asymptotischer Eigenschaften geordneter Statistiken beruht. Hierbei werden insbesondere die Verteilungen lokalisierter Merkmale wie z.B. Konturpunkte betrachtet.

Die Idee, die diesem Ansatz zugrunde liegt, ist die Schätzung der Lage- und Skalierungsparameter auf Basis der inversen Verteilungsfunktion (icdf). Die Koordinaten der Merkmale werden hierbei als eine Realisierung einer zweidimensionalen Zufallsvariablen interpretiert. Um geordnete Statistiken einsetzen zu können, beschränken wir uns hier auf die Marginalverteilung in x -Richtung. Die Ergebnisse lassen sich jedoch leicht auf jede beliebige Projektion verallgemeinern. Es kann gezeigt werden, daß in einem Intervall der Zusammenhang der icdf

der Merkmale F_n^{-1} zum Zeitpunkt t_n mit der icdf F_0^{-1} zum Zeitpunkt t_0 durch

$$F_n^{-1}(x) = \theta_1 + \theta_2 F_0^{-1}(x)(c'_1 x + c'_2),$$

beschrieben werden kann. Hierbei sind θ_1 und θ_2 die Lage- und Skalierungsparameter, die Parameter c'_1 und c'_2 hängen vom Hintergrund ab. Die Schätzung der icdf basiert auf asymptotischen Eigenschaften der geordneten Statistiken der Merkmale [ABN92][Dav80], die Parameterschätzung auf der Optimierung einer quadratischen Fehlerfunktion.

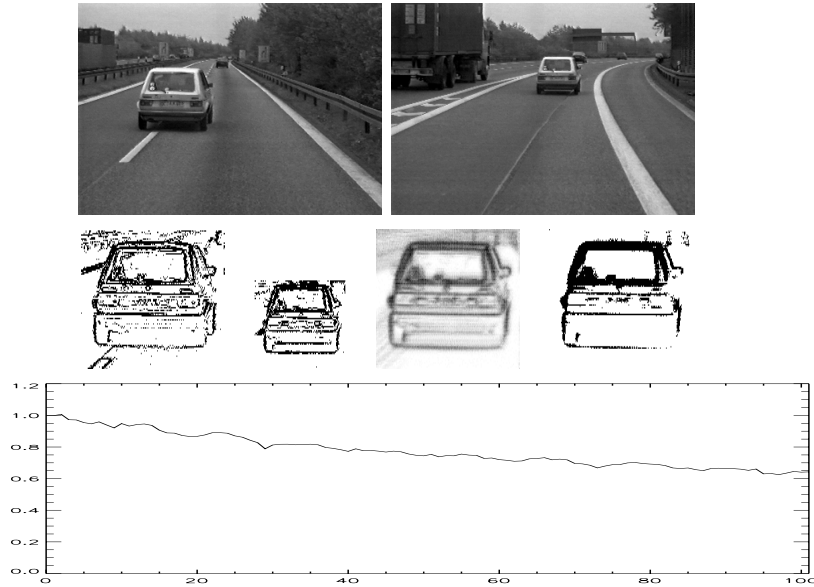


Abbildung 12: Bild 1 einer Objektverfolgungssequenz, Bild 100 der Sequenz ($\Delta t = 50ms$), Merkmale des Objektes in Bild 1, Merkmale des Objekts in Bild 100, Ergebnis der Akkumulation der reskalierten retranslatierten Merkmale, ein Binärbild, das durch Schwellwertbildung der Akkumulation entsteht, Ergebnis der Skalierungsschätzung.

1.8 Die Hausdorff Distanz zur Objektverfolgung und -klassifikation

Christos Tzomakas, Martin Werner, Thomas Kalinke

Der geometrische Vergleich von Formen ist für die modellbasierte Erkennung von Objekten fundamental. Die meisten Methoden in Objekterkennungssystemen beziehen sich auf eine Distanzmessung [Pag91] zwischen den Modellmerkmalen und den Bildmerkmalen. Die *Hausdorff Distanz* [Hut93] mißt die Abweichung einer Menge von Merkmalen gegenüber einer Referenzmenge von Merkmalen, die im betrachteten Anwendungsfall in der Regel Objektkonturen beschreiben. Der Vergleich ähnlicher Objektkonturen führt hierbei zu kleinen Distanzwerten, Objekte unterschiedlicher Form ergeben größere Distanzen.

Die gerichtete Hausdorff Distanz h einer Menge A zu einer Menge B wird als das Maximum über alle minimalen Distanzen der Punkte der Menge A zu den Punkten der Menge B

definiert. Die gesamte Hausdorff Distanz H der zwei Mengen A und B ist dann das Maximum der zwei gerichteten Distanzen.

$$\begin{aligned} h(A, B) &= \max_{p \in A} \min_{q \in B} \|p - q\| \\ H(A, B) &= \max(h(A, B), h(B, A)) \end{aligned}$$

Eine Modifikation davon ist die partielle Hausdorff Distanz, bei der alle minimalen Distanzen ihrer Größe nach geordnet werden und nur ein prozentueller Anteil anstelle des Maximums berücksichtigt wird.

Im Gegenteil zu den klassischen Korrelationsverfahren benutzt die Hausdorff Distanz Min-Max Operationen anstelle von Multiplikationen und ist daher effizienter. Die partielle Hausdorff Distanz ist robust gegenüber Ausreißern, die durch Rauschen, mangelhafte Merkmal-extraktion und im Fall von teilverdeckten Objekten entstehen können.

Neben der Eignung der Hausdorff Distanz zur Objektdetektion, kann die geometrische Eigenschaft der Hausdorff Distanz als Basis zur Klassifikation von Verkehrsteilnehmern in komplexen Szenen eingesetzt werden. Aufgrund der dazu notwendigen Definition einer Referenzkontur handelt es sich hierbei um ein modellbasiertes Verfahren.

Zunächst wurde die Hausdorff Distanz zur Klassifikation von PKWs und LKWs erprobt. Anhand der unterschiedlichen Form der Rückansicht eines PKWs bzw. eines LKWs kann man erwarten, daß nach Erzeugung generischer Modelle für jede Klasse die Objekte beider Klassen trennbar sind.

Bei dieser Klassifikation wird jede vorgegebene Region mit einem PKW- und einem LKW-Modell verglichen. Hierbei wird zur Merkmalextraktion die lokale Orientierungskodierung (Abschnitt 1.1, [Goe94] [GNW96]) verwendet. Da die Mächtigkeit der Menge der Modellmerkmale, die horizontale Strukturen beschreiben, in der Regel weitaus größer ist als die Mächtigkeit der Menge, die zugehörige vertikale Strukturen beschreibt, werden diese beiden Teilmengen separat bewertet.

Die Hausdorff Distanz wird für jedes Modell unter Beachtung der im vorgegebenen Bildausschnitt erlaubten Translationen und einem zu definierenden Bereich von Skalierungen



Abbildung 13: Jede Region wird mit zwei Modellen verglichen

berechnet. Die Mächtigkeiten der Teilmengen von Modell- und Bildmerkmalen, die eine festzulegende Distanzschwelle nicht überschreiten, gelten für das jeweilige Modell als Kriterien für die Klassifikation. Diese Werte bilden einen Merkmalvektor der mittels eines Multilayer Perzeptron (MLP, [Hay94], Seite 138ff) Netzes klassifiziert wird.

1.9 Parametrische Optimierung

Uwe Handmann, Detlev Noll, Martin Werner

In diesem Abschnitt wird ein merkmalsbasierter Ansatz zur Objekterkennung und parametrischen Optimierung beschrieben, welcher sich an den Optimierungsansatz des elastischen Netzes [DW87] anlehnt. Das Korrespondenzproblem wird hierbei mit der Optimierung einer Energiefunktion gelöst, wobei das Verfahren des *deterministic annealing* angewendet wird. Dadurch wird iterativ das globale Optimum gefunden.

Bei Verfahren des deterministic annealing beeinflusst der Temperaturkoeffizient die Kurve der Energiefunktion. Am Anfang der Optimierung ist die Form der Kurve konvex und das globale Optimum kann schnell gefunden werden. Bei allmählicher Reduzierung des Temperaturkoeffizienten entstehen zunehmend lokale Optima und das globale Optimum ändert seine Lage. Mit Hilfe eines lokalen Optimierungsansatzes kann das globale Optimum jedoch verfolgt werden. Die Analyse dieses Ansatzes zeigt [Nol96], daß sich die im folgenden definierte Funktion F ,

$$F = -\frac{1}{\beta} \sum_j \ln \sum_i s_{ij} e^{-\beta d_{ij}^2} + \frac{\alpha}{\beta} \sum_j |\hat{y}_j - \vec{y}_j|^2,$$

unter Variation des Temperaturkoeffizienten β wie oben gefordert verhält. Die Parameter s_{ij} beschreiben die Ähnlichkeit der Merkmale, d_{ij} beschreibt die Distanz zwischen den zu vergleichenden Merkmalen und α ist ein Gewichtungsfaktor. Ein Maß für die lokale Deformation ist definiert durch $E(\{\vec{y}_j\}) = \sum_j |\vec{y}_j - \hat{y}_j|^2$, wobei $\{\vec{y}_j\}$ die Positionen der Modellmerkmale beschreibt und $\{\hat{y}_j\}$ das Referenzmodell ist.

Ein erstes Beispiel des Verhaltens des Algorithmus wird in Abbildung 14 gezeigt. Linienelemente sind hier die Basismerkmale. Ein Flaschenmodell (dicke Linien) wird in dem gegebenen Bild gesucht. Das Modell darf in der Ebene skaliert, versetzt und rotiert werden. Oben wird die Position des Modells bei kleinen und großen Werten von β gezeigt. Zusätzlich sind Schnitte durch die Energiefunktion F dargestellt. Unten werden Ergebnisse globaler (links) und lokaler (rechts) Optimierung verdeutlicht. Eine Erweiterung des Algorithmus auf 3-dimensionale Objekte ist einfach zu realisieren [Nol96].

In dem in Abbildung 14 vorgestellten Beispiel wurde β exponentiell erhöht. Die initiale Position des Modells ist abhängig von a priori Wissen. Die initiale Größe von β kann entsprechend der Aufgabenstellung, beispielsweise bei Objektverfolgung, angepaßt werden. Die Abbruchbedingung muß entsprechend der Genauigkeitsanforderungen gesetzt werden. Wenn die Funktionalität dieses Ansatzes unabhängig von der entsprechenden Wahl des lokalen Optimierungsalgorithmus ist, ist die Konvergenzgeschwindigkeit direkt abhängig von der obigen Auswahl. Wird beispielsweise ein EM-Algorithmus (*expectation-maximization*) für die lokale

Optimierung eingesetzt [Nol96, YSU94], benötigt die Implementierung auf einer SUN Sparc 20 typischerweise 50-100ms, um Rückansichten von Fahrzeugen in Straßenverkehrsszenen (Abbildung 15) zu finden und zu verfolgen.

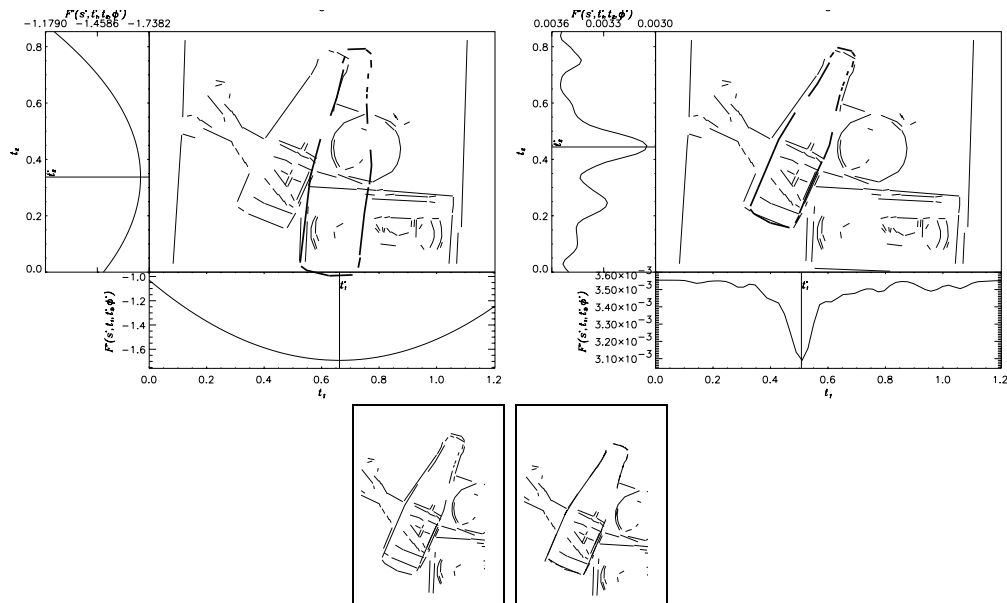


Abbildung 14: Oben: Das Modell (dick) mit den Bildmerkmalen unterlegt, für $\beta = 1$ (links) und für $\beta = 1540$ (rechts). Zusätzlich sind Schnitte durch F dargestellt. Unten: Das Modell nach globaler Optimierung (links) und nach lokaler Optimierung (rechts).



Abbildung 15: Objektverfolgung: Sequenz (dargestellt ist jedes zehnte Bild) mit einem Modell der Rückansicht eines PKW (weiß)

1.10 Fusion zur initialen Segmentierung

Uwe Handmann

Ziel der Fusion in der Bildverarbeitung ist die Verbesserung von Speziallösungen und Einzelmethode durch Kopplung im Hinblick auf ihren Einsatzbereich. Insbesondere wird die modulare Kopplung einzelner Verarbeitungsschritte durch die Erzeugung der für die Objekterkennung wichtigen Redundanz motiviert. Durch diese Vorgehensweise soll einerseits eine Erhöhung der Flexibilität und Robustheit der Bildverarbeitungsmodule erreicht und

andererseits eine höhere Adaptivität der Module bezüglich der Problemstellungen erzielt werden.

1.10.1 Neuronale Kopplung

Ein Ansatz zur robusten und fehlertoleranten Segmentierung wird in Abbildung 16 verdeutlicht. Es koppeln Bildverarbeitungsmodule, hier *Linienbild* (Polygonzugapproximation der Kontur), *Entropie* [KvS96] und die *Lokale Orientierungskodierung* (LOC, [Goe94] [GNW96]), in ein neuronales Netz ein (Abbildung 16). Eine Rückkopplung über der Zeit ist ebenso möglich, wie die Integration von zusätzlichen Sensorinformationen.

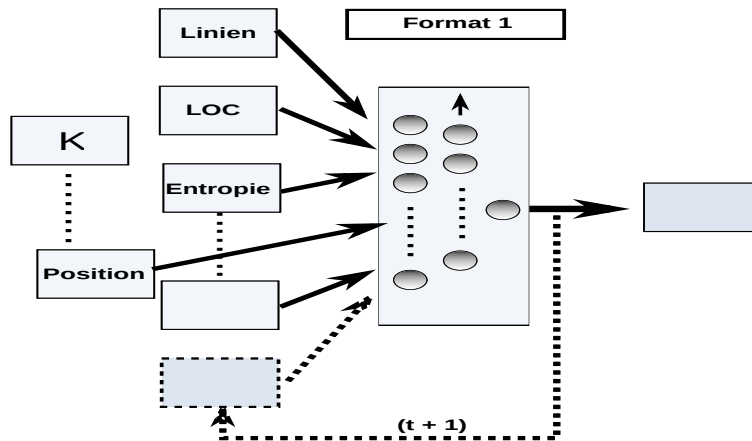


Abbildung 16: Kopplungsmodell

1.10.2 Segmentierung durch Fusion

Basierend auf den integrativen (Entropie) und differentiellen (LOC, Linien) Repräsentationen wird für jedes Pixel ein 12-dimensionaler Eingangsvektor

$$\mathbf{u}(x, y)^T = (\mathbf{u}_1(x, y)^T, x, y)^T$$

für das Kopplungsnetz erzeugt. Der 10-dimensionale Vektor $\mathbf{u}_1(x, y)$ beinhaltet die Kontur- und Texturinformationen, die Variablen x und y repräsentieren die Pixelkoordinaten. Vektor $\mathbf{u}_1(x, y)$ ist definiert durch

$$\mathbf{u}_1(x, y) = \sum_{(i,j) \in R} \mathbf{v}(i, j),$$

wobei R die lokale Nachbarschaft (z.B. 9×9) von (x, y) und $\mathbf{v}(i, j)$ einen binären Vektor beschreibt. $(v_1(x, y), \dots, v_4(x, y))^T$ kodieren reduzierte LOC-Merkmale, $(v_5(x, y), \dots, v_9(x, y))^T$ kodieren die Entropiewerte und $v_{10}(x, y)$ wird gesetzt, wenn (x, y) Teil eines Linienelements ist.

Als Kopplungsstruktur wird ein *Multilayer Perzeptron* (MLP, [Hay94], Seite 138ff) eingesetzt. Es findet eine 12-5-1-Struktur der Neuronen Verwendung. Die zwölfdimensionalen Merkmalsvektoren werden über zwölf Eingangsneuronen und fünf versteckte Neuronen zum Ausgangsneuron vorwärts durch das Netz propagiert. Die Aktivität des Ausgangsneurons o beschreibt die Zugehörigkeit der Pixel zu den initialen Segmenten ($o > 0.5$).

Das Training des Kopplungsnetzes findet überwacht mit Hilfe des *Back Propagation Algorithmus* ([Hay94], Seite 142ff) statt. Mit einer vorsegmentierten Datenbank wird die Fehlerfunktion zwischen Ausgangsaktivität und dem Sollzustand am Ausgang ermittelt und rückwärts durch das Netz propagiert.

Diese Kopplungsstruktur ist in der Erweiterbarkeit flexibel und auf neue Probleme leicht adaptierbar.

Abbildung 17 verdeutlicht das robuste Segmentierungsergebnis von Fahrzeugen an einer Sequenz von 250 Bildern einer Straßenverkehrsszene, wobei jedes fünfzigste Bild dargestellt ist.

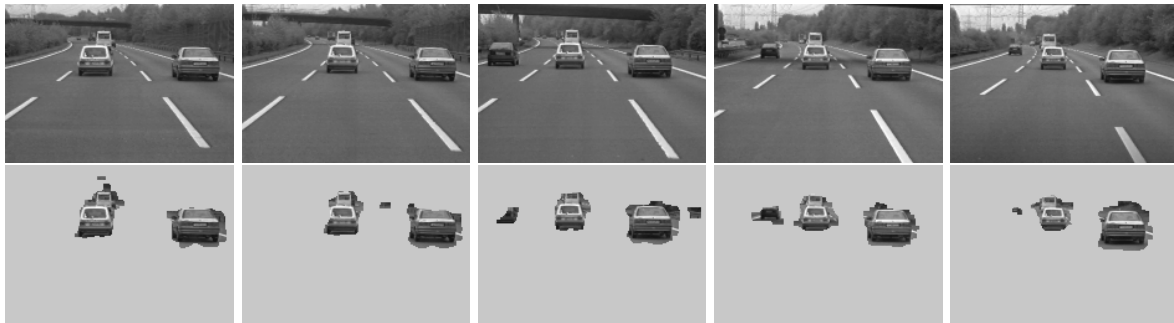


Abbildung 17: Initiale Segmentierung mit Hilfe eines MLP-Kopplungsnetzwerkes

1.11 LOC-Klassifikation

Christian Goerick, Uwe Handmann, Thomas Kalinke, Martin Werner

In diesem Abschnitt wird ein Klassifikator beschrieben, der mit Hilfe der lokalen Orientierungskodierung (LOC, Abschnitt 1.1, [Goe94] [GNW96]) eine Klassifikation von Fahrzeugen durchführt. Der Klassifikator muß mit partiellen Verdeckungen, unterschiedlichen Kamerapositionen sowie Objektrotationen und -translationen stabil arbeiten. Verrauschte Rohdaten, unterschiedliche, von der Entfernung abhängige Auflösungen und Fahrzeuge mit unterschiedlichster Form und Farbe müssen ebenfalls bewältigt werden. Zusätzlich sollte der Klassifikator mit Hilfe einer relativ geringen Anzahl von Trainingsbeispielen die wesentlichen Merkmale von Fahrzeugen lernen, um eine Generalisierbarkeit zu erreichen. Aus diesem Grund wird ein neuronales Netz zur Lösung des Klassifikationsproblems ausgewählt. Ein neuronales *feed forward* Netz mit einer Zwischenschicht, welches mit Hilfe des *back propagation* Algorithmus trainiert wird, kommt zum Einsatz [HKP91] [RMG86] (Abbildung 18).

Dieser Netztyp ist als universeller Approximator für kontinuierliche Funktionen bekan-

nt [HSW89]. Außerdem wurde gezeigt, daß diese Strukturen, mit geringfügigen Änderungen in der Art eines Bayesklassifikators a posteriori Wahrscheinlichkeiten approximieren [FM93].

Die Eingangsdaten des neuronalen Netzes bestehen aus Histogrammen einer Untermenge der LOC-Kodes. Der Netzausgang beschreibt die Klasse der übergebenen Region.

Das komplette System wurde auf einem Testfahrzeug (Mercedes Benz VITA II) implementiert und intensiv getestet [BGGZ94]. Unterschiedliche Fahrzeugklassen wurden trainiert. Eine genauere Beschreibung kann in [BZF95] nachgelesen werden.

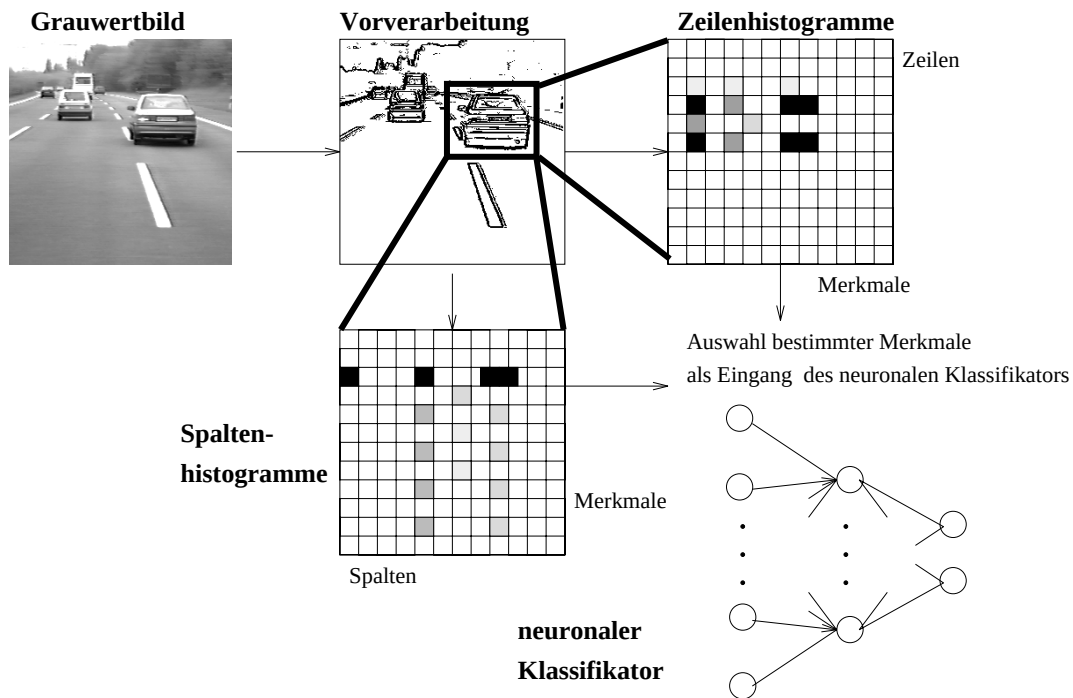


Abbildung 18: Schema des Klassifikators

1.12 Ein modulares neuronales Netzwerk zur Klassifikation von Fahrzeugen

Iris Leefken, Thomas Kalinke

Neben den Aufgabenstellungen Objektdetektion und -verfolgung über der Zeit ist die Klassifikation der Verkehrsteilnehmer von Interesse. Die große Anzahl verschiedener Fahrzeugtypen und störender Umweltbedingungen wie Witterungsverhältnisse (Beleuchtungsintensität, Schlagschatten, u.s.w.) erfordern ein hohes Maß an Flexibilität und Robustheit des Klassifikators. A-priori Wissen fließt durch eine Dekomponierung der Klassifikationsaufgabe in einzelne Teilmodule ein. Ein modulares neuronales Netzwerk ist zur Klassifikation von Fahrzeugszenen konzipiert worden, damit zum einen der Dateneingangsraum für jedes Teilmodul klein bleibt und zum anderen eine gezielte Analyse der Leistungsfähigkeit des Gesamtklassifikators möglich ist. Aus diesem Grund werden Module so entworfen, daß sie

in einem schmalbandigen Anwendungsspektrum optimal operieren. Die einzelnen Module stellen Experten dar, die in einer geeigneten Koppelstruktur ihre Leistungsfähigkeit komplementieren und zu einem breitbandigen System zusammenwachsen. Basierend auf dem LOC-Klassifikator (Abschnitt 1.11) und dem Hausdorff-Klassifikator (Abschnitt 1.8) ist ein modulares Netzwerk entwickelt worden. Als Koppelstruktur beider Teilklassifikatoren ist ein Multilayer Perzeptron trainiert worden, welches die Leistung der Module effizient fusioniert. Abbildung 19 zeigt Klassifikationsergebnisse des Gesamtsystems PKW (car), LKW (truck) und Hintergrund (no object) für eine typische Verkehrsszene.

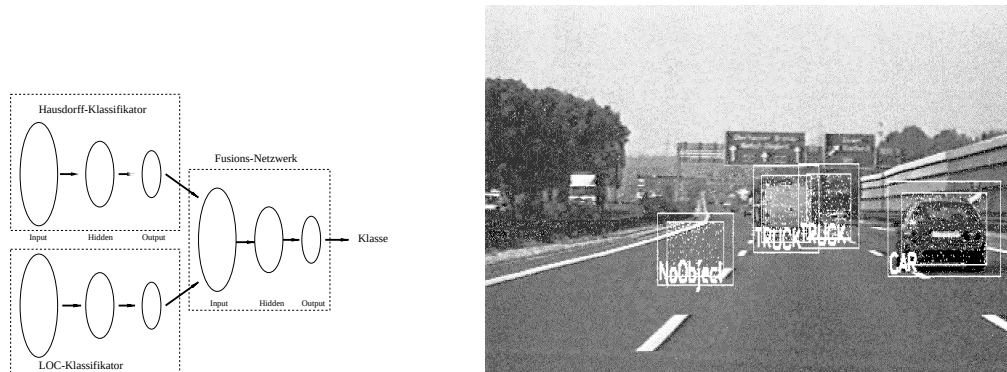


Abbildung 19: Netzwerkmodell und Klassifikationsergebnisse

2 Erkennung von Sprengstoffzündern in Fluggepäck

David Kastrup, Bernd Völpel

Aufgabe dieses Projektes ist die Unterstützung von Verfahren zur röntgenbasierten Erkennung von Plastiksprengstoffen in Fluggepäck durch Detektion der üblicherweise eingesetzten Detonatoren. Die zum Einsatz kommenden Röntgenanlagen sind Produkte des Projektpartners Heimann Systems. Zur Entlastung der Operateure an den Röntgenanlagen sollen Bereiche, in denen potentiell Sprengstoff oder zugehörige Detonatoren detektiert wurden, in den Röntgenbildern farblich markiert werden. Die nötige Rechenleistung wird durch Parallelrechnersysteme mit wenigen Prozessoren bereitgestellt. Aufgabe des Instituts für Neuroinformatik war die Entwicklung der Detektionssoftware für die Zünder.

Die Anfangsphase des Projektes konzentrierte sich auf die Suche nach geeigneten detektierbaren Merkmalen, was in der Simulationssprache IDL geschah. Eine grobe Vordetektion kann durch Schwellwertoperationen in Verbindung mit der Bildung morphologischer Kriterien (Größen zusammenhängender Regionen, charakteristische lokale Extrema) erzielt werden. Die Umrisse werden mit LOCs (lokale Orientierungskodierung [Goe94]) markiert, so daß es möglich ist, ihre charakteristische Form durch Statistiken über in feinem Raster rotierte und skalierte Masken zu detektieren. Das Innere eines Zünders wird durch einen anschließenden Matching-Prozeß mit vorgegebenen Profilen verglichen.

In der zweiten Phase wurden die Algorithmen auf die Sprache C portiert und Optimierungen mit dem Ziel, den Echtzeitanforderungen der Aufgabe gerecht zu werden, vorgenommen. Abgesehen von der Entwicklung schneller Varianten einiger verwendeter Algorithmen (etwa zum Markieren zusammenhängender Bereiche, und zur schnellen Rotation von Binärbildern) wurden verschiedene, effiziente Lookupstrukturen für Teilprobleme entwickelt, eingesetzt und anhand von Laufzeitmessungen optimiert. Wichtig war auch die laufzeitorientierte Kaskadierung der Bildverarbeitungsoperationen. Durch die vorgezogene Detektion solcher Merkmale aus dem Inneren der Detonatoren, die sich mit weitgehend skalierungs- und richtungsunabhängigen Filteroperationen detektieren ließen, konnte die Zahl der zu überprüfenden Maskenorientierungen und -skalierungen stark reduziert werden. Dadurch ließ sich auch die Fehlerkennungsrate deutlich senken, ohne daß die Erkennungsrate merklich beeinträchtigt wurde.

Der Projektpartner Heimann Systems beabsichtigt, bei entsprechendem Interesse das entwickelte Verfahren als Ergänzung ihrer Linie von Sicherheitsprodukten anzubieten.

3 Meeresumweltüberwachung

Carsten Winkel

Die Verschmutzung von offenen Gewässern hat gravierende ökologische Folgen auch auf die Gesundheit der Menschen. Eine schnelle Erkennung sowie die Quantifizierung von Art und Menge der eingebrachten Stoffe sind die Voraussetzung dafür, auf derartige Verschmutzungen geeignet reagieren zu können. Nur so ist es möglich, die Verursacher zu identifizieren und mögliche Schäden zu begrenzen. Deshalb wird die Meeresoberfläche der Nord- und Ostsee hinsichtlich Öl- und anderen Verschmutzungen im Auftrag des BMV¹ ständig vom Flugzeug aus mit verschiedenen Sensoren überwacht. Neben Standardsensoren wie SLAR (side-looking airborne radar) und UV/IR-Kameras werden durch die Unterstützung des BMBF² auch Neuentwicklungen wie ein Laserfluorosensor (LFS) [GHWW96] und ein Mikrowellenradiometer (MWR) mit drei Kanälen [Grü96] eingesetzt. Abb. 20 zeigt als Beispiel Aufnahmen eines Ölfleckes mit dem MWR und ein UV-Bild. Aus dieser vielfältigen Sammlung verschiedener multispektraler Sensordaten sollen durch geschickte Kopplung Art, Menge und Ausdehnung der Verschmutzung ermittelt werden.

Vergleicht man die Hauptkomponenten der Meßwerte aus den 12 Kanälen des LFS mit einer Datenbank, in der eine Vielzahl an Spektren verschiedener möglicher Verschmutzungen abgelegt sind, so kann man deren Art (z. B. die Ölsorte) klassifizieren. Die Benutzung neuronaler Netze soll die Güte der Klassifikation und der damit verbundenen Schätzung der Materialparameter, insbesondere der Dielektrizitätszahl verbessern und deren Anwendbarkeit auf sich ändernde Witterungsbedingungen erweitern.

Wegen der verschiedenen Eindringtiefen bei den benutzten Frequenzen geben die drei Mikrowellenkanäle des Radiometers Aufschluß über unterschiedliche Öldicken. Um aus den drei

¹Bundesministerium für Verkehr

²Bundesministerium für Bildung, Wissenschaft, Forschung und Technologie

Kanälen eine Dickenverteilung der Ölverschmutzung zu schätzen, können diese Informationen aber nicht einfach addiert werden, da die einzelnen Kanäle verschiedene radiometrische und geometrische Auflösungen besitzen. Vielmehr müssen diese zuerst angeglichen werden. Dies geschieht durch die Maximierung der Entropie [Mei88] der beiden niederfrequenten Aufnahmen unter Berücksichtigung diverser Randbedingungen sowie Nutzen von a priori-Wissen. Aufgrund der nicht meßbaren Zielwerte für die Zuordnung der Meßwerte zu Schichtdicken beim Training neuronaler Netze sind klassische überwachte neuronale Verfahren nicht einsetzbar. Deshalb erlaubt die Kenntnis des jeweiligen Gesamtölvolumens einer Messung nur die Benutzung eines integralen Fehlermaßes. Durch ein lineares Verfahren, basierend auf der SVD (singular value decomposition), werden die Parameter eines Modells bestimmt, welches die Abhängigkeit der gemessenen Strahlung zur Dicke der Ölschicht beschreibt. Weiterführendes Ziel ist jedoch die Entwicklung einer rein datengetriebenen Methode, welche diese Zusammenhänge auch bei unterschiedlichen äußeren Einflüssen, vor allem des Windes und der Himmelsstrahlung, extrahiert. Dazu existiert ein Ansatz von Wei et al. [WH96]. In Zusammenarbeit mit dem Institut für Hochfrequenztechnik der DLR soll sich das Meßsystem aufgrund der Ergebnisse der Datenanalyse und -vorverarbeitung selbst kalibrieren, um damit die Güte der Mikrowellenaufnahmen zu erhöhen. Ein zentrales Datenregister multispektraler Umweltdaten, das für weitere Auswertungen und Kopplungen mit anderen Informationen genutzt werden soll, befindet sich ebenfalls im Aufbau.

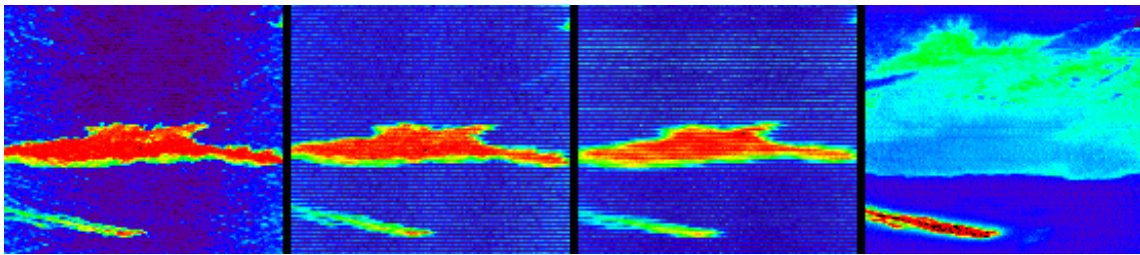


Abbildung 20: MWR-Bilder eines Ölfleckes (aufgenommen bei 89, 36 und 18 GHz, von links nach rechts) sowie ein UV-Bild

Literaturverzeichnis

- [ABN92] B.C. Arnold, N. Balakrishnan und H.N. Nagaraja. *A First Course in Order Statistics*. John Wiley & Sons, 1992.
- [BG85] V. Bruce und P.R. Green. *Visual Perception Physiology, Psychology and Ecology*. Lawrence Erlbaum Associates, London, 1985.
- [BGGZ94] M.E. Brauckmann, C. Goerick, J. Groß und T. Zielke. Towards all around automatic visual obstacle sensing for cars. In *Proceedings of the Intelligent Vehicles '94 Symposium, Paris, France*, Seiten 79–84, 1994.
- [BS89] M. Bichsel und P. Seitz. Minimum Class Entropy: A Maximum Information Approach to Layered Networks. *Neural Networks*, 2:133–141, 1989.

- [BT90] A. Blake und T. Troscianko. *AI and the eye*. John Wiley & Sons, 1990.
- [BZF95] S. Bohrer, T. Zielke und V. Freiburg. An Integrated Obstacle Detection Framework for Intelligent Cruise Control on Motorways. In *Proceedings of the Intelligent Vehicles Symposium, Detroit*, Seiten 276–281, 1995.
- [Dav80] H.A. David. *Order Statistics*. John Wiley & Sons, second. Auflage, 1980.
- [DW87] R. Durbin und D. Willshaw. An analogue approach to the travelling salesman problem using an elastic net approach. *Nature*, 326(6114):689–691, 1987.
- [FM93] M. Finke und K.-R. Müller. Estimating A-Posteriori Probabilities Using Stochastic Network Models. In *Proceedings of the Summer School on Neural Networks, Bolder, Colorado*, Seiten 276–281, 1993.
- [GHW96] K. Grüner, T. Hengstermann, P. Wagner und M.B. Wischnewsky. Recent Results Obtained from a 2nd Generation Airborne Surveillance System for Identification of Maritime Pollution. In *2nd International Airborne Remote Sensing Conference and Exhibition*, Jgg. III, Seiten 791–802, Ann Arbor, Michigan, USA, 24.-27.6. 1996. ERIM.
- [GNW96] C. Goerick, D. Noll und M. Werner. Artificial Neural Networks in Real Time Car Detection and Tracking Applications. *Pattern Recognition Letters*, 1996.
- [Goe94] C. Goerick. Local Orientation Coding and Adaptive Thresholding for Real Time Early Vision. Internal Report IR-INI 94-05, Institut für Neuroinformatik, Ruhr-Universität Bochum, D-44780 Bochum, Germany, Juni 1994.
- [Grü96] K. Grüner. The Three Frequency Microwave Radiometer of a Second Generation Airborne Surveillance System for Remote Sensing of Maritime Oil Pollution. In *IEEE-Workshop 'RF and Microwave Noise'*, Seiten 66–69, Germany, 9.-10.6. 1996. Techn. Univ. Ilmenau.
- [GS97] W.J. Gillner und M. Schröter. 2D-Ortszeit-Funktionen zur Schätzung von Relativbewegung in dynamischen Bildsequenzen. In *Proceedings 9. Aachener Kolloquium „Signaltheorie“ Bild- und Sprachsignale*, Seiten 47–50, RWTH Aachen, 1997.
- [Hay94] S. Haykin. *Neural Networks*. Macmillan College Publishing Company, Inc., 1994.
- [HKP91] J.A. Hertz, A.S. Krogh und R.G. Palmer. *Introduction to the Theory of Neural Computation*. Addison Wesley, 1991.
- [HSD73] R.M. Haralick, K. Shanmugan und I. Dinstein. Textual Features for Image Classification. *IEEE Transactions on Systems, Man and Cybernetics*, 3(6), 1973.

- [HSW89] K. Hornik, M. Stinchcombe und H. White. Multilayer Feedforward Networks are Universal Approximators. *Neural Networks*, 2:359–366, 1989.
- [Hut93] D.P. Huttenlocher. Comparing Images Using the Hausdorff Distance. *IEEE Transactions on PAMI*, 15(9), September 1993.
- [JLS94] W. Joung, C. Latimer und C. Stevens. Modelling Symmetry Detection with Backpropagation Networks. In *Proc. of the fifth Australian Conference on Neural Networks, ACNN 1994, Brisbane*, Seiten 33–36, 1994.
- [KCC89] J.M. Keller, S. Chen und R.M. Crownover. Texture description and segmentation through fractal geometry. *Computer Vision, Graphics, and Image Processing*, 45(2):150–166, 1989.
- [KvS96] T. Kalinke und W. von Seelen. Entropie als Maß des lokalen Informationsgehalts in Bildern zur Realisierung einer Aufmerksamkeitssteuerung. In *Mustererkennung 1996*, Seiten 627–634, Berlin, Heidelberg, 1996. Springer-Verlag.
- [KvS97] T. Kalinke und W. von Seelen. Kullback-Libler Distanz als Maß zur Erkennung nicht rigider Objekte. In *Mustererkennung 1997*, 1997.
- [Lis92] P.J.G. Lisboa. Der Begriff der Entropie in der Wahrscheinlichkeitsrechnung. In P.J.G. Lisboa, Hrsg., *Neural Networks: current applications*, Kapitel 1, Seiten 163–184. Chapman & Hall London, 1992.
- [Mar89] G. Marola. Using symmetry for detecting and locating objects in a picture. *Computer Vision, Graphics, and Image Processing*, 2(46):179–195, 1989.
- [MBS93] T.M. Martinetz, S.G. Berkovich und K.J. Schulten. Neural-Gas Network for Vector Quantization and its Application to Time-Series Prediction. *IEEE Transactions on Neural Networks*, 4(4):558–569, 1993.
- [Mei88] E.S. Meinel. Maximum-entropy image restoration: Lagrange and recursive techniques. *Journ. Opt. Soc. Am.*, 5(1):25–29, January 1988.
- [Nol96] D. Noll. *Ein Optimierungsansatz zur Objekterkennung*. Number 454 in Fortschrittberichte, Reihe 10. VDI-Verlag, Düsseldorf, 1996. Dissertation, Ruhr-Universität Bochum.
- [Pag91] D.W. Paglieroni. Distance Transforms: Properties and Machine Vision Applications. *CVGIP*, 54(1):56–74, January 1991.
- [Pen84] A.P. Pentland. Fractal-based description of natural scenes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6, 1984.
- [PNHA84] S. Peleg, J. Naor, R. Hartley und D. Avnir. Multiple resolution texture analysis and classification. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6, 1984.

- [RMG86] D. Rumelhart, J.L. McClelland und PDP Research Group. *Parallel Distributed Processing*. MIT Press, Cambridge, Massachusetts, USA, 1986.
- [Roc84] I. Rock. *Perception*. Scientific American Library New York, 1984.
- [Ros86] A. Rosenfeld. Axial representation of shape. *Computer Vision, Graphics, and Image Processing*, 33:156–173, 1986.
- [Roy81] F.R. Royer. Detection of symmetry. *Journal of Experimental Psychology: Human, Perception and Performance*, 7(46):1186–1210, 1981.
- [SK92] E. Schulze-Krüger. *Analyse von Augenbewegungen des Menschen zur Symmetrie- und Raumwahrnehmung und Vergleich zu einem aktiven Kameratelesystem*. VDI Verlag, Reihe 17 Biotechnik Nr.83, 1992.
- [WH96] G.Q. Wei und G. Hirzinger. Learning Shape from Shading by a Multilayer Network. *IEEE Trans. Neural Networks*, 7(4):985–995, July 1996.
- [YdSB93] G. Yahiaoui und M. de Saint Blancard. Texture-based image segmentation for road recognition with neural networks. In *Neural Networks and their Industrial & Cognitive Applications*, Seiten 321–328, 269-287, Rue de la Garenne, 92024 Nanterre Cedex, France, 1993. EC2.
- [YSU94] A.L. Yuille, P. Stolorz und J. Utans. Statistical Physics, Mixtures of Distributions, and the EM Algorithm. *Neural Computation*, 6:334–340, 1994.
- [ZBvS93] T. Zielke, M. Brauckmann und W. von Seelen. Intensity and Edge-Based for Symmetry Detection with an Application to Car-Following. *Computer Vision, Graphics, and Image Processing*, 2(58):177–190, 1993.