

Kapitel 3

Verkehrszeichenerkennung

3.1 Architektur des Systems

Das System zur Klassifikation von Verkehrszeichen basierend auf einer symbolischen Bildbeschreibung kann nicht für sich alleine stehen. Es ist angebracht auch die Bildverarbeitungsschritte kurz zu beschreiben, die für die Bereitstellung der symbolischen Bildbeschreibung verantwortlich sind. Erst aus den beiden Systemen gemeinsam ergibt sich eine Bildverarbeitungskette, die sich von der Kamera aus bis hin zum klassifizierten Verkehrszeichen erstreckt.

Um einen Überblick über das Gesamtsystem zu bekommen, sollen zunächst die Komponenten, sowie die zwischen ihnen ausgetauschten Informationen und Daten kurz vorgestellt werden, bevor eine detailliertere Beschreibung gegeben wird.

3.1.1 Überblick über das Gesamtsystem

3.1.1.1 Komponenten der Bildvorverarbeitung

Bis eine symbolische Beschreibung einer Verkehrsszene vorliegt, sind mehrere Schritte einer Bildverarbeitungskette zu durchlaufen. Zunächst wird ein digitalisiertes Bild der Umgebung aufgezeichnet, welches mit Methoden der ikonischen Bildverarbeitung behandelt wird. Anschließend wird aus der Pixelinformation eine symbolische Darstellung des Bildes in Form von Formprimitiven abgeleitet. Die Formprimitive stellen die Schnittstelle zwischen Bildvorverarbeitung und symbolischer Bildverarbeitung dar.

Kamera : Im Fahrzeug wird eine CCD-Kamera installiert, die das Blickfeld des Fahrers aufnimmt. Die Kamera wird hier nicht näher beschrieben, technische Daten über CCD-Kameras können in [Fe92a] nachgeschlagen werden.

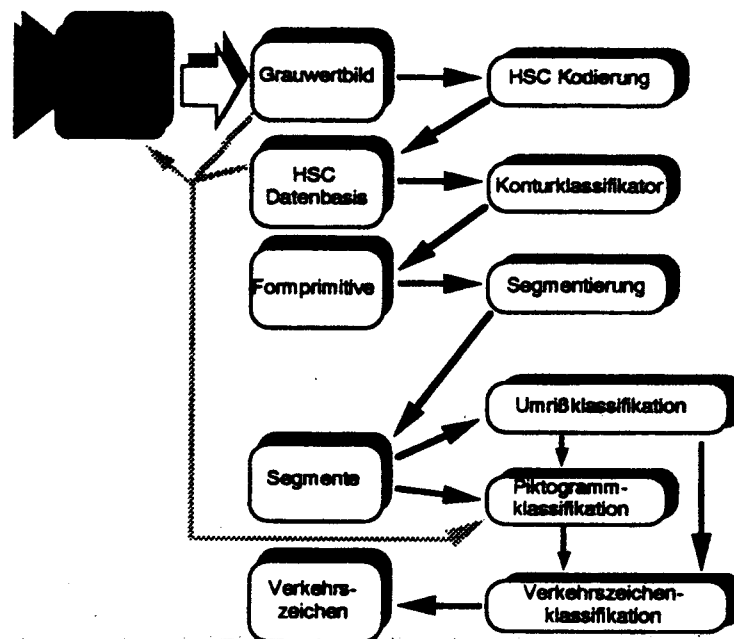


Abbildung 3.1: Konzeption des Bilderkennungssystems

Frame-Grabber : Das Kamerabild, bzw. ein Ausschnitt des Bildes wird zu einem bestimmten Zeitpunkt eingefroren und auf eine 512 x 512 Bildpunkt-Matrix als Grauwertbild abgebildet. Auch auf diese Komponente wird nicht näher eingegangen. Für Informationen über eine Frame-Grabber-Karte sei auf [Fe90] verwiesen.

HSC-Kodierung : Zusammenhängende Kanten werden aus dem Grauwertbild herausgefiltert und mit Hilfe eines hierarchisch aufgebauten Bildanalyse-systems (HSC) kodiert. Hier wird bereits Information über offene oder geschlossenen Strukturen bereitgestellt.

Erkennung von Formprimitiven : Die vom HSC gefundenen Strukturen werden untersucht. Dabei werden Entscheidungen über Ort und Zahl der Eckpunkte in Linienzügen sowie über Mittelpunkt und Radius eines Kreises oder Kreissegments gefällt, die zur symbolischen Beschreibung der Szene durch Formprimitive benötigt werden.

3.1.1.2 Symbolische Bildverarbeitungskomponenten

Die symbolischen Bildverarbeitungskomponenten, die im Folgenden angeführt werden sind Bestandteil der vorliegenden Arbeit. Während zur Bildvorverarbeitung nur ein kurzer Überblick gegeben wird, wird auf die symbolischen Bildverarbeitungskomponenten noch näher eingegangen.

Segmentierung : Die Segmentierungskomponente gruppiert die Szenenelemente zu interessanten Regionen, die einzeln klassifiziert werden sollen. Die Segmente können während der Klassifikation noch ergänzt und aufgespalten werden.

Umrißklassifikation : Bestimmung der Verkehrszeichenklasse für ein Segment, wobei noch keine Details des möglichen Verkehrszeichens (Piktogramme) untersucht werden. Das Ergebnis ist eine Verkehrszeichenklasse, bzw. eine Menge von möglichen Verkehrszeichenklassen.

Piktogrammklassifikation : Die möglichen Verkehrszeichenklassen werden näher untersucht, wobei durch die Umrißklassifikation bereits ein interessanter Bildbereich bestimmt und die Menge der möglichen Piktogramme eingeschränkt wurde. Hier kann eine Rückkopplung erfolgen indem der interessante Bereich vergrößert und die Vorverarbeitung noch einmal durchlaufen wird.

Verkehrszeichenklassifikation : Die Ergebnisse der beiden vorhergehenden Klassifikationsstufen werden kombiniert und das Verkehrszeichen (bzw. die Verkehrszeichenklasse) mit dem höchsten Ähnlichkeitswert zu einem in der Wissensbasis modellierten Zeichen wird dem jeweiligen Szenensegment zugeordnet.

3.1.2 Architektur des symbolischen Klassifikationssystems

Das Klassifikationssystem besteht aus zwei Komponenten, der Wissensbasis, die strukturelle Informationen über Verkehrszeichen bereitstellt, sowie dem Klassifikationsmechanismus, der den Ablauf einer Klassifikation steuert. Durch die Struktur der in Form eines Netzwerkes organisierten Wissensbasis wird die Klassifikation bereits stark unterstützt, so daß - ähnlich wie bei konnektionistischen Ansätzen - eine exakte Trennung zwischen Verarbeitung und Speicherung schwer durchzuführen ist.

3.1.2.1 Modellierung der Verkehrszeichen

Um Verkehrszeichen klassifizieren zu können, muß dem System deren Struktur bekannt sein. Dazu wird eine Wissensbasis in Form eines Semantischen Netzwerkes angelegt, die Beschreibungen einer Menge von Verkehrszeichen enthält. Dieses Netzwerk besteht aus zwei Netzwerkebenen. In einer Ebene wird strukturelle Information

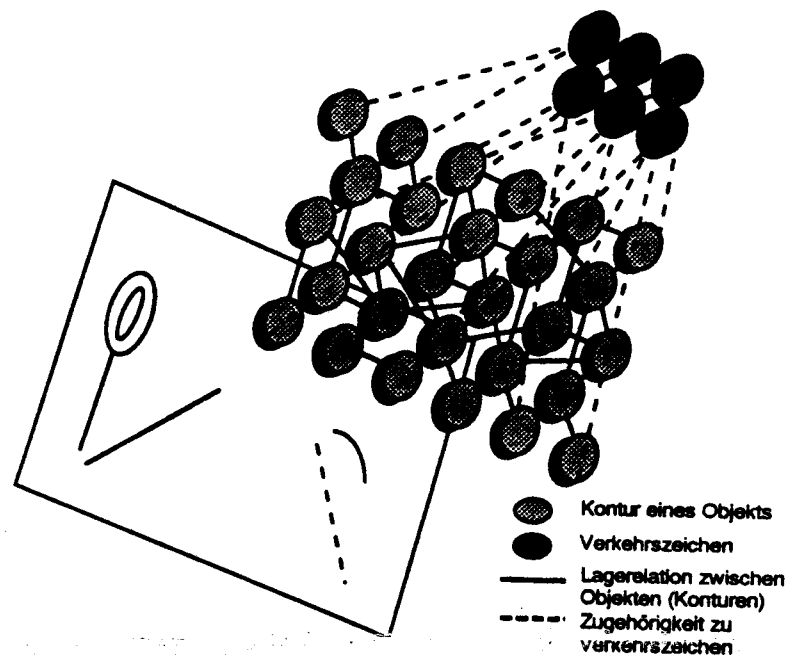


Abbildung 3.2: Schema des Klassifikationsnetzwerks

zu den modellierten Verkehrszeichen gespeichert. Knoten dieser Ebene repräsentieren die Konturinformation der geometrischen Primitive, sowie eine geeignete Beschreibung der komplexen, als Piktogramm bezeichneten geometrischen Objekte. Kanten repräsentieren Lagerrelationen zwischen den einzelnen geometrischen Formen. Die zweite Ebene besteht aus einer Menge von Knoten, die Verkehrszeichen repräsentieren. Jeder dieser Knoten ist mit Knoten und Kanten der ersten Ebene durch *HAS-PART* Kanten verknüpft, die mit Gewichten versehen sind. Abbildung 3.2 soll die Struktur des Netzwerks noch einmal schematisch darstellen. Dabei stehen die helleren Knoten für Piktogramme und geometrische Primitive, die Verbindungen zwischen ihnen für Lagerrelationen. Die dunkleren Knoten stellen die Knoten dar, die Verkehrszeichen repräsentieren. Diese sind durch gestrichelte Linien mit den Knoten und Kanten der unteren Ebene verbunden, die deren Zugehörigkeit zu Verkehrszeichen bestimmen.

3.1.2.2 Klassifikation eines Bildes

Durch einen Segmentierungsschritt wird die durch eine Menge von Formprimitiven gegebene Bildbeschreibung in - anhand der Lagebeziehung *enthält* aufgespannte -

Baumstrukturen organisiert, die jeweils ein mögliches Verkehrszeichen repräsentieren. Damit ist eine Struktur der Eingangsdaten gegeben, die mit der Wissensbasis verglichen werden kann.

Die gefundenen Bäume werden durch einen ähnlichkeitsbasierten Match auf die erste Ebene des Netzwerks abgebildet. Dabei werden den Netzwerkknoten Ähnlichkeitswerte zu den Formprimitiven zugeordnet. Ein Knoten, dem ein Ähnlichkeitswert zugeordnet wurden, der über einem Schwellwert liegt, wird als *erkannt* eingestuft. In einem nächsten Schritt werden die im Netzwerk modellierten Lagereaktionen überprüft, die zwischen zwei *erkannten* Knoten definiert sind. Die Ähnlichkeitswerte, die nach dieser Abbildung Knoten und Kanten zugeordnet worden sind, werden den modellierten Verkehrszeichenknoten durch die gewichteten *HAS-PART* Kanten übermittelt. Nach der Verknüpfung von Gewichten und Ähnlichkeiten an den Verkehrszeichenknoten, wird eine weitere Schwellwertoperation durchgeführt, durch die eine Verkehrszeichenklasse bestimmt werden kann. Dieser Schritt wird in der obigen Unterteilung als *Umrißklassifikation* bezeichnet.

Es folgt eine gezielte Untersuchung von durch die Verkehrszeichenklasse bestimmten Bereichen in denen Piktogramme erwartet werden. Die Suchbereiche und die in Frage kommenden Piktogramme sind durch den ersten Klassifikationsschritt und das strukturelle Wissen über Verkehrszeichen bereits festgelegt, bzw. eingeschränkt. Ergebnis der Piktogrammklassifikation ist die Zuordnung von Ähnlichkeitswerten zu den Piktogrammknoten der ersten Netzwerkebene.

Als letzter Schritt werden jetzt die vollständig vorliegenden Ähnlichkeitswerte analog zur *Umrißklassifikation* durch Gewichtung und Schwellwertbildung zur endgültigen Klassifikation von Verkehrszeichen verwendet.

3.2 Verkehrszeichen

Die Modellierung der Domäne, also die Speicherung von Wissen über den Anwendungsbereich ist eines der Hauptprobleme, die sich bei der automatischen Interpretation von natürlichen Szenen ergeben. Bei der Wahl der Wissensrepräsentationsform ist darauf zu achten, daß eine möglichst präzise und strukturierte Modellierung der Domäne vorgenommen wird, die sich auch an den Eingangsdaten für die gestellte Klassifikationsaufgabe orientiert.¹ Hierzu ist eine die Kenntnis der Struktur der

¹Die Eingabe an das System soll möglichst wenig umgeformt werden müssen, um mit der Repräsentationsform verglichen werden zu können. Gleiches gilt für die erforderliche Ausgabe (Klasse) und der Form in der sie gespeichert ist. Bei der Erstellung des Modells, ist zu beachten, daß sich die wirkliche Erscheinungsform von der als Eingabe vorliegenden Form unterscheiden kann. (z.B. Verzerrung und Aufreißen von Konturen)

Domäne von großer Wichtigkeit, deshalb soll der Problembereich Verkehrszeichen in diesem Abschnitt näher betrachtet werden.

Die Menge der Verkehrszeichen, ihre Bedeutung, sowie ihre Form, Farbe, Größe, Beschriftung und Anbringung ist durch die deutsche Straßenverkehrsordnung (StVO) genau festgelegt. Daraus ergeben sich semantische und syntaktische Strukturierungsmöglichkeiten.

3.2.1 Semantische Einteilung von Verkehrszeichen

Die StVO unterteilt die Verkehrszeichen anhand ihrer *Bedeutung* für den Straßenverkehr in vier Gruppen:

Gefahrenzeichen : Warnung vor einer Gefahrensituation, die vom Fahrer nicht rechtzeitig erkannt werden kann.

Vorschriftzeichen : Bei dieser Gruppe handelt es sich um Gebote und Verbote. Zu dieser Klasse zählen Verkehrsverbote, Streckenverbote, Haltegebote, Halteverbote und vorgeschriebene Fahrtrichtung.

Richtzeichen : Das Vorfahrtsschild, Ortsnamensschilder, Verkehrslenkzeichen und Hinweisschilder zählen zu dieser Klasse von Verkehrszeichen.

Zusatzzeichen : Diese Schilder werden unterhalb eines Schildes der oben genannten Klassen angebracht und können u.a. Gefahren näher bezeichnen oder den Verlauf der Vorfahrtsstraße angeben.

Vorschriftzeichen und Richtzeichen unterteilen sich weiter in verschiedene Untergruppen, so daß man von einer hierarchischen semantischen Ordnung (siehe Abbildung 3.3) der Verkehrszeichen sprechen kann.

3.2.2 Syntaktische Einteilung von Verkehrszeichen

Für die Klassifikation von Verkehrszeichen aufgrund von Kanteninformationen, wie sie hier durchgeführt werden soll, hat die Erscheinungsform eines Verkehrszeichens große Bedeutung. Die Erscheinungsform eines Verkehrszeichens ist bestimmt durch seine Form und seine Farbe².

²Bei diesem Ansatz wurde auf die Berücksichtigung der Farbe verzichtet und ein Verkehrszeichen lediglich anhand von Kanteninformation dargestellt. Ein Grund für diese Einschränkung ist der höhere technische Aufwand, den ein System mit sich bringen würde, welches Farbinformation verarbeitet.

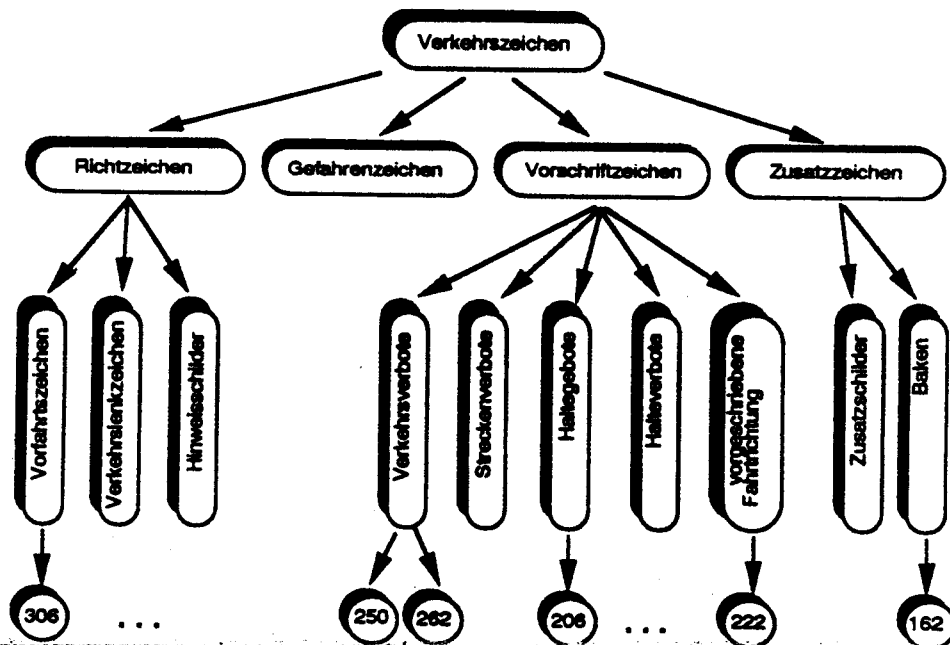


Abbildung 3.3: Semantischer Hierarchiebaum der Verkehrszeichen

Verkehrszeichen enthalten hauptsächlich die Farben Schwarz³, Weiß⁴, Rot⁵, Blau⁶ und Gelb⁷. Durch diese Einschränkung kann eine Vorfilterung der Verkehrsszene vorgenommen werden, wobei Flächen mit bestimmter Farbe (*Area of Interest*) als Kandidaten für Verkehrszeichen selektiert werden können⁸.

Betrachtet man die Form eines Verkehrszeichens, so lassen sich einige charakteristische Umrißformen unterscheiden.

1. Kreis
2. Dreieck
 - (a) Ecke nach oben
 - (b) Ecke nach unten

³für Schriften und Piktogramme

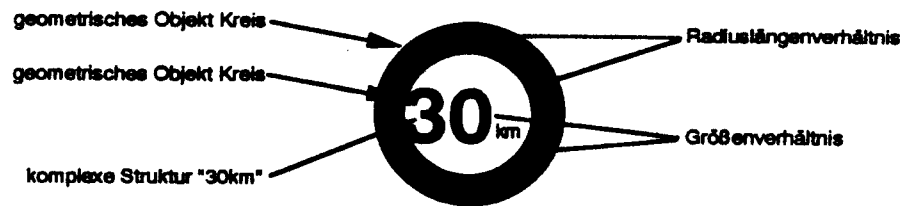
⁴für Schriften und Hintergrund

⁵für Rahmen von Verbots- und Warnschildern

⁶zur Kennzeichnung von Autobahnschildern

⁷hauptsächlich Hinweisschilder

⁸siehe dazu auch Anhang

Abbildung 3.4: Aufbau des Zeichens *Geschwindigkeitsbeschränkung*

3. Rechteck

(a) Quadrat

- i. Ecke nach unten
- ii. Seite nach unten

(b) breiter als hoch

(c) höher als breit

4. regelmäßiges Achteck

5. Pfeil (Fünfeck)

6. Kreuz (Andreaskreuz)

Diese geometrischen Objekte beschreiben den Umriß eines möglichen Verkehrszeichens. Die Beschreibung eines Verkehrszeichens besteht nun aus dieser äußeren Form sowie weiteren darin enthaltenen geometrischen Objekten. Hier kann zwischen einfachen und komplexen geometrischen Formen unterschieden werden, aus denen sich das Verkehrszeichen aufbaut. Komplexe geometrische Formen sollen hier als *Piktogramme*⁹ bezeichnet werden.

Abbildung 3.4 zeigt den Aufbau eines Verkehrszeichens am Beispiel des Zeichens für eine Geschwindigkeitsbeschränkung. Dieses Zeichen besteht aus zwei ineinanderliegenden Kreisen und einer komplexen Struktur bestehend aus den Zahlen 30 und dem Schriftzug km.

Die Gesamtmenge der Verkehrszeichen soll jetzt in einer Form strukturiert werden, die, im Hinblick auf die konkrete Anwendung, keine Farbinformation enthält. Verkehrszeichen bauen sich aus geometrischen Formen auf, die ineinander liegen.

⁹ Diese nehmen sowohl bei der Repräsentation als auch bei der Klassifikation eine Sonderstellung ein. (Siehe dazu auch 3.4.2 und 3.4.1)

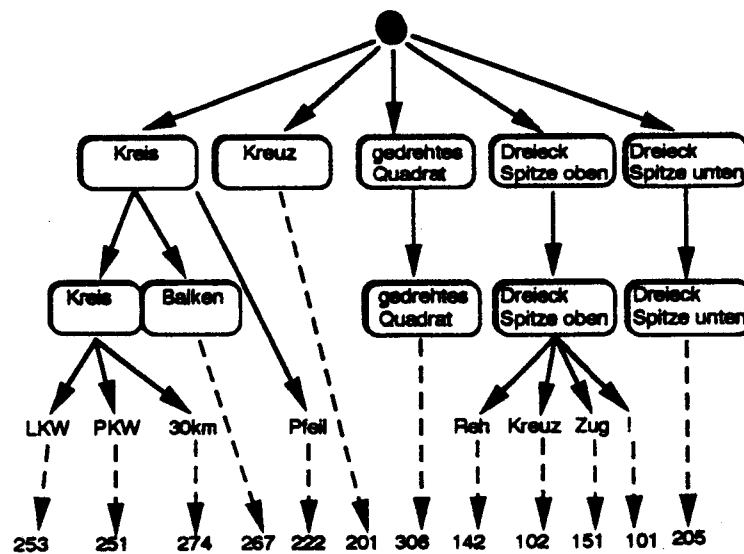


Abbildung 3.5: Strukturierung der Menge der Verkehrszeichen

Als Strukturierung der Menge der Verkehrszeichen bietet sich also ein Baum an, der anhand der *enthält* Relation aufgespannt wird. Dabei befinden sich auf der ersten Ebene die oben angeführten geometrischen Objekte, die den Umriß eines Verkehrszeichens beschreiben. Alle Verkehrszeichen mit gleicher Umrißform werden unter dem Knoten, der diese beschreibt, zusammengefaßt. Der Umriß wird aus der Verkehrszeichenbeschreibung entfernt und die Unterbäume in gleicher Weise unter diesem Knoten erzeugt. Die Blätter des Baumes werden mit den Verkehrszeichen verknüpft, deren Bestandteil sie sind. Es ergibt sich eine Struktur, wie sie anhand einer Teilmenge der Verkehrszeichen in Abbildung 3.5 dargestellt ist.

3.3 Bildvorverarbeitung

Bevor eine Szenenbeschreibung vorliegt, die symbolisch verarbeitet werden kann, müssen einige Vorverarbeitungsschritte durchlaufen werden. Zunächst muß das Kamerabild eingefroren, d.h. digitalisiert und abgespeichert werden. Dies wird von einer speziellen Hardware¹⁰ erledigt. Nachdem nun ein Grauwertbild vorliegt, muß dieses eine *ikonische*¹¹ Bildverarbeitungskomponente durchlaufen. Ergebnis der

¹⁰Framegrabber

¹¹Als ikonische Bildverarbeitung wird eine Verarbeitungskomponente bezeichnet, deren Eingabe reine Pixelinformation ist.

Bildvorverarbeitung ist eine Menge von Formprimitiven, die das Bild symbolisch beschreiben. Diese Formprimitive stellen die Schnittstelle zwischen Bildvorverarbeitung und der symbolischen Bildverarbeitung dar. Diese Bildverarbeitungskette, beginnend bei der Kamera und endend bei einer symbolischen Szenenbeschreibung durch Formprimitive, wird von B. Besserer im Hause der Daimler-Benz AG Esslingen, im Rahmen seiner Doktorarbeit erstellt. Da beide Arbeiten zusammen erst eine vollständige Bildverarbeitungskette darstellen, soll hier kurz auf die nötigen Schritte zur Erzeugung einer symbolischen Szenenbeschreibung eingegangen werden.

3.3.1 Die HSC Kodierung

Der HSC (Hierarchical Structure Code) stellt eine spezielle Form der Kodierung von Bildinformation dar.¹² Er liefert eine Datenstruktur, die Informationen über erkannte Formelemente wie Linien, Flächen oder Kanten auf verschiedenen Auflösungsebenen sowie die Verknüpfung dieser Formelemente zur Verfügung stellt.

Ein von einer Kamera aufgenommenes Bild wird in eine 512×512 Pixel umfassende Matrix von Grauwerten digitalisiert und in weitere, gröbere Auflösungen (256×256 , 128×128 , ...) transformiert. Auf jeder Auflösungsebene wird eine Laplace-Transformation gefolgt von einer Schwellwertoperation durchgeführt, so daß jedem Pixel eines Bildes genau ein Wert +, - oder 0 zugeordnet werden kann. Dabei bedeutet ein +, daß es sich um ein helles Pixel handelt, welches von dunklen umgeben ist. Analog dazu wird ein dunkles Pixel, welches sich in der Nähe von hellen Pixeln befindet, mit einem - bezeichnet. Ein als 0 eingestuftes Pixel hebt sich von seiner Umgebung nicht stark ab.

Auf der niedrigsten Auflösungsebene wird von einem hexagonal angeordneten Gitter von 512×512 Pixeln ausgegangen. Mit jeder höheren Auflösungsebene verringert sich die Auflösung um die Hälfte (256×256 , 128×128 , ...). Ausgehend von der ausgewählten Auflösungsebene werden verschiedene Verknüpfungsebenen aufgebaut. Dabei werden jeweils 7 benachbarte Inseln (Auf der untersten Verknüpfungsebene wird eine Insel durch nur ein Pixel beschrieben) einer Verknüpfungsebene zu einer Insel der übergeordneten Verknüpfungsebene zusammengefaßt (siehe Abbildung 3.6).

Jeder Insel wird nun eine Beschreibung zugeordnet. Bei diesem Vorgang werden die 6 Nachbarinseln der Insel mit in Betracht gezogen und anhand dieser Gesamtinformation (Verteilung der +, -, 0 Werte in angrenzenden Pixelringen) werden der Insel ein Typ (Fleck, Linie, Kante und jeweilige Helligkeit), Informationen wie

¹²Der HSC wurde unter der Leitung von Prof. Dr. G. Hartmann im Fachbereich Elektrotechnik an der Universität-Gesamthochschule Paderborn entwickelt.

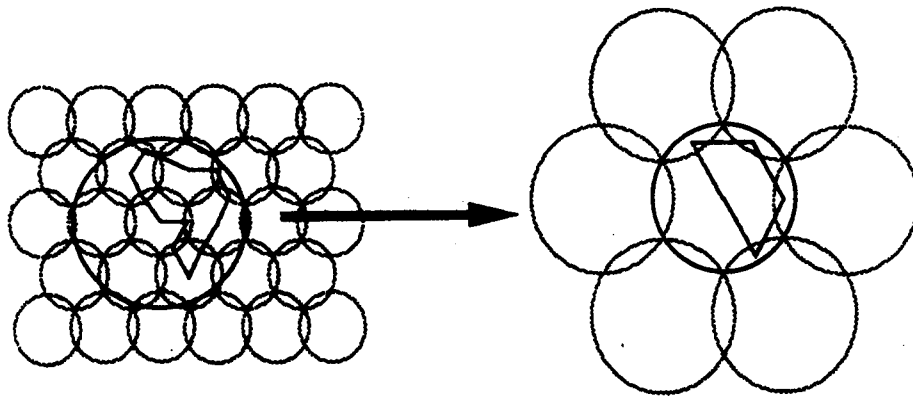


Abbildung 3.6: Inselbildung des HSC

Linienende und Anschlußpunkte zu den Nachbarinseln, sowie die referenzierten Pixel zugeordnet.

Auf der nächst höheren Verknüpfungsebene werden jetzt wieder 7 benachbarte Inseln zu einer neuen zusammengefaßt, wobei die zusammenhängenden Strukturen verknüpft werden (siehe Abbildung 3.6). Dabei ist der Überlappungsbereich zweier Inseln Kriterium dafür, ob zwei Elemente verknüpft werden können. Dies ist nur möglich, wenn beide Elemente im Überlappungsbereich dieselbe Substruktur aufweisen. Jedes neue Element trägt Indizes für jede Substruktur aus der es zusammengesetzt wurde, so daß auf jeder Verknüpfungsebene mittels dieser Indizes (über Hash-Funktionen) die Elemente der untergeordneten Verknüpfungsebene, bis hin zur Pixelinformation der Grundkodierung, angesprochen werden können.

Nach der Generierung des HSC Struktur können die einzelnen gefundenen Formelemente untersucht werden. Die vom HSC gelieferte Repräsentation der Bildinformation kann als eine Menge von Bäumen interpretiert werden. Jeder Baum stellt ein Formelement dar und seine Wurzel befindet sich auf der Verknüpfungsebene, auf der das Formelement in einer einzigen Insel enthalten ist. Die zu dem Formelement gelieferte Information ist an der Wurzel nur sehr grob und wird von Sohn zu Sohn feiner. Abbildung 3.7 zeigt die Wurzeln von zwei ineinanderliegenden Dreiecken. Die Wurzel des inneren Dreiecks ist auf einer niedrigeren Ebene zu finden als die des äußeren Dreiecks. Zur ausführlicheren Beschreibung des HSC sei auf

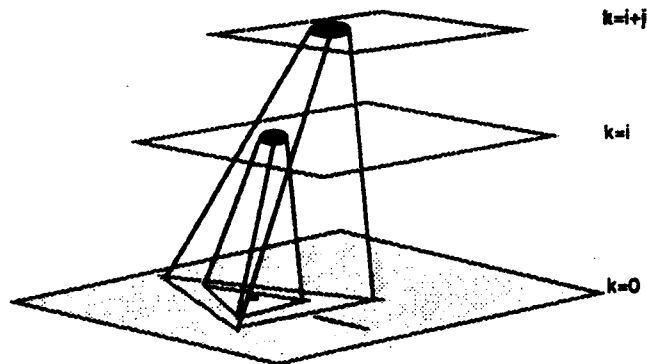


Abbildung 3.7: Wurzeln des HSC auf verschiedenen Ebenen

[Har83, DH86, MBH90] und [MH88] verwiesen.

3.3.2 Benutzung des HSC zur Generierung von symbolischen Formprimitiven

Die in diesem System verwendeten Methoden nutzen den hierarchischen Aufbau der HSC Datenbasis nicht aus¹³. Ein zu untersuchendes Objekt wird anhand seiner Wurzel ausgewählt. Danach wird direkt auf die zugrundeliegende Pixelstruktur zugegriffen, ohne die dazwischenliegenden Verknüpfungsebenen bei der Untersuchung zu nutzen.

Durch Begrenzung der zu betrachtenden Verknüpfungs- und Auflösungsebenen kann eine Vorauswahl relevanter Strukturen getroffen werden. Betrachtet man nur Objekte, deren Wurzel auf höheren Ebenen liegen, so fallen kleine Objekte heraus, da deren Wurzel auf niedrigeren Ebenen zu finden sein wird. Betrachtet man in Abbildung 3.7 beispielsweise nur die Ebenen die größer sind als i , so würde das innere Dreieck herausfallen und nicht weiter betrachtet werden. Es wird also bereits hier eine Vorauswahl von zu untersuchenden Strukturen getroffen, so daß deren Zahl erheblich eingeschränkt werden kann.

Mit Hilfe verschiedener Methoden werden die verbleibenden Strukturen auf ihre Form hin untersucht. Ziel ist eine Zuordnung von zusammenhängenden Pixeln zu einem Objekt aus einer Menge von Formprimitiven wie sie in 3.3.3 definiert ist. Dabei

¹³Der HSC wurde für industrielle Anwendungen konzipiert. Einfache, etwa formatfüllende Objekte sollten erkannt werden, wobei die Abstraktion, die durch betrachten des Objekts auf einer höheren Verknüpfungs- oder Auflösungsebene auftritt, genutzt werden kann. In der hier vorgestellten Anwendung erweist sich die Abstraktion als störend.

besteht die Aufgabe in der Detektion von Eckpunkten in Kantenzügen, der Erkennung eines Kreises, der Bestimmung von Fläche und Umfang, sowie Mittelpunkt und Radius eines zu untersuchenden Objekts.

Die Ergebnisse verschiedener Testfunktionen (Messungen) bestehen aus einem Vektor, der jeder Formprimitivklasse und jeder (Oder-)Verknüpfung von Formprimitivklassen eine Evidenz zuordnet. Diese Evidenzen werden mit Hilfe der Dempster-Shafer Regel verknüpft, wodurch das Objekt klassifiziert werden kann. Die durch Anwendung der Dempster-Shafer Regel erhaltenen Evidenzen werden dem klassifizierten Formprimitiv als Zuverlässigkeitswert zugeordnet.

3.3.3 Szenendarstellung durch Formprimitive

Nachdem die Vorverarbeitungsschritte durchlaufen sind, wird die transformierte Szene in geeigneter Form¹⁴ an die Komponente zur symbolischen Bildverarbeitung übertragen. Die obige Tabelle beschreibt die Formprimitive, die von der Bildvorverarbeitung zur Verfügung gestellt werden. Diese Formprimitive beschreiben die Originalszene in stark abstrahierter Form, so daß Lagerrelationen und Konturen der einzelnen Objekte nur in verzerrter, bzw. vereinfachter Form vorliegen. Auch hat die symbolische Bildverarbeitungskomponente keine Möglichkeit, auf Pixelinformation zuzugreifen, um eventuell benötigte genauere Informationen über bestimmte Formprimitive zu erhalten.

Zur Vereinfachung sollen für die nachfolgenden Kapitel folgende Bezeichnungen eingeführt werden: Die Menge der Formprimitive wird mit Γ bezeichnet. Auf die einem Formprimitiv $a \in \Gamma$ zugeordneten Attribute kann durch den Namen des Attributs zugegriffen werden. Beispielsweise liefert *contour.length(a)* die Länge der Kontur des Formprimitivs a . Bezeichnet t einen Formprimitivtyp (circular, triangle, rectangle, ...), so wird die Menge der Formprimitive des Typs t mit Γ_t bezeichnet.

3.4 Symbolische Bildverarbeitung

3.4.1 Die Wissensrepräsentation

Unter wissensbasierter Bildanalyse wird die Generierung einer symbolischen Beschreibung eines Bildes, oder einer Bildfolge verstanden, die auf der Nutzung explizit repräsentierten problemspezifischen Wissens beruht. In den letzten Jahren ist die Wissensrepräsentation zu einem zentralen Forschungsgebiet der Künstlichen Intelligenz geworden und oft wird die Repräsentation von Wissen in KI-System als

¹⁴In der aktuellen Implementierung bildet eine Datei die Verbindung zwischen Vorverarbeitung und symbolischer Bildverarbeitung.

Bezeichnung des Formprimitivs	Beschreibung
triangle	ID reliability bounding box list of corners contour length surface
circular	ID reliability bounding box estimated center mean radius contour length surface
rectangle	ID reliability bounding box list of corners contour length surface
polygonal	ID reliability bounding box number of corners list of corners contour length surface
line	ID reliability list of corners contour length
open polygonal	ID reliability number of corners list of corners contour length
circular line	ID reliability endpoint 1 endpoint 2 estimated center mean radius contour length

Abbildung 3.8: Die Menge der definierten Formprimitivtypen

definierendes Merkmal verstanden. Es gibt zwei zusammengehörige Methoden der Wissensrepräsentation: die Repräsentation von Wissen über Fakten, die deklarativ dargestellt werden, und das Wissen über die Verwendung des Faktenwissens, das prozedural formuliert wird.

Der in dieser Arbeit verwendete Repräsentationsformalismus wird als ein *Semantisches Netzwerk* bezeichnet. Ein Semantisches Netzwerk besteht aus markierten Knoten und Kanten sowie einer semantischen Interpretation, die Knoten und Kanten eine Bedeutung zuordnet.

Die meisten Semantischen Netzwerke haben die grundlegende Netzwerkstruktur gemein, unterscheiden sich aber in der zugeordneten Semantik. Wichtig ist die Tatsache, daß ein Netzwerk alleine noch keine Bedeutung hat. Diese wird erst durch die Prozeduren bestimmt, die auf das Netzwerk zugreifen. Nähere Informationen über Semantische Netzwerke sind [Ric89], [Cha82] und [BF81] zu entnehmen.

Dem System muß Wissen über den Aufbau von Verkehrszeichen zur Verfügung stehen. Da der Aufbau von Verkehrszeichen weitestgehend durch die deutsche Straßenverkehrsordnung (StVO) festgelegt ist, ist damit eine exakte Beschreibung der Domäne möglich. In Abschnitt 3.2 wird eine Strukturierungsform der Domäne angegeben, die in ähnlicher Form zur Wissensrepräsentation übernommen werden soll.

3.4.1.1 Das Verkehrszeichennetzwerk

Die Verkehrszeichen sollen in einem Semantischen Netzwerk abgelegt werden. Da die Farbinformation fehlt, werden nur Konturen und Lagerrelationen zur Beschreibung eines Verkehrszeichens benutzt¹⁵. Die Konturen der einzelnen Bestandteile eines Verkehrszeichens werden als Knoten, während die Lagerrelationen als Verbindung zwischen zwei Knoten in das Netzwerk eingetragen werden.

Definition: *Verkehrszeichennetzwerk*

Das Verkehrszeichennetzwerk soll in mehreren Schritten definiert werden. Zunächst wird die im Netzwerk gespeicherte Information vorgestellt. In einem zweiten und dritten Schritt werden die Knoten und Kanten und zuletzt das Netzwerk selbst definiert.

1. Gespeicherte Informationen

Das Netzwerk speichert Information über eine Menge von Verkehrszeichen und deren strukturellen Aufbau. Dabei sind die Konturen,

¹⁵Überlegungen zur Integration der Farbinformation wurden bereits angestellt. Siehe dazu auch 4.2.2.

Piktogramme und Lagerrelationen aufzunehmen, die ein Verkehrszeichen beschreiben, sowie das Verkehrszeichen selbst.

Verkehrszeichen In der deutsche Straßenverkehrsordnung sind die Verkehrszeichen mit einer dreistelligen Zahl kodiert. Diese Kodierung wird hier übernommen und zur Referenz auf ein Verkehrszeichen benutzt. Die Menge der Verkehrszeichen sei mit VZ bezeichnet.

Konturen Die Konturen eines Verkehrszeichens werden mit Hilfe eines Winkel/Längen-Kodes beschrieben. Diese Darstellung wird in 3.11 noch näher beschrieben. Eine Ausnahme bildet die Kreiskontur sowie eine Kontur in Form eines Kreisbogens. Diese stellen spezielle Konturen dar und werden gesondert betrachtet. Die Menge der Konturbeschreibungen wird im Folgenden mit Θ bezeichnet.

Piktogramme Ein Piktogramm kann aus einem einzigen, aber auch aus mehreren Objekten bestehen¹⁶. Die Menge der Piktogramme wird im Folgenden mit Ω bezeichnet.

Lagerrelationen Eine Lagerrelation ist zwischen zwei Bildelementen (Kontur oder Piktogramm) definiert. Sie bestimmt beispielsweise das Größenverhältnis zwischen den Bildelementen. Eine Lagerrelation wird durch eine Funktion $f: \Gamma \times \Gamma \rightarrow [0,1]$ dargestellt. Die Funktion berechnet die Ähnlichkeit der Lagerrelation von zwei Formprimitiven, zu einer implizit durch die Funktion dargestellten Musterrelation. Eine genauere Erläuterung dieser Funktionen folgt in 3.4.1.3. Die Menge der Lagerrelationen wird im Folgenden mit Δ bezeichnet.

2. Netzwerkknoten

Es gibt drei Knotentypen sowie einen ausgezeichneten Knoten s , der keinem der Knotentypen angehört und keine Information speichert. Der Knoten s wird als Startknoten bezeichnet und bildet den Ausgangspunkt für jede bei der Klassifikation durchgeführte Abbildung eines Segments auf das Netzwerk.

¹⁶Piktogrammerkennung wurde bei der Implementierung nicht durchgeführt, eine Schnittstelle zu einer Piktogrammklassifikationskomponente ist jedoch vorhanden. Die Piktogramme werden so behandelt, als existiere ein Piktogrammklassifikator, der diesen Ähnlichkeitswerte zuweist. Näheres zu Piktogrammen in 3.4.1.3.

Der Knotentyp CN ¹⁷ Ein Knoten des Typs CN speichert Wissen über die Kontur eines Objekts. Jedem Knoten $x \in \text{CN}$ ist eine Kontur $k \in \Theta$ zugeordnet.

Der Knotentyp PN ¹⁸ Ein Knoten des Typs PN speichert strukturelles Wissen über ein Piktogramm. Jedem Knoten $x \in \text{PN}$ ist ein Piktogramm $p \in \Omega$ zugeordnet.

Der Knotentyp SN ¹⁹ Ein Knoten des Typs SN steht für ein Verkehrszeichen. Jedem Knoten $x \in \text{SN}$ ist ein Verkehrszeichen $v \in \text{VZ}$ zugeordnet.

Durch eine Funktion V wird jedem Knoten seine Bedeutung, also jeweils das repräsentierte Verkehrszeichen, die Kontur oder das Piktogramm, zugeordnet.

3. Netzwerkkanten

Das Netzwerk enthält Kanten verschiedenen Typs, die sich durch die an den Kanten gespeicherte Information, sowie durch die Knotentypen, die durch die Kanten verbunden werden dürfen, unterscheiden.

Der Kantentyp E ²⁰ Die Kanten des Typs E verbinden den Startknoten s mit Netzwerkknoten des Typs CN. Wie der Startknoten tragen auch sie keine Information, bestimmen aber durch die Verbindung des Knotens s mit dem übrigen Netzwerk den Abbildungsvorgang.

Der Kantentyp RL ²¹ Eine Kante vom Typ RL verbindet zwei Knoten vom Typ CN oder PN. Jeder Kante $x \in \text{RL}$ ist eine Lagerrelation $f_x \in \Delta$ zugeordnet.

Der Kantentyp CL ²² Eine Kante vom Typ CL verbindet einen Knoten des Typs SN mit einem Knoten des Typs CN oder PN. Außerdem verbindet eine solche Kante einen Knoten des Typs SN mit einer Kante des Typs RL²³. Jedem CL-Knoten ist ein Gewicht $w \in \text{Nat}^+$ ²⁴ zugeordnet.

¹⁷ component node

¹⁸ picture node

¹⁹ sign node

²⁰ entrynode

²¹ relation link

²² component link

²³ Eine CL-Kante verbindet einen Knoten, der ein Verkehrszeichen repräsentiert mit den Netzwerkkomponenten, die dessen strukturelle Beschreibung darstellen. Dazu gehören sowohl Knoten als auch die RL-Kanten, wodurch es zu der ungewöhnlichen Kante als Verbindung zwischen Knoten und Kante kommt.

²⁴ Menge der natürlichen Zahlen

Auch den RL-Kanten wird durch die Funktion V die durch sie repräsentierte Lagerrelation zugeordnet.

4. Das Netzwerk

Ein Verkehrszeichennetz N besteht aus einer Knotenmenge N_N und einer gerichteten Kantenmenge L_N . Dabei untergliedert sich die Knotenmenge N_N in disjunkte Teilmengen verschiedener Knotentypen.

$$N_N = \{s_N\} \cup SN_N \cup CN_N \cup PN_N$$

Die Kantenmenge untergliedert sich ebenfalls in drei disjunkte Teilmengen verschiedener Kantentypen.

$$L_N = E_N \cup RL_N \cup CL_N$$

Der Knotentyp bestimmt die Art der Information, die in einem Knoten abgelegt werden soll. Dabei wird jedem modellierten Verkehrszeichen ein Knoten der Menge SN zugeordnet, jeder in irgendeinem Verkehrszeichen vorkommenden Kontur ein Knoten der Menge CN und jedem Piktogramm ein Knoten aus PN . Für die Kantentypen gilt:

$$\begin{aligned} E &= \{(s, x) \mid x \in CN\} \\ RL &= \{(x, y) \mid x, y \in CN \cup PN\} \\ CL &= \{(x, y) \mid x \in SN, y \in CN \cup PN \cup RL\} \end{aligned}$$

Wie bereits in 3.1 angesprochen, kann von einer Aufteilung des Netzes in zwei Ebenen gesprochen werden. Die untere Ebene wird durch das Strukturnetz dargestellt, die obere Ebene durch eine Menge von Verkehrszeichenknoten, die untereinander nicht verbunden sind.

Definition : Strukturnetz

Ein Strukturnetz S_N ist das Subnetz eines Verkehrszeichennetzes N , welches durch die Knotenmenge $CN_N \cup PN_N$ und die Kantenmenge RL_N aufgespannt wird.

Repräsentation eines Verkehrszeichens Ein einzelnes Verkehrszeichen wird durch ein Subnetz des gesamten Verkehrszeichennetzes dargestellt. Die Beschreibung eines Verkehrszeichens besteht aus den Konturen, aus denen es sich zusammensetzt, sowie den Lagerrelationen, die zwischen einzelnen Konturen bestehen. Als eigenes Objekt wird das bei den meisten Verkehrszeichen vorhandene Piktogramm

behandelt, das hier vorerst nur mit seinem Namen referenziert werden soll.

Definition : Verkehrszeichensubnetz

Sei $x \in VZ$, so bezeichnet $VS(x)$ das zu x gehörige Verkehrszeichensubnetz. Ein Verkehrszeichensubnetz VS eines Verkehrszeichennetzes N besteht aus einer Knotenmenge N_{VS} und einer Kantenmenge L_{VS} .

$$\begin{aligned} N_{VS} &= PN_{VS} \cup CN_{VS} \cup \{v\} \\ L_{VS} &= CL_{VS} \cup RL_{VS} \end{aligned}$$

Dabei gilt: $v \in SN_N$, $CN_{VS} \subseteq CN_N$, $PN_{VS} \subseteq PN_N$, $RL_{VS} \subseteq RL_N$, $CL_{VS} \subseteq CL_N$. Durch den Verkehrszeichenknoten v , mit $V(v) = x$, werden die Knoten und Kanten des Verkehrszeichensubnetzes $VS(x)$ festgelegt. Das Subnetz besteht aus allen CL-Kanten die mit v verbunden sind und allen Knoten und Kanten, die durch diese CL-Kanten mit v verbunden sind. Die Knoten- und Kantenmengen CL_{VS} , RL_{VS} , CN_{VS} und PN_{VS} sind demnach definiert durch:

$$\begin{aligned} CL_{VS} &= \{e \in CL \mid e = (a, b) \wedge a = v\} \\ CN_{VS} &= \{n \in CN \mid \exists e \in CL_{VS}, e = (a, b) : b = n\} \\ PN_{VS} &= \{n \in PN \mid \exists e \in CL_{VS}, e = (a, b) : b = n\} \\ RL_{VS} &= \{r \in RL \mid \exists e \in CL_{VS}, e = (a, b) : b = r\} \end{aligned}$$

Das Subnetz von VS , bestehend aus den Knoten CN_{VS} und PN_{VS} sowie den Kanten RL_{VS} ist zusammenhängend, d.h. je zwei Knoten sind voneinander erreichbar. Bereits das Subnetz von VS , bestehend aus den Knoten CN_{VS} und PN_{VS} sowie der Kantenmenge $\{k \mid k \in RL_{VS}, V_{RL}(k) = \text{enthält}\}$, ist zusammenhängend und erfüllt insbesondere die Eigenschaften eines Baumes.

Da die CL-Kanten für jeden Knoten $v \in SN$ eine eindeutige Zuordnung der durch v repräsentierten Struktur geben, kann ein Verkehrszeichen auch in Form einer Liste seiner Komponenten dargestellt werden. Dabei wird die Reihenfolge der referenzierten Komponenten bei der Deklaration des Zeichens zu dem Netzwerk festgelegt. Es ergibt sich folgende Darstellung eines Verkehrszeichens:

$$v = ((k_1, \dots, k_{n+m}), (w_1, \dots, w_{n+m}))$$

Dabei gilt: $k_1, \dots, k_{n+m} \in CN_{VS} \cup PN_{VS} \cup RL_{VS}$, $w_1, \dots, w_{n+m} \in Nat^+$