

Institut für Photogrammetrie und Kartographie

Automatische Extraktion urbaner Straßennetze aus Luftbildern

Stefan Hinz

Vollständiger Abdruck
der von der Fakultät für Bauingenieur– und Vermessungswesen
der Technischen Universität München
zur Erlangung des akademischen Grades eines
Doktor–Ingenieurs (Dr.–Ing.)
genehmigten Dissertation.

Vorsitzender:

Univ.-Prof. Dr. phil. nat. Markus Rothacher

Prüfer der Dissertation:

1. Univ.-Prof. Dr.-Ing. Heinrich Ebner
2. Univ.-Prof. Dr.-Ing. Helmut Mayer,
Universität der Bundeswehr München

Die Dissertation wurde am 30. Oktober 2003 bei der Technischen Universität München eingereicht und durch die Fakultät für Bauingenieur– und Vermessungswesen am 10. Dezember 2003 angenommen.

für Daniel

The eye sees only what the mind is prepared to comprehend.

(Robertson Davies)

Zusammenfassung

In dieser Dissertation wird ein neuer Ansatz zur automatischen Extraktion urbaner Straßennetze aus Luftbildern entwickelt. Um die hohe Komplexität urbaner Szenen handhaben zu können, wird ein explizit formuliertes Straßenmodell aufgestellt, das sowohl detailliertes als auch abstrahiertes Wissen über Straßen integriert. Die Modellierung von lokalem und globalem Kontext ergänzt das Straßenmodell.

Das Wissen, in welchem Stadium der Extraktion bestimmte Modellkomponenten den größten Nutzen für die Extraktion bringen, ist in einer Extraktionsstrategie zusammengefasst. Die Grobgliederung der Strategie umfasst drei Phasen: Im Rahmen der Kontextanalyse (Phase 1) werden die globalen Kontextgebiete „Wald“, „Siedlung“ und „offene Landschaft“ segmentiert und lokale Kontextinformation wie z.B. Schattenregionen und Verdeckungen durch Fahrzeuge analysiert. Die Extraktion setzt sich in den Siedlungsgebieten mit der Bestimmung von sicheren Straßenstücken fort (Phase 2). Die Bestimmung dieser Teilstücke des Straßennetzes beginnt mit der Detektion von homogenen Bändern in auflösungsreduzierten Versionen der Luftbilder und der Extraktion von Markierungen und Straßenrändern in der hohen Auflösung. Aus dieser Eingangsinformation werden durch mehrere Gruppierungsoperationen Fahrspuren und Fahrbahnen gebildet. Die Straßennetzbildung in (Phase 3) verfolgt schließlich das Ziel, Fahrbahnen aus mehrfach überlappten Bildern zu fusionieren und daraus ein zusammenhängendes Netz zu erzeugen. Zu diesem Zweck kommt ein iterativer Prozess zum Einsatz, der zwischen den fusionierten Fahrbahnen Verknüpfungshypothesen generiert, die anschließend einer Verifikation im Bild unterzogen werden. Das Resultat dieser Gruppierung sind Straßen, die Kreuzungs- und Endpunkte des Netzes miteinander verbinden.

Ein besonderes Merkmal der vorliegenden Arbeit ist die Entwicklung und systematische Nutzung eines Schemas zur internen Bewertung von Extraktionsergebnissen. Die interne Bewertung ermöglicht es, während der Extraktion auf die momentan am erfolgversprechendsten Hypothesen zu fokussieren.

Die Ergebnisse zeigen, dass das implementierte System in der Lage ist, Straßen in komplexen Umgebungen automatisch zu extrahieren. Anhand einer externen Bewertung mit manuell erfassten Referenzdaten wird die Leistungsfähigkeit des Systems validiert sowie Vorteile und Defizite diskutiert.

Summary

In this dissertation, a new approach to automatic road extraction from high resolution aerial imagery taken over urban areas is developed. In order to deal with the high complexity of this type of scenes, the approach integrates detailed knowledge about roads and their context using explicitly formulated scale-dependent models.

The knowledge about how and when certain parts of the road and context model are optimally exploited is expressed by an extraction strategy. It is subdivided into three levels: Context-based data analysis (Level 1) comprises the segmentation of the scene into the urban, rural, and forest area as well as the analysis of context relations (e.g. the determination of shadow areas and the detection of vehicles). Processing continues in urban areas. To determine salient roads (Level 2) the extraction includes the detection of homogeneous ribbons as preliminary road segments in coarse scale, collinear grouping thin bright road markings in fine scale, and the construction of lanes from groups of road markings and road sides. Then, the lanes are further grouped into larger road objects, so-called road segments. During road network completion (Level 3), road segments detected in overlapping images are fused and gaps in the extraction are iteratively closed by hypothesizing and verifying connections between previously extracted road segments. To this end, local as well as global criteria exploiting the network characteristics are used. The result of network completion are roads that connect junctions or endpoints.

A key feature of the presented approach is the incorporation of a scheme for self-diagnosis. With this scheme each hypothesis generated during extraction is internally evaluated so that its relevance for further processing can be assessed. This facilitates decisions inherently appearing during the extraction process.

The results achieved with the approach show that the implemented system is able to extract roads in complex environments, i.e. the extraction is also possible when the appearance of roads is heavily affected by other objects. Based on an external evaluation of the results, the system is validated, and advantages but also remaining deficiencies are discussed.

Inhaltsverzeichnis

1	Einführung	1
1.1	Praktische und wissenschaftliche Relevanz der Thematik	1
1.2	Wissenschaftliches Umfeld, Zielsetzung und Schwerpunkte der Arbeit	2
1.3	Gliederung und Hinweise zum Lesen der Arbeit	3
 Teil I: Konzept zur Extraktion urbaner Straßennetze		5
2	Grundlagen und verwandte Forschungsarbeiten	5
2.1	Grundlagen der Bildanalyse	5
2.1.1	<i>Zielsetzung der Bildanalyse</i>	5
2.1.2	<i>Komplexität von Luftbildern</i>	7
2.1.3	<i>Modellbildung</i>	8
2.1.4	<i>Strategie</i>	11
2.2	Verfahren zur automatischen Straßenextraktion	14
2.2.1	<i>Straßenextraktion aus niedrig aufgelösten Bilddaten</i>	14
2.2.2	<i>Straßenextraktion aus mittel und hoch aufgelösten Bilddaten</i>	15
2.3	Diskussion und Folgerungen	19
2.3.1	<i>Analyse der Modelle</i>	19
2.3.2	<i>Analyse der Strategien</i>	21
3	Modellierung	25
3.1	Datengrundlage und Sensormodell	25
3.2	Straßenmodell	27
3.2.1	<i>Komponenten und Eigenschaften des Objektmodells</i>	29
3.2.2	<i>Komponenten und Eigenschaften des Bildmodells</i>	30
3.2.3	<i>Repräsentation des Straßenmodells</i>	35
3.3	Kontextmodell	38
3.3.1	<i>Lokaler Kontext</i>	39
3.3.2	<i>Globaler Kontext</i>	40
4	Interne Bewertung von Extraktionsergebnissen	42
4.1	Bedeutung und Ziele	42
4.1.1	<i>Interne Bewertung zur Nutzerführung</i>	42

4.1.2	<i>Interne Bewertung zur Steuerung der Extraktion</i>	43
4.2	Ansätze zur internen Bewertung	44
4.3	Rahmenkonzept	46
4.3.1	<i>Modellierung und interne Bewertung</i>	46
4.3.2	<i>Nutzung interner Bewertung</i>	48
5	Extraktionsstrategie – Überblick	50
 Teil II: Realisierung der Straßenextraktion		55
6	Kontextanalyse	55
6.1	Globaler Kontext	56
6.2	Lokaler Kontext	57
6.2.1	<i>Gebäude und hohe Vegetation</i>	57
6.2.2	<i>Fahrzeuge und Fahrzeugkolonnen</i>	60
7	Straßenobjekte	68
7.1	Vorläufige Straßenabschnitte	68
7.2	Fahrspurbegrenzungen	72
7.2.1	<i>Durchgezogene Markierungen</i>	72
7.2.2	<i>Strichlierte Markierungen</i>	73
7.2.3	<i>Straßenränder</i>	75
7.3	Fahrspuren	76
7.3.1	<i>Konstruktion von Fahrspuren</i>	77
7.3.2	<i>Fusion und Gruppierung von Fahrspuren</i>	81
7.4	Fahrbahnen	84
7.4.1	<i>Konstruktion von Fahrbahnen</i>	88
7.4.2	<i>Gruppierung von Fahrbahnen</i>	96
8	Fusion und Netzbildung	100
8.1	Fusion	100
8.2	Netzbildung	104
8.2.1	<i>Generierung und Verifikation von Fahrbahnverknüpfungen und Kreuzungen</i>	105
8.2.2	<i>Netzanalyse</i>	107

Teil III: Diskussion der Ergebnisse und Ausblick	115
9 Externe Bewertung	115
9.1 Bedeutung der Referenzdaten	115
9.2 Bewertungsschema	116
10 Ergebnisse und Diskussion	119
10.1 Beispieldatensätze und Analyse der Extraktionsergebnisse	119
10.2 Allgemeine Diskussion und Vergleich mit anderen internationalen Arbeiten	129
11 Zusammenfassung und Ausblick	134
Literatur	137
Danksagung	143

1 Einführung

Thema dieser Dissertation ist die *automatische Extraktion urbaner Straßennetze aus Luftbildern*. Diese Thematik erfährt sowohl von praktischer als auch von wissenschaftlicher Seite aus eine außerordentliche Relevanz:

Straßen sind ein wesentlicher Bestandteil unseres Lebensraumes. Durch ihre generelle Funktion, Verbindungen herzustellen, erschließen sie einen Großteil des menschlichen Lebensraumes. Sie ermöglichen Kommunikation, Transport und Mobilität, was sie zu einer tragenden Säule der heutigen Wirtschafts- und Gesellschaftsform macht. Detaillierte und aktuelle Daten über das Straßennetz – seien sie analog in Form von Karten oder digital in Geoinformationssystemen (GIS) gespeichert – sind daher als grundlegende Voraussetzung für die Erhaltung oder gar Steigerung der wirtschaftlichen Leistungsfähigkeit und der im Zeitgeist des 20. und 21. Jahrhunderts verankerten Mobilität zu sehen (Konecny, 1995). Dies spiegelt sich z.B. in modernen Anwendungen wie Routenplanern, Fahrzeugnavigationssystemen oder LKW-Flottenmanagementsystemen wider. Neben diesen verkehrsbezogenen Anwendungen sind Straßendaten oft auch integrativer Bestandteil von Analysen und Simulationen in der Raumplanung. Dies gilt in besonderer Weise für Daten, die das *urbane* Straßennetz beschreiben: In keinem anderen Raum ist das Straßennetz so engmaschig wie in der Stadt. Dies hat zur Folge, dass Begleitscheinungen von Straßen wie Luftverschmutzung und Lärmbelastung in urbanen Gebieten verstärkt auftreten und den Menschen direkt betreffen. Daher sind viele Projekte der Raum- und Stadtplanung auf Informationen über das urbane Straßennetz angewiesen. Die Relevanz solcher Daten für praktische Anwendungen wurde in einer von der OEEPE (Organisation Européenne d'Etudes Photogrammétriques Expérimentales) durchgeführten Umfrage über 3D-Stadtmodelle quantitativ erfasst. Etwa 85% der Teilnehmer gaben an, dass Straßendaten für ihre jeweilige Anwendung von maßgeblicher Bedeutung sind (Fuchs et al., 1998).

1.1 Praktische und wissenschaftliche Relevanz der Thematik

Das in der Praxis vorrangig eingesetzte Verfahren zur (Neu-)Erfassung und Aktualisierung von Straßendaten ist die Auswertung von Luftbildern durch einen Operateur (Englisch und Heipke, 1998). Die hohe Zuverlässigkeit der Ergebnisse wird jedoch durch einen großen Aufwand an Arbeitszeit erkaufte. Insbesondere wenn hohe Ansprüche hinsichtlich Aktualität, Detaillierungsgrad und Konsistenz der Daten zu erfüllen sind, wie dies für die oben skizzierten Anwendungsfelder der Fall ist, wird die manuelle Auswertung zu einem Hemmnis für wirtschaftliche Erfassungs- und Aktualisierungszeiten. Daher besteht seitens der Praxis der Wunsch nach einer möglichst weitreichenden Automatisierung der Erfassung und Aktualisierung von Straßendaten (Illert, 1993). Dies setzt jedoch voraus, dass das interessierende Objekt „Straße“ vom Computer möglichst selbständig im Bild erkannt bzw. aus dem Bild *extrahiert* wird.

Aus wissenschaftlicher Sicht fällt die Thematik der Straßenextraktion in das Forschungsgebiet der Bildanalyse. Deren weitgefasstes Ziel ist es, die dem Menschen ureigene Fähigkeit, seine Umwelt mit den Augen zu erfassen, sie nach Objekten zu strukturieren und die Objekte mit Semantik zu versehen, mit technischen Mitteln zu realisieren. Je geringer dabei die Interaktion des Menschen sein soll, desto umfangreicheres Wissen über die reale Welt muss im Computer in Form eines *Modells* repräsentiert und durch entsprechende Algorithmen genutzt werden. Die Aufgabe der vollautomatischen Objektextraktion hat sich vor allem dann als sehr anspruchsvoll erwiesen, wenn keine laborähnlichen Bedingungen wie bei vielen industriellen Anwendungen¹ vorausgesetzt werden können, sondern die abgebildete Szene durch hohe Komplexität gekennzeichnet ist, wie das bei urbanen Szenen in Luftbildern der Fall ist. Zur Zeit kann für solche, aber größtenteils auch für einfachere Szenen, die automatische Objektextraktion aus Luftbildern nicht als gelöst betrachtet werden. Trotz vieler erfolgversprechender Forschungsarbeiten müssen Fragestellungen wie die computergerechte Formalisierung und

¹ Hier steht dafür der hohe Anspruch einer echtzeitfähigen Realisierung im Vordergrund.

Nutzung des Wissens über Objekte und ihre Beziehungen untereinander oder wie die automatische Steuerung des Extraktionsprozesses wissenschaftlich weiter durchdrungen werden. In der vorliegenden Arbeit soll dies für den speziellen Fall des urbanen Straßennetzes erfolgen.

1.2 Wissenschaftliches Umfeld, Zielsetzung und Schwerpunkte der Arbeit

Wie man aus den Pionierarbeiten (Bajcsy und Tavakoli, 1976) ersehen kann, ist die automatische Straßenextraktion aus Luft- und Satellitenbildern seit mehr als 25 Jahren Gegenstand von intensiven Forschungsanstrengungen. Die Entwicklungen über diesen Zeitraum zeigen zum einen die starke Vertiefung des Wissens. Zum anderen wurde aber die enorme Komplexität dieser Aufgabe sichtbar. Deswegen ist mittelfristig nicht zu erwarten, dass Extraktionsverfahren ohne menschliche Interaktion die Ansprüche eines operationellen Einsatzes hinsichtlich Vollständigkeit und Zuverlässigkeit der Ergebnisse erreichen werden. Auch bei weniger komplexen Szenen als Stadtgebieten scheint zur Zeit eine Überwachung oder Nachbearbeitung der Extraktion durch einen Operateur unvermeidbar (Mayer, 1998).

Als Konsequenz dieser Problematik haben sich zwei grundsätzlich verschiedene Entwicklungsstränge in der Erforschung der Straßenextraktion herauskristallisiert: Auf der einen Seite steht die praxisorientierte Forschung, deren Ziel die schnelle Entwicklung von hocheffizienten und operationell einsetzbaren Verfahren ist. Hierbei handelt es sich um nutzergeführte Ansätze², die dadurch charakterisiert sind, dass der Computer Aufgaben mit einem geringeren Anteil an Semantik durchführt und semantisch anspruchsvollere Aufgaben, wie Entscheidungen über den Verfahrensablauf, dem Operateur vorbehalten bleiben. Typischerweise benötigt solch ein Verfahren eine manuelle Initialisierung durch Vorgabe von ausgewählten Punkten auf dem Straßennetz. Aufgabe des Computers ist es dann, unter Berücksichtigung straßentypischer Kriterien die Straße von diesen Startpunkten aus automatisch zu verfolgen (*Tracking*), bzw. die Startpunkte optimal zu verbinden (*Pfadoptimierung*). Können die Kriterien im Laufe der Berechnungen nicht eingehalten werden, bricht das Verfahren ab und erwartet vom Operateur Vorgaben für weitere Aktionen. Neben einer Minimierung der nötigen Anzahl an Nutzerinteraktionen wurde in den letzten Jahren einer möglichst effizienten Gestaltung der Schnittstelle Mensch versus Computer zunehmend Aufmerksamkeit geschenkt. Bekannte Vertreter für Tracking-Ansätze sind (McKeown und Denlinger, 1988; Vosselman und de Knecht, 1995; Geman und Jedynek, 1996; Baumgartner et al., 2002), und für Pfadoptimierung (Grün und Li, 1997; Merlet und Zerubia, 1996).

Im Gegensatz zu diesen Arbeiten behalten Ansätze der grundlagenorientierten Forschung den Anspruch vollständiger Automation bei. Die hier erzielten Ergebnisse können mit jenen der nutzergeführten Ansätze nicht konkurrieren. Der wissenschaftliche Erkenntnisgewinn liegt vielmehr in der eingehenden Analyse, welche Teile der Straßenextraktion ohne Eingriff eines Operateurs zu lösen sind und welche Gründe für ein Scheitern der Extraktion an bestimmten Stellen verantwortlich sind. Die Identifikation, Eingrenzung und Dokumentation solcher Problemstellen bereitet eine fundierte Basis für neue Forschungsarbeiten und letztlich für neue, leistungsfähigere nutzergeführte Ansätze. So konnten in den letzten Jahren für die Straßenextraktion in offenen, außerstädtischen Bereichen wichtige Anhaltspunkte gewonnen werden, welche Komponenten das Straßenmodell enthalten muss und welche Strategie zur Extraktion verwendet werden muss, um relativ vollständige und zuverlässige Ergebnisse zu erhalten (Ruskoné, 1996; Wiedemann, 2001; Baumgartner, 2003). Diese Fragen sind für die Extraktion des Straßennetzes in urbanen Gebieten bisher nur rudimentär bearbeitet worden.

Im Licht der grundlagenorientierten Forschung ist auch die **Zielsetzung** der vorliegenden Arbeit zu sehen. Es wird ein Verfahren zur Extraktion von Straßen aus Luftbildern entwickelt und analysiert, das folgende Randbedingungen erfüllt:

² Häufig wird dies im Fachjargon auch „semi-automatisch“ oder „halb-automatisch“ bezeichnet.

- ◇ Das Verfahren ist vollautomatisch.
- ◇ Das Verfahren ist zur Behandlung komplexer urbaner Szenen geeignet.
- ◇ Das Verfahren liefert eine detaillierte Beschreibung des urbanen Straßennetzes inklusive der Topologie des Netzes.

Dominierender Gesichtspunkt der Arbeit ist die Handhabung der hohen Komplexität urbaner Szenen. Diese entsteht vor allem durch die Vielfalt der Objekte und ihr Beziehungsgeflecht untereinander. Von entscheidender Bedeutung ist in diesem Zusammenhang der *Straßenkontext*. Durch ihn werden einerseits Beziehungen zwischen dem Objekt „Straße“ und den unmittelbar benachbarten Objekten erfasst, andererseits wird auch die Abhängigkeit der Gestalt und Erscheinungsform einer Straße von der großräumigen Umgebung berücksichtigt. Als Konsequenz ergeben sich folgende drei **Schwerpunkte** der Arbeit:

- ◇ *Modellbildung*: Erstellung eines Modells, das urbane Straßen in ausreichendem Detaillierungsgrad beschreibt und ihren Kontext integriert.
- ◇ *Strategie*: Konzipierung und Implementierung einer Ablaufsteuerung für den Extraktionsprozess. Sie muss ein hohes Maß an Robustheit gegenüber störenden Einflüssen gewährleisten und gleichzeitig Flexibilität hinsichtlich des Kontexts bieten.
- ◇ *Methodik*: Entwicklung und Anpassung von Algorithmen, die in spezialisierter Weise Komponenten des Modells aus dem Bild extrahieren und damit die Grundelemente des Extraktionsprozesses bereitstellen.

Die Interaktion des entwickelten Ansatzes mit GIS-Daten wird zwar im Laufe der Arbeit mehrmals angeschnitten, eine tiefer gehende Untersuchung wird jedoch nicht durchgeführt. Sowohl die Integration von GIS-Daten als Vorinformation für die Straßenextraktion als auch die Aktualisierung von existierenden GIS-Straßendaten durch die Ergebnisse der Extraktion sind hochaktuelle Problemstellungen (de Gunst, 1996; Bordes et al., 1997; Doucette et al., 1999; Zhang et al., 2001; Willrich, 2002). Im Kern sind jedoch beide Aufgaben direkt von den Möglichkeiten und Grenzen einer automatischen Extraktion abhängig. Daher ist eine Fokussierung der vorliegenden Arbeit auf die Thematik der Extraktion naheliegend und gerechtfertigt.

1.3 Gliederung und Hinweise zum Lesen der Arbeit

Der Aufbau der Arbeit ist so strukturiert, dass der Hauptteil einer Dreigliederung unterliegt:

- ◇ **Teil I** entwickelt in vier Kapiteln das Konzept zur Extraktion urbaner Straßennetze. Aufbauend auf der Betrachtung relevanter Ansätze zur Straßenextraktion in Kapitel 2 wird in Kapitel 3 das Straßen- und Kontextmodell erläutert. In Kapitel 4 wird das für die Extraktionsstrategie wichtige Konzept zur internen Bewertung von Extraktionsergebnissen eingeführt, bevor in Kapitel 5 ein Gesamtüberblick über die Extraktionsstrategie gegeben wird. Eine kurze Zusammenfassung schließt Teil I ab.
- ◇ **Teil II** geht auf die Realisierung der Straßenextraktion ein. Die Kapitel 6, 7 und 8 beschreiben die entwickelten Verfahren zur Kontextanalyse, zur Extraktion und Fusion von Straßen(teil)objekten sowie zur Straßennetzbildung. Die Darstellung hält sich größtenteils an die zugrunde liegende Extraktionsstrategie.
- ◇ **Teil III** befasst sich mit der Bewertung und Diskussion der erzielten Ergebnisse. Als Grundlage wird in Kapitel 9 kurz ein etabliertes Schema zur externen Bewertung von Ergebnissen einer Straßenextraktion erläutert. Daran schließt sich in Kapitel 10 die Diskussion und Bewertung der Leistungsfähigkeit des Verfahrens anhand ausgewählter Beispiele an. Auf Grundlage der erreichten Ziele und der noch offenen Punkte wird der Stand des entwickelten Verfahrens auch im internationalen Vergleich erörtert. Im letzten Kapitel (Kapitel 11) werden die wesentlichen Aspekte der Arbeit zusammengefasst. Ein Ausblick auf mögliche Erweiterungen schließt die Arbeit ab.

Diese Form der Gliederung wurde gewählt, um dem Leser eine Führung durch die Dissertation an die Hand zu geben. Der konzeptionell interessierte Leser befaße sich vor allem mit Teil I, wobei der in Bildanalyse und Straßenextraktion kundige Leser direkt mit Kapitel 3 beginnen kann. Während der an der technischen Realisierung interessierte Leser mit Teil II fortfährt, kann der eilige Leser direkt zu Teil III übergehen. Dort ist die Leistungsfähigkeit des Verfahrens dargestellt.

Teil I:

Konzept zur Extraktion urbaner Straßennetze

2 Grundlagen und verwandte Forschungsarbeiten

Die Publikationen von Forschungsarbeiten zur Straßenextraktion aus Luft- und Satellitenbildern reichen bis in die Mitte der 70er Jahre des vergangenen Jahrhunderts zurück. Trotz der Vielzahl und Heterogenität der publizierten Ansätze ist der überwiegenden Mehrheit ein wesentliches Merkmal gemeinsam: Sie wurden für Bildmaterial entwickelt und getestet, das hauptsächlich offene, ländlich geprägte Gebiete darstellt. Die Gründe hierfür sind vielschichtig: Höhere Erfolgsaussichten auf Grund der geringeren Szenenkomplexität, Vorgaben seitens der Forschungsförderung, spezielle Interessenslagen der Forschergruppen, usw. Da die vorliegende Arbeit auf die Extraktion des urbanen Straßennetzes ausgerichtet ist, verfolgt dieses Kapitel neben der Darstellung und Charakterisierung bestehender Ansätze das Ziel, die Frage nach der Übertragbarkeit der verwendeten Modelle, Strategien und Methoden auf die Straßenextraktion in urbanen Gebieten zu beantworten. Entsprechend soll geklärt werden, in welchen Bereichen Weiterentwicklungen nötig sind.

Im ersten Abschnitt dieses Kapitels werden einige wichtige und im Verlauf dieser Arbeit wiederkehrende Prinzipien und Grundbegriffe der Bildanalyse erläutert. Der zweite Abschnitt gibt einen Überblick über die momentan am weitesten entwickelten Ansätze zur Straßenextraktion. Im dritten Abschnitt werden spezielle Aspekte der Ansätze näher beleuchtet und hinsichtlich ihrer Eignung für die Straßenextraktion in urbanen Gebieten analysiert und diskutiert. Zugunsten der Übersichtlichkeit konzentriert sich die Darstellung auf repräsentative und im Rahmen dieser Arbeit wesentliche Ansätze. Aus anderen Verfahren, z.B. nutzergeführten Ansätzen und Verfahren, die Vorinformation aus GIS-Daten heranziehen, werden nur bestimmte Gesichtspunkte aufgegriffen, sofern sie für die Extraktion des urbanen Straßennetzes relevant sind. Ein umfassender Überblick über automatische Objektextraktion im Allgemeinen und Straßenextraktion im Speziellen ist in (Mayer, 1998) zu finden. Sammlungen der wichtigsten neuen Arbeiten auf diesen Gebieten sind z.B. (Grün und Nevatia, 1998; Ebner et al., 1999; Baltsavias et al., 2001; Leberl und Kalliany, 2002).

2.1 Grundlagen der Bildanalyse

2.1.1 Zielsetzung der Bildanalyse

Unter Bildanalyse³ versteht man „*the automatic derivation of an explicit meaningful description of physical objects in the real world from images*“ (Rosenfeld, 1982). Die Zielsetzung der Bildanalyse ist also die Erzeugung einer expliziten Beschreibung von Objekten aus der im Speichermedium „Bild“ implizit enthaltenen Information. Diese Aufgabe ist mit der menschlichen Fähigkeit der Bildinterpretation vergleichbar. Die Fragestellungen, wie die menschliche Vorgehensweise der Bildinterpretation aussieht und wie sie sich mittels Computer imitieren lässt, spielen jedoch bei der überwiegenden Mehrheit der Forschungsarbeiten nur eine untergeordnete Rolle.

³ Die Begriffe „Bildanalyse“ (image analysis), „Bildverstehen“ (image understanding), „Computersehen“ (computer vision) oder „Robotersehen“ (robot vision) werden in der Literatur praktisch synonym verwendet.

Auch wenn die menschliche Vorgehensweise nützliche Hinweise geben kann, ist das primäre Ziel das Erreichen des gleichen Interpretationsergebnisses.

Neben den Herausforderungen für die Bildanalyse, die durch die komplexe Struktur der Welt bedingt sind (Abschnitt 2.1.2), ergeben sich weitere Herausforderungen aus den Aufnahme- und Speicherungsprozessen im Zuge der Bildgenerierung. Entscheidend für diese ist zum einen der Sensor, z.B. eine digitale Luftbildkamera, und zum anderen der Informationsträger. Letzterer ist in der Regel die elektromagnetische Strahlung (EMS). Vom Informationsträger hängt es ab, welche Eigenschaften der Objekte der realen Welt überhaupt zum Sensor übermittelt werden können, also im Fall der EMS die Strahlungseigenschaften der Objekte. Der Sensor hingegen bestimmt, wie diese Informationen abgebildet und gespeichert werden. Hierbei findet eine Reihe von geometrischen und physikalischen Transformationen statt, die die ursprüngliche Information grundlegend verändern. Diese umfassen z.B. die geometrische Transformation von 3D nach 2D, die räumliche Diskretisierung der Abbildung durch die CCD-Sensorelemente, die physikalische Wandlung der EMS in Ladung und schließlich deren Quantisierung in natürliche Zahlen.

Überträgt man dies auf den in dieser Arbeit relevanten Fall von digitalen Luftbildern⁴, so ergeben sich hinsichtlich der in den Bildern enthaltenen Informationen folgende Charakteristika:

- ◇ Radiometrische Information: Die gespeicherten Intensitätswerte sind näherungsweise proportional zu den Strahlungseigenschaften der Objekte, allerdings eingeschränkt auf die radiometrische und spektrale Auflösung des Sensors.
- ◇ Geometrische Information: Die ursprünglich kontinuierliche Geometrie der Objekte liegt nur mehr diskret vor. Ihre Dimension ist durch die zentralperspektivische Abbildung von 3D auf 2D reduziert. Dies hat insbesondere zur Folge, dass in bestimmten Bildbereichen die Topologie verändert ist: Im Fall von Verdeckungen sind z.B. zwei benachbarte Bildelemente keine benachbarten Flächenelemente in der realen Welt.
- ◇ Semantische Information: Objekte und ihre Semantik sind allein vom Menschen definiert. Die Semantik „haftet“ nicht an den Objekten im physikalischen Sinne, weshalb die gespeicherten Daten auch keine Semantik der abgebildeten Szene enthalten können.

Die Aufgabe der Bildanalyse ist es einerseits, die Bildgenerierung zu invertieren und andererseits anhand der gespeicherten Abbildung die in der Szene enthaltenen Objekte und Objektrelationen hinsichtlich ihrer Semantik und Eigenschaften zu bestimmen. Als große Hilfe bei der automatischen Bildinterpretation hat sich die vorausgehende oder simultane Rekonstruktion der 3D-Information durch die Verwendung von mehreren Bildern, die verschiedene Sichten auf die Szene liefern, herausgestellt. Bei bekannter Orientierung der Bilder kann – zumindest theoretisch – jeder Objektpunkt, der in zwei oder mehr Bildern enthalten ist, mittels des photogrammetrischen Messprinzips dreidimensional festgelegt werden. Diese Aufgabe ist im Allgemeinfall nicht-semantischer Natur und kann daher relativ zuverlässig automatisiert werden. Noch vorteilhafter ist es, wenn eine zeitgleiche Aufzeichnung der Szene durch einen Sensor mit direkten Distanz- und Richtungsmessungen vorliegt, wie z.B. mit einem Laserscanner.

Die Bestimmung der Semantik der Objekte ist nur durch Integration von zusätzlichem Wissen möglich. Dieses Wissen wird durch ein *Modell* repräsentiert, das eine Spezifizierung des allgemeinen Aufnahme- und Speicherungsprozesses bezüglich der zu extrahierenden Objekte darstellt. Das im Modell enthaltene Wissen bildet die Grundlage für einen Rückschluss von den rohen Intensitätswerten auf die Objekte. Der Formalismus, der einzelne Bildverarbeitungsoperationen und die Reihenfolge ihrer Ausführung festlegt, um ein Objekt anhand des Modells zu extrahieren, wird als *Extraktionsstrategie* oder nur *Strategie* bezeichnet. Da die Strategie die Nutzung des Modell-Wissens bestimmt, hat sie ebenso hohe Relevanz wie das Modell. Die automatische Objektextraktion ruht somit auf zwei Säulen: Auf der einen Seite die computergerechte Formalisierung des Wissens (Modellbildung) und auf der anderen Seite die Nutzung des Wissens für die Extraktion (Strategie).

⁴ Auf eine Unterscheidung zwischen digital aufgezeichneten und gescannten analogen Luftbildern wird hier verzichtet.

Neben Modell und Strategie ist beim Design eines Bildanalyse-Systems auch die Komplexität der zu behandelnden Datengrundlage von entscheidender Bedeutung. Um überhaupt eine erfolgreiche und effiziente Objektextraktion zu ermöglichen, müssen Modell und Strategie auf die Komplexität abgestimmt sein. Daher wird vor der Darstellung der Prinzipien von Modellbildung und Strategie der Begriff der Komplexität von Luftbildern umrissen.

2.1.2 Komplexität von Luftbildern

Der Begriff „Komplexität“ ist im Falle von Bildern eng mit dem visuellen Eindruck des individuellen Betrachters verbunden. Abbildung 1 zeigt hierzu exemplarisch einige Ausschnitte aus Luftbildern. Der Betrachter empfindet die Ausschnitte von ländlichen Gebieten eindeutig als leichter zu interpretieren und damit als weniger komplex als jene von Stadtgebieten. Das subjektive „Empfinden“ von Komplexität macht ihre Quantifizierung – idealerweise mit einer einzigen Maßzahl – außerordentlich schwierig (Shufelt, 1999). Um jedoch zumindest eine qualitative Einordnung zu ermöglichen, können eine Reihe von Faktoren angeführt werden, die mit mehr oder weniger starkem Gewicht zur Komplexität beitragen. Hierbei sind zuallererst der *Inhalt* der Szene in der realen Welt und die vom Sensor vorgegebene *Beobachtbarkeit* des Inhalts zu unterscheiden (Mayer, 1998).

Wichtige Faktoren, die bezogen auf eine Flächeneinheit zu einem komplexen Inhalt einer Szene der realen Welt führen, sind:

- ◇ Anzahl der Objekte und Objektklassen: Je mehr Objekte aus verschiedenen Objektklassen wie z.B. Gebäude, Straßen und Gewässer in der Szene vorkommen, desto inhaltsreicher und variabler ist die Szene.
- ◇ Variabilität innerhalb einer Objektklasse: Unter Variabilität werden hier unterschiedliche Ausprägungen der zu *einer* Klasse gehörigen Objekte verstanden. Die Dachlandschaft von Gebäuden in einem alten Stadtzentrum hat z.B. einen wesentlich höheren Detailreichtum als jene von Neubausiedlungen der Nachkriegszeit.
- ◇ Relationen zwischen Objekten: Je mehr unterschiedliche Relationen zwischen Objekten auftreten, desto variabler ist auch das Erscheinungsbild von Objekten.

Daneben wird die Komplexität durch sensorbedingte Faktoren beeinflusst. Nachfolgend sind die Besonderheiten für den Luftbildfall zusammengestellt. Effekte durch zeitlich versetzte Aufnahmen werden vernachlässigt:

- ◇ Abbildungsgeometrie: Sie bestimmt, welche Teile der Szene aufgenommen werden. So erzeugen Luftbildkameras mit Flächensensoren Bilder mit zentralperspektiver Projektion in alle Richtungen, wohingegen Bilder von Zeilensensoren nur orthogonal zur Flugrichtung eine zentralperspektive, aber in Flugrichtung eine parallele Projektion besitzen.
- ◇ Bildmaßstab/Bodenpixelgröße: Je kleiner die Bodenpixelgröße, umso kleinere Objekte bzw. Details sind beobachtbar und umso höher ist die Komplexität. Im Zusammenhang mit der Straßenextraktion bezeichnet man üblicherweise Bilder mit einer Bodenpixelgröße größer als 1m als *niedrig*, zwischen 0.25m und 1m als *mittel* und kleiner als 0.25m als *hoch* aufgelöst (Mayer, 1998).
- ◇ Spektrale Empfindlichkeit: Je mehr Kanäle zur Verfügung stehen, desto mehr Information über die spektralen Eigenschaften der Objekte sind in der Abbildung enthalten.
- ◇ Radiometrische Auflösung: Sie gibt vor, wie differenziert die Energie der EMS aufgezeichnet werden kann. Während zur Zeit noch 8bit (= 256 Grauwerte) pro Kanal die übliche Quantisierung darstellt, erlaubt die radiometrische Auflösung der meisten neueren Sensoren bereits eine Aufzeichnung mit 12bit und mehr.

In erster Linie erschwert eine höhere Komplexität der Datengrundlage die Bildanalyse. Vielzahl und Variabilität von unterschiedlichen Objekten und Objektklassen sowie Relationen zwischen ihnen führen leicht zu Lücken im Extraktionsergebnis oder zu Fehlinterpretationen. Eine hohe Komplexität ist jedoch nicht zwangsläufig ein Hindernis für die Objektextraktion. Oft sind es gerade die Details eines Objektes und die Relationen zu anderen Objekten, die es sicher und eindeutig identifizierbar machen. Das Teilobjekt „Schornstein“ gibt z.B.

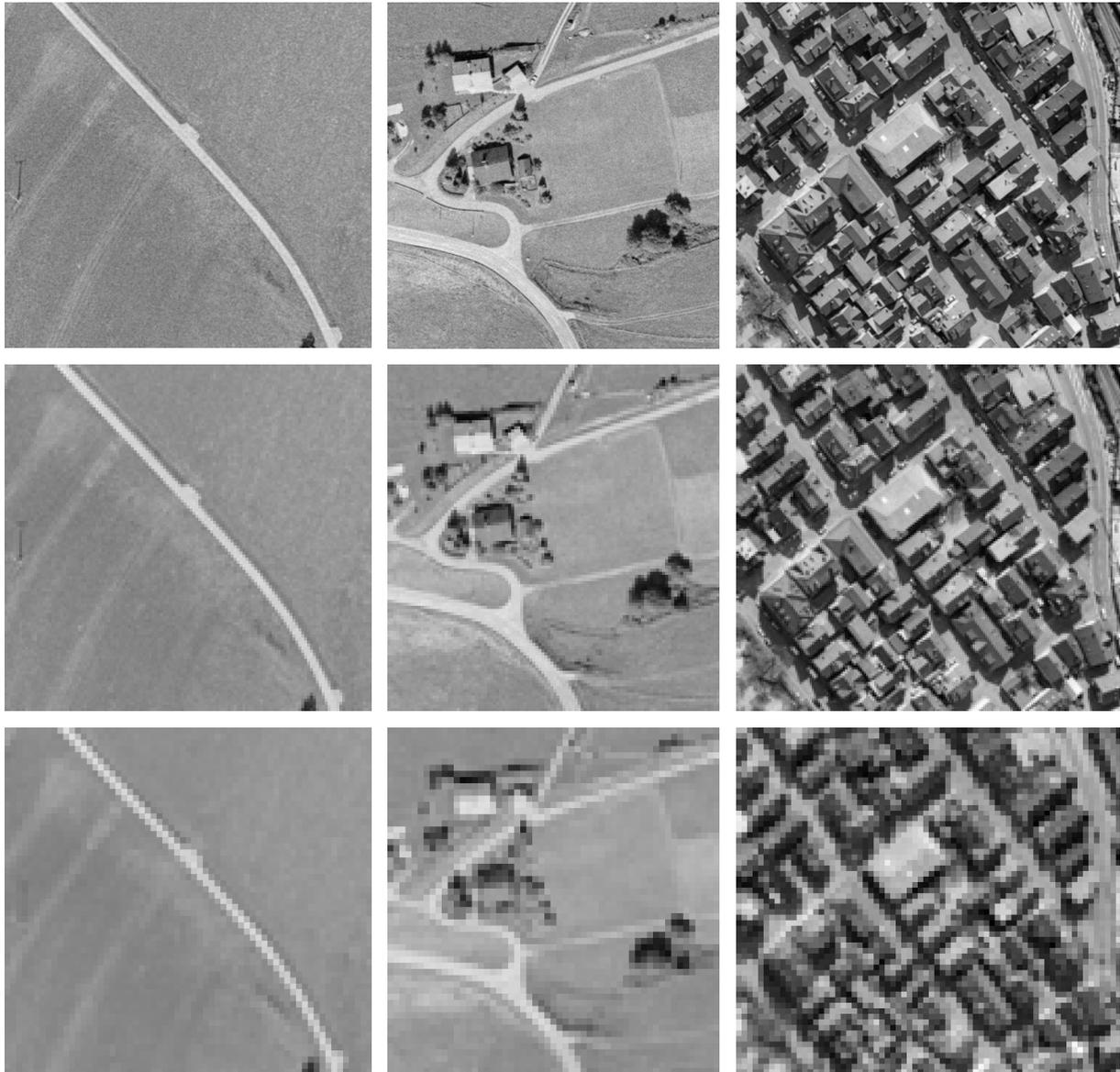


Abbildung 1. Komplexität von Bildern in Abhängigkeit ihrer Auflösung:
 Zeilen: von links nach rechts zunehmende Komplexität des Inhalts der Szene.
 Spalten: von oben nach unten abnehmende Beobachtbarkeit (Auflösung: *hoch*, *mittel*, *niedrig*)

einen eindeutigen Hinweis auf ein Hausdach bzw. ein Gebäude, und die Relation „Fahrzeug fährt auf Straße“ impliziert eine verdeckte Fahrspur und damit eine Straße. Aus diesen Beispielen wird ersichtlich, dass sich ein leistungsfähiges Bildanalyse-System durch eine entsprechende Abstimmung von Modell und Strategie auf die Komplexität der Datengrundlage auszeichnet.

2.1.3 Modellbildung

Ziel der Modellbildung ist es, eine strukturierte, möglichst widerspruchsfreie und im Computer formal repräsentierbare Beschreibung eines Objektes zu liefern. Generell ist eine explizite Beschreibung des *Modell-Wissens* über Objekte und ihre Relationen vorteilhaft, da hierdurch die vom Computer generierten Ergebnisse leichter analysiert und damit für andere Datensätze prädiert werden können. Eine Strukturierung des Modell-Wissens

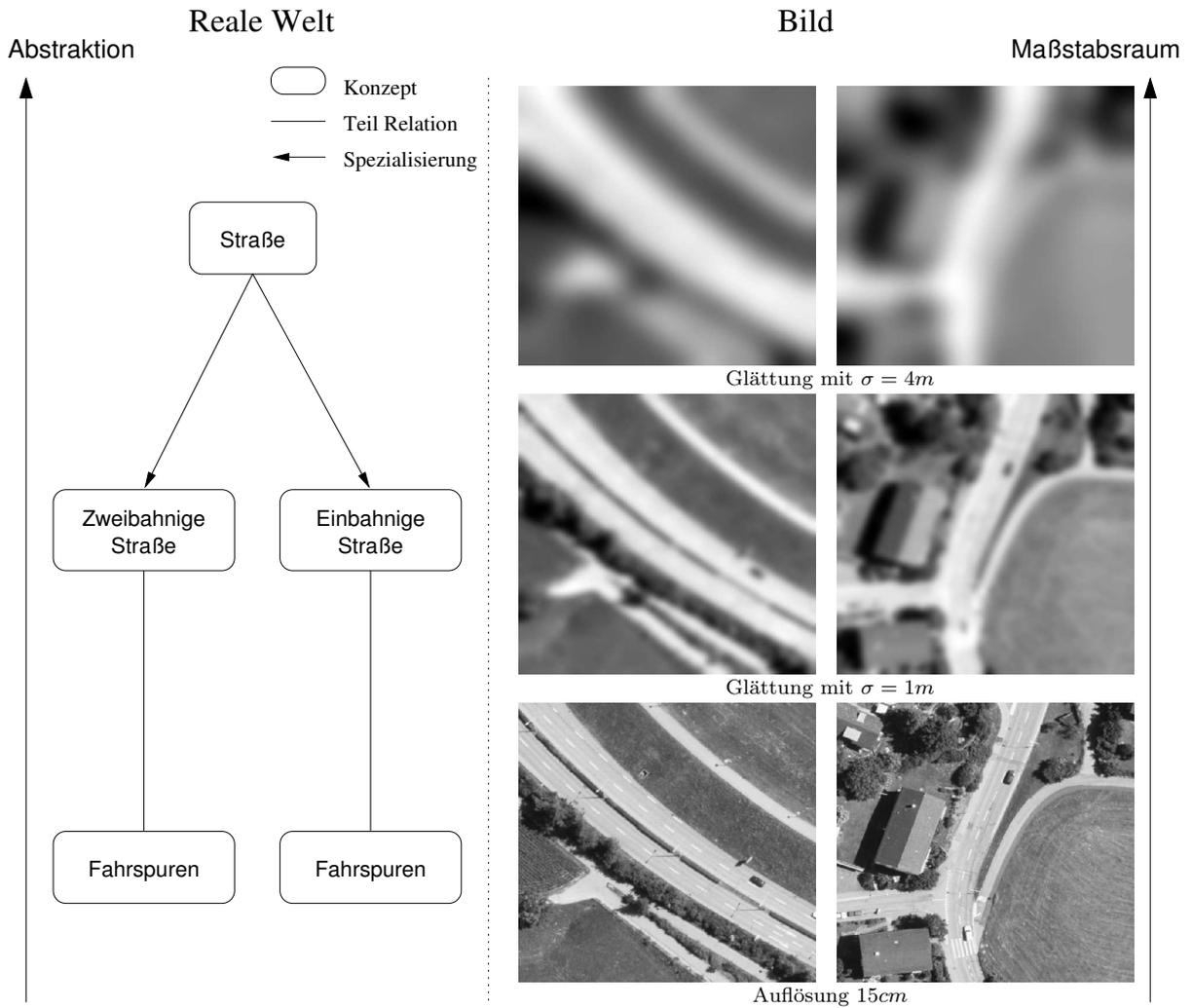


Abbildung 2. Abstraktionsstufen (links) und Maßstabsraum (rechts): Mit zunehmender Glättung σ des Bildes werden zunächst Details wie Markierungen und Zebrastreifen eliminiert und später größere Strukturen wie Fahrzeuge und die bauliche Trennung der Fahrbahnen. Die Ergebnisse dieser Ereignisse können verschiedenen Abstraktionsstufen in der realen Welt zugeordnet werden.

lässt sich dadurch erreichen, dass man das Modell in ein *Objektmodell*, das einer Beschreibung durch Konzepte der realen Welt entspricht, und ein *Bildmodell*, das die Erscheinung des Objektes im Bildraum wiedergibt, unterteilt. Die Übertragung des Objektmodells in den Bildraum geschieht anhand des *Sensormodells*, das die geometrischen und physikalischen Transformationen der Bildgenerierung erfasst. Für einige Anwendungen reicht es aus, keinen expliziten Übergang vom Objektmodell in den Bildraum zu vollziehen, sondern ausschließlich ein Bildmodell vorzugeben. Man spricht in diesem Fall von aussehensbasierter Modellierung.

Das im Modell enthaltene Wissen lässt sich anhand verschiedener, sich ergänzender Gesichtspunkte strukturieren:

Strukturierung in Form von Ebenen:

Eine erste Untergliederung geschieht oft in Form von *Ebenen*. Die beiden wichtigsten Ebenen sind die „Reale-Welt“-Ebene und die „Bild“- bzw. „Daten“-Ebene, weil sie das Objekt- bzw. Bildmodell repräsentieren (siehe

Abbildung 2). Auf jeder Ebene kann ein Objekt anhand bestimmter Kriterien durch Teile-Hierarchie und Spezialisierung weiter gegliedert werden. Zur Strukturierung auf Bildebene dienen in aller Regel Kriterien, die sich auf die Möglichkeiten und die Effizienz von Bildverarbeitungsalgorithmen beziehen. Im Unterschied hierzu spielt bei der Gliederung auf „Realer-Welt“-Ebene die Funktion der Objekte eine entscheidende Rolle. So müssen z.B. Autobahnen, um hohe Geschwindigkeiten und sichere Überholmanöver zu erlauben, mehrere Fahrspuren haben und bezüglich Krümmung und Neigung vorgegebenen Bauvorschriften entsprechen. Derartige funktionale Eigenschaften lassen sich sehr vorteilhaft für die Strukturierung des Objektes in seine Teile und die Festlegung von Parameterbereichen für geometrische und physikalische Eigenschaften einsetzen.

Strukturierung gemäß Auflösung:

Zusätzlich kann ein Modell in *verschiedene Auflösungsstufen* gegliedert werden. Bezogen auf den Bildraum bedeutet dies, dass das Bildmodell in verschiedenen Maßstäben vorliegt, also in diskreten Realisierungen des an sich kontinuierlichen Maßstabsraumes. Der gängigste und auch in dieser Arbeit verwendete Maßstabsraum ist der lineare oder Gauß'sche Maßstabsraum. Hierin werden niedrigere Auflösungsstufen eines n -dimensionalen Originalsignals durch Faltung mit dem Gauß'schen Glättungsoperator

$$g(\mathbf{x}, t) = \frac{1}{(2\pi t)^{n/2}} \cdot e^{-\frac{\mathbf{x}^T \mathbf{x}}{2t}}$$

erzeugt, wobei \mathbf{x} bei herkömmlichen Bildern zweidimensional ist ($n = 2$) und der Glättungs- bzw. Maßstabsraumparameter t für die Varianz σ^2 steht. Aus speichertechnischen Gründen erfolgt oft nach der Reduktion der Auflösung unter Beachtung des Nyquist-Theorems eine Neuabtastung, so dass sich die geringere Auflösung der Bilder bzw. ein kleinerer Maßstab durch größere Bodenpixel ausdrückt.

Wie in Abbildung 2 dargestellt, entspricht eine auflösungsbasierte Modellierung auf der Ebene der realen Welt einer Unterteilung des Objektmodells in mehrere Abstraktionsstufen. Diese Art der Strukturierung betont die Generalisierung des Wissens auf höheren Abstraktionsstufen und hebt die unterschiedlichen Eigenschaften des Objektes in verschiedenen Maßstäben hervor. Die Vorteile dieser Modellierung wurden besonders in neueren Arbeiten zur Objektextraktion herausgestellt, u.a. in (Mayer, 1998; Mayer und Steger, 1998; Baumgartner et al., 1999; Couloigner und Ranchin, 2000; Straub und Heipke, 2001). Es zeigt sich nämlich, dass eine Extraktion im größtmöglichen Maßstab alleine nicht zwangsläufig zu den besten Ergebnissen führt. Jede Objekteigenschaft besitzt einen bestimmten Maßstabsbereich, in dem sie besonders gut extrahierbar ist. Somit trägt eine auflösungsbasierte Modellierung nicht nur zu einer besseren Strukturierung und effizienteren Nutzung des Wissens bei. Aufgrund der Möglichkeiten zur Fusionierung von Ergebnissen aus verschiedenen Maßstäben ist sie vielmehr eine wichtige Grundlage für eine erfolgreiche Extraktion. Die Verknüpfung einer Abstraktionsstufe mit dem zugehörigen Maßstab geschieht meist im Voraus durch die Fixierung des Maßstabs. Einige Forschungsarbeiten beschäftigen sich auch damit, Verknüpfungen durch Analyse von so genannten Maßstabsraumereignissen analytisch zu begründen (Mayer und Steger, 1998). Im Falle von komplexen topographischen Objekten gelingt die in diesem Zusammenhang nötige automatische Maßstabsschätzung in aller Regel nur exemplarisch. Für die Extraktion einfacher Bildmerkmale wie Linien und Kanten kann sie jedoch mit einer größeren Allgemeingültigkeit formuliert werden (Lindeberg, 1998).

Strukturierung durch Kontext:

Da in natürlichen Umgebungen ein Objekt oft den Einflüssen seiner Nachbarobjekte bzw. seiner gesamten Umgebung unterliegt, ist es meist nicht ausreichend, das Objekt als isolierte Einheit zu modellieren. Straßen erscheinen z.B. in Luftbildern oft als helle graue Fläche. Im Falle von Schattenwurf durch angrenzende Gebäude oder Bäume wird der helle Belag aber stark abgedunkelt und entspricht dadurch nicht mehr der Modellvorstellung. Das Wissen über solche Einflüsse lässt sich durch ein so genanntes *Kontextmodell* formulieren. Das

heißt, das Objektmodell wird um Relationen zu anderen Objekten erweitert oder im Falle einer umfassenderen Modellierung in ein Modell für die gesamte Szene eingebettet. In diesem Zusammenhang steht der Begriff „globaler Kontext“ eines Objektes für die Abhängigkeit von Eigenschaften und Erscheinungsformen des Objektes von seiner großräumigen Umgebung. Relationen zu den unmittelbar benachbarten Objekten werden dementsprechend als „lokaler Kontext“ bezeichnet (Mayer, 1998; Baumgartner et al., 1999). Bei einer detaillierten Kontextmodellierung ist zu beachten, dass Interaktionen von Objektrelationen auftreten können. Ein Auto, das den Schatten eines Gebäudes durchfährt, wirft praktisch keinen zusätzlichen Schatten auf die Straße. Das heißt, die Schatten-Relation zwischen der Straße und dem Fahrzeug wird durch die Schatten-Relation zwischen der Straße und dem Gebäude ausgelöscht.

Repräsentation des Modell-Wissens:

Ein für die Modellbildung wesentlicher Punkt ist die Darstellung des im Modell enthaltenen Wissens. Die Form der *Wissensrepräsentation* sollte so gewählt sein, dass das Modell-Wissen ohne große Einschränkungen hinsichtlich Umfang und Strukturierung computergerecht darstellbar ist. Gleichzeitig sollte sie aber eine hohe Transparenz für eine detaillierte Untersuchung und Erprobung des Bildanalysesystems gewährleisten. In (Brachman, 1979) werden fünf aufeinander aufbauende Wissensebenen vorgestellt: die linguistische, die konzeptuelle, die epistemologische, die Logik- und die Implementierungsebene. Die linguistische Ebene basiert auf natürlichsprachlicher Repräsentation und erscheint in Ermangelung einer klaren Formalisierung für die Wissensrepräsentation zur Objektextraktion ungeeignet. Die Logik- und Implementierungsebene hingegen unterliegen diesbezüglich sehr starken Zwängen, was bei einem umfangreicheren Bildanalysesystem leicht zu einer unübersichtlichen und teils widersprüchlichen Wissensrepräsentation führen kann, wie das z.B. (McKeown et al., 1989) zeigt. Während die epistemologische Ebene eine anwendungsunabhängige Repräsentation von Wissen erlaubt und daher für die vergleichende Darstellung verschiedener Ansätze wie in (Mayer, 1998) geeignet ist, hat die konzeptionelle Ebene – bezogen auf die Objektextraktion – den Vorteil, dass sie eine objektzentrierte und somit problemnahe Repräsentation erlaubt, gleichzeitig aber eine grundlegende Formalisierung erzwingt. Deshalb erscheint sie für diese Arbeit am besten geeignet. Ein zur Umsetzung besonders geeigneter und häufig eingesetzter Repräsentationsformalismus ist ein so genanntes *Semantische Netz* (Woods, 1975; Brachman, 1977, 1979; Sagerer, 1982; Niemann et al., 1990; Liedke et al., 2001).

Als unvermeidbare Problematik tritt bei der Formalisierung von Wissen die Handhabung von unscharfen Beschreibungen und unsicheren Aussagen auf. Hierfür steht eine Reihe von mathematisch und stochastisch fundierten Theorien wie Bayes-, Evidenz- oder Fuzzy-Theorie zur Verfügung (Zadeh, 1989; Shafer und Pearl, 1990), die mit Formalismen für die Wissensrepräsentation gekoppelt werden können. Bei der Verwendung dieser Theorien zur Extraktion von Objekten aus Bildern ist besonders darauf zu achten, dass Modellierung und Extraktion die grundlegenden Annahmen dieser Theorien nicht verletzen.

2.1.4 Strategie

Im Rahmen der Objektextraktion wird unter Strategie „*ein Plan für die Erkennung von Objekten in Bildern durch Zuordnung von Daten zu einem Modell unter Nutzung verschiedener Methoden*“ verstanden (Mayer, 1998). Diese Aussage macht deutlich, dass die Strategie aus zwei grundlegenden, miteinander verkoppelten Elementen besteht: einerseits die Planung des Erkennungsprozesses und andererseits die Wahl der Methoden zu dessen Verwirklichung. Unter diese Methoden fallen auch Verfahren der Bildverarbeitung, da durch sie die Information, die den Modellen zugeordnet werden muss, entscheidend beeinflusst wird. Die Wahl der Methoden ist meist stark anwendungsabhängig, weshalb die für die Straßenextraktion relevanten Verfahren erst später beschrieben werden.

Die Planung des Extraktionsprozesses ist – präziser definiert – die Steuerung, wann welches Wissen über das Objekt zur Extraktion genutzt wird und wie Ergebnisse von Teil-Extraktionen in den weiteren Verfahrensab-

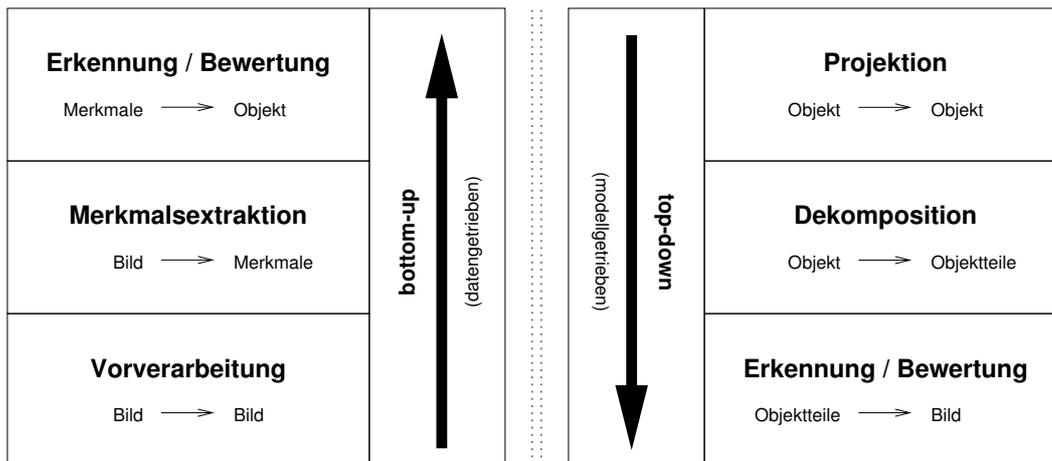


Abbildung 3. Vorgehensweisen der Bildanalyse in idealisierter Form: datengetrieben (links) und modellgetrieben (rechts)

lauf eingebunden werden. Man unterscheidet grundsätzlich die *datengetriebene*, die *modellgetriebene* und die *hybride / gemischte* Steuerung. In Abbildung 3 sind die grundlegenden Schritte der datengetriebenen und modellgetriebenen Vorgehensweise gegenübergestellt. Ebenso wie bei der folgenden Beschreibung handelt es sich um eine idealisierte Darstellung. In der Praxis ist in aller Regel nur die Verwendung der hybriden Vorgehensweise realistisch. Die Unterschiede bestehen in der Gewichtung von daten- und modellgetriebenen Anteilen.

- ◇ Die datengetriebene Vorgehensweise startet mit rein ikonischen Vorverarbeitungsschritten. Im nächsten Schritt werden Merkmale aus den Bildern extrahiert und zu höherwertigen Strukturen zusammengefasst. Dieser Prozess wird solange in 2D und 3D fortgesetzt, bis Strukturen aufgebaut worden sind, die den Vorgaben des Modells ausreichend entsprechen und somit eine Zuordnung zum Modell möglich ist. Die Voraussetzung für den Erfolg dieser Vorgehensweise ist, dass bei jedem Schritt neue Information gewonnen wird, die insgesamt zur Extraktion des Objektes ausreicht. Vorteilhaft ist, dass das Modell wenig spezifisch formuliert sein muss, da seine Bestandteile durch die Merkmalsextraktion direkt aus der Datengrundlage gewonnen werden.
- ◇ Bei einer rein modellgetriebenen Steuerung wird das (3D-)Objektmodell mittels Sensormodell ins Bild projiziert und daraus eine Hypothese für das Objekt gebildet. Die Qualität einer Objekthypothese wird – meist nach Dekomposition in einzelne Objektteile – durch Zuordnung zu entsprechenden Bildregionen bewertet. Eine Voraussetzung dieser Vorgehensweise ist, dass das Modell und seine Erscheinung im Bild sehr spezifisch beschrieben werden kann, damit eine Hypothese im Bildraum gut mit der tatsächlichen Erscheinungsform des Objektes übereinstimmt. Der Vorteil der spezifischen Beschreibung besteht darin, dass einerseits der Suchraum für Zuordnungen stark eingeschränkt werden kann und andererseits Störungen in den Daten weniger gravierende Auswirkungen als bei der datengetriebenen Steuerung haben.
- ◇ Eine hybride Steuerung kombiniert datengetriebene und modellgetriebene Anteile. Im Falle der Interpretation von komplexen Luftbildern ist dies die einzig sinnvolle Vorgehensweise. Entscheidend für die Effizienz einer hybriden Steuerung ist die Wahl der „Wechselpunkte“, d.h. wann von datengetriebener Steuerung auf modellgetriebene übergegangen wird bzw. umgekehrt. Eine typische Vorgehensweise besteht darin, erst mit einer (datengetriebenen) Merkmalsextraktion zu beginnen, um daraus zuverlässige Hypothesen für leicht extrahierbare Objektteile zu gewinnen. Diese werden modellgetrieben auf Plausibilität untersucht und zur Generierung weiterer Hypothesen genutzt, die dann durch Zuordnung zu gezielt neu extrahierten Merkmalen verifiziert oder falsifiziert werden. Ein großer Vorteil der hybriden Steuerung liegt darin begründet, dass durch die Wahl der Wechselpunkte zusätzliches Wissen über erwartete Störungen und die erwartete Extrahierbarkeit von Objektteilen einfließen kann. So ist z.B. das Umringspolygon eines Flachdaches vergleichsweise einfach aus Luftbildern zu extrahieren, weshalb mit einer stark datengetriebenen Steuerung

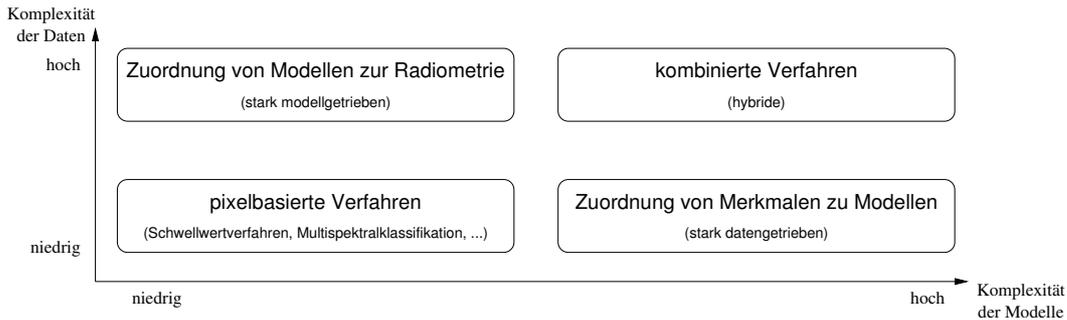


Abbildung 4. Effiziente Verfahren für verschiedene Komplexitätsstufen von Daten und Modellen (in Anlehnung an (Suetens et al., 1992; Mayer, 1998))

zuverlässige Dachhypothesen generiert werden können. Mit Hilfe der Dachhypothesen kann der Suchraum für weitere Teilobjekte wie senkrechte Haus- oder Schattenkanten modellgetrieben eingeschränkt werden. Extrahierte Teilobjekte verifizieren einerseits die Dachhypothese und ermöglichen andererseits aber auch eine detaillierte Gebäuderekonstruktion (Kim et al., 2001).

Die Eignung einer Strategie lässt sich in Abhängigkeit der Komplexität von Bilddaten und Modell folgendermaßen einordnen (siehe Abbildung 4): Bei einfachen Daten und einfachen Modellen eignen sich Verfahren der Segmentierung oder Klassifikation zur Erkennung von Objekten. Daher sind z.B. Verfahren der Multispektralklassifikation für die Auswertung von ländlich geprägten Satellitenbildszenen geeignet und werden dort auch erfolgreich angewendet. Die Zuordnung von Modellen zur Radiometrie – also eine stark modellgetriebene Strategie – lässt sich dann einsetzen, wenn sich das Datenmaterial durch hohe Komplexität kennzeichnet, das Modell aber einfach und spezifisch zu beschreiben ist oder sich der Suchraum stark einschränken lässt. Dieser Gesichtspunkt spielt in der vorliegenden Arbeit bei der Extraktion von Fahrzeugen eine wesentliche Rolle. Im Fall einer geringen Daten- aber hoher Modellkomplexität eignet sich eine Strategie, bei der die datengetriebene Zuordnung symbolischer Merkmale zum Modell überwiegt. Denn aufgrund der einfachen Daten werden schon durch Merkmalsextraktion und Segmentierung gesicherte Hinweise auf das Objekt gewonnen. Somit lassen sich auch komplexere Modelle den Merkmalen erfolgreich zuordnen. Sind sowohl Daten als auch Modell komplex, sollte eine hybride Strategie zur Objektextraktion verwirklicht werden, die auf die speziellen Charakteristika von Daten und Modell ausgerichtet ist (Suetens et al., 1992).

Bei der Realisierung der Strategie kommen verschiedene Verfahren zum Einsatz, die sich vereinfachend in anwendungsunabhängige Methoden und – insbesondere die Parametrisierung betreffend – anwendungsabhängige Methoden aufteilen lassen. Zu den anwendungsabhängigen Methoden gehören einerseits verschiedenste Algorithmen der digitalen Bildverarbeitung wie radiometrische, geometrische oder morphologische Transformationen, Segmentierung, Merkmalsextraktion, Polygonapproximation etc. Andererseits sind hierunter auch Verfahren zur (multisensoriellen) Datenfusion und Gruppierung von Teilstrukturen zu sehen, da diese auf die von der Anwendung vorgegebene Problemstellung anzupassen sind. Zu den anwendungsunabhängigen Methoden hingegen zählen generische Suchalgorithmen (Vosselman, 1992; Geman und Geman, 1984) und Bewertungsschemen (Shafer und Pearl, 1990).

Bei der Realisierung der Strategie kommen verschiedene Verfahren zum Einsatz. Hierunter fallen einerseits relativ anwendungsunabhängige Algorithmen der digitalen Bildverarbeitung wie radiometrische, geometrische oder morphologische Transformationen, Segmentierung, Merkmalsextraktion, Polygonapproximation, etc., andererseits auch stark anwendungsabhängige Verfahren zur (multisensoriellen) Datenfusion und Gruppierung von Teilstrukturen. Eine weitere Gruppe von Verfahren stellen generische Suchalgorithmen (Geman und Geman, 1984; Vosselman, 1992) und Bewertungsschemen (Zadeh, 1989; Shafer und Pearl, 1990) dar, die vordergründig überhaupt keine Abhängigkeit von der Anwendung besitzen. Allerdings fließt auch hier – wie bei allen anderen Methoden – allein durch die Parametrisierung des jeweiligen Algorithmus der Bezug zur Anwendung mit ein.

Nachdem die wichtigsten Grundlagen der Bildanalyse angesprochen wurden, werden im nächsten Abschnitt ausgesuchte Verfahren der Straßenextraktion vorgestellt.

2.2 Verfahren zur automatischen Straßenextraktion

Aufgrund der Komplexität von Luftbildern ist die automatische Extraktion des Straßennetzes ohne Mehrdeutigkeiten oder Lücken eine sehr schwierige Aufgabe. Ein System zur Straßenextraktion hat typischerweise zwei zentrale Aufgaben zu lösen: 1.) Die *Detektion* von Straßenabschnitten, die mit hoher Sicherheit zum Straßennetz gehören (Phase 1), und 2.) deren *Verknüpfung* zu einem möglichst vollständigen und topologisch korrekten Netz bei simultaner Elimination von eventuellen Fehlextraktionen (Phase 2). Beide Aufgaben zerfallen in der Regel in weitere Teilschritte, die sequentiell oder in Schleifen, d.h. mit Rückkopplungen, abgearbeitet werden. Die folgende Darstellung der Ansätze richtet sich nach der Auflösung der Eingangsdaten (*niedrig* oder *mittel/hoch*), wobei die zweite Gruppe noch einmal nach der Art der Netzbildung (*Tracking* oder *Verknüpfung*) untergliedert ist.

2.2.1 Straßenextraktion aus niedrig aufgelösten Bilddaten

Eines der ältesten Verfahren ist das bei SRI International, Menlo Park, entwickelte und in mehreren Etappen verbesserte System zur Straßenextraktion aus Bildern niedriger Auflösung (Fischler et al., 1981; Fischler, 1994; Fischler und Heller, 1998). Aufgrund der geringen Auflösung werden Straßen als ein Netz aus dünnen, hellen Linien modelliert. Die Extraktion basiert auf der Verwendung mehrerer Linienoperatoren, die sich grob in zwei Klassen einteilen lassen. Operatoren vom Typ I zeichnen sich durch eine hohe Zuverlässigkeit⁵ aber durch geringere Vollständigkeit in der Extraktion aus. Sie sind also hauptsächlich zur Detektion von Teilstücken geeignet (Phase 1). Operatoren vom Typ II erreichen dagegen eine wesentlich höhere Vollständigkeit, worunter jedoch die Zuverlässigkeit der Extraktion leidet. Diese Operatoren sind dann einsetzbar, wenn durch die Detektion schon eine gewisse Evidenz für Straßen vorliegt, aber noch Lücken zu schließen sind (Phase 2). Die Strategie besteht darin, zunächst aus den Ergebnissen der Typ I Operatoren Liniencluster zu bilden, die anschließend mit Hilfe der Ergebnisse der Typ II Operatoren verbunden werden. Die Generierung der Liniencluster wird dadurch erreicht, dass nach einer gewichteten Kombination der Typ I Ergebnisse in einem rekursiven Prozess kurze Lücken geschlossen und simultan minimale Abstandsbäume aufgebaut werden. In diesen Bäumen werden Pfade maximaler Länge gesucht, die als Hypothesen für zuverlässig extrahierte Straßenabschnitte ausgesucht werden. Zur Vervollständigung des Netzes werden nun die Ergebnisse der Typ II Operatoren herangezogen. Sie dienen zur Berechnung von „Kosten“, die ein Pfad zwischen zwei vorgegebenen Punkten verursacht. Die letztendliche Verknüpfung der Lücken geschieht durch Auswahl von Start- und Endpunkten aus den Linienclustern und Berechnung des jeweils besten („billigsten“) Pfades mittels des so genannten F*-Algorithmus innerhalb vorgegebener Suchbereiche. Um die als sehr zuverlässig eingeschätzten Ergebnisse der Typ I Operatoren mit in die Pfadberechnung einfließen zu lassen, werden diese mit minimalen Kosten – bezogen auf den jeweiligen Typ II Operator – angesetzt.

Weiterentwicklungen des ursprünglichen Ansatzes (Fischler et al., 1981) betreffen vor allem die Verknüpfung der Liniencluster. In (Fischler, 1994) wird die Bestimmung der besten Pfade durch einen Algorithmus, der allgemeine Gruppierungskriterien wie Nähe, Kollinearität, Kontinuität usw. zugrunde legt, ergänzt. Die aktuellste Veröffentlichung (Fischler und Heller, 1998) beinhaltet die Erweiterung um eine straßenspezifische Gruppierungskomponente. Diese integriert für Bilder einer Auflösung von 1m und höher Linien-Attribute wie Länge, Breite und Krümmung. In (Heller et al., 1998) wird ein Modul zur effizienten manuellen Korrektur der automatisch erzielten Ergebnisse vorgestellt, mit dem sich die Erfassungszeit verglichen mit einer manuellen

⁵ „Zuverlässigkeit“ bedeutet, dass ein hoher Prozentsatz der gefundenen Linien auch tatsächlich Straßen sind.

Digitalisierung um den Faktor zehn für den gezeigten Bildausschnitt verringern ließ. Die automatisch generierten Ergebnisse zeigen, dass dank der verschiedenen und zum Teil komplementären Operatoren im offenen Gelände gute Resultate erzielt werden können. Während eine detailliertere Analyse von Störungen aufgrund der niedrigen Auflösung nur schwer möglich erscheint, bleibt jedoch als offensichtlicher Schwachpunkt des Ansatzes, dass der Netzcharakter von Straßen nur in relativ geringem Ausmaß modelliert und zur Extraktion genutzt wird.

Die Verwendung des Wissens über den Netzwerkcharakter von Straßen steht im Fokus des an der Technischen Universität München (TUM) entwickelten Ansatzes zur Straßenextraktion aus multispektralen Bildern einer Auflösung von 2m – 4m (Wiedemann und Hinz, 1999; Wiedemann und Ebner, 2000; Wiedemann, 2001). Zunächst werden in jedem Kanal mit einem hochqualitativen Operator (Steger, 1998b) lineare Strukturen extrahiert, dann anhand von Attributen wie z.B. Länge, Breite, Krümmung, Kontrast, usw. bewertet und schließlich zu einem einheitlichen Datensatz fusioniert. Basierend auf dieser Bewertung wird mittels Schwellwerten entschieden, welche Linien mit hoher Wahrscheinlichkeit Teile des Straßennetzes repräsentieren (Phase 1). Die Verknüpfung dieser als *seeds* bezeichneten Linien geschieht in mehreren Schritten (Phase 2): Zur Generierung eines anfänglichen Straßennetzes wird die Verbindungsfunktion von Straßen explizit ausgenutzt, indem zwischen sämtlichen extrahierten Linien Verknüpfungshypothesen aufgebaut und anhand lokaler geometrischer Kriterien bewertet werden (Steger et al., 1997). Die Auswahl von Linien und Verknüpfungshypothesen, die Straßen repräsentieren, erfolgt über die Berechnung optimaler Pfade zwischen ausgewählten Paaren von *seeds*. Die Kombination aller Pfade stellt das anfänglich extrahierte Straßennetz dar, das im nun folgenden Schritt vervollständigt wird. Hierzu wird das Wissen verwendet, dass Straßen nicht nur ein zusammenhängendes Netz bilden, sondern dass sie speziell im offenen Gelände eine effiziente Verbindung zwischen zwei beliebigen Punkten auf dem Netz gewährleisten müssen. Das Prinzip zur Generierung neuer Verknüpfungshypothesen besteht darin, dass die jeweilige Wegstrecke zwischen ausgewählten Punktepaaren auf dem Netz oder über Zusammenhangskomponenten hinweg jeweils zur optimalen euklidischen Distanz ins Verhältnis gesetzt wird. Daraus wird ein so genannter „Umweg-“ bzw. „Verbindungsfaktor“ berechnet, der die Signifikanz einer neuen Verknüpfungshypothese widerspiegelt. Eine Hypothese gilt dann als verifiziert, wenn innerhalb eines Suchbereiches eine Verbindung zwischen Anfangs- und Endpunkt der Hypothese über lineare Strukturen hinweg gefunden werden kann. Prinzipiell ähneln die erzielten Ergebnisse jenen des oben vorgestellten SRI-Ansatzes. Dies gilt zumindest in Blick auf Vollständigkeit und Zuverlässigkeit. Die ausführliche Analyse des Verfahrens in (Wiedemann, 2001) zeigt jedoch, dass durch die Verwendung von Wissen über die Netzwerkeigenschaften insbesondere topologisch wichtige Straßenabschnitte extrahiert werden können, was in einer wesentlich besseren Verknüpfung des Straßennetzes resultiert. Ein weiterer Vorteil dieses Ansatzes ist, dass auf lokaler Ebene wenig harte Schwellwerte zum Einsatz kommen. Straßenhypothesen werden hier lediglich bewertet und erst anhand der globaler Netzwerkeigenschaften selektiert.

2.2.2 Straßenextraktion aus mittel und hoch aufgelösten Bilddaten

Da obige Verfahren für Straßenextraktion aus niedrig aufgelösten Bildern entwickelt wurden, können sie lokale Straßenevidenz bestenfalls in Form von attributierten Linien einsetzen. Im Gegensatz hierzu stehen bei Verfahren, die Bilder mittlerer und hoher Auflösung als Eingangsdaten verwenden, lokale Eigenschaften der Straße wie die Parallelität der Straßenränder und die Homogenität des Belages im Vordergrund. Dies spiegelt sich z.B. in den langjährigen Arbeiten zur Straßendetektion und -verfolgung, die an der Carnegie Mellon University (CMU) durchgeführt wurden, wider (McKeown und Denlinger, 1988; Aviad und Carnine, 1988; Zlotnick und Carnine, 1993; McKeown, 1990; Harvey, 1999). Charakteristisch für das dort verfolgte Konzept ist die Kooperation verschiedener Module, die entweder die Aufgabe der Straßendetektion oder die der Verfolgung lösen (McKeown et al., 1997). Das in (Aviad und Carnine, 1988; Zlotnick und Carnine, 1993) vorgestellte Modul zur Detektion von Straßenmittelachsen (Phase 1) basiert auf der Extraktion von Kantenpixel als potentielle Straßenränder und deren Gruppierung zu parallelen, anti-polarisierten Pixelpaaren. Die Mittelpunkte der Pixelpaare werden unter Glattheitsbedingungen zu Linien verknüpft, die als Startsegmente für eine anschließende Straßenverfolgung herangezogen werden können. Eine Alternative zu diesem Verfahren stellt der ursprünglich zur Detektion von

Kreuzungen entwickelte Ansatz (Harvey, 1999) dar. Hier stützt sich die Detektion von Mittelachsen und Kreuzungspunkten auf eine Variante der Hough-Transformation. Die Verwendung der Hough-Transformation ist insofern für komplexere Daten vorteilhaft, da sie gegenüber dem vorherigen Verfahren weniger störanfällig auf eine fragmentierte Kantenextraktion reagiert. Die Straßenverfolgung (Phase 2) wird mit zwei komplementären Algorithmen zur Straßenverfolgung realisiert (McKeown und Denlinger, 1988). Beginnend an vorher detektierten Startsegmenten verfolgt der so genannte *Kantentracker* iterativ parallele Kanten des Startsegmentes, indem bei jeder Iteration aus Kantenstärke und Kantenrichtung die neue Laufrichtung und ein Qualitätsmaß abgeleitet werden. Wenn die Qualität als zu schlecht bewertet wird, weil z.B. beide Kanten unterbrochen sind, wird eine Störung angenommen und abgebrochen. Simultan zum Kantentracker wird ein *Profiltracker* gestartet, der bei jeder Iteration die neue Laufrichtung durch robuste Korrelation eines Referenzprofils mit dem Grauwertprofil senkrecht zur aktuellen Laufrichtung berechnet und bewertet. Das Referenzprofil wird aus dem Startsegment oder durch Mittelung zurückliegender Profile bestimmt. Der Profiltracker trägt kleinräumigen Störungen und Änderungen des Straßenbelags dadurch Rechnung, dass im Falle schlechter Korrelation nur dann abgebrochen wird, wenn das Referenzprofil auch nach mehreren überbrückenden Schritten nicht mehr korrelierbar ist, bzw. wenn auch die Generierung eines neuen Referenzprofils innerhalb dieser Schritte keine tolerierbaren Ergebnisse erzeugt. Die beiden Tracking-Verfahren kooperieren miteinander, indem die Verfolgung solange aufrecht erhalten wird, wie mindestens einer der beiden Tracker gültige Ergebnisse liefert. Läuft einer der beiden Tracker fehl, wird er auf der Spur des anderen neu initialisiert. Die besondere Stärke dieses Ansatzes ist, dass solche Ereignisse während des Tracking als Störungen mit verschiedenen Ursachen auf der Ebene der realen Welt modelliert werden. So deuten z.B. fehlende Kanten bei homogenen Grauwerten in der Straßenmitte auf eine Kreuzung hin. Umgekehrt kann ein Fahrzeug eine Unterbrechung des Profiltrackers verursachen, wohingegen der Kantentracker dadurch kaum beeinflusst wird und eine Fortsetzung des Trackings erlaubt. In (McKeown, 1990) wird gezeigt, dass die beiden Tracker in Kombination wesentlich stabilere Ergebnisse liefern als bei alleinigem Einsatz.

In (Barzohar und Cooper, 1996; Barzohar et al., 1997) wird ein integriertes System zur Detektion und Verfolgung von Straßen vorgestellt. Beide Phasen werden mit einem Modell realisiert, das Straßen als homogene Flächen mit begrenzenden parallelen Kanten beschreibt. Im Gegensatz zu den vorher genannten Ansätzen wird das Modell durch ein Markov-Zufallsfeld repräsentiert. Straßendetektion und -verfolgung basieren auf einer Schätzung der maximalen a posteriori Wahrscheinlichkeit innerhalb eines lokalen Bereiches. Im Falle der Detektion muss dieser rechenaufwändige Prozess – ähnlich einer Filterung – auf den gesamten Bildausschnitt angewendet werden, während dies im Falle der Straßenverfolgung durch Verschiebung von Fenstern realisiert ist. Vorteilhaft an diesem Ansatz ist die theoretisch fundierte Handhabung von Unsicherheiten durch eine konsequente statistische Modellierung. Ein unübersehbares Defizit ist jedoch die Verwendung des selben, relativ einfachen und unflexiblen Modells für beide Phasen. Die gezeigten Ergebnisse machen deutlich, dass Störungen durch Hintergrundobjekte lediglich dann überbrückt werden können, wenn sie durch die in der statistischen Modellierung erfassten Unschärfen erklärbar sind. Eine Feststellung der möglichen Ursache einer Störung wie in (McKeown und Denlinger, 1988) fehlt gänzlich.

Der an der TUM entwickelte Ansatz zur Straßenextraktion aus mittel- bis hochauflösten Luftbildern (Baumgartner et al., 1997, 1999; Baumgartner und Hinz, 2000; Baumgartner, 2003) zeichnet sich durch eine maßstababhängige und durchgängig explizite Modellierung von Straßen und ihrem Kontext aus. Im kontextunabhängigen Teil des Modells werden Straßen als homogene Flächen mit parallelen Rändern in der hohen Auflösung (ca. 0.25m) und als helle dünne Linien in der niedrigen Auflösung (ca. 2m) beschrieben. Die Verwendung mehrerer Auflösungsstufen hat den Vorteil, dass qualitativ hochwertige lokale Evidenz mit abstrahierter globaler Evidenz verbunden wird. So können aus der hohen Auflösung Straßensegmente mit hoher Genauigkeit und aussagekräftigen geometrischen und radiometrischen Attributen extrahiert werden. Durch die simultane Extraktion in der niedrigeren Auflösung können zudem die fundamentalen Netzwerkeigenschaften bei gleichzeitiger Unterdrückung lokaler Störungen einfließen. Durch die Erweiterung des Straßenmodells um das Kontextmodell wird das Erscheinungsbild von Straßen in Abhängigkeit ihrer Umgebung erfasst. Der *globale Kontext* teilt das Bild in die Regionen „Wald“, „offene Landschaft“ und „Siedlung“ ein. Der *lokale Kontext* beschreibt jene Beziehungen zwischen Hintergrundobjekten und Straßen, die das erwartete Erscheinungsbild der Straßen

grundlegend stören. Da dieser Ansatz für die offene Landschaft konzipiert ist, spielt Schattenwurf von angrenzenden Baumgruppen oder isolierten Gebäuden die wichtigste Rolle. Störungen durch vereinzelte Fahrzeuge werden meist durch die Integration der niedrigen Auflösung abgefangen, weil solch kleinräumige Strukturen im Maßstabsraum eliminiert werden können (Mayer und Steger, 1998).

Die Extraktionsstrategie besteht zuerst darin, nach einer texturbasierten Segmentierung der globalen Kontextgebiete sichere Hypothesen für Straßenabschnitte durch Fusionierung der beiden Auflösungsstufen zu detektieren (Phase 1) und diese anschließend durch ein iteratives, modellgetriebenes Gruppierungsverfahren zu verbinden (Phase 2). In Phase 1 werden helle Linien in der niedrigen und Kanten in der hohen Auflösung extrahiert und mittels eines Regelwerkes fusioniert und zu Rechteckstrukturen gruppiert. Die Regeln geben z.B. vor, dass Kantenpaare parallel sein müssen, dass sie eine homogene Fläche einschließen müssen und dass eine korrespondierende Linie in der niedrigen Auflösung extrahiert sein muss. Die resultierenden Rechteckstrukturen repräsentieren die anfänglichen Straßenabschnitte, die im Folgeschritt zu verknüpfen sind. Hierzu werden basierend auf geometrischen Kriterien Verknüpfungshypothesen aufgestellt und unter Hinzunahme radiometrischer Kriterien in einem Drei-Schritt-Verfahren verifiziert. Zunächst wird eine Verknüpfungshypothese und ihre angrenzenden Straßenabschnitte auf radiometrische Konsistenz geprüft. Schlägt dieser Schritt fehl, wird durch den Einsatz einer Ziplock-Ribbon-Snake (Kass et al., 1988; Neuenschwander et al., 1995; Laptev et al., 2000) versucht, schwache und fragmentierte Straßenevidenz auszunutzen. Als schwächste und daher letzte Möglichkeit der Verifikation dient die Erklärung der Störung durch eine Schattenregion, also über lokalen Kontext. Durch iteratives Abarbeiten der Verknüpfungshypothesen werden sukzessiv längere Straßenabschnitte gebildet, wobei gleichzeitig kurze, isolierte Straßenabschnitte eliminiert werden. Abschließend erfolgt die Netzbildung durch Rekonstruktion einfacher Kreuzungen und Einmündungen.

Insgesamt zeigt dieser Ansatz eine große Zuverlässigkeit und Vollständigkeit in der offenen Landschaft. Schwächen in der topologischen Korrektheit resultieren vor allem aus der geringen Verwendung globaler Straßeneigenschaften (Hinz et al., 2000). In (Baumgartner und Hinz, 2000) wird eine Erweiterung der Generierung von Verknüpfungshypothesen durch die Integration von globalen Gruppierungskriterien vorgestellt. Diese erweist sich hauptsächlich bei stärker fragmentierten initialen Straßenabschnitten als hilfreich. Eine ähnlich leistungsfähige Variante des Systems wird in (Laptev et al., 2000) beschrieben. Hier ersetzen Snake-basierte Verfahren die Schritte der Parallelen-gruppierung und der Kreuzungsextraktion. Kontextinformation wird nicht explizit ausgeschöpft, weshalb dieser Ansatz sensibler auf starke Abschattungen und Verdeckungen reagiert.

Ein vergleichbar umfassendes System zur Straßenextraktion aus Luftbildern mittlerer Auflösung wurde am Institut Géographique National (IGN), Paris, entwickelt (Airault et al., 1994; Ruskoné et al., 1994; Ruskoné, 1996). Die Modellierung ist im Vergleich zu (Baumgartner, 2003) weniger explizit und bezieht sich nur auf eine einzige Auflösungsstufe. Sie integriert aber umfangreicheres Wissen über den lokalen Kontext von Straßen. Phase 1 startet mit der Bestimmung von Straßenabschnitten durch Detektion länglicher und homogen heller Regionen über eine Wasserscheidenextraktion im kantengefilterten Luftbild. Diese werden im Folgeschritt durch ein Homogenitäts-Tracking noch um ca. 100m Meter ausgedehnt. Hierzu wird ein Suchbaum aufgespannt und derjenige Pfad selektiert, der ein Maximum an Homogenität und Geradlinigkeit bzw. Rechtwinkligkeit besitzt. Die simultane Extraktion von parallelen Kanten hilft, den Pfad in seiner Lage zu verbessern. Die so entstandenen Straßenachsen werden im Zuge der Netzbildung (Phase 2) durch einen iterativen Gruppierungsalgorithmus miteinander verknüpft und zu Kreuzungen verschnitten. Im Gegensatz zu (Baumgartner, 2003) geschieht die Gruppierung lediglich unter geometrischen und topologischen Gesichtspunkten. Die Verifikation des extrahierten Straßennetzes stützt sich jedoch stark auf die Interpretation von lokalem Kontext. Dazu wird das Netz in kleine Segmente unterteilt, denen in einer überwachten Klassifikation neben der Bedeutung „Straße“ auch die Bedeutungen „Kreuzung“, „Schatten“, „Baum“ oder „Feld“ zugewiesen werden können. Die Validierung der Klassifikation geschieht über einfache Modelle wie z.B. durch die Berechnung typischer Texturparameter für Schatten- und Baum-Regionen oder die Untersuchung zirkulärer Profile in der Umgebung von Kreuzungsmittelpunkten. Komplexer ist dagegen die Validierung von verkehrsreichen Straßen. Mittels eines Neuronalen Netzes werden Fahrzeughypothesen generiert, die, wenn sie sich zu längeren Fahrzeugreihen gruppieren lassen, als Teilstrukturen von Straßen interpretiert werden.

Die Analyse der Ergebnisse in (Ruskoné, 1996) zeigt, dass vor allem die Integration von Wissen über den lokalen Straßenkontext einen sehr positiven Einfluss auf das Endergebnis ausübt. Die ersten Schritte des Verfahrens stützen sich sehr stark auf die Homogenität des Straßenbelags, was den Ansatz z.B. im Vergleich zu (Baumgartner, 2003) weniger anfällig für Störungen in den Straßenrändern macht. Defizite des Ansatzes bestehen jedoch darin, dass das Maßstabsraumverhalten von Straßen ungenutzt bleibt und dass während der Netzbildung auf wertvolle Straßenevidenz aus radiometrischen Merkmalen verzichtet wird.

Ein zu allen obigen Verfahren sehr komplementärer Ansatz wird in (Price, 1999, 2000) vorgestellt. Dieser Ansatz wurde speziell zur Extraktion gitterförmiger Straßennetze, wie sie in amerikanischen Städten oder Militärbasen zu finden sind, entwickelt. Global wird das Straßennetz als eine Kombination von unterschiedlich orientierten und nur geringfügig verformbaren Gittern modelliert. Einige Verbindungen zwischen Gitterknoten (Straßensegmente) dürfen jedoch fehlen und das resultierende Netz muss nicht unbedingt zusammenhängend sein. Durch die explizite Modellierung von Kreuzungen als Gitterknoten, die die Straßensegmente miteinander verknüpfen, wird eine korrekte Topologie der extrahierten (Teil-)Netze sichergestellt. Auf lokaler Ebene wird ein Straßensegment als eine fast horizontale Verknüpfung zweier Kreuzungen mit annähernd parallelen Straßenrändern beschrieben. Ob kleinere Objekte im Innenbereich eines Segments liegen, spielt – sofern sie sich nicht im Digitalen Oberflächenmodell (DOM) abbilden – keine Rolle.

Die Straßenextraktion startet mit der Vorgabe der ersten Masche eines jeden Gitters durch manuelle Initialisierung. Basierend auf Größe und Orientierung der Initialisierung werden iterativ die angrenzenden Maschen hypothetisiert und durch Anpassung an Bildkanten und lokale Minima des DOM verbessert, bewertet und in Güteklassen inklusive einer Ablehnungsklasse eingeteilt. Liegen überlappende Luftbilder vor, wird dieser Schritt für jedes Luftbild separat ausgeführt, wobei im Falle von überlappenden Straßensegmenten dasjenige mit der besten Bewertung maßgebend ist. Nach Abschluss der Netzausdehnung folgt die Überprüfung aus einem globaleren Blickwinkel auf Basis von so genannten „ausgedehnten Straßenabschnitten“, d.h. Straßensegmente werden nun unter Hinzunahme ihrer direkten Nachbarsegmente neu bewertet. So kann z.B. ein schlecht bewertetes Segment durch zwei angrenzende sehr gut bewertete Segmente *auf-* oder ein isoliert stehendes, sehr gut bewertetes Segment *abgewertet* oder aussortiert werden.

Dieses relativ stark modellgetriebene Verfahren erreicht in komplexen, stereotypen Stadtgebieten wie Washington D.C. oder Purdue sehr gute Ergebnisse, was vor allem durch das strenge Modell zu erklären ist. Die manuelle Initialisierung dürfte nur sehr geringen Einfluss auf die Leistungsfähigkeit dieses Systems haben, da diese relativ grobe Startinformation auch über einfache automatische Verfahren gewonnen werden könnte, z.B. mittels der in (Faber und Förstner, 2000) vorgestellten Histogrammanalyse von Linien- und Kantenrichtungen. Neben der Einschränkung auf gitterförmige Straßennetze ist die wenig detaillierte lokale Modellierung ein weiteres Defizit dieses Ansatzes. Die Ergebnisse in (Price, 2000) verdeutlichen, dass die Straßenextraktion nur dann erfolgreich ist, wenn die Straßenränder über ausreichend große Distanzen extrahierbar sind. Ist dies nicht der Fall, treten zwangsläufig Lücken oder Fehlinterpretationen auf. So „schmiegen“ sich teilweise extrahierte Straßen an Dachkanten an.

Zusammengefasst spiegeln die hier dargestellten Ansätze den momentanen Stand der Forschung in der automatischen Straßenextraktion wider. Daneben gibt es noch eine Fülle weiterer Arbeiten, die jedoch abgesehen von Details die theoretische Grundlage nicht erweitern. Umfangreiche Referenzen für Straßenextraktion aus Luftbildern finden sich z.B. in (Mayer, 1998; Baumgartner, 2003) und für Straßenextraktion aus Satellitenbildern in (Wiedemann, 2001).

2.3 Diskussion und Folgerungen

In diesem Abschnitt werden die Kernpunkte der vorgestellten Verfahren zusammengestellt und werden mit Blick auf Datenmaterial und Zielsetzung der vorliegenden Arbeit analysiert. Darauf aufbauend werden Folgerungen für Modellierung und Strategie gezogen. Daher wird die Analyse obiger Verfahren getrennt nach Modellierung und Strategie durchgeführt.

2.3.1 Analyse der Modelle

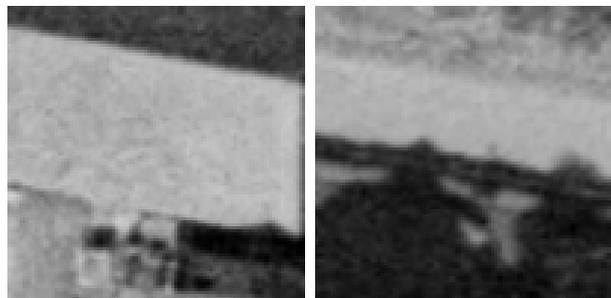
Straßen: Wie in Kapitel 2.1 ausgeführt, ist eine explizite Modellierung des Objekts in der realen Welt und im Bild Voraussetzung für die Transparenz eines Bildanalyse-System. Vor allem bei älteren Arbeiten zur Straßenextraktion ist solch eine Modellierung nicht immer erfolgt, jedoch kann das grundlegende Modell aus den verwendeten Algorithmen und Parametern bzw. deren Beschreibung abgeleitet werden. Betrachtet man zunächst die Modellierung von Straßen ohne Kontext, so gehen im *radiometrischen Teil* des Modells alle Ansätze mit Ausnahme von (Price, 2000) von relativ homogenem Straßenbelag aus, der im Bild je nach verwendeten Auflösungen als einheitlich helle Region (McKeown und Denlinger, 1988; Barzohar und Cooper, 1996; Ruskoné, 1996; Baumgartner, 2003) oder als kontrastreiche, dünne Linie (Fischler et al., 1981; Wiedemann, 2001; Baumgartner, 2003) wiedergegeben wird. In den *geometrischen Teil* des Modells finden vor allem Annahmen über die ausgeprägte Parallelität der Straßenränder, die begrenzte Krümmung und – sofern Höheninformation vorhanden ist – die geringe Steigung der Straße Eingang. Da in Stadtgebieten parallele Straßenränder schwerer als im offenen Gelände zu extrahieren sind, wird diese Eigenschaft in (Ruskoné, 1996) und (Price, 2000) auf lokaler Ebene weniger stark gewichtet. Anstelle dessen müssen die Straßenränder nur über eine größere Distanz parallel sein und dürfen lokale Unterbrechungen aufweisen. Die konsequente Nutzung von geometrischen und radiometrischen Modellteilen stellt bei allen Ansätzen die Grundlage der Extraktion dar, auch wenn ihre jeweilige Gewichtung zwischen den Ansätzen differiert. Im Vergleich hierzu finden *topologische Eigenschaften* des Straßennetzes meist nur bei der Gruppierung kürzerer Straßensegmente zu längeren Einheiten Eingang in das Modell. Lediglich in (Wiedemann, 2001) und (Price, 2000) wird die globale Netztopologie inklusive Kreuzungspunkte explizit modelliert und für die Extraktion genutzt. Durch die Fokussierung auf gitterförmige Straßennetze integriert (Price, 2000) auch eine globale geometrische Eigenschaft. Zwar ist dieses strenge Modell auf Europäische Innenstädte nicht übertragbar, jedoch stellt die globale Netzgeometrie eine wichtige und in den meisten Ansätzen ungenutzte Eigenschaft dar, die für eine abschließende Konsistenzprüfung des extrahierten Straßennetzes wertvolles Potential erschließen würde (Hinz und Wiedemann, 2003). Für Europäische Innenstädte wäre z.B. eine Netzgeometrie denkbar, die aus regelmäßigen Mustern wie Ringen, Rauten oder allgemeinen Vierecken zusammengesetzt ist.

Im Hinblick auf die Straßenextraktion in Stadtgebieten ist es wichtig zu analysieren, wie die Modellierung der Ansätze eine höhere Komplexität der Daten berücksichtigt. Hierbei sind zum einen die störenden Einflüsse von Kontextobjekten zu beachten und zum anderen die größere Vielfalt an Objekten sowie insbesondere die Variabilität des Objektes „Straße“. Damit steigt die Gefahr von Verwechslungen mit anderen Objekten und Fehlinterpretationen. Abbildung 5 greift einige typische Fälle dieser Problematik auf.

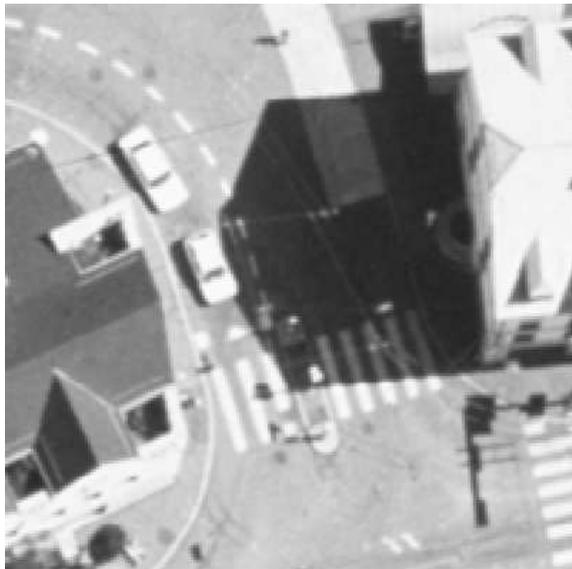
Störungen: Räumlich sehr begrenzte Störungen, wie sie z.B. durch Fahrbahnausbesserungen auftreten können, werden bei allen Ansätzen implizit modelliert. Das heißt, es werden *lokale Abweichungen vom Modell* zugelassen, ohne die eigentliche Ursache der Störung auf Ebene der realen Welt zu interpretieren. So können Störungen während des Trackings in (McKeown und Denlinger, 1988; Ruskoné, 1996) durch das Übergehen einiger schlechter Abtastwerte überbrückt werden. Ähnliches gilt für Detektion sicherer Straßenabschnitte durch die *Fusion verschiedener Abstraktionsstufen* in (Baumgartner, 2003). Auch hier spielt die Art der Störung keine Rolle, sie muss lediglich durch die niedrigere Auflösung eliminiert werden. Die Ansätze (Price, 2000; Baumgartner, 2003) sind darüber hinaus in der Lage, großräumigere Störungen zu behandeln, deren unüberprüfte Überbrückung zu unsicher wäre. Durch die Verwendung von längeren Straßensegmenten bzw. Ziplock-Ribbon-Snakes führen sie



(a) Verdeckungen durch hohe Objekte (hier Vegetation)



(c) Mehrdeutigkeiten: Hausdach (links) und Straße (rechts) nur durch Umgebungsinformation zu differenzieren (oben)



(b) Substrukturen und lokaler Kontext: Markierungen, Gebäudeschatten und Fahrzeuge



(d) Variabilität: Dunkle Straße mit eindeutig definierten Fahrspuren (links), helle Straße ohne Substrukturen (rechts)

Abbildung 5. Charakteristika von Straßen in Stadtgebieten

Zwänge für den Straßenverlauf über eine größere Distanz ein und können dadurch auch *fragmentierte oder lokal schwache Straßenevidenz* ausnutzen. Gravierendere Einflüsse von Hintergrundobjekten auf das Erscheinungsbild von Straßen kann eine solche Modellierung jedoch nicht mehr in den Griff bekommen. Eine *Erweiterung des Modells um Kontextwissen* ist hierfür unumgänglich. Die hierin am weitesten entwickelten Ansätze (Ruskoné, 1996; Baumgartner, 2003) können durch die Analyse von lokalem Kontext beispielsweise Schattenbereiche oder partielle Verdeckungen überbrücken. Für die Straßenextraktion in Stadtgebieten ist jedoch zu beachten, dass Relationen zwischen Kontextobjekten und Straßen im Vergleich zur offenen Landschaft wesentlich häufiger vorkommen und der gegenseitige Einfluss auf das Erscheinungsbild wesentlich stärker wird. Daher sollten nicht nur jene Relationen modelliert werden, die Störungen verursachen, sondern auch jene, die die Straßenextraktion unterstützen. Ein Beispiel hierfür ist in (Ruskoné, 1996; Ruskoné et al., 1996) gegeben, wo gruppierte Fahrzeuge als die Teilstruktur „Fahrspur“ einer Straße herangezogen werden.

Variabilität und Fehlinterpretationen: Die Problematik, dass durch die Vielfalt der Objekte mit Mehrdeutigkeiten gerechnet werden muss, wird in den bisherigen Arbeiten nur ansatzweise behandelt. Oft wird die Variabilität implizit durch die Verwendung unterschiedlicher Bildverarbeitungsoperatoren und Parametereinstellungen oder durch stochastische Modelle abgefangen (Fischler et al., 1981; Barzohar und Cooper, 1996). Dagegen werden in (McKeown und Denlinger, 1988) signifikante Variationen einer Straße, z.B. die Änderung des Fahrbahnbelags, durch *automatische Schlussfolgerung* aufgrund bestimmter Ereignisse in der Kanten- und Profilverfolgung detektiert und explizit beschrieben. Wie auch bei den zuvor genannten Ansätzen besteht aufgrund der relativ einfachen Modellierung von Straßen die Gefahr, dass es zu Fehlinterpretationen kommt und ein falsches Objekt extrahiert wird. In (Ruskoné, 1996) wird versucht, Fehlinterpretationen durch die abschließende Analyse von lokalem Kontext zu identifizieren. Dies setzt voraus, dass die anfängliche Straßenextraktion auf einem so allgemein gehaltenen Modell basiert, dass auch stark variierende Straßenabschnitte extrahiert werden können, um sie anschließend überhaupt analysieren zu können. Bei der Anwendung eines sehr allgemeinen Modells auf komplexe Szenen entsteht jedoch unweigerlich eine unübersehbare Anzahl von Hypothesen, deren Bearbeitung auch mit sehr leistungsfähigen Computern nicht machbar ist. Geeigneter ist ein Modell, das einerseits einen großen Detailreichtum aufweist, so dass Mehrdeutigkeiten von vorneherein weniger wahrscheinlich sind. Andererseits sollte es ausreichend flexibel sein, so dass die hohe Variabilität von Straßen gehandhabt werden kann. Einen guten Anhaltspunkt für eine detaillierte Straßenmodellierung gibt der GIS-gestützte Ansatz (de Gunst, 1996; Vosselman und de Gunst, 1997). Auch wenn die Extraktionskomponente dieses Ansatzes aufgrund der einfachen Bildverarbeitungsmethoden lediglich für Hauptverkehrsstraßen und Autobahnen in der offenen Landschaft und nicht für Stadtgebiete geeignet ist, erscheint die Modellierung von Straßen inklusive ihrer Fahrspuren, Abbiegespuren und Markierungen wegen des hohen Detaillierungsgrads besonders für komplexere Szenen vielversprechend. Als Zusatz oder Alternative zu Markierungen und Fahrspuren liegt es nahe, wie in (Ruskoné, 1996) über das Kontextmodell Fahrzeuge zu integrieren.

Zusammenfassend lassen sich als **Konsequenzen für das Modell** folgende Punkte festhalten:

- ◇ Das Modell sollte geometrische, radiometrische und topologische Straßeneigenschaften detailliert erfassen, um Mehrdeutigkeiten bestmöglich ausschließen zu können. Dies impliziert auch die Ausschöpfung von Höheninformation, sofern diese in der Datengrundlage enthalten ist.
- ◇ Das Modell sollte flexibel sein. Es sollte verschiedene Alternativen bieten, um gleichzeitig Detailreichtum und Variabilität der Straßen erfassen zu können.
- ◇ Das Modell sollte robust gegenüber Störungen und Einflüssen anderer Objekte sein. Neben dem Potential, auch schwächere Straßenevidenz auszuschöpfen, sollte das Modell vor allem um Kontextinformation erweitert werden. Lokaler Kontext sollte so modelliert werden, dass er nicht nur zur Analyse von Extraktionslücken dienen kann, sondern auch als Hinweis auf potentielle Straßenstücke genutzt und zur Adaption von Modellparametern herangezogen werden kann.
- ◇ Das Modell sollte mehrere Abstraktionsstufen beinhalten. Diese Art der Modellierung unterstützt eine effiziente Steuerung und ermöglicht eine fokussierte Extraktion der Straßeneigenschaften.
- ◇ Das Modell sollte konsequent in all seinen Bestandteilen explizit formuliert werden, um bestmögliche Transparenz des Bildanalysesystems für Analysen zu gewährleisten.

2.3.2 Analyse der Strategien

Generell gesehen zeichnen sich die Strategien der oben besprochenen Verfahren mit Ausnahme von (Price, 2000) durch ein relativ einheitliches Konzept aus:

- (1) Die Extraktion startet mit der Detektion von sicheren Straßenabschnitten (Phase 1).
- (2) Die sicheren Straßenabschnitte werden durch Tracking oder Verknüpfung zu einem Netz verbunden (Phase 2).

Detaillierter betrachtet unterscheiden sich die Ansätze jedoch neben algorithmischen Aspekten vor allem darin, wann welches Modellwissen zur Bearbeitung bestimmter Aufgaben einfließt und welche Bewertungs- und Kontrollmechanismen eingesetzt werden, um Erkenntnisse aus vorausgegangenen Schritten in Planung und Ausführung von Folgeschritten einzubringen.

In Phase 1 ist es wichtig, sichere aber auch genügend viele Startabschnitte zu detektieren, da sonst Phase 2 schwieriger oder unter Umständen gar nicht lösbar ist. Daher startet die Extraktion in der Regel mit Algorithmen, die Modellwissen über geometrische und radiometrische Straßeneigenschaften relativ umfassend ausschöpfen. Meist handelt es sich hierbei um die Detektion von längeren, homogenen Straßenflächen inklusive ihrer Ränder mittels Verfahren der Linien- und/oder Kantenextraktion, Parallelengruppierung sowie Homogenitätsanalyse (Fischler et al., 1981; Zlotnick und Carnine, 1993; Barzohar und Cooper, 1996; Ruskoné, 1996; Wiedemann, 2001; Baumgartner, 2003). In Stadtgebieten sind diese Verfahren jedoch weniger robust als im offenen Gelände, da sie nicht nur Straßenflächen sondern z.B. auch viele Hausdächer extrahieren. Dies unterstreicht die prinzipielle Ausrichtung obiger Ansätze auf die offene Landschaft. Das Wissen über die eingeschränkte Tauglichkeit solcher Verfahren wird lediglich in (Baumgartner, 2003) explizit zur Steuerung der Extraktion in einer sehr frühen Phase genutzt, indem vor der eigentlichen Extraktion eine texturbasierte Klassifikation globaler Kontextgebiete stattfindet. Damit in Stadtgebieten dennoch sichere Straßenabschnitte extrahiert werden können, ist es nötig, bereits in der ersten Phase sehr detailliertes Wissen über Straßen zu nutzen. Dies ist angesichts der hohen Komplexität solcher Szenen nur möglich, wenn im Gegensatz zu (Ruskoné, 1996; Baumgartner, 2003) bereits *zu Beginn* neben globaler auch lokale Kontextinformation extrahiert wird und für Folgeschritte zur Verfügung steht. Dies stellt im Vergleich zu den bisherigen Ansätzen eine grundlegend komplementäre Vorgehensweise dar.

Zur Verbindung der Startabschnitte (Phase 2) kommen in der Regel Tracking- und straßenspezifische Verknüpfungsverfahren zum Einsatz. Tracking-Verfahren wie in (McKeown und Denlinger, 1988; Barzohar und Cooper, 1996; Ruskoné, 1996) besitzen den Vorteil, dass das Bildmodell des Trackings, d.h. das Referenzprofil, aus dem jeweiligen Startsegment gewonnen werden kann. Es ist somit besser an das Erscheinungsbild der aktuellen Straße angepasst, als wenn es von vorneherein vorgegeben ist. Nachteilig gegenüber einer Verknüpfung mit *beidseitigem* Anschluss wie in (Wiedemann, 2001; Baumgartner, 2003) ist jedoch, dass die elementare Funktion von Straßen, nämlich eine effiziente Verbindung zwischen zwei Punkten herzustellen, nicht ausgenutzt wird. Desweiteren können solche Verfahren ebenfalls an die lokalen Eigenschaften der Straße angepasst werden, wie dies z.B. bei Ziplock-Ribbon-Snakes der Fall ist. Daher sind mit einigen wenigen Ausnahmen⁶ Gruppierungsverfahren zur Verknüpfung von Startabschnitten vorzuziehen. Das hierfür in (Baumgartner, 2003) vorgestellte Verfahren, verschiedene Arten von Modell- und Kontextwissen für Hypothesenbildung und Verifikation einzusetzen, ist auch für die Straßenextraktion in Stadtgebieten als außerordentlich vielversprechend anzusehen. Durch eine Kombination mit globalen Gruppierungskriterien, die die Netzcharakteristik von Straßen berücksichtigen, ließe sich nochmals eine Verbesserung erreichen. Für die Adaption eines solchen Verfahrens auf Stadtgebiete ist jedoch zu bedenken, dass vor allem die Methoden zur Verifikation wegen der zusätzlichen Einflussfaktoren erheblich erweitert werden müssen. (Price, 2000) umgeht diese Problematik durch eine wesentlich strengere, modellgetriebene Hypothesenbildung für die Straßensegmente. Die Strategie unterscheidet sich insofern von allen anderen Arbeiten, dass im Zuge der Netzausdehnung eine Übersegmentierung – also viele Straßensegmente aber teilweise auch falsche – in Kauf genommen wird. Die Übersegmentierung wird anschließend anhand von einfachen Bewertungskriterien reduziert. Der Gesichtspunkt, auch für initiale Straßenabschnitte zunächst viele Hypothesen aufzustellen, diese anschließend unter Hinzunahme von bislang ungenutztem Wissen zu bewerten und zu selektieren, ist angesichts des starken Einflusses von Hintergrundobjekten sehr wertvoll. Allerdings ist das in (Price, 2000) eingesetzte Verfahren der Hypothesenbildung durch maschenweise Netzausdehnung für weniger stereotype Stadtgebiete nicht erfolgversprechend, und die involvierten Bewertungskriterien sind zu wenig an die Komplexität angepasst.

⁶ Ausnahmen sind Sackgassen oder der Fall, wenn das Straßennetz bis an den Bildrand ausgedehnt werden soll.

Ein weiterer Gesichtspunkt, der die Leistungsfähigkeit einer Strategie wesentlich bestimmt, ist die Bereitstellung eines geeigneten Kontrollmechanismus. Da bei der Bearbeitung komplexer Szenen eine Vielzahl von alternativen Hypothesen zu testen ist, hat hierbei die interne Bewertung von Zwischenergebnissen besonders hohe Bedeutung (Förstner, 1996). Die Bewertungsmaße stellen die Grundlage dar, um geeignete Schlussfolgerungen über die nächsten Schritte zu ziehen, und sind somit der Schlüssel für eine flexible Steuerung der Extraktion. In Arbeiten zur Straßenextraktion wurde diese so genannte *interne Bewertung* bislang nicht einheitlich und durchgängig behandelt sondern allenfalls in Teilaspekten, z.B. in (McKeown und Denlinger, 1988; Barzohar und Cooper, 1996; Geman und Jedynek, 1996). Verschiedene andere Forschungsarbeiten zur automatischen Objektextraktion greifen stärker auf theoretisch fundierte Bewertungsschemata zurück, allen voran Bayes-, Evidenz- und Fuzzy-Netze (Fischer et al., 1998; Tönjes und Growe, 1998; Tupin et al., 1999; Kim et al., 2001; Heuel und Förstner, 2001). Bei der praktischen Anwendung solcher anwendungsunabhängigen Schemata erweist es sich jedoch oft als schwierig zu zeigen, dass Datengrundlage und Modellierung alle Anforderungen der Theorie erfüllen. Aus den genannten Gründen ist bei der Entwicklung der Kontrollkomponente vor allem der *anwendungsabhängige* Teil von Bewertung und Steuerung kritisch zu beleuchten. Defizite bestehen bisher hauptsächlich in der Formalisierung, in der Frage, welches Wissen zur Extraktion und welches zur Bewertung eingesetzt wird, und in einer durchgängigen und konsistenten Umsetzung für die Extraktion. Wegen der hohen Relevanz für die Straßenextraktion in Stadtgebieten wird auf die interne Bewertung von Extraktionsergebnissen in einem separaten Kapitel (Kapitel 4) ausführlich eingegangen.

Folgende Punkte lassen sich als **Konsequenzen für die Strategie** festhalten:

- ◊ Die grundlegende Strategie, zunächst sichere Straßenabschnitte zu detektieren und daraus ein Straßennetz aufzubauen, erscheint im Vergleich zur streng modellbasierten Netzausdehnung von (Price, 2000) als die bessere Alternative. So können auch geometrisch unregelmäßige Straßennetze extrahiert werden.
- ◊ Kontextwissen sollte nach Möglichkeit bereits zu Beginn genutzt werden, um es für Folgeoperationen verfügbar zu machen und dadurch eine flexiblere Strategie zu ermöglichen.
- ◊ Zur Straßendetektion und Straßennetzbildung sollten alternative, hochspezialisierte Methoden bereitgestellt werden, um der jeweiligen kontextuellen Situation Rechnung zu tragen.
- ◊ Bei der Straßennetzbildung sind straßenspezifische Verknüpfungsverfahren Trackingverfahren vorzuziehen.
- ◊ Die Kontrolle über die Strategie sollte sich auf interne Bewertungsmaße stützen. Um die Gültigkeit der internen Bewertung zu gewährleisten, sollten anwendungsabhängige und -unabhängige Teile der Kontrolle so gewählt werden, dass das zur Bewertung eingebrachte Wissen transparent und ohne Verletzung der theoretischen Grundlagen des Bewertungsschemas genutzt werden kann.

Die **Literaturanalyse abschließend** können folgende vier Defizite bisheriger Arbeiten hinsichtlich der Straßenextraktion in urbanen Gebieten identifiziert und dementsprechende Herausforderungen für diese Arbeit abgeleitet werden:

- (1) Der Detaillierungsgrad der verwendeten Straßenmodelle ist für Stadtgebiete nicht ausreichend hoch. Um Fehlinterpretationen zu minimieren und der Variabilität von urbanen Straßen Rechnung zu tragen, müssen mehr Teilobjekte wie Fahrspuren und Markierungen in das Modell integriert werden.
- (2) Die systematische Nutzung von Kontextinformation wurde bisher vorwiegend im Hinblick auf die Straßenextraktion in der offenen Landschaft vorangetrieben. Insbesondere die Modellierung von lokalem Kontext ist für eine zuverlässige Straßenextraktion in urbanen Gebieten erheblich auszudehnen. Dies beinhaltet auch die konsequente Nutzung von Höheninformation.
- (3) Große Defizite herrschen bei der systematischen Modellierung interner Bewertungskriterien und der durchgängigen Nutzung von internen Qualitätsmaßen während der Extraktion. Vor allem mit Blick auf die hohe Komplexität urbaner Szenen und der daraus resultierenden Notwendigkeit einer ausreichend flexiblen Extraktionsstrategie sollte dieser Bereich deutlich vertieft werden.
- (4) Die grundlegende Strategie, zunächst auf leicht und sicher zu extrahierende Teile des Straßennetzes zu

fokussieren und diese anschließend zu verbinden, ist auch in Stadtgebieten anwendbar. Zur Extraktion sicherer Teilstücke ist jedoch der beträchtliche Einfluss des lokalen Kontexts auf das Erscheinungsbild der Straßen zu berücksichtigen. Daher muss im Unterschied zu den bisherigen Ansätzen sehr frühzeitig und robust lokale Kontextinformation extrahiert werden, was zum Teil die Entwicklung nahezu autarker Extraktionsansätze für Kontextobjekte impliziert.

Das folgenden Kapitel geht nun auf die ersten beiden Punkte ein, d.h. die Entwicklung eines Modells für das urbane Straßennetz inklusive seines Kontexts.

3 Modellierung

Grundlage jeder Objekterkennung ist ein Modell des zu extrahierenden Objekts. Nur mit Hilfe eines Modells kann den implizit enthaltenen Informationen eines Bildes eine Bedeutung in der realen Welt zugewiesen werden. Daher muss das Ziel der Modellbildung sein, das verfügbare Wissen über das Objekt so zu strukturieren, zu formalisieren und zu repräsentieren, dass es allein auf Basis von Algorithmen gelingt, Semantik und Eigenschaften des Objekt anhand der gespeicherten Intensitätswerte zu rekonstruieren. Wie im vorangegangenen Kapitel ausgeführt, ist es dabei sehr hilfreich, ein Objektmodell auf Ebene der realen Welt aufzustellen und es mittels eines Sensormodells in ein Bildmodell zu überführen. Dies hat den Vorteil, dass die Erscheinungsform des Objektes und seiner Eigenschaften im Bild mit der dem Menschen vertrauten Form der hierarchischen Strukturierung und symbolischen Beschreibung von Objekten auf explizite Weise verbunden werden kann. Darüber hinaus ist es besonders bei der Behandlung komplexerer Szenen nötig, Wissen über die umgebenden Einflüsse auf das Objekt und seine Erscheinungsform in das Modell einzubringen. Dies geschieht durch ein Kontextmodell.

In den folgenden Abschnitten wird das Modell für die Straßenextraktion in Siedlungsgebieten entwickelt. Als Basis dafür werden zunächst in Abschnitt 3.1 die Datengrundlage und das Sensormodell sowie deren relevante Charakteristika mit Blick auf das Straßenmodell vorgestellt. Sie definieren in gewissem Sinne die Randbedingungen dieses Ansatzes, da sich die Extraktion auf solche Straßeneigenschaften stützen muss, die in den Daten zuverlässig zu beobachten sind. Anschließend wird in Abschnitt 3.2 Wissen über Straßeneigenschaften in Form von einzelnen Komponenten des Objekt- und Bildmodells zusammengestellt, strukturiert und in eine einheitliche Repräsentationsform gebracht. Das so aufgebaute Straßenmodell stellt den kontextunabhängigen „Standardfall“ des Modells dar. Abschnitt 3.3 erweitert das Straßenmodell schließlich um das Kontextmodell.

3.1 Datengrundlage und Sensormodell

Als primäre Datenquelle des hier entwickelten Ansatzes dienen qualitativ hochwertige panchromatische Luftbilder, wie sie üblicherweise in einem photogrammetrischen Bildflug über Stadtgebieten aufgenommen werden. Aufgrund des starken „Reliefs“ von Stadtgebieten werden hier Bildverbände oft mit einer Längsüberdeckung von mindestens 60% und einer Querüberdeckung von 30% – 45% bzw. teilweise auch 60% geflogen (Eidenbenz et al., 2000; Zimmermann, 2001). Dank automatisierter Flugnavigation und stabilisierter Kameraplattform kann der gesamte Bildverband annähernd im Nadirblick aufgenommen werden. Daher stehen neben günstigen Blickwinkeln auch in vielen Fällen mehrere Sichten auf die abgebildeten Objekte zur Verfügung. Der Maßstab der Luftbilder liegt im Bereich von 1 : 3 000 – 1 : 10 000, so dass sich durch nachträgliches Scannen einer analogen Photographie mit $7.5\mu\text{m} - 15\mu\text{m}$ eine Bodenpixelgröße von etwa 5cm – 15cm ergibt. Unter der Annahme, dass entsprechend hochqualitative Filme bzw. Scanner verwendet wurden, garantiert dies eine ausreichende Auflösung zur Abbildung von Detailstrukturen wie Markierungen und Fahrzeuge (siehe Abbildung 6). Weiterhin kann bei einem Bildflug neben der Bestimmung aller relevanten Aufnahme- und Orientierungsparameter auch eine genaue Kalibrierung der Kamera ohne Einschränkung vorausgesetzt werden. Daher darf als grundlegendes geometrisches Sensormodell die zentralperspektive Abbildung gelten. Die alternative Verwendung von 3-Zeilenscanner-Bildern würde lediglich eine veränderte Transformationvorschrift zwischen Objekt- und Bildraum nach sich ziehen; sie stellt aber keine prinzipielle Limitierung des Ansatzes dar.

Die radiometrische Auflösung der digitalen Bilder beträgt meist noch 8bit pro Kanal. Daher wurde die Modellierung auf ein Sensormodell mit dieser Bittiefe abgestimmt. Weiterhin wird von lediglich einem einzigen panchromatischen Kanal ausgegangen. Mit weiteren Farb- oder Infrarotkanälen stünden zwar zusätzliche Informationen über die Reflexionseigenschaften der Objekte zur Verfügung, der Gewinn – im Hinblick auf die Straßenextraktion in Stadtgebieten – hält sich jedoch in Grenzen. Dies liegt zum einen an der Tatsache, dass sich Straßen, wie auch viele andere Objekte, typischerweise durch strukturelle Bildmerkmale wie Linien und Kanten auszeichnen, die in panchromatischen Bildern ähnlich robust detektiert werden können wie in Farbbil-



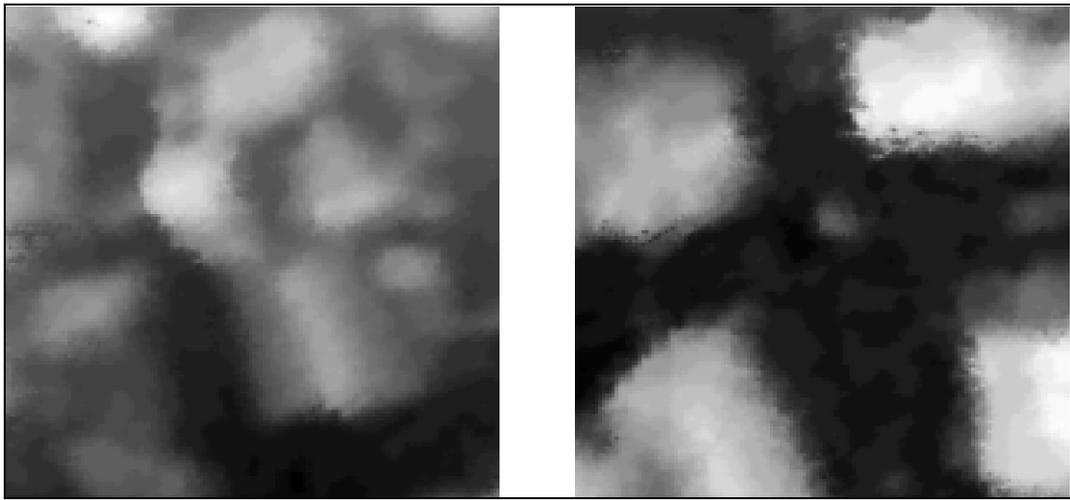
Abbildung 6. Detailerkennbarkeit: Markierungen und Fahrzeuge bei einer Bodenpixelgröße von 14cm ; auch Teilobjekte von Fahrzeugen werden abgebildet (Dach, Windschutzscheibe, etc.)

dern (siehe hierzu auch (Steger, 2000)⁷). Zum anderen bilden sich viele Objekte wie Gebäude und Vegetation, deren Farbinformation zur Abgrenzung von Straßen hilfreich wäre, auch in einem Digitalen Oberflächenmodell (DOM) ab. Die Verwendung eines DOM ermöglicht darüber hinaus über eine Projektion vom Bild- in den Objektraum die Integration mehrerer Sichten sowie die Ableitung weiterer Kontextinformation wie z.B. potentielle Verdeckungs- und Schattenbereiche. Da diese Aspekte besonders für die Objektextraktion in Stadtgebieten von großer Bedeutung sind, wird in dem vorliegenden Ansatz auf die Verwendung von Farbinformation verzichtet und stattdessen auf die Nutzung von Höheninformation fokussiert.

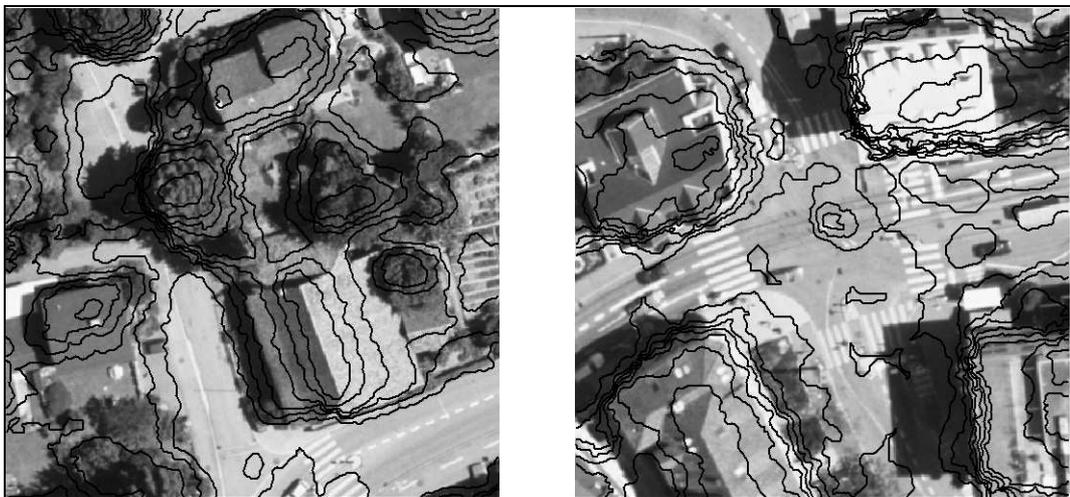
Die beiden wesentlichen Gesichtspunkte bei der Verwendung eines DOM für die Bildanalyse sind seine Genauigkeit und seine inhaltliche Konsistenz mit den Bildern. Um letztere von vorneherein sicherzustellen, ist entweder eine zeitgleiche Aufnahme von Bild- und Höhendaten nötig – z.B. mit einem integrierten System zur Bild- und Laserdaten-Aufzeichnung – oder die Höheninformation muss aus den Bildern nachträglich abgeleitet werden. Diese Möglichkeit führt zwar im Vergleich zu direkten Aufnahmen mit integrierten Sensorsystemen zu einem Qualitätsverlust des DOM, sie verursacht jedoch geringere Kosten, und dementsprechend sind solche Datensätze besser verfügbar. Daher wird im vorliegenden Ansatz von den Qualitätsmerkmalen eines DOM ausgegangen, wie sie typischerweise durch kommerzielle automatische Bildzuordnungsverfahren erreicht werden können. Solche Höhenmodelle erreichen in Stadtgebieten eine Höhen- und Lagegenauigkeit von etwa 1m – 2m . Wie auch aus Abbildung 7 ersichtlich, hat dies zur Folge, dass größere Objekte wie Bäume, Gebäude und zum Teil auch parkende Fahrzeuge im DOM enthalten sind, dass aber präzise geometrische Informationen kaum abzuleiten sind. Weiterhin ist zu beachten, dass üblicherweise die Höhe Z als Funktion der Lagekoordinaten X, Y beschrieben wird ($Z = f(X, Y)$) und somit keine echte 3D-Beschreibung des Objektraumes verfügbar ist. In Abbildung 8 sind das Höhenmodell eines Innenstadtbereiches und ein korrespondierender Bildausschnitt dargestellt. Man erkennt deutlich, dass die Höhenverhältnisse zwar einigermaßen realistisch wiedergegeben werden. Der Verlauf eines exemplarischen Höhenprofils senkrecht zu einer Häuserschlucht weist jedoch eindeutig nach, dass exakte geometrische Informationen wie Bruchkanten und vertikale Wände nicht enthalten sind.

Zusammenfassend lässt sich festhalten, dass sich der Ansatz auf qualitativ hochwertiges, in der Praxis aber bereits etabliertes Bildmaterial stützt. Um die inhaltliche Konsistenz zu wahren, wird lediglich auf Daten zurückgegriffen, die im Zusammenhang mit einem Bildflug aufgenommen und abgeleitet werden. Dies hat aus operationeller Sicht den günstigen Nebeneffekt, dass die Kette der Datenvorverarbeitung bereits zu einem hohen

⁷ Das Ergebnis einer Linien- und Kantenextraktion in Grauwert- und Farbbildern zeigt nur in wenigen Ausnahmefällen signifikante Unterschiede, z.B. wenn eine intensiv rot gefärbte Fläche an eine ebenso intensiv grün gefärbte Fläche angrenzt (Steger, 2000).



(a) Ausschnitte des DOM als höhencodierte Grauwertbilder (korrespondierende Bildausschnitte: siehe b)



(b) Wohngebiet und Vegetation (links) und Kreuzungsbereich (rechts) mit Höhenlinien (Äquidistanz $2m$).

Abbildung 7. Eigenschaften von Bild- und Höhendaten:

(a) und (b) beschränkte Genauigkeit des DOM; man beachte vor allem die unpräzise geometrische Abbildung von Gebäuden und den fehlerhaften „Hügel“ im Kreuzungsbereich.

Grad automatisiert ist. Die mindere Qualität des Höhenmodells muss aber bei der Modellierung berücksichtigt werden.

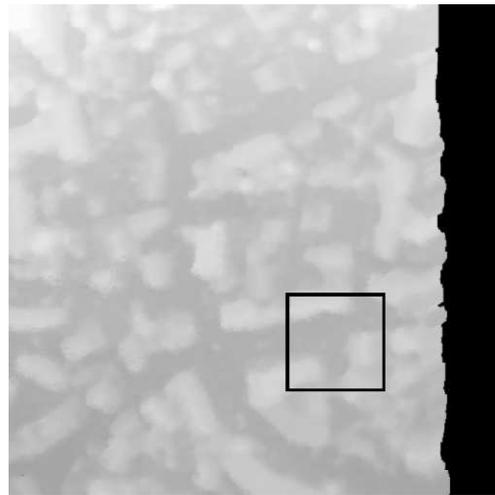
3.2 Straßenmodell

In diesem Abschnitt wird das der Extraktion zugrunde liegende Straßenmodell entwickelt. Es beschreibt Straßen unabhängig von ihrer Umgebung und stellt somit den „Standardfall“ des Modells dar. Wissen über umgebende Einflussfaktoren wird separat durch das Kontextmodell beschrieben (Abschnitt 3.3).

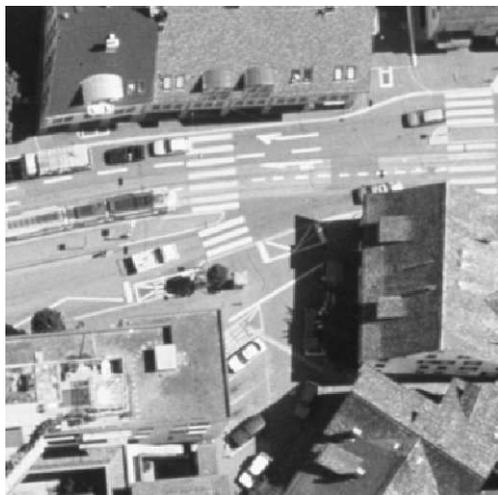
Wie in Abschnitt 2.3 festgestellt, sollte das Straßenmodell aufgrund der hohen Komplexität urbaner Szenen möglichst viel des verfügbaren Wissens über Straßen beinhalten. Hierzu ist es insbesondere nützlich, die Funktion von Straßen detaillierter zu beleuchten, weil sich die meisten physikalischen, geometrischen und topologischen Eigenschaften direkt aus ihr ableiten lassen. Deshalb geht Abschnitt 3.2.1 zunächst auf das Wissen über Straßen in der realen Welt ein und stellt die Eigenschaften des Objektmodells zusammen. Diese werden



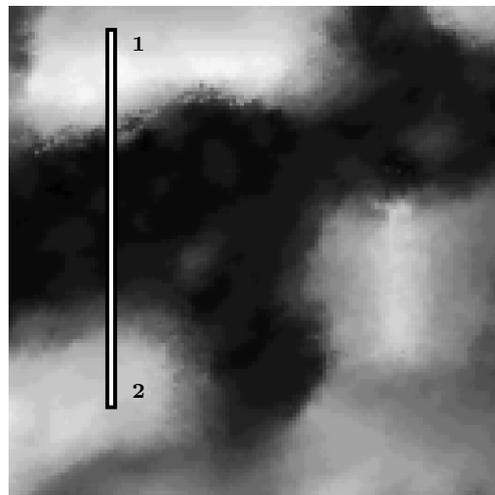
(a) Luftbildausschnitt



(d) DOM-Ausschnitt zu (a), Höhen grauwertcodiert



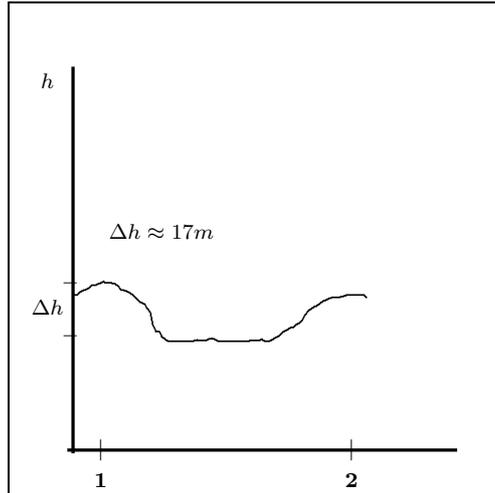
(b) z.T. verdeckter Straßenabschnitt (Detail aus a)



(e) korrespondierende Häuserschlucht im DOM



(c) Straßenabschnitt in anderem Luftbild (vgl. b)



(f) Höhenprofil (siehe e)

Abbildung 8. Charakteristika der Datengrundlage bedingt durch das „Stadtreief“: (a) – (c) Bildinformation: partielle Verdeckungen; (d) – (f) Höheninformation: geometrische Ungenauigkeiten.

anschließend in Abschnitt 3.2.2 in Form eines Bildmodells auf die Datenebene übertragen. Der Begriff „Bildmodell“ steht hier als Oberbegriff für die *Abbildung* des Objektmodells auf eine beliebige Datengrundlage, also auch auf das DOM oder reduzierte Auflösungsstufen des Originalbildes. In Abschnitt 3.2.3 werden schließlich die Komponenten von Objekt- und Bildmodell mit Hilfe eines Semantischen Netzes strukturiert und in eine einheitliche, im Computer implementierbare Repräsentationsform gebracht.

3.2.1 Komponenten und Eigenschaften des Objektmodells

Schon (Stark und Bowyer, 1991) stellen heraus, dass für eine ausgefeilte Modellierung eine detaillierte Betrachtung der *Funktion* der Objekte sinnvoll ist. Straßen dienen in erster Linie zur Erschließung des menschlichen Lebensraumes. Sie stellen Verbindungen zwischen besiedelten Bereichen her und erschließen für den Menschen bedeutsame Gebiete wie Wohnsiedlungen, Agrarland, Wälder, Rohstoffabbauregionen und Freizeitgebiete. Dies hat zur Folge, dass Straßen – abgesehen von vernachlässigbaren Ausnahmen – ein topologisch komplett verknüpftes Netz ergeben. Das Netz ist um so dichter, je intensiver die Nutzung eines Gebietes durch den Menschen ist. Dazu kommt besonders in hochentwickelten Ländern die Bildung einer Straßenklassenhierarchie, die in vielen Fällen durch rechtliche Widmungen genormt ist (z.B. Autobahnen, Überlandstraßen, Ortsstraßen, ...). Die jeweilige Widmung zieht wiederum bestimmte Bauvorschriften nach sich und beeinflusst dadurch wichtige Eigenschaften für die Extraktion wie minimale Breite, maximale Krümmung, maximale Steigung usw. Daneben spielen vor allem die Materialeigenschaften eine maßgebende Rolle, da durch sie das Reflexionsverhalten von Straßen entscheidend geprägt ist. Wegen der Verwendung von Sensoren im sichtbaren EMS Bereich ist jedoch nur der aufgetragene Straßenbelag und nicht der Zustand von tieferliegenden Schichten von Bedeutung. Eine letzte Kategorie bilden schließlich Objekte, die eine eigene Semantik besitzen, aber praktisch nur im Zusammenhang mit Straßen auftreten. Diese so genannten *Teilobjekte* sind z.B. Fahrbahnen, Fahrspuren, Markierungen, Leitplanken und weitere Bestandteile der Straßenmöblierung.

Im Folgenden wird das für die Extraktion relevante Wissen für Straßen in Siedlungsgebieten in Form von Modellkomponenten spezifiziert und in vier Kategorien eingeteilt: *Geometrische Eigenschaften*, *physikalische Eigenschaften*, *topologische Eigenschaften* und *Teilobjekte*. Die Modellkomponenten sind in den folgenden Absätzen kursiv gekennzeichnet.

Kategorie geometrische Eigenschaften: Aus der grundlegenden Funktion von Straßen, eine für Kraftfahrzeuge (KFZ) taugliche Verbindung zwischen zwei Punkten herzustellen, ergibt sich in erster Linie die *langgestreckte, bandähnliche Struktur* von Straßen. Aus dem gleichen Grund können Straßen nur eine *beschränkte Steigung und Querneigung* besitzen. Um einen zügigen Verkehrsfluss zu ermöglichen, müssen Straßen zumindest *abschnittsweise geradlinig* verlaufen. Abrupte Richtungsänderungen treten hauptsächlich an Kreuzungspunkten oder bei Zufahrten in Wohn- und Gewerbegebieten auf. Im Gegensatz zur offenen Landschaft haben Straßen in Siedlungsgebieten nur *stückweise konstante Breite bzw. parallele Straßenränder*. Änderungen in der Breite werden u.a. durch Abbiegespuren vor Kreuzungen, durch seitliche Parkbuchten, durch Vegetation oder Grünstreifen sowie durch andere verkehrsberuhigende Maßnahmen verursacht.

Kategorie physikalische Eigenschaften: Straßen müssen unter Umständen tonnenschweren KFZ eine dauerhaft befahrbare Unterlage bieten. Neben dem üblichen *Straßenbelag aus Asphalt oder Beton* kommt bei Straßen in Stadtgebieten unter anderem auch *Kopfsteinpflaster* zum Einsatz. Andere Beläge wie Lehm oder Kies finden zwar weltweit gesehen eine große Verbreitung, sie spielen jedoch in Stadtgebieten hochentwickelter Länder, die im Zentrum dieser Arbeit stehen, keine Rolle. Dafür ist aber zu beachten, dass sich auf vielbefahrenen Straßen *streifenartige Partikelrückstände* durch Reifenabrieb und Auspuffabgase absetzen, die zu einer signifikanten Verfärbung des Straßenbelags beitragen.

Kategorie topologische Eigenschaften: Straßen bilden in Städten ein ausnahmslos *zusammenhängendes Netz*. Ein abruptes Ende einer Straße muss durch ein anderes Objekt erklärt werden, das dieser Straße eine Bedeutung als Zufahrtsstraße zuweist, z.B. ein Wohn- oder Industriegebäude. Einzelne Straßen des Netzes werden durch *Kreuzungen* miteinander verknüpft. Die Verknüpfung erfolgt in aller Regel *höhengleich*. Lediglich im Fall von Stadtautobahnen kommen Überführungen mit speziellen Auf- und Abfahrten zum Einsatz. Prinzipiell ist wie in der offenen Landschaft auch in Stadtgebieten eine Hierarchie von Straßenklassen zu beobachten (Stadtautobahnen, Stadtstraßen, Wohnstraßen, . . .). Aufgrund von Zwängen aus der bebauten Umgebung – Europäische Innenstädte sind in der Regel über Jahrhunderte gewachsen – sind jedoch wesentlich mehr Ausnahmeregelungen als in der offenen Landschaft anzutreffen, so dass eine Nutzung des Wissens über Straßenhierarchien hier weniger erfolgversprechend scheint und im weiteren Verlauf der Arbeit außer Acht bleibt.

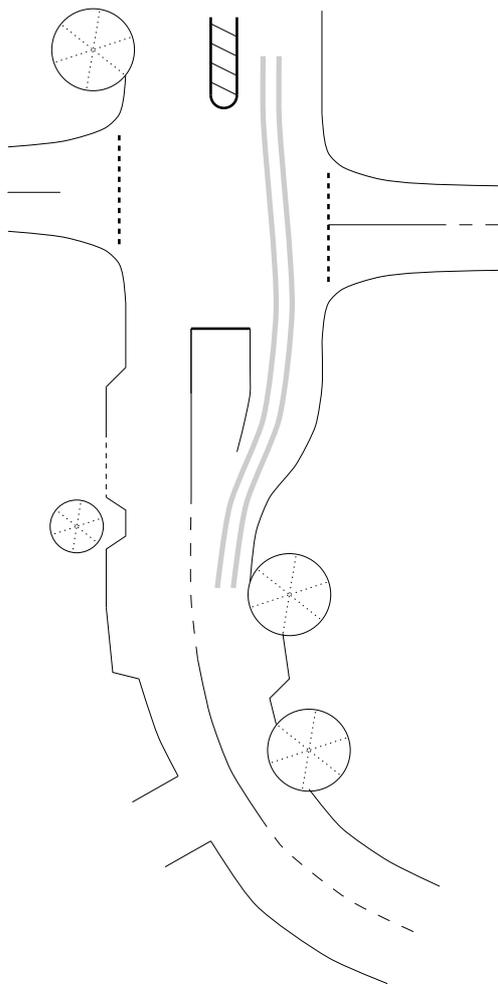
Kategorie Teilobjekte: Straßen bestehen in den allermeisten Fällen aus *ein bis zwei Fahrbahnen*, die sich ihrerseits aus parallel verlaufenden *Fahrspuren* zusammensetzen. Kreuzungsbereiche bilden hier insofern eine Ausnahme, als dass Fahrbahnen und Fahrspuren gegenseitig nicht parallel ausgerichtet sein müssen. Oft werden zur Abgrenzung einzelner Fahrspuren *lineare, durchgezogene oder strichlierte Markierungen* aufgetragen. Dies ist in Stadtgebieten besonders häufig der Fall, weil der Verkehrsfluss hier wesentlich intensiver als in außerstädtischen Bereichen ist und daher aus Sicherheitsgründen stärker gelenkt werden muss. Der *Straßenrand* wird in aller Regel durch eine *bauliche Trennung* von Fahrbahn und Umgebung festgelegt, wobei die Trennelemente unterschiedlichster Natur sein können. Oft wird der Straßenrand durch eine Bordsteinkante mit angrenzendem Geh-, Radweg oder Grünstreifen definiert. In einigen Fällen dienen auch Markierungen oder Leitplanken zur Begrenzung. Einen für die Luftbilddauswertung besonders problematischen Fall stellen seitliche Parkstreifen dar. Sofern sie nur durch Schilder, aber nicht durch eingelassene Bordsteine, Markierungen mit bestimmter Symbolik oder Ähnliches gekennzeichnet sind, ist eine eindeutige Identifikation – auch für einen menschlichen Operateur – praktisch unmöglich.

Neben diesen vier Kategorien wird in (Wiedemann, 2001) eine weitere Kategorie mit den funktionalen Komponenten des Straßenmodells eingeführt. Dies geschieht dort vor allem unter dem Gesichtspunkt, dass das Straßennetz in außerstädtischen Gebieten relativ direkte Verbindungen zwischen zwei Punkten auf dem Netz gewährleisten muss. Das städtische Straßennetz ist jedoch wegen der dichten Bebauung wesentlich weniger stark hinsichtlich solcher Eigenschaften optimiert, weshalb die verbleibenden funktionalen Komponenten in die anderen Kategorien integriert wurden. Abbildung 9 stellt die Komponenten des Objektmodells noch einmal zusammenfassend dar.

3.2.2 Komponenten und Eigenschaften des Bildmodells

Nachfolgend werden die im vorherigen Abschnitt zusammengestellten Komponenten und Eigenschaften des Objektmodells mit Hilfe des Sensormodells (Abschnitt 3.1) auf die Datengrundlage abgebildet. Dies wird nicht nur für das Originalbild und das DOM durchgeführt, sondern auch für reduzierte Auflösungsstufen, um das Verhalten der Straßeneigenschaften im Maßstabsraum miteinzubeziehen. Die Kategorisierung der Komponenten des Bildmodells erfolgt in gleicher Weise wie im Falle des Objektmodells.

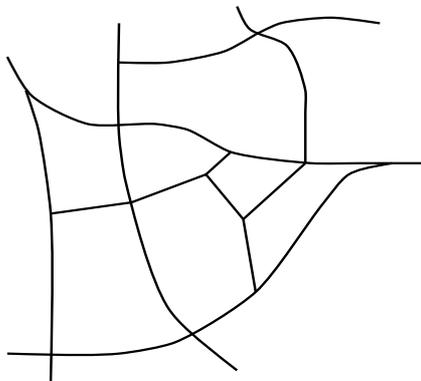
Kategorie geometrische Eigenschaften: Wegen der Bauvorschriften für Straßen und der speziellen Aufnahmebedingungen bei einem Bildflug können bei der Übertragung der geometrischen Straßeneigenschaften in den Bildraum einige erleichternde Grundannahmen getroffen werden. Wie in Abschnitt 3.1 erläutert, sind Luftbilder relativ stabil im Nadirblick aufgenommen. Demzufolge werden Objekte in horizontaler Richtung mit vergleichsweise geringer perspektiver Verzerrung abgebildet, was aufgrund der beschränkten Steigung auch auf Straßen zutrifft. Geometrische Komponenten des Objektmodells wie Länge, Breite und Krümmung dürfen daher über den gesamten Bildraum hinweg *als unverzerrt angesehen* werden und müssen lediglich mit dem



Kategorie: Geometrische Eigenschaften
Langgestreckte, bandähnliche Struktur
Beschränkte Steigung und Querneigung
Geringe Krümmung
Abschnittsweise konstante Breite

Kategorie: Physikalische Eigenschaften
Belag meist aus Asphalt, Beton oder Pflastersteinen
z.T. streifenartige Partikelrückstände

Kategorie: Teilobjekte
Fahrbahn(en)
Mehrere Fahrspuren
Durchgezogene oder strichlierte Markierungen
Straßenrand gekennzeichnet durch Materialwechsel



Kategorie: Topologische Eigenschaften
Zusammenhängendes Netz
Verknüpfungen durch meist höhengleiche Kreuzungen

Abbildung 9. Komponenten des Objektmodells

Bildmaßstab skaliert werden. Gleiches gilt natürlich auch für reduzierte Auflösungsstufen des Originalbildes. Allerdings ist dort die genaue Lage von Straßen schlechter definiert, weshalb auch die Messung von geometrischen Straßenparametern wie Breite und Krümmung nur mit geringerer Genauigkeit möglich ist. Dagegen kommt durch die auflösungsbedingte Auslöschung kleinerer Objekte und Störungen die fundamentale *längliche Struktur* von Straßen oft besser zum Ausdruck als im Originalbild. Um die Vorteile der verschiedenen Auflösungen bestmöglich auszuschöpfen, sollten daher die Straßeneigenschaften in der jeweils günstigsten Auflösung extrahiert und anschließend fusioniert werden. Abbildung 10 zeigt dies anhand der Extraktion von Linien als Hypothesen für Straßenachsen und Kanten als Hypothesen für korrespondierende Straßenränder in zwei

verschiedenen Auflösungsstufen. In einem mit $\sigma = 4m$ stark auflösungsreduzierten Bild können die Achsen der zweispurigen Straße über eine Linienextraktion nach (Steger, 1998b) gut bestimmt werden. Die Kantenextraktion in der gleichen Auflösungsstufe weist hingegen geometrische Defizite auf. Genau das umgekehrte Bild ergibt sich bei geringerer Glättung $\sigma = 1m$: Die Kanten sind präzise am Straßenrand lokalisiert, aber die Linienextraktion liefert lediglich die Achse der schmalen Zufahrtsstraße. Eine Fusionierung von Linien- und Kantenextraktion aus verschiedenen Auflösungsstufen kann folglich die jeweiligen Vorteile vereinen.

Das DOM repräsentiert die Oberfläche des Objektraumes *prinzipiell unverzerrt*, weil bei der DOM-Generierung typischerweise nur eine Abbildung des 3D Objektraumes auf eine Flächenfunktion ohne Transformation der Lagekoordinaten X, Y stattfindet:

$$f(X, Y, Z) = 0 \quad \mapsto \quad f(X, Y) = Z$$

Zu beachten ist jedoch, dass ein durch Bildzuordnung erzeugtes DOM bei dem stark ausgeprägten „Relief“ von Stadtgebieten wegen Filterung und Interpolation geringere Genauigkeit besitzt (siehe auch Abbildungen 7 und 8). Folglich kann die Berechnung von Straßeneigenschaften wie Steigung oder Querneigung für kleine Bereiche der Straßensfläche grob fehlerhaft sein und sollte nur für ausgedehnte Straßenabschnitte erfolgen. Die Bestimmung von Position, Breite und Krümmung einer Straße kann grundsätzlich nur dann auf Basis des DOM durchgeführt werden, wenn sich die Straßensoberfläche in ihrer Höhe deutlich von umgebenden Objekten unterscheidet. Solches Wissen wird später ins Kontextmodell eingebracht (Abschnitt 3.3).

Kategorie physikalische Eigenschaften: Der Straßenbelag aus Asphalt oder Beton weist in aller Regel einen hohen Reflexionsgrad im sichtbaren EMS-Spektrum auf, weshalb Straßen in Luftbildern meist auffällig hell erscheinen. In hohen Auflösungen bilden sich daher Straßen als *relativ homogene, helle Flächen* ab, wogegen in niedrigen Auflösungen der *bandähnliche Charakter* besser zum Tragen kommt (siehe auch Abbildung 10). Im Vergleich zur Umgebung sind Straßen jedoch nicht zwangsläufig die hellsten Objekte, insbesondere wenn es sich um frisch asphaltierte Straßenabschnitte oder Kopfsteinpflaster handelt. Weiterhin entstehen bei besonders viel befahrenen Straßen durch Reifenabrieb und Abgasrückstände *in Fahrtrichtung dunkle Streifen*, die dem Grauwertprofil in der hohen Auflösung einen typischen Wellencharakter verleihen. Abbildung 11 zeigt solch eine Straße und illustriert, wie sich ein exemplarisches Grauwertprofil dieser Straße im Maßstabsraum verhält.

In einem durch Bildzuordnung erzeugten DOM spiegeln sich keine Materialeigenschaften der Objekte wider. Beim Einsatz von Laserdaten mit Aufzeichnung multipler Reflexionen lässt sich aber auf die Materialzusammensetzung rückschließen, was z.B. zur Unterscheidung von bebauten Flächen und Vegetation sehr nützlich ist.

Kategorie topologische Eigenschaften: Durch die Abbildung der 3D Szene in den 2D Bildraum können Teile des Objektraumes in topologisch veränderter Form wiedergegeben werden, d.h. benachbarte Bildelemente sind u.U. keine benachbarten Flächenelemente im Objektraum. Im Falle der Abbildung des urbanen Straßennetzes werden topologische Inkonsistenzen hauptsächlich durch Verdeckungen von Straßenabschnitten aufgrund von Gebäuden, Brücken und hoher Vegetation verursacht. Daher erscheint das urbane Straßennetz im Bild *meist nur lückenhaft*. Handelt es sich nur um relativ kleinräumige Verdeckungen, besteht die Möglichkeit, dass solche Verdeckungen in niedrigeren Auflösungsstufen unterdrückt werden. Daher kann unter Umständen die Netztopologie in Bildern mit reduzierter Auflösung besser erhalten bleiben als im Originalbild. Dagegen sind Kreuzungsbereiche größerer Straßen meist so weitläufig, dass Verdeckungen nur eine untergeordnete Rolle spielen. Bei einer maßstabsabhängigen Modellierung von Kreuzungen ist jedoch zu beachten, dass eine korrekte Zuordnung von Bildmerkmalen aus niedrigeren Auflösungen zu solchen aus höheren Auflösungen möglicherweise schwierig ist – z.B. wenn durch die Elimination von Verkehrsinseln im Maßstabsraum eine *Generalisierung der lokalen Kreuzungstopologie* stattfindet. Da es für die Bildung des Straßennetzes jedoch in erster Linie nur wich-



(a) Auflösungsreduziertes Bild ($\sigma = 4m$)



(b) Auflösungsreduziertes Bild ($\sigma = 1m$)



(c) Linien- und Kantenextraktion aus (a)

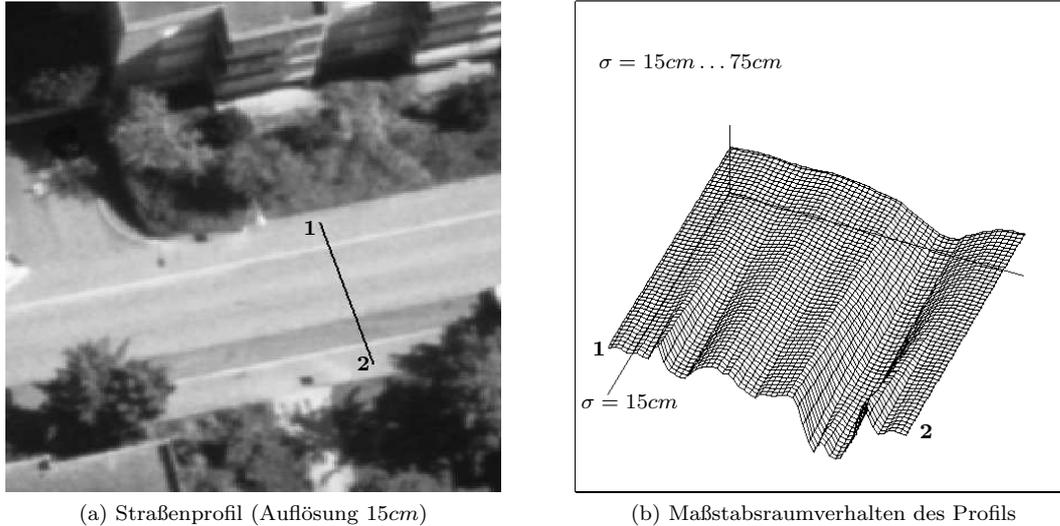


(d) Linien- und Kantenextraktion aus (b)



(e) Liniextraktion aus (a) und Kantenextraktion aus (b)

Abbildung 10. Liniextraktion (weiß) und Kantenextraktion (schwarz) mit den in (Steger, 2000) beschriebenen Verfahren: Im Vergleich zur Extraktion in einer einzigen Auflösungsstufe (a/c bzw. b/d) liefert die aus Linien der niedrigeren Auflösung und Kanten der höheren Auflösung fusionierte Information das beste Ergebnis zur Extraktion der Hauptstraße.



(a) Straßenprofil (Auflösung 15cm)

(b) Maßstabsraumverhalten des Profils

Abbildung 11. Maßstabsraumverhalten eines typischen wellenartigen Straßenprofils.

tig ist, dass zwei oder mehr Straßen miteinander verknüpft sind, erscheint die Extraktion einer gegebenenfalls generalisierten Kreuzungstopologie für die Zwecke dieser Arbeit ausreichend.

Eine Möglichkeit zur zuverlässigen Rekonstruktion der Netztopologie im Falle von verdeckten Bereichen bietet auch die gezielte Integration von mehreren Sichten. Weil in den meisten Fällen die Straßenfläche von zwei oder mehr Bildern einsehbar ist, würde dies zunächst die Verwendung eines mit modernen Verfahren generierten Orthobildes nahe legen (Mayr, 2002). Der entscheidende Negativaspekt eines allein aus Luftbildern abgeleiteten Orthobildes von Stadtgebieten besteht jedoch in der geringeren geometrischen Genauigkeit des zugrunde liegenden DOM. Für die Straßenextraktion hilfreiche geometrische Beziehungen wie Parallelität und Kollinearität von Markierungen und Straßenrändern verlieren bei der Extraktion im Orthobild an Bedeutung. Dies deutet bereits an dieser Stelle darauf hin, dass zur topologisch korrekten Straßennetzbildung eine geeignete Fusionierung von präzisen geometrischen Informationen aus mehreren Originalbildern durch Unterstützung der topologischen Information aus dem DOM zu entwickeln ist.

Kategorie Teilobjekte: Markierungen werden wegen ihrer verkehrslenkenden Funktion meist gut in Stand gehalten. Sie heben sich überwiegend als *helle lineare Strukturen* von der Fahrbahnfläche ab. Dies hat zur Folge, dass Markierungen trotz ihrer geringen Breite erst relativ spät im Maßstabsraum eliminiert werden und zum Teil bis in Auflösungen von 50cm und mehr erhalten bleiben. Eine zuverlässige Extraktion ist jedoch nur in der höchsten Auflösung von ca. 15cm möglich. Teilobjekte, die den Fahrbahnrand begrenzen, wie Bordsteine und Leitplanken, können sich dagegen auf unterschiedlichste Weise abbilden. Entscheidend für die Straßenextraktion ist jedoch, dass diese Objekte aus einem anderen Material als der Straßenbelag bestehen. Der Materialwechsel impliziert einen *signifikanten Grauwertübergang*, der auch in mittlerer Auflösung über eine Kantendetektion bestimmbar ist. Im Gegensatz zu den vorherigen Objekten ist das Teilobjekt „Fahrspur“ nicht direkt, sondern nur indirekt über Markierungen, Straßenrand, Fahrzeuge und zum Teil auch durch Reifenabriebspuren erkennbar⁸. Ähnliches gilt für Fahrbahnen, die – sofern sie nicht mit einbahnigen Straßen gleichgesetzt werden – nur über ihre Fahrspuren definiert sind und sich wie auch alle anderen modellierten Teilobjekte nicht im DOM abbilden. Dies hebt die besondere Funktion von Markierungen hervor, weil sie neben dem Straßenrand die am zuverlässigsten zu extrahierenden Teilobjekte sind. Um eine sichere Rekonstruktion der auf Markierungen aufbauenden Teilobjekte zu gewährleisten, sollte eine „kritische“ Auflösung von etwa 15cm für ihre Extraktion vorliegen. Abbildung 12 zeigt ein typisches Ergebnis einer Extraktion dünner und heller Linien in einer hohen

⁸ Objekte, die nur über Teilobjekte identifizierbar sind, werden in (Mayer, 1998) als „Makrostrukturen“ definiert.



Abbildung 12. Extraktion von Kanten in einem auflösungsreduzierten Bild (weiß, siehe auch Abbildung 10) und von dünnen hellen Linien in einer hohen Auflösung von 15cm (schwarz)

Auflösung von 15cm kombiniert mit einer Kantenextraktion in mittlerer Auflösung von 1m .

Zusammenfassend ergibt sich, dass neben Bild und DOM auch reduzierte Auflösungen des Bildes zusätzliche Informationen beisteuern. Erfahrungsgemäß ist zur Extraktion von Markierungen und demnach auch zur präzisen Rekonstruktion von Fahrbahnen und Fahrspuren eine hohe Auflösung von ca. 15cm von Nöten. In einer mittleren Auflösung von ca. 75cm – 1m ist der Reifenabrieb größtenteils eliminiert worden, so dass sich Straßen als flächenhafte homogene Objekte beschreiben lassen. Hier können auch Straßenränder, die durch einen deutlichen Materialwechsel zwischen Straßenbelag und Umgebung gekennzeichnet sind, extrahiert werden. In einer niedrigen Auflösung von ca. 4m kommt der längliche, linien- oder bandähnliche Charakter von Straßen zum Tragen. Die Elimination von kleinräumigen Störungen im Maßstabsraum hat zur Folge, dass die Topologie des Straßennetzes oft besser erkennbar wird als in höheren Auflösungen. Die Netztopologie ist auch im DOM kaum beeinträchtigt, wobei der Hauptnutzen aber in der Abbildung des groben Höhenverlaufes der Straßen liegt. Aufgrund der geringeren Genauigkeit des DOM ist jedoch die Bestimmung präziser geometrischer Parameter – wie auch aus einem damit abgeleiteten Orthobild – nicht sinnvoll. Solche Parameter sollten aus höheren Auflösungen gewonnen werden und mit den Informationen aus niedrigeren Auflösungen und dem DOM fusioniert werden.

Im nächsten Abschnitt werden die Komponenten von Objekt- und Bildmodell strukturiert und in eine einheitliche Repräsentationsform gebracht.

3.2.3 Repräsentation des Straßenmodells

Zur Formalisierung des im Straßenmodell enthaltenen Wissens wird ein Semantisches Netz eingesetzt. Die Repräsentationsform der Semantischen Netze unterstützt einerseits eine hierarchische, stark objektbezogene Modellierung, so dass Wissen auch in verschiedenen Abstraktionsstufen wiedergegeben werden kann. Anderer-

seits erzwingt ein Semantisches Netz durch die Bildung von Konzepten eine computergerechte Formalisierung des Wissens. Konzepte entsprechen dabei Teilobjekten, deren Eigenschaften, Funktionalitäten und Relationen sich z.B. auf eine C++ – Klassenhierarchie abbilden lassen.

Eine erste Strukturierung des Straßenmodells erfolgt ähnlich wie in (Tönjes et al., 1999) durch die Beschreibung der Straßeneigenschaften auf drei unterschiedlichen Ebenen: *Reale Welt*, *Geometrie/Material* und *Daten*. Während auf der Geometrie/Material- und der Realen Welt-Ebene die Komponenten des Objektmodells formalisiert werden, dient die Datenebene zur Darstellung der Komponenten des Bildmodells. Auf Ebene der Realen Welt ist das Objekt „Straßennetz“ hierarchisch in seine Teilobjekte gegliedert. Diese Gliederung sollte so detailliert sein, dass einzelne Teilobjekte Merkmalen auf der Datenebene entsprechen, die mit Methoden der Bild- bzw. Signalverarbeitung extrahierbar sind. Da jedoch Daten mit unterschiedlichen Charakteristiken zur Verfügung stehen, ist zwischen die Reale Welt- und Datenebene die Geometrie/Material-Ebene eingefügt. Sie beschreibt 3D-Gestalt und Materialeigenschaften der Teilobjekte unabhängig von Datengrundlage oder Sensor. Durch die Verknüpfung der Geometrie/Material-Ebene mit den beiden anderen Ebenen wird eine Konkretisierung von (Teil-)Objekten der Realen Welt auf Datenebene möglich. Im Umkehrschluss bedeutet dies, dass Merkmale, die auf Datenebene extrahiert werden, eine für den Menschen sinnvolle Bedeutung zugewiesen bekommen. Erst durch diese Verknüpfung wird die Grundlage für die Rekonstruktion des abgebildeten Objektes aus Merkmalen – also die Invertierung des Modells – möglich.

Als Konsequenz der Analysen zum Maßstabraumverhalten von Straßen wird neben der Gliederung in die oben angesprochenen Ebenen eine weitere Strukturierung des Modells in drei Abstraktionsstufen bzw. Maßstäbe vorgenommen. Auf der höchsten Abstraktionsstufe sind Straßen durch stark generalisiertes Wissen beschrieben, nämlich dass sie ein durch Kreuzungen verknüpftes Netz bilden. Die mittlere Abstraktionsstufe umfasst dagegen detaillierteres Wissen, indem hier Straßen auf Basis von Fahrbahnen modelliert werden. Die detaillierteste Beschreibung findet schließlich auf der niedrigsten Abstraktionsstufe statt, wo Straßen aus Fahrspuren und deren Bestandteilen aufgebaut werden.

Abbildung 13 zeigt das Straßenmodell als Semantisches Netz. Die horizontalen Trennstriche kennzeichnen die Unterteilung des Netzes in die drei Ebenen Reale Welt, Geometrie/Material und Daten. Die Datenebene umfasst sowohl das Originalbild und reduzierte Auflösungen als auch das DOM. Das DOM wurde seiner Auflösung entsprechend unter den Bildern gruppiert, um zu verdeutlichen, dass es hauptsächlich zur Bestimmung *zusätzlicher* Attribute für Bildmerkmale dient. Die vertikalen Trennstriche deuten die drei Abstraktionsstufen bzw. Maßstäbe an. Im Vergleich zu den Abschnitten 3.2.1 und 3.2.2 wurden der Übersichtlichkeit halber in der Graphik teilweise verkürzte Begriffe verwendet.

Auf der **höchsten Abstraktionsstufe** besteht das Straßennetz aus Kreuzungen und Straßen. Weil es aufgrund von Störungen in der Regel nicht möglich ist, in einem einzigen Extraktionsschritt eine ununterbrochene Straßenverbindung zwischen zwei Kreuzungen zu extrahieren, werden Straßen in kleinere Einheiten, so genannte Straßenabschnitte, zerlegt. In der niedrigen Auflösung ist hiervon nur der lineare Verlauf der Straßenachse extrahierbar. Auf Geometrie/Material-Ebene entspricht diese einer limitiert gekrümmten und geneigten Asphaltspur und im niedrig aufgelösten Bild einer limitiert gekrümmten Linie mit relativ konstantem Höhenverlauf. Kreuzungen sind in zwei Kreuzungstypen eingeteilt: Einerseits einfache Kreuzungen mit einem Kreuzungsbereich ohne Markierungen und Verkehrsinseln und andererseits komplexe Kreuzungen, die sich aus verschiedenen Fahrbahnen zusammensetzen. Ein Kreuzungsbereich kennzeichnet sich auf der Geometrie/Material-Ebene durch eine kompakte Asphaltfläche, die in niedrig aufgelösten Daten als eine kompakte homogene und annähernd horizontale Region wiedergegeben wird. Die Modellierung von Straßenabschnitten – wie auch von komplexen Kreuzungen – setzt sich auf der **mittleren Abstraktionsstufe** auf Basis von Fahrbahnen⁹ fort. Fahrbahnen bestehen aus Fahrbahnrand und Fahrbahnbelag, wobei auf der Geometrie/Material-Ebene der Fahrbahnbelag als limitiert geneigte, längliche Asphaltfläche und der Fahrbahnrand als Materialwechsel zwischen Fahrbahn-

⁹ Streng genommen handelt es sich auf der mittleren Abstraktionsstufe um *Fahrbahnabschnitte* bzw. auf der niedrigsten Abstraktionsstufe um *Fahrspurabschnitte*. Im Folgenden wird aber jeweils die verkürzte Form verwendet.

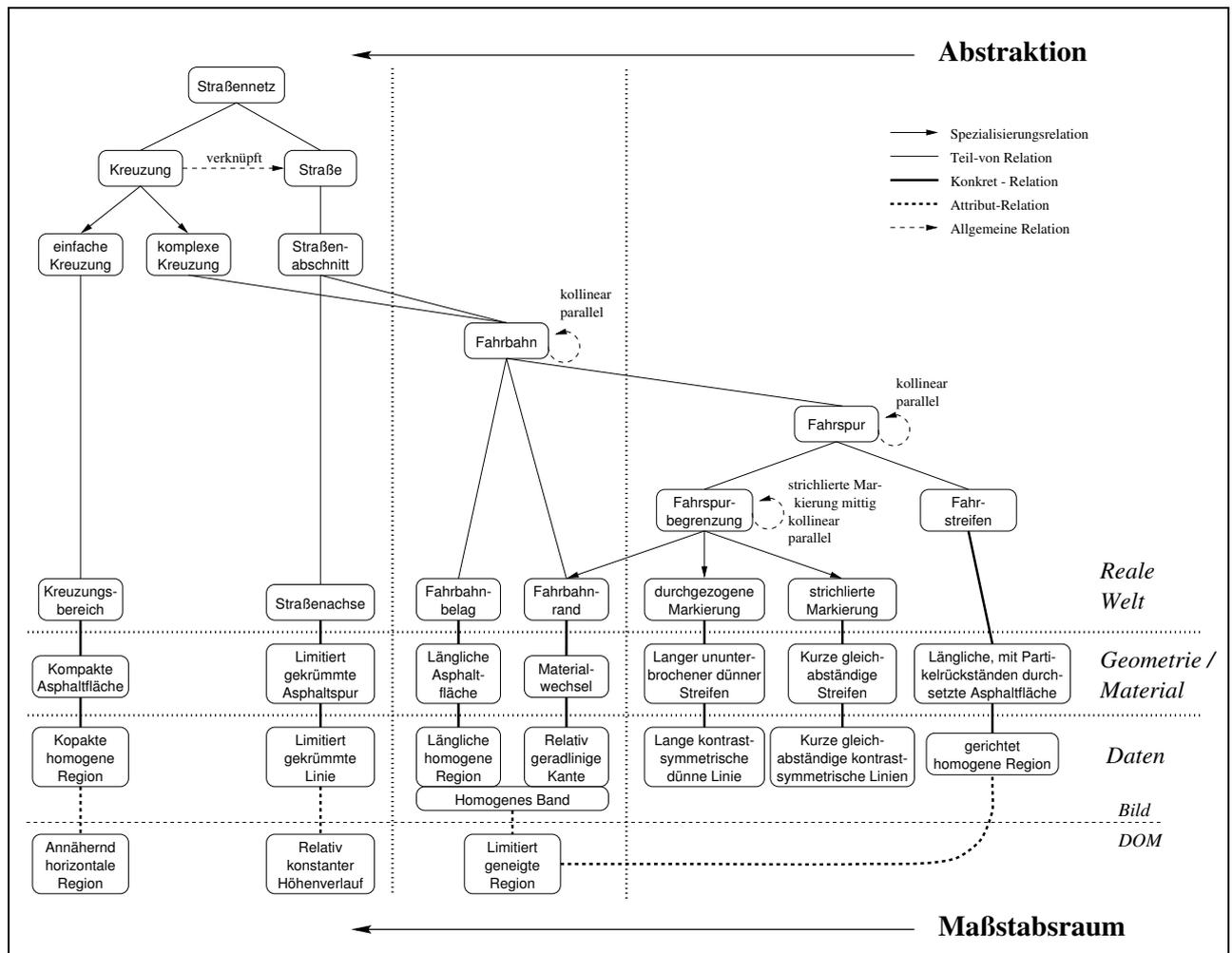


Abbildung 13. Straßenmodell als Semantisches Netz.

belag und unmittelbarer Umgebung modelliert ist. Der Fahrbahnbelag bildet sich dann in Daten mittlerer Auflösung als eine homogene, längliche sowie limitiert geneigte Region und der Fahrbahnrand als relativ geradlinige Kante ab. Beide zusammengenommen ergeben ein so genanntes homogenes Band. Auf der **niedrigen Abstraktionsstufe** werden Fahrbahnen aus einzelnen Fahrspuren aufgebaut, die sich wiederum aus Fahrspurbegrenzung und Fahrstreifen zusammensetzen. Der Fahrstreifen ist eine mit Partikelrückständen durchsetzte, horizontale Asphaltfläche, die auf Datenebene als gerichtet homogene und limitiert geneigte Region wiedergegeben wird. Für die Fahrspurbegrenzung gibt es drei Spezialisierungen: Zum einen durchgezogene oder strichlierte Markierungen und zum anderen der Fahrbahnrand, der bereits auf der mittleren Abstraktionsstufe modelliert wurde. Strichlierte Markierungen dürfen aber nicht als Fahrbahnrand auftreten (Relation „strichlierte Markierung mittig“). Markierungen sind auf der Geometrie/Material-Ebene je nach Spezialisierung gleichabständige kurze oder durchgezogene lange helle Streifen, die sich als dünne helle Linien mit symmetrischem Kontrast aus dem Bild extrahieren lassen.

Das Straßenmodell zusammenfassend ist festzuhalten, dass auf der höchsten Abstraktionsstufe die Globalstruktur, auf der mittleren Abstraktionsstufe lokale Eigenschaften und auf der niedrigen Abstraktionsstufe Details bzw. Teilobjekte modelliert werden. Für die Extraktion bedeutet dies, dass alle drei Auflösungen entscheidend zum Evidenzgewinn beitragen. Die niedrige Auflösung trägt die fundamentalen Eigenschaften wie Netzcharakter, glatter Verlauf und lineare Struktur bei. Da lineare Strukturen relativ einfach aus Bildern zu extrahieren

sind, lässt sich die niedrige Auflösung auch gut zur Generierung initialer Straßenhypothesen einsetzen. Andererseits können durch die Analyse topologischer Eigenschaften eines bereits extrahierten Straßennetzes auch Hypothesen für fehlende Straßenabschnitte gebildet werden. Präzisere Informationen über Straßeneigenschaften liefert die mittlere Auflösung. Durch die Extraktion von homogenen Bändern kann hier Evidenz über die lokalen Straßeneigenschaften wie exakte Position, Breite, Krümmung und Belagszustand gewonnen werden. Die höchste Auflösung schließlich liefert Information über Teilobjekte. Solche Information ist vor allem bei der Bearbeitung komplexer Szenen von Bedeutung, da oft erst durch diese zusätzliche Information eindeutig über die Präsenz einer Straße entschieden werden kann.

Der wesentliche Unterschied des hier vorgestellten Straßenmodells im Vergleich zu anderen Arbeiten – vor allem (Mayer, 1998; Baumgartner, 2003) – besteht im erhöhten Detaillierungsgrad des Modells und der durchgängigen Modellierung von Fahrspuren als Bestandteile von Straßen und Kreuzungen. Wie in Abschnitt 2.3 erarbeitet, ist dies als entscheidende Grundlage für eine robuste Bildanalyse in komplexen Szenen zu betrachten. Allerdings ist anzumerken, dass weder für Fahrbahnen noch für Kreuzungen bauliche Trennelemente modelliert wurden (Grünstreifen, Verkehrsinseln, etc.), die z.B. eine Aggregation paralleler Fahrbahnen zu einer mehrbahnigen Straße ermöglichen. Sind solche Trennelemente in der zu bearbeitenden Szene vorhanden, führt dies zur Extraktion zweier separater Straßen bzw. mehrerer Kreuzungen. Im Falle eines Kreisverkehrs würden z.B. vier voneinander unabhängige Einmündungen extrahiert werden.

3.3 Kontextmodell

Bei der Bildanalyse natürlicher Szenen ist es meist nicht ausreichend, die zu extrahierenden Objekte als isolierte Einheiten zu modellieren. Sowohl ihre Gestalt in der realen Welt als auch ihre Erscheinungsform in den Daten hängt in großem Maße von der jeweiligen Umgebung ab. Derartiges Wissen, das über das eigentlich zu extrahierende Objekt hinausgeht, wird vor allem aufgrund der Arbeiten von Strat (Strat, 1992, 1995) als „Kontext“ bezeichnet und in vielen Ansätzen als unabdingbar für eine zuverlässige Extraktion angesehen. Bei detaillierterer Betrachtung impliziert die Nutzung von Kontext jedoch oft die Verwendung sehr unterschiedlichen Wissens: So verstehen z.B. (Burlina et al., 1995) unter Kontext Wissen über Aufnahmeparameter und -bedingungen, (Hoogs und Bajcsy, 1995) führen durch Kontext a priori Wissen über die Extrahierbarkeit von Bildmerkmalen ein, (Schistad-Solberg, 1999) nutzt Kontextinformation in Form von Pixelnachbarschaften für die Fusionierung verschiedener Datensätze und (Bordes et al., 1995, 1997; de Gunst und Vosselman, 1997) schließlich modellieren mit Kontext den Bezug zwischen extrahierten Straßenachsen und ungenauen bzw. veralteten GIS-Daten.

In Anlehnung an die originale Auffassung von (Strat, 1992) beschreibt in der vorliegenden Arbeit Kontext das *Wissen über die Umgebung des zu extrahierenden Objektes und die daraus resultierenden Einflüsse auf seine Gestalt und Erscheinungsform*. Die Interaktion von Objekt und Umgebung ist besonders bei Einflüssen aus der unmittelbaren Nachbarschaft offensichtlich. Da jedoch auch bei großräumiger Betrachtung von Objekteigenschaften und Objektbeziehungen Abhängigkeiten von der Umgebung auftreten, wird analog zu (Mayer, 1998; Baumgartner, 2003) eine Unterteilung in *lokalen* und *globalen Kontext* vorgenommen. Lokaler Kontext beschreibt die Interaktionen einer Straße mit Objekten ihrer unmittelbaren Umgebung. Im Gegensatz dazu erfasst der globale Kontext die großräumige Prägung von Straßen und ermöglicht dadurch eine Fokussierung auf die im jeweiligen globalen Kontext typischen Charakteristika, Erscheinungsformen und lokalen Kontexte von Straßen. Diese Art der Kontextmodellierung unterstützt hochgradig die Flexibilisierung der Extraktionsstrategie, da in Abhängigkeit der unterschiedlichen lokalen und globalen Kontexte die Extraktionsstrategie verschiedene Alternativen bereitstellen kann und somit eine dynamische Steuerung des Extraktionsprozesses ermöglicht¹⁰.

¹⁰In der englischsprachigen Literatur wird dies auch als *contextual reasoning* bezeichnet.

3.3.1 Lokaler Kontext

Die Notwendigkeit einer lokalen Kontextmodellierung ergibt sich aus der Tatsache, dass sich in Stadtgebieten relativ viele Objekte auf engem Raum befinden und dadurch lokale Beziehungen zwischen Straßen und ihren so genannten *Kontextobjekten* die Interpretation der Szene stark beeinflussen. Die Modellierung und Extraktion solcher Beziehungen kann einerseits zur Erklärung von Abweichungen vom Straßenmodell dienen, andererseits können solche Beziehungen aber auch zur Gewinnung zusätzlicher Evidenz herangezogen werden. Durch die Extraktion von Fahrzeugen lässt sich z.B. nicht nur die Verdeckung von Fahrspuren erklären, sondern – besonders im Falle von längeren Fahrzeugkolonnen – erst ein eindeutiger Hinweis auf die Präsenz einer Straße ableiten.

Die Kontextobjekte mit den engsten Beziehungen zu Straßen sind in Stadtgebieten hauptsächlich Gebäude, Vegetation und Fahrzeuge. Daher beschränkt sich im Folgenden die Modellierung des lokalen Straßenkontexts auf die Relationen zu diesen Kontextobjekten sowie die daraus resultierenden Schlussfolgerungen. Abbildung 14 zeigt die modellierten Relationen zwischen den Kontextobjekten und den Straßenobjekten „Fahrspur“ und „Straße“ bzw. „Fahrbahn“. Die entsprechenden Schlussfolgerungen sind in Tabelle 1 zusammengefasst.

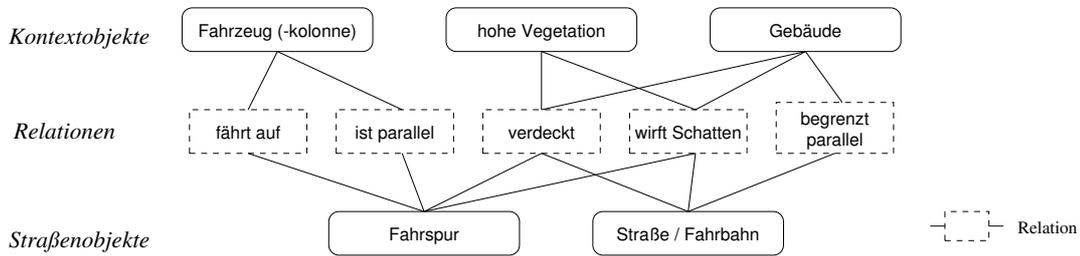


Abbildung 14. Modell für den lokalen Kontext.

Kontextrelation	Schlussfolgerung
„fährt auf“	<i>Position und Länge einer Fahrspurhypothese</i>
„ist parallel“	<i>Richtung einer Fahrspurhypothese</i>
„verdeckt“	<i>Auswahl passender Bildausschnitte</i>
„wirft Schatten“	<i>Modifikation von Parametersätzen</i>
„begrenzt parallel“	<i>Position und Richtung einer Straßenhypothese</i>

Tabelle 1. Schlussfolgerungen aus den Kontextrelationen.

Je näher Gebäude und Vegetation an den Straßenrand rücken, umso stärkeren Einfluss üben sie auf das lokale Erscheinungsbild einer Straße aus. Von Bedeutung für die Straßenextraktion sind daher die Relationen „verdeckt“ und „wirft Schatten“. Während die Verdeckungsrelation Hinweise für eine geeignete automatische Auswahl der Bildausschnitte gibt, können im Falle von Schattenbereichen z.B. modifizierte Parametereinstellungen für weitere Extraktionen in diesen Regionen verwendet werden. Die vor allem in Innenstädten sehr dichte und hohe Bebauung hat darüber hinaus zur Folge, dass Straßen oft nur durch Häuserschluchten definiert sind (Relation „begrenzt parallel“). In einem DOM bilden sich Häuserschluchten als langgestreckte Täler ab und können mit entsprechenden Algorithmen detektiert werden. Dies hilft bei einer folgenden detaillierten Straßenextraktion den Suchbereich hinsichtlich Position und Richtung der Straße deutlich einzuschränken. Eine weitere nützliche Relation, die hier allerdings nur implizit modelliert wird, ergibt sich aus der Tatsache, dass Straßen in aller Regel nicht komplett unter Gebäuden oder Vegetation verlaufen. Bildbereiche, die potentielle Gebäude- oder Vegetationsregionen darstellen, dürfen folglich für die Straßenextraktion als Ausschlussregionen herangezogen werden.

Fahrzeuge oder Fahrzeugkolonnen verdecken das Straßenobjekt „Fahrspur“ häufig zur Gänze. Theoretisch besteht zwar die Möglichkeit, dass solche Fahrspuren aufgrund des Verkehrsflusses in einem anderen Bild nicht mehr verdeckt ist, in Stadtgebieten mit sehr dichtem Verkehr ist diese Option jedoch nur teilweise erfolgversprechend. Vielmehr sollte die positive Evidenz von Fahrzeugen für die Straßenextraktion explizit genutzt werden. Daher werden Fahrzeuge in dem hier entwickelten Modell nicht wie in den meisten anderen Ansätzen implizit als Störobjekte, sondern explizit über die Relation „fährt auf“ als Kontextobjekte eingeführt. Diese Relation ist ausdruckskräftiger als die Verdeckungsrelation, weil ein detektiertes Fahrzeug dadurch direkt eine Hypothese für eine Fahrspur liefert. Weiterhin wird angenommen, dass Fahrzeuge immer annähernd parallel zur Straße ausgerichtet sind (Relation „ist parallel“). Parkplätze, auf denen Fahrzeuge unter Umständen völlig anders orientiert stehen können, werden im Kontextmodell nicht weiter berücksichtigt.

An dieser Stelle sei darauf hingewiesen, dass mit dem lokalen Kontextmodell eine geeignete Schnittstelle zur Integration von GIS-Daten bereitgestellt wird. Je nach Genauigkeit und inhaltlicher Konsistenz mit den Bildern könnten zusätzliche Kontextobjekte und Relationen modelliert und dadurch weiteres Wissen in das Modell integriert werden. Dies ist vor allem im Falle von solchen Kontextobjekten und Relationen hilfreich, die kaum oder gar nicht aus Bildern extrahierbar sind.

3.3.2 Globaler Kontext

Die Zielsetzung, globales Kontextwissen zu modellieren, stammt von der Beobachtung, dass in Abhängigkeit der jeweiligen großräumigen Umgebung bestimmte Eigenschaften und lokale Kontexte des Straßenmodells besonders auffällig sind und andere in den Hintergrund rücken. Z.B. gibt es in unbebauten Gebieten zwischen zwei Punkten auf dem Straßennetz meist nur eine einzige, unter ökonomischen Gesichtspunkten optimale Verbindung, während in Stadtgebieten aufgrund der dichten Bebauung oft mehrere, annähernd gleichwertige Alternativen bestehen (van Cleynenbreugel et al., 1990). Modelleigenschaften, die sich aus der Verbindungsfunktion von Straßen ergeben, haben also in unbebauten Gebieten eine deutlich höhere Relevanz als in Stadtgebieten. Ähnliches gilt für die Bedeutung verschiedener lokaler Kontexte: So tritt die oben angesprochene Relation zwischen Fahrzeugen und Fahrspuren in Stadtgebieten wegen des dichteren Verkehrs recht häufig auf und hat daher dort höhere Relevanz als in unbebauten Gebieten, wo der Verkehr weitaus geringer ist. Durch die Einführung solcher *globaler Kontextgebiete* in das Modell können abhängig vom jeweiligen globalen Kontext aus einem komplexen Modell bestimmte Komponenten herausgezogen oder zumindest gewisse Modellparameter modifiziert werden (van Cleynenbreugel et al., 1990). Dies kann soweit führen, dass für jedes globale Kontextgebiet ein individuelles Extraktionssystem mit speziell angepasstem Modell und eigens entwickelter Strategie zu Verfügung steht, so dass auch sehr umfangreiche Szenen mit überschaubaren Modellen und effizienten Strategien bearbeitet werden können.

Die Modellierung des globalen Kontexts sieht eine Einteilung der Szene in die drei Kontextgebiete *Siedlung*, *offene Landschaft* und *Wald* vor. Dadurch kann die Verwendung des umfangreichen Straßenmodells aus Abschnitt 3.2 auf die Bearbeitung der inhaltlich komplexen Bereiche einer Szene eingeschränkt werden. Für die Straßenextraktion in der offenen Landschaft bietet sich dementsprechend an, auf schlankere Modelle und Strategien zurückzugreifen, wie sie in (Baumgartner, 2003; Wiedemann, 2001; Ruskoné, 1996) entwickelt wurden. Im Unterschied zu Siedlungsgebieten treten in der offenen Landschaft lokale Kontexte wie die Zufahrt zu einem Feld oder zu einem Gehöft in den Vordergrund (Baumgartner, 2003). Dagegen herrschen in Waldgebieten starke Verdeckungen vor, so dass dort eine Straßenextraktion allein aus Luftbildern kaum möglich ist. Abhilfe würden hier Datenquellen schaffen, die besseren Aufschluss über die Volumenstreuung von Objekten geben, z.B. Laserdaten mit Aufzeichnungen multipler Reflexionen oder tomographische SAR-Daten wie sie in (Reigber, 2002) vorgestellt werden. Ansätze, die speziell auf Waldgebiete fokussieren, wurden bisher aber noch nicht entwickelt. Abbildung 15 illustriert abschließend die typischen Charakteristiken der globalen Kontextgebiete.

Insbesondere durch die Modellierung der Relationen zu Gebäuden und Vegetation wird deutlich, dass neben der Bildinformation auch Information über die Objekthöhen von Nöten ist. Daher wird dem DOM bei der Extrak-



(a) „Siedlung“



(b) „Offene Landschaft“



(c) „Wald“ (Beispiel 1)



(d) „Wald“ (Beispiel 2)

Abbildung 15. Beispiele zum globalen Kontext.

tion der Relationen entscheidende Bedeutung zukommen. Anders verhält es sich bei Relationen zu Fahrzeugen: Wenn die Höhendaten durch automatische Bildzuordnungsverfahren nachträglich abgeleitet wurden, lassen sich Fahrzeuge aufgrund der beschränkten Genauigkeit dieser Verfahren oder der Bewegung von Fahrzeugen zwischen den Aufnahmezeitpunkten in aller Regel nicht mehr identifizieren. Wie in (Briese et al., 2001) gezeigt wird, ist dies aber im Falle einer zeitgleichen Aufzeichnung von Laserdaten mit mittleren Punktabständen unter einem Meter durchaus möglich.

Ruft man sich die Anforderungen aus dem letzten Abschnitt von Kapitel 2 in Erinnerung, so wurde dem geforderten Detaillierungsgrad vor allem durch die maßstabsabhängige Modellierung von Teilobjekten Rechnung getragen. Auf die geforderte Integration lokaler Kontextinformation wird durch die systematische und explizite Modellierung von Beziehungen zwischen Teilobjekten und Kontextobjekten eingegangen. Dabei ist insbesondere die Nutzung von positiver Evidenz des lokalen Kontexts von Bedeutung.

Das folgende Kapitel geht auf die interne Bewertung von Extraktionsergebnissen ein. Es führt einerseits die – bislang statische – Modellierung im Hinblick auf die Handhabung von Unsicherheiten und unscharfen Begriffen weiter. Andererseits ist die interne Bewertung von Extraktionsergebnissen eine wichtige Voraussetzung für eine flexible Extraktionsstrategie, weil die Einführung von Qualitätsmaßen eine dynamische Steuerung des Extraktionsprozesses hochgradig unterstützt. Die interne Bewertung kann folglich als Bindeglied zwischen Modell und Strategie angesehen werden, und daher wird ihr hier ein separates Kapitel gewidmet.

4 Interne Bewertung von Extraktionsergebnissen

Im Zuge der immer größeren Vielfalt an Ansätzen zur Objektextraktion, die in den letzten Jahren entwickelt wurden, hat auch die formale Charakterisierung und Bewertung der Ansätze zunehmende Bedeutung gewonnen. Dies gilt einerseits für die theoretische Konzeption der Ansätze, wie es z.B. in (Mayer, 1998) angegangen wurde, und andererseits für die Bewertung der praktischen Leistungsfähigkeit der Ansätze und ihrer Teilkomponenten (Förstner, 1996; Heipke et al., 1998). Bei Letzterem wird üblicherweise zwischen *interner Bewertung* – oft auch Selbst-Bewertung oder Selbst-Diagnose genannt – und *externer Bewertung* unterschieden. Die externe Bewertung hat zum Ziel, die Leistungsfähigkeit verschiedener Ansätze von einem möglichst objektiven Standpunkt aus vergleichend zu bewerten. „Objektivität“ unterstellt dabei in aller Regel die Verwendung von Referenzdaten weit höherer Qualität als die der Extraktion, weshalb die externe Bewertung vor allem zur nachträglichen Analyse von Extraktionsergebnissen hilfreich ist (siehe auch Kapitel 9). Aufgabe der internen Bewertung hingegen ist es, *simultan* zur Extraktion Zwischen- oder Endergebnisse allein auf Basis derjenigen Datenquellen zu evaluieren, die auch zur Objektextraktion zur Verfügung stehen.

In Abschnitt 4.1 werden zunächst Bedeutung und Ziele der internen Bewertung von Extraktionsergebnissen umrissen. Abschnitt 4.2 analysiert exemplarisch Arbeiten aus dem Bereich der Objektextraktion, die Ansätze und Komponenten von interner Bewertung beinhalten, und stellt nachfolgend elementare Anforderungen an ein Konzept zur internen Bewertung zusammen. Abschnitt 4.3 schließlich erläutert die Kernpunkte des in dieser Arbeit entwickelten Konzeptes zur internen Bewertung.

4.1 Bedeutung und Ziele

Die Zielsetzung der internen Bewertung besteht darin, Zwischen- oder Endergebnisse der Extraktion mit einer überschaubaren Anzahl an *Konfidenzwerten* zu versehen, die Auskunft über die Qualität der Resultate geben. Im Bereich der Objektextraktion stellt dieses der Fehler- und Ausgleichsrechnung verwandte Prinzip ein noch relativ junges Forschungsgebiet dar. Wie unlängst auch im Editorial von (Baltsavias et al., 2001) festgehalten wurde, werden die Entwicklung und der konsequente Einsatz interner Bewertung bislang nur von wenigen Forschergruppen vorangetrieben, obgleich sie inzwischen als absolut notwendig für entscheidende Fortschritte in der automatischen Bildanalyse angesehen wird. Die womöglich größte Herausforderung bei der Entwicklung eines Bildanalyse-Systems mit interner Bewertung besteht in der Modellierung und Nutzung von *unabhängigen Kriterien* („Beobachtungen“), die die nötige Redundanz für verlässliche Qualitätsaussagen liefern.

Generell gesehen ist die interne Bewertung von Extraktionsergebnissen in zweierlei Hinsicht bedeutsam:

- (1) Zur effizienteren Kontrolle und Editierung des Endergebnisses einer Extraktion, vor allem mit Blick auf einen operationellen Einsatz eines Bildanalyse-Systems.
- (2) Zur automatischen Überwachung und dynamischen Steuerung des Extraktionsprozesses.

Der erste Punkt spielt für die vorliegende Arbeit nur eine untergeordnete Rolle, weshalb die interne Bewertung zur Nutzerführung in Abschnitt 4.1.1 nur exemplarisch angeführt wird. Der zweite Punkt ist für diese Arbeit von sehr großer Bedeutung und wird im Anschluss vertieft behandelt.

4.1.1 Interne Bewertung zur Nutzerführung

Das Endergebnis einer Extraktion sollte mit einer überschaubaren Anzahl an Konfidenzwerten versehen sein, die dem Nutzer eine schnelle Einschätzung der Qualität des Ergebnisses ermöglichen. In diesem Zusammenhang wird oft auf das so genannte „Ampelparadigma“ hingewiesen (Förstner, 1996). Ein als grün eingestuftes Resultat

tat bedeutet, dass die Ergebnisqualität zufriedenstellend ausgefallen ist, wohingegen ein als rot klassifiziertes Ergebnis eine fehlerhafte Extraktion signalisiert. Ein gelber Wert bedeutet, dass während der Extraktion möglicherweise Probleme aufgetreten sind und eine sichere Entscheidung über Annahme oder Ablehnung auf Basis der involvierten Bewertungskriterien nicht gewährleistet werden kann. In diesem Fall sollte deshalb eine Überprüfung durch den Operateur vorgenommen werden. Geht man von einer idealisierten, fehlerfreien internen Bewertung aus, hat dies für den Operateur die erleichternde Konsequenz, dass nur noch die als gelb eingestufteten Teile des Ergebnisses überprüft werden müssen. Darüber hinaus kann auf Basis der Konfidenzwerte eine Strategie zur effizienten Führung des Operateurs während Ergebnisüberprüfung erstellt werden.

In (Hinz und Wiedemann, 2003) wird hierzu ein zweistufiger Ablauf vorgeschlagen. Zunächst wird dem Operateur ein Überblick über das Gesamtergebnis der internen Bewertung vermittelt, woraufhin er entscheiden kann, ob eine komplette Neuberechnung mit modifizierten Parametern oder eine Editierung einzelner Ergebnisteile sinnvoller ist. Anschließend wird das Ergebnis in Teilstücke zerlegt und kann vom Operateur näher inspiziert bzw. als korrekt oder fehlerhaft eingestuft werden. Hierbei erhält der Operateur detailliertere Einsicht in die Bewertungen durch die einzelnen Kriterien. Dies erleichtert insbesondere die Analyse für das Zustandekommen der jeweiligen Bewertung sowie die Einschätzung der Zuverlässigkeit einzelner Bewertungskriterien und entsprechende Schlussfolgerungen für deren Modifikation oder Gewichtung. Die Definition der Bewertungskriterien in (Hinz und Wiedemann, 2003) stützt sich auf das in Abschnitt 4.3 näher erläuterte Rahmenkonzept zur internen Bewertung.

Neben der Art der Nutzerführung ist eine wesentliche Quintessenz der Analysen in (Hinz und Wiedemann, 2003), dass bei alleiniger Überprüfung der als gelb eingestuften Ergebnisteile und selbst bei restriktiver Einstellung der internen Bewertungskriterien im Durchschnitt nur ein Drittel des Straßennetzes editiert werden muss und trotzdem eine Zuverlässigkeit im Gesamtergebnis von rund 95% erreicht wird. Ein alternativer Ansatz der Nutzerführung während der Editierung eines Extraktionsergebnisses ist in (Heller et al., 1998) beschrieben. Hier wird der Operateur ähnlich wie beim Straßentracking durch das gesamte Endergebnis der Extraktion geführt. Trotz dieser relativ aufwendigen Vorgehensweise wird eine Zeitersparnis um Faktor 10 gegenüber einer manuellen Erfassung angegeben.

4.1.2 Interne Bewertung zur Steuerung der Extraktion

Eine entsprechende Berücksichtigung der internen Bewertung spielt nicht nur aus operationeller Sicht, sondern insbesondere auch bei der Konzeption eines Bildanalyse-Systems eine wichtige Rolle. Üblicherweise werden im Zuge des Extraktionsprozesses eine Vielzahl von Zwischenergebnissen, Hypothesen und Alternativen aufgestellt, über deren Relevanz das System selbständig entscheiden muss. Solche Entscheidungen werden erheblich erleichtert, wenn jedes Modul, das zur Bearbeitung einer Problemstellung aufgerufen wird, nicht nur das Ergebnis selbst, sondern auch eine Abschätzung über die Qualität des Ergebnisses liefert. Auf diese Weise lässt sich eine beträchtliche Steigerung der Effizienz des Bildanalyse-Systems erreichen, weil irrelevante Zwischenergebnisse frühzeitig verworfen werden können und auf vielversprechende Hypothesen fokussiert werden kann.

Daneben wird durch einen derartigen Einsatz von interner Bewertung eine wichtige Basis für eine flexible Extraktionsstrategie gelegt. Je nach Bewertung einer Hypothese kann ein effizienter Ablauf an Folgeaktionen abgeleitet werden. Wird z.B. eine Verknüpfungshypothese zwischen zwei kollinearen Fahrspuren als sehr homogen in ihrem Grauwertverlauf bewertet, ist es zur Verifikation der Verknüpfung sinnvoller, zunächst ein Modul zur Überbrückung mittels Snakes aufzurufen als ein Modul zur Detektion von Fahrzeugen.

Ein weiterer Gesichtspunkt des Einsatzes von interner Bewertung ist die Integration von Kriterien, die eine Bewertung aus verschiedenen Blickwinkeln durchführen. Die ersten Extraktionsschritte können in aller Regel nur in lokal begrenzten Bildausschnitten wie dem Einzugsbereich einer Filtermaske durchgeführt werden, so dass sich auch die Bewertung auf eine eher lokale Sichtweise bezieht. Zu einem späteren Zeitpunkt der Prozessierung – z.B. nach einer Gruppierung von Merkmalen – kann aber die Relevanz der Zwischenergebnisse in einem größeren

Zusammenhang analysiert werden. Die gegenseitige Ergänzung lokaler Bewertungen mit globaleren Aspekten resultiert vor allem, wenn eine Rückkopplung der globalen Bewertung zurück auf lokale Ebene vorhanden ist, in einer deutlichen Steigerung der Vollständigkeit und Zuverlässigkeit des Endergebnisses (Baumgartner und Hinz, 2000).

Der folgende Abschnitt beleuchtet Ansätze zur internen Bewertung hauptsächlich im Hinblick auf eine automatische Überwachung und dynamische Steuerung der Straßenextraktion.

4.2 Ansätze zur internen Bewertung

Aufgrund der Tatsache, dass sich die interne Bewertung auf dem Gebiet der Bildanalyse noch in der Anfangsphase ihrer Entwicklung befindet, lassen sich noch keine klar definierten Konzepte oder gar Theorien identifizieren, deren Alternativen vergleichend gegenüber gestellt werden könnten. Eine Reihe von Publikationen beinhalten aber Ansätze und Komponenten von interner Bewertung, deren Nutzen hinsichtlich eines umfassenderen Konzeptes analysiert werden kann.

Auch unter den bisherigen Publikationen zur Straßenextraktion gibt es einige Arbeiten, die Ansätze von interner Bewertung integrieren. Meist findet hier die Bewertung in Form von Gewichtungsfunktionen auf Merkmalsebene statt. So verwenden z.B. (McKeown und Denlinger, 1988) die Korrelation aufeinanderfolgender Grauwertprofile zur internen Bewertung des profilbasierten Trackings, (Wiedemann, 2001) benutzt geometrische und radiometrische Attribute zur Bewertung von extrahierten Linien, und (Price, 2000) schließlich die Neigung und Kantenunterstützung von Rechteckprimitiven zur Bewertung von Straßensegmenten. Ähnliche Kriterien kommen in vielen weiteren Arbeiten zum Einsatz; wie auch in den erwähnten Arbeiten werden die abgeleiteten Konfidenzwerte aber nur zur *einmaligen* Selektion von Primitiven herangezogen und nicht in weiteren Schritten verwendet.

Eine konzeptionelle Erweiterung stellen Ansätze dar, die sich auf eine durchgängige Modellierung von Unsicherheiten z.B. mit Bayes-Netzen stützen. Auf dem Gebiet der Straßenextraktion sind hier vor allem (Geman und Jedynak, 1996; Barzohar und Cooper, 1996) anzuführen (siehe Kapitel 2) oder aus dem Bereich der Gebäudeextraktion (Fischer et al., 1998; Kulschewski, 1999; Brunn, 2001; Scholze et al., 2001; Kim et al., 2001) und dem der kompletten Szenenmodellierung (Köster, 1995; Tupin et al., 1999; Hellwich et al., 2001; Liedke et al., 2001)¹¹. Typischerweise werden in diesen Arbeiten die Konfidenzwerte von Zwischenergebnissen im Sinne der Fehlerfortpflanzung miteinander verknüpft und in einer „bottom-up“-Strategie auf höhere Ebenen propagiert. Am Ende – oder manchmal auch an ausgewählten Stellen – des Extraktionsprozesses wird dann mittels Teststatistik über Annahme und Ablehnung der Hypothesen entschieden.

Ein wesentlicher Punkt, der bei der Propagierung von Unsicherheiten jedoch nicht berücksichtigt wird, besteht in der Ausschöpfung von Wissen, das auf höheren Ebenen *zusätzlich* zur Verfügung steht und somit eine ähnlich unabhängige Kontrollfunktion erfüllen kann wie die zu Beginn von Abschnitt 4.1 erwähnte Redundanz. Als sehr fortschrittlich sind diesbezüglich Ansätze wie z.B. (Köster, 1995; Tupin et al., 1999) einzuordnen. Hier werden mit relativ einfachen Bildverarbeitungsverfahren Objekte extrahiert, denen zunächst auf Basis lokaler Kriterien eine vorläufige Semantik zugewiesen wird. Anschließend wird in Form von a priori Wahrscheinlichkeiten zusätzliches Wissen über Nachbarschaftsrelationen von Objektklassen eingeführt, um die vorläufige Semantik aus einer anderen Sichtweise zu bewerten und gegebenenfalls neu zuzuweisen. Da sich die Veränderung der Semantik eines einzigen Objektes über die Nachbarschaftsrelationen hinweg wiederum auf andere Objekte auswirkt, muss die endgültige Bestimmung der Objektsemantik mittels globaler Optimierungsverfahren gelöst werden. Eine gewisse Einschränkung besteht bei solchen Verfahren allerdings darin, dass die Optimierung auch

¹¹ Oft basieren derartige Arbeiten auch auf allgemeinen Bildinterpretationssystemen wie z.B. (Geo-)AIDA (Liedke et al., 2001) oder ERNEST (Niemann et al., 1990).

bei einer relativ geringen Anzahl an Objekten und Relationen unverhältnismäßig aufwändig werden kann. Deshalb erscheint es im Hinblick auf eine hohe Szenen- und Modellkomplexität sinnvoller, solche Optimierungsprobleme nur zur Bearbeitung von Teilaufgaben heranzuziehen und diese in den festen Rahmen einer Extraktionsstrategie einzubetten (Hinz und Baumgartner, 2000).

Neben der Nutzung von Wissen über Objekte und ihre Relationen zur internen Bewertung ist noch der mathematische Rahmen, der zur Beschreibung von unscharfen Aussagen und zur Auswertung von Konfidenzwerten dienen soll, kritisch zu beleuchten. Dies soll hier vor allem unter dem Gesichtspunkt geschehen, welche bestehende Theorie als ein geeignetes „Werkzeug“ für die typischen konzeptionellen und praktischen Problemstellungen der Bildanalyse zu betrachten ist:

Am häufigsten eingesetzt und sicherlich auch am weitesten entwickelt ist die Wahrscheinlichkeits- bzw. Bayes-Theorie und, sehr verwandt damit, die Informations- und Evidenz-Theorie (Shafer und Pearl, 1990; Vosselman, 1992). Diese Theorien haben den Vorteil, dass es ihre mathematische Formulierung erlaubt, funktionale Zusammenhänge gleichermaßen für Bewertung *und* Prädiktion – dem inversen Problem der Bewertung – heranzuziehen. Beispiele für beide Problemstellungen finden sich z.B. in (Kulschewski, 1999; Hellwich et al., 2001). Im Hinblick auf die Objektextraktion unterliegen diese Theorien jedoch der Einschränkung, dass zu ihrem korrekten Einsatz einige Grundvoraussetzungen erfüllt sein müssen, die vor allem bei Bewertung semantisch höherwertiger Hypothesen schwer einzuhalten sind. So müssten z.B. bei der oben angesprochenen Verifikation einer Verknüpfung von kollinearen Fahrspuren im Falle der Bayes-Theorie die Auftrittswahrscheinlichkeiten von Straßenrändern, Markierungen und Fahrzeugen in Abhängigkeit der Grauerthomogenität als bedingte Wahrscheinlichkeitsdichtefunktionen formuliert werden. Die Bestimmung einer solchen Funktion ist zwar über die Auswertung von Trainingsdaten prinzipiell möglich, jedoch stellt sich die berechtigte Frage, ob die hierfür nötigen *repräsentativen* Daten mit vertretbarem Aufwand akquiriert werden können, ohne vorher die tatsächliche Aussagekraft der Funktion genau abschätzen zu können. Diese Problematik führt häufig zu einer Verletzung des theoretischen Fundaments, indem entweder keine repräsentativen Trainingsdaten zum Einsatz kommen oder die Funktionen auf Basis von Erfahrungswerten „manuell“ definiert werden. Weil letzteres tatsächlich oft den einzig praktikablen Weg bei der Entwicklung eines komplexen Bildanalyse-Systems darstellt, erscheint die Verwendung einer Theorie mit weniger restriktiven Grundannahmen sinnvoller, auch wenn dies unter Umständen zu Lasten der statistischen Aussagekraft geht.

Aus diesen Gründen wird in der vorliegenden Arbeit die Verwendung der Fuzzy-Theorie bevorzugt (Dubois und Prade, 1988; Zadeh, 1989). Ihr theoretisches Fundament wurde vor allem für eine leichtere Transformation unscharfer Beschreibungen und Definitionen in mathematische Ausdrücke entwickelt (Zadeh, 1989; Mendel, 1995). Solche Beschreibungen kommen typischerweise auch im Straßenmodell vor wie z.B. „Fahrspur mit großer Grauerthomogenität“. Zur Bewertung der „großen Grauerthomogenität“ wird eine so genannte Fuzzy-Funktion definiert, die mittels Parameter aus dem Straßenmodell, aber ebenso auch aus Beispieldaten abgeleitet werden kann. Das Ergebnis der Bewertung ist ein Fuzzy-Wert, der die Übereinstimmung zwischen Modellvorstellung und den aus der Datengrundlage ermittelten Werten angibt. Ebenso bietet die Fuzzy-Theorie Möglichkeiten zur Verstärkung und Abschwächung von Kriterien („*sehr* groß“ bzw. „*mäßig* groß“) sowie zur Aggregation verschiedener Kriterien mittels Fuzzy-„und“ bzw. Fuzzy-„oder“ Operationen.

Ein gewisser Nachteil der Fuzzy-Theorie besteht darin, dass der gleiche funktionale Zusammenhang nicht simultan für Bewertung und Prädiktion eingesetzt werden kann. Dies ist für die vorliegende Arbeit jedoch zweitrangig, da hier die Verwendung eines Bewertungsschemas für Zwischen- und Endergebnisse im Vordergrund steht. Eine weitere typische Problematik beim Einsatz der Fuzzy-Theorie ist die mit der Vielfalt an eingeführten Begriffen und Definitionen schnell ansteigende Anzahl an Fuzzy-Funktionen. Dieses Defizit wird hier zumindest teilweise durch die formalisierte Repräsentation des Straßenmodells abgefangen.

Fasst man die Erkenntnisse der vorangegangenen Abschnitte zusammen, ergibt dies folgenden Anforderungskatalog für ein Konzept zur internen Bewertung von Extraktionsergebnissen:

- ◊ Die interne Bewertung sollte nicht nur punktuell, sondern durchgängig und konsistent im gesamten Bildanalyse-System eingesetzt werden.
- ◊ Die interne Bewertung jeder Hypothese sollte sich auf komplementäre („orthogonale“) Kriterien beziehen, die *nicht* zur Extraktion der Hypothese herangezogen wurden. Konsequenterweise sollten Kriterien zur Bewertung in gleicher Weise wie Kriterien zur Extraktion aus dem Modell abgeleitet werden.
- ◊ Zur internen Bewertung semantisch höherwertiger Hypothesen in späteren Schritten der Extraktion sollten nicht nur propagierte Konfidenzwerte aus vorherigen Schritten verwendet werden. Vielmehr sollten *zusätzliche* Kriterien herangezogen werden, die entsprechend der höheren Semantik der Hypothesen neues Wissen in die Bewertung einfließen lassen.
- ◊ Die Voraussetzungen des mathematischen Rahmens der internen Bewertung sollten in Theorie und Praxis erfüllbar sein.

Im Folgenden wird ein Rahmenkonzept zur Modellierung und Nutzung interner Bewertungskriterien vorgestellt, das diesen Forderungen Rechnung trägt. Es wird zwar im Hinblick auf die Aufgabenstellung der Straßenextraktion in urbanen Gebieten entwickelt, die Grundzüge sind jedoch allgemeinerer Natur und auf andere Anwendungen der Bildanalyse übertragbar.

4.3 Rahmenkonzept

Das Rahmenkonzept gliedert sich in zwei Komponenten: Die Definition interner Bewertungskriterien auf Basis des Straßenmodells (Abschnitt 4.3.1) und – eingebettet in das Straßenmodell – deren strategische Nutzung während der Extraktion (Abschnitt 4.3.2).

4.3.1 Modellierung und interne Bewertung

Ausgangspunkt für die Definition von internen Bewertungskriterien ist das Straßenmodell. In der in Kapitel 3.2.3 gezeigten Repräsentationsform als Semantisches Netz ist das Straßenmodell in einzelne Konzepte gegliedert, für die im Zuge der Extraktion Instanzen hypothetisiert werden. Jedes Konzept wird dabei durch eine Reihe von Kriterien beschrieben. Diese Kriterien werden in zwei Klassen eingeteilt: Kriterien *K I*, die zur Extraktion des Konzeptes¹² dienen, und Kriterien *K II*, die zur Bewertung des Konzeptes¹³ herangezogen werden. Um eine möglichst große Unabhängigkeit der internen Bewertung zu erreichen, sollten die Kriterien beider Klassen weitestgehend unabhängig sein. Abbildung 16 und Tabelle 2 zeigen dies beispielhaft für das Konzept „strichlierte Markierungen“. Kriterien zu deren Extraktion sind die Kamm-ähnliche Struktur der Linien im Grauwertgebirge, die minimal bzw. maximal begrenzte Länge der Linien und der dazwischen liegenden Lücken, deren (paarweise) Kollinearität und die Gesamtlänge der strichlierten Markierung. Als Bewertungskriterien dienen die Grauerthomogenität einer Linie, der symmetrische Kontrast beiderseits einer Linie, die Konstanz von Linien- und Lückenlängen sowie die mittlere Krümmung entlang der gesamten strichlierten Markierung.

Bei der Frage der Zuordnung von Kriterien zu einer der beiden Klassen lassen sich aus Erfahrungswerten einige Richtlinien mit relativ großer Allgemeingültigkeit formulieren:

Kriterien K I: Aus praktischen Gesichtspunkten dienen zur Extraktion eines Konzeptes in erster Linie seine generischen Grundelemente. Kriterien für diese Grundelemente beziehen sich meist auf eine begrenzte Umgebung wie den Einzugsbereich einer Filtermaske oder auf topologische Relationen zu einer geringen Anzahl an

¹²Eigentlich handelt es sich um die Instanz eines Konzeptes. Im Folgenden wird aber auf diese strenge begriffliche Trennung verzichtet.

¹³dito

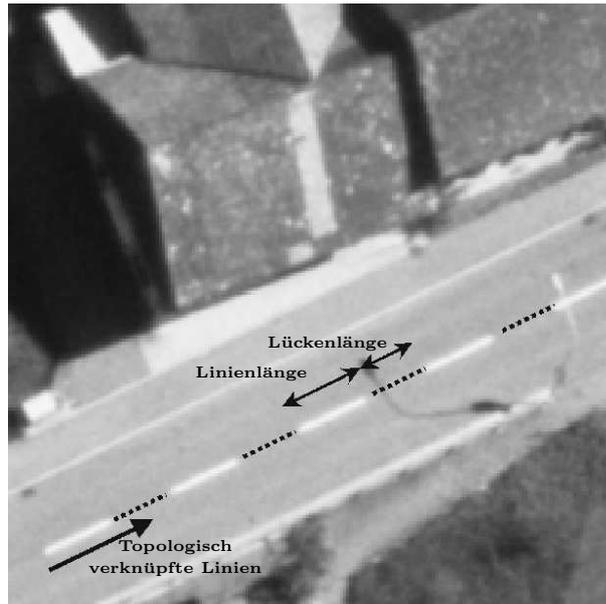


Abbildung 16. Strichlierte Markierungen.

Kriterien K I	Kriterien K II
Kamm-ähnliche Struktur im Grauwertgebirge	Große Grauwerthomogenität entlang der Linien
Minimale und maximale Linienlänge	Große Kontrastsymmetrie der Linien
Minimale Kollinearität von Linienpaaren	Konstante Linien- und Lückenzängen
Maximale Lückenzängen	Geringe mittlere Krümmung
Minimale Gesamtlänge	

Tabelle 2. Kriterien zur Extraktion und Bewertung von strichlierten Markierungen.

Nachbarobjekten. Weiterhin sollten für die Extraktion Kriterien eingesetzt werden, deren Erscheinungsbild – sofern überhaupt möglich – vom lokalen Kontext weniger stark beeinflusst wird. Kontextinformation kann oft nur mit geringerer Zuverlässigkeit und Vollständigkeit aus den Daten gewonnen werden, so dass bei fehlender oder falscher Kontextinformation die Extraktion einer Hypothese nicht grundsätzlich ausgeschlossen sein sollte. Im Falle einer Markierung wird daher z.B. nur der relative Kontrast einer Linie zur Umgebung als Kriterium herangezogen und nicht die absolute Helligkeit. Somit besteht zumindest bei ausreichendem Umgebungslicht die Chance, auch dann in Schattengebieten Linien zu detektieren, wenn die Extraktion der Kontextinformation „Schatten“ gescheitert ist. Grundsätzlich ist es sinnvoll, bei der Extraktion großzügige Parametereinstellungen zu verwenden und dadurch tendenziell eine Übersegmentierung zu erzielen, denn in vielen Fällen ist die anschließende Analyse mit Hilfe der internen Bewertung leichter zu bewerkstelligen als die Extraktion fehlender Hypothesen.

Kriterien K II: Kriterien, die zur Bewertung von Hypothesen dienen, können im Gegensatz zu K I-Kriterien objektspezifische Details beinhalten, weil allein auf Basis einer generierten Hypothese zusätzliche Information aus dem Bild extrahierbar ist. So kann z.B. nach der Generierung von Linienhypothesen das objektspezifische Wissen genutzt werden, dass auf beiden Linienseiten relativ ähnlicher Kontrast herrschen sollte. Gleiches gilt für Kriterien, die eine globalere Sichtweise auf das Konzept verkörpern als jene zur Extraktion. So können nach Gruppierung einzelner Linien zu einer strichlierten Markierung nun Kriterien wie mittlere Krümmung und Konstanz der Linien- und Lückenzängen verwendet werden. Bei der Bewertung kommen Fuzzy-Funktionen

zum Einsatz. Jede Fuzzy-Funktion bemisst dabei die Qualität der Hypothese bezüglich *eines* Kriteriums. Die Gesamtqualität der Hypothese ergibt sich schließlich aus der Aggregation der Einzelbewertungen mittels Fuzzy-„und“ Operation. Nur als Hinweis sei hier angeführt, dass bei einer großen Anzahl an Hypothesen auch statistische Aussagen über die Bewertungskriterien getroffen werden können, die wiederum einen Rückschluss auf die Vertrauenswürdigkeit einzelner Kriterien erlauben.

Diese Art der Unterteilung des Modells ließe sich auch auf die Extraktion und Bewertung von Kontextinformation ausdehnen. Hiervon wird im Weiteren aber abgesehen, weil sich die Kontextextraktion in großem Umfang auf ein DOM stützt, dessen eingeschränktes Genauigkeitsniveau von vorneherein bekannt ist. Anstelle von Konfidenzmaßen wird daher der jeweilige lokale und globale Kontext einer Hypothese als Kontextattribut angefügt. Entsprechend der Kontextattribute – z.B. „Schatten“ – wird im Rahmen der Extraktionsstrategie dann entschieden, welche Modellparameter bzw. Fuzzy-Funktionen zur Bewertung einer Hypothese heranzuziehen sind.

Die Definition von internen Bewertungskriterien stützte sich bisher allein auf die Konzepte des Semantischen Netzes, ohne deren Bezug untereinander zu berücksichtigen. Im folgenden Kapitel werden nun die hierarchischen Beziehungen des Straßenmodells dazu verwendet, um eine möglichst effiziente Nutzung der internen Bewertung für die Extraktionsstrategie vorzubereiten.

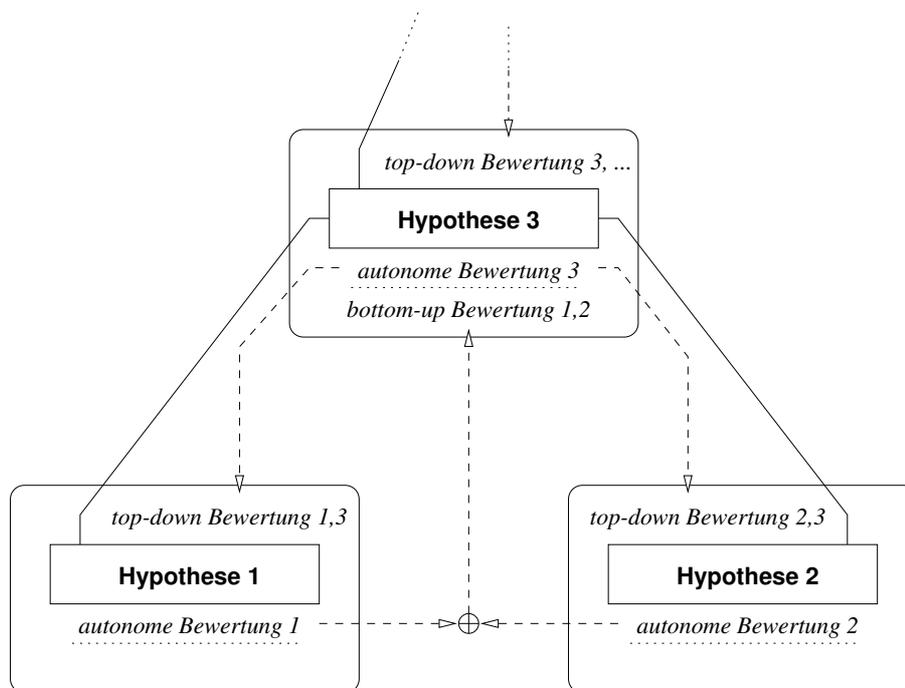


Abbildung 17. Schema zur Nutzung der internen Bewertung basierend auf der Hierarchie des Objektmodells. „⊕“ steht für die Fuzzy-„und“ Operation.

4.3.2 Nutzung interner Bewertung

Für jedes Konzept der untersten Hierarchieebene des Semantischen Netzes werden, wie im vorangegangenen Abschnitt erläutert, Kriterien zur Extraktion und Bewertung definiert, so dass jede Hypothese mit einem Konfidenzmaß durch eine *autonome Bewertung* versehen ist. Werden im Zuge der Extraktion diese Hypothesen dazu verwendet, auf der nächsthöheren Hierarchieebene weitere Hypothesen zu konstruieren, können die autonomen Bewertungen im Sinne von „Genauigkeiten“ kombiniert und fortgepflanzt werden, wobei die Kombination hier nach den Regeln der Fuzzy-Theorie geschieht. Diese so genannte *bottom-up Bewertung* beschreibt die Qualität einer Hypothese, wie sie auf Basis der Qualitätsmaße ihrer Bestandteile hergeleitet wird. In Ergänzung

dazu können aber in den allermeisten Fällen auf einer höheren Hierarchieebene weitere Bewertungskriterien formuliert werden, die Objektwissen repräsentieren, das in tieferliegenden Ebenen noch nicht genutzt werden konnte. Es findet hier also eine zusätzliche autonome Bewertung statt, die aufgrund ihrer Unabhängigkeit von der bottom-up Bewertung wiederum als kontrollierende *top-down Bewertung* für Hypothesen der tieferliegenden Hierarchieebene herangezogen werden darf. In gleicher Weise wird für die nächsthöheren Hierarchieebenen verfahren.

In Abbildung 17 wird die Nutzung der internen Bewertung für den Allgemeinfall verdeutlicht. Für das konkrete Beispiel der Konstruktion von Fahrspuren aus strichlierten Markierungen bedeutet dieses Schema, dass die autonomen Bewertungen zweier paralleler Markierungen aus Abschnitt 4.3.1 zur einer bottom-up Bewertung einer Fahrspurhypothese kombiniert werden. Als unabhängige autonome Bewertung der Fahrspurhypothese könnte z.B. nun die Grauwerthomogenität des eingelagerten Fahrstreifens dienen, die wiederum als kontrollierende top-down Bewertung an die Markierungen weitergereicht wird. Dadurch erhält die interne Bewertung einer Markierung eine zusätzliche globale Sichtweise, weil nicht nur die Konsistenz zwischen Modellvorstellung und Daten sondern auch die Relevanz einer jeden Hypothese für die weitere Extraktion involviert ist. Für die praktische Implementierung bedeutet dies, dass eine Hypothese – sofern ihr weiterer Nutzen aufgrund der autonomen Bewertung nicht völlig ausgeschlossen werden kann – solange aufrecht erhalten werden muss, bis die nächsthöhere Hierarchieebene komplett prozessiert wurde.

Das Kapitel der internen Bewertung abschließend seien noch einmal die Kernpunkte des Rahmenkonzeptes zusammengestellt:

- ◇ Einteilung der Kriterien für jedes Konzept des Semantischen Netzes in zwei unabhängige Klassen: Kriterien zur Extraktion und Kriterien zur Bewertung. Praktische Überlegungen helfen dabei zur Ableitung von Richtlinien für die Zuordnung der jeweiligen Kriterien.
- ◇ Nutzung der Hierarchie des Objektmodells durch drei supplementäre Bewertungen: autonome Bewertung, bottom-up Bewertung, top-down Bewertung.
- ◇ Verwendung der Fuzzy-Theorie als mathematischer Rahmen.

5 Extraktionsstrategie – Überblick

In diesem Kapitel wird ein Überblick über die Extraktionsstrategie gegeben, ohne dabei auf algorithmische Details näher einzugehen. Der anschließende Teil II dieser Arbeit (Kapitel 6 bis 8) behandelt Methodik und Implementierung der Straßenextraktion in detaillierterer Weise und illustriert anhand von Zwischenergebnissen einzelne Schritte des Verfahrens.

Die Extraktionsstrategie legt die wesentlichen Grundzüge für die Steuerung des Extraktionsprozesses fest. Durch sie wird der Rahmen vorgegeben, wann welches Wissen über das Objekt zur Extraktion genutzt wird und wie Ergebnisse von Teil-Extraktionen in den weiteren Verfahrensablauf eingebunden werden. Die Strategie beinhaltet also Wissen über die Extrahierbarkeit von Merkmalen bzw. Teil-Objekten sowie die Effizienz von Algorithmen. Die Anpassung der Strategie an die Komplexität von Daten und Modell ist daher ein wichtiger Einflussfaktor für die Leistungsfähigkeit eines Bildanalyse-Systems.

Aufgrund der hier zu behandelnden Komplexität kommt nur eine hybride Strategie – also eine wechselseitig daten- und modellgetriebene Vorgehensweise – in Frage (siehe Abschnitt 2.1.4). Zu Beginn werden Bildprimitive datengetrieben extrahiert und daraus initiale Hypothesen für Teil-Objekte, z.B. für Markierungen und Fahrspuren, generiert. Diese werden anschließend intern bewertet und verifiziert. Basierend auf den verifizierten Hypothesen können nun modellgetrieben weitere Hypothesen gebildet werden, z.B. für Verknüpfungen von Fahrspuren aber auch für Konzepte von höheren Hierarchiestufen wie z.B. Fahrbahnen, die dann wiederum die Grundlage für die nächsten Schritte der Extraktion bilden. Diese sequentielle Abfolge an Operationen führt zu einem schrittweisen Aufbau des Straßennetzes.

Die interne Bewertung spielt dabei in zweierlei Hinsicht eine Rolle: Zum einen dient sie der *Verifikation* von Hypothesen, also der Entscheidung über Annahme oder Ablehnung einer Hypothese. Eine Hypothese gilt dann als verifiziert, wenn die in Abschnitt 4.3 eingeführten Typen an Bewertungen durchgeführt und insgesamt positiv ausgefallen sind. Zum anderen ermöglicht die interne Bewertung eine effizientere Extraktion, da verifizierte Hypothesen gemäß ihrer Bewertung geordnet werden können und somit bei Folgeschritten die erfolgversprechenden Hypothesen zuerst bearbeitet werden können. Ein positiver Nebeneffekt der Bewertung ist, dass der Einfluss von harten Schwellwerten auf das Endergebnis verringert wird, weil schlechtere aber unter Umständen korrekte Hypothesen nur geringer gewichtet, jedoch nicht vorzeitig verworfen werden (Steger et al., 1997).

Wie in Abschnitt 2.3.2 erläutert, erscheint das feste Rahmenkonzept, zunächst auf die sicher und leichter zu extrahierenden Teile des Straßennetzes zu fokussieren und anschließend das Netz zu vervollständigen, für den groben Ablauf der Extraktion als sehr erfolgversprechend. Durch die Vorgabe einer groben Struktur kann die Bearbeitung typischer, immer wiederkehrender Problemstellungen standardisiert und die Bearbeitung vieler Alternativen und das damit verbundene Testen unnötiger Hypothesen vermieden werden. Für die Behandlung von Szenen hoher Komplexität ist jedoch ein starr vorgegebener Ablauf aufgrund der damit verbundenen Einschränkung der Flexibilität nachteilig. Abhilfe wird im vorliegenden Ansatz dadurch geschaffen, dass zur Bearbeitung von bestimmten Teilaufgaben, die innerhalb des festen Rahmens auftreten, verschiedene Alternativen bereitgestellt werden. Welche der Alternativen im Zuge der Extraktion zum Einsatz kommt, wird auf der einen Seite über den Straßenkontext gesteuert. Zu diesem Zweck ist der eigentlichen Straßenextraktion eine initiale Phase zur Extraktion von globaler und lokaler Kontextinformation vorgeschaltet. Auf der anderen Seite dienen die Kriterien der internen Bewertung zur Auswahl von alternativen Algorithmen. Eine Fahrspurhypothese mit sehr inhomogenen Grauwerten lässt z.B. auf die Präsenz eines Fahrzeugs schließen, so dass in diesem Fall die Fahrzeugextraktion als die bevorzugte Alternative zur Verifikation dieser Hypothese herangezogen wird. Ist die Hypothese jedoch zu kurz für ein Fahrzeug, wird eine unmodellierete Störung angenommen und die Auslöschung dieser Störung in der niedrigen Auflösung überprüft.

Die konkrete Vorgehensweise der Extraktion ist in Abbildung 18 dargestellt. Sie sieht eine Grobgliederung in die Phasen (1) *Kontextanalyse*, (2) *Extraktion von sicheren Straßenobjekten* und (3) *Straßennetzbildung* vor.

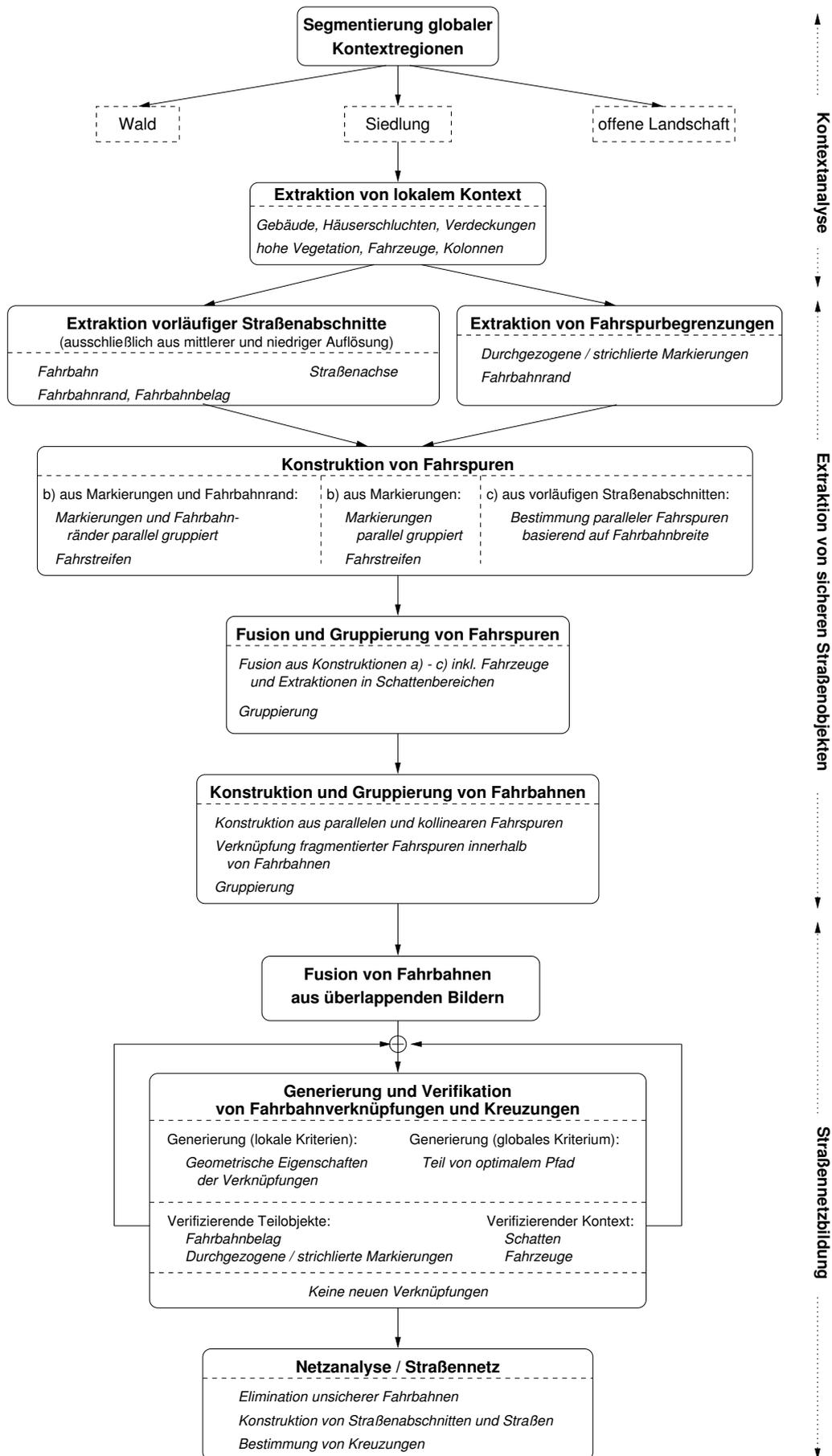


Abbildung 18. Ablauf der Extraktion.

Phase 1: Kontextanalyse

Die Extraktion startet mit der Segmentierung der globalen Kontextgebiete „Wald“, „Siedlung“ und „offene Landschaft“. Zur Unterscheidung dieser Gebiete fungieren Struktur- und Textur-Merkmale, die aus dem Bild und DOM extrahiert werden. Alle folgenden Operationen beschränken sich dann auf die Bildregionen mit der Bedeutung „Siedlung“. Die nächsten Teilschritte der Kontextanalyse umfassen die Gewinnung lokaler Kontextinformation. Hierzu werden zunächst aus dem DOM Täler extrahiert, die als potentielle Häuserschluchten interpretiert werden. Simultan dazu werden DOM-Regionen segmentiert, die sich von ihrer lokalen Umgebung durch eine deutliche Höhendifferenz abheben. Diese Regionen bilden für die Straßenextraktion als Ausschlussregionen, unabhängig davon, ob es sich hierbei um Gebäude, hohe Vegetation oder andere hohe Objekte handelt. Mit Kenntnis des Sonnenstandes bei der Bildaufnahme können sie darüber hinaus zur Prädiktion von Schattenregionen herangezogen werden. Die endgültige Detektion von Schattenregionen erfolgt dann im Bild. Weil primär abgeschattete Straßen gesucht werden, dient neben der geringen Grauwertintensität auch die Homogenität der Grauwerte als Merkmal einer Schattenregion. Anschließend wird das DOM von jedem Bild aus einer Sichtbarkeitsanalyse unterzogen, um Bereiche zu identifizieren, die in bestimmten Bildern verdeckt sind. Der abschließende Teilschritt der Kontextanalyse besteht aus der Extraktion von Fahrzeugen und Fahrzeugkolonnen. Hierzu wird ein autarker Ansatz herangezogen, dessen Fahrzeugmodell detaillierte Merkmale wie Fronthaube, Dach und Windschutzscheibe mit Wissen über die typische Kolonnenbildung auf verkehrsreichen Straßen integriert.

Phase 2: Extraktion von sicheren Straßenobjekten

Zentraler Punkt der Extraktion von sicheren Straßenobjekten ist die Konstruktion von Fahrspuren, da sie in der niedrigen Abstraktionsstufe die Objekte mit der größten Semantik sind (siehe Abbildung 13). Als Eingangsinformation zu ihrer Konstruktion dienen einerseits Fahrspurbegrenzungen (durchgezogene / strichlierte Markierungen und Fahrbahnränder), die durch die Detektion und Gruppierung heller, dünner Linien bzw. Kanten extrahiert werden. Andererseits werden vorläufige Straßenabschnitte aus auflösungsreduzierten Versionen des Originalbildes extrahiert, um das Maßstabsraumverhalten von Straßen zu nutzen. Eine Linie in der niedrigen Auflösung entspricht einer Straßenachse eines vorläufigen Straßenabschnitts. Dagegen wird seine Fahrbahn in der mittleren Auflösung in Form eines homogenen Bandes bestimmt, dessen Region dem Fahrbahnbelag und dessen seitliche Kanten den Fahrbahnrändern entsprechen. Anschließend erfolgt die Konstruktion von Fahrspuren in drei verschiedenen Varianten: a) Aus parallelen Markierungen und Fahrbahnrändern, die einen homogenen Fahrstreifen beinhalten, b) aus parallelen Markierungen mit eingeschlossenem, homogenem Fahrstreifen und c) auf Basis der vorläufigen Straßenabschnitte, indem die Anzahl der parallelen Fahrspuren aus der Fahrbahnbreite geschätzt. Zur Straßenextraktion in Schattenbereichen werden die gleichen Verfahren eingesetzt, jedoch mit besser an die Radiometrie in Schattenbereichen angepassten Parametern.

Im Rahmen der folgenden Fusion und Gruppierung von Fahrspuren werden die Fahrspuren aus beschienenen und abgeschatteten Bereichen inklusive der im Zuge der Kontextanalyse extrahierten Fahrzeuge und Fahrzeugkolonnen zusammengeführt. Da in diesem Stadium der Extraktion jedes Bild einzeln bearbeitet wird, kann die Fusion im Bildraum stattfinden. Die resultierende eindeutige Menge an Fahrspuren dient nachfolgend zum Aufbau mehrspuriger Fahrbahnen durch paralleles und kollineares Gruppieren von Fahrspuren. Lücken zwischen kollinear gruppierten Fahrspuren können überbrückt werden, wenn sich die Störung entweder über die lokalen Kontexte Schatten oder Fahrzeug erklären lässt oder wenn sie in der niedrigen Auflösung eliminiert wird. Konnte im Rahmen der Kontextanalyse kein Fahrzeug entdeckt werden, wird nun versucht, ein Fahrzeug mit einem weniger detaillierten Modell als zuvor zu extrahieren. Dies ist an dieser Stelle gerechtfertigt, weil im Gegensatz zu Phase 1 nun Wissen über die Präsenz von Straßen in der direkten Umgebung des Fahrzeugs vorliegt.

Das Ergebnis dieser Phase sind Fahrbahnen, die mit relativ großer Sicherheit extrahiert wurden. Die Fahrbahnen ergeben aber in der Regel noch kein zusammenhängendes Straßennetz. Außerdem ist bis zu diesem Stadium der Extraktion vorwiegend lokales – wenn auch sehr detailliertes – Modell-Wissen in die Extraktion eingeflossen. Daher ist mit Fehlextraktionen zu rechnen, die unter Umständen erst durch die Verwendung von globalen Netzwerk-Kriterien eliminiert werden können.

Phase 3: Straßennetzbildung

Die Extraktionen und Bewertungen der vorherigen Phasen werden für alle überlappenden Bilder getrennt durchgeführt. Die Fusion der Fahrbahnen geschieht nun im Objektraum durch Überlagerung aller Fahrbahnen auf dem DOM und der Elimination überlappender Anteile mit ähnlicher Orientierung. Vor der Fusion werden alle Fahrbahnen anhand ihrer Bewertungen sortiert, so dass bei der Elimination redundanter Anteile immer der Teil der schlechter bewerteten Fahrbahn entfernt werden kann.

Um die Grundlage für die Nutzung globaler Netzwerk-Kriterien zu bereiten, wird in den folgenden Teilschritten das Straßennetz als Graph repräsentiert. Die Fahrbahnen bilden die Kanten des Graphen und die Endpunkte seine Knoten. Nach Aufbau des Graphen geschieht die Straßennetzbildung in einem iterativen Prozess durch Generierung und Verifikation von Verknüpfungen zwischen Fahrbahnen (siehe Abbildung 18). Um eine erste Auswahl an potentiellen Fahrbahnverknüpfungen zu erhalten, werden die Verknüpfungen zunächst hinsichtlich ihrer Geometrie und hinsichtlich der Übereinstimmung an zu verbindenden Fahrspuren überprüft, bewertet und als vorläufige Kante in den Graph eingefügt. Die Verknüpfung wird aber nur dann weiter analysiert, wenn sie Teil eines optimalen Pfades ist, der zwei weit entfernt voneinander liegende Fahrbahnen miteinander verbindet. Durch dieses Kriterium wird auf Verknüpfungen fokussiert, die aus globaler Sicht besonders wichtig für die Topologie des Straßennetzes sind.

Die Verifikation einer Verknüpfung erfolgt durch die Analyse der korrespondierenden Bildregion in Bezug auf Teilobjekte und lokalen Kontext. Verifizierende Teilobjekte sind der Fahrbahnbelag und Markierungen, sofern diese nicht erfolgreich zur Konstruktion von Fahrspuren eingesetzt werden konnten (z.B. mangels Fahrbahnrand). Verifizierende Kontexte sind zum einen Schattenbereiche und zum anderen Fahrzeuge. Eine verifizierte Fahrbahn wird schließlich in das Straßennetz eingefügt, der Graph entsprechend aktualisiert und die nächste Iteration gestartet. Im Falle, dass keine neue Verknüpfung mehr generiert werden kann, bricht die Iteration ab.

Das Resultat am Ende dieses iterativen Prozesses besteht in der Regel aus einer großen Anzahl an verknüpften und aus einer geringen Anzahl an isolierten Fahrbahnen, die nicht an das restliche Netz angeschlossen werden konnten. Diese werden im Rahmen der Straßennetzbildung eliminiert, so dass gewährleistet ist, dass das endgültige Straßennetz aus einer einzigen bzw. einigen wenigen großen Zusammenhangskomponenten besteht. Anschließend werden die Fahrbahnen zu Straßenabschnitten und weiter zu Straßen mit jeweils konstanter Anzahl an Fahrspuren zusammengefasst. Abbiegespuren o.ä. sind nicht modelliert und bleiben daher unberücksichtigt. Der abschließende Schritt der Extraktion besteht aus der Bestimmung von Kreuzungen durch Analyse des Graphen hinsichtlich Knoten mit Grad 3 oder höher.

Abschluss Teil I

Den **konzeptionellen Teil I abschließend** lässt sich zusammenfassen, dass die am Ende von Kapitel 2 herausgestellten vier Defizite der bisherigen Ansätze durch

- ◇ die detaillierte Modellierung von Straßen (Abschnitt 3.2),
- ◇ die detaillierte Modellierung von lokalem und globalem Kontext (Abschnitt 3.3),
- ◇ die Entwicklung eines Konzepts zur internen Bewertung (Kapitel 4) und
- ◇ die Entwicklung einer Extraktionsstrategie mit flexiblen Komponenten (Kapitel 5)

in die Konzeption des vorliegenden Ansatzes einbezogen wurden. Im folgenden Teil II dieser Arbeit wird die Realisierung der einzelnen Teile der Extraktion detaillierter beschrieben und der Funktionsnachweis des implementierten Bildanalyse-Systems erbracht.

Teil II:

Realisierung der Straßenextraktion

In diesem Teil der Arbeit werden die einzelnen Schritte zur Realisierung der Straßenextraktion erläutert. Die Zielsetzung, das Extraktionssystem anhand der Beschreibung exakt re-implementieren zu können, würde jede Lesbarkeit dieses Teils zerstören. Vielmehr soll der in Bildanalyse kundige Leser konkrete Anhaltspunkte für die Grundprinzipien der Extraktion bekommen. Unter dem Gesichtspunkt, den „rote Faden“ der Konzeption möglichst präsent zu halten, werden manche Argumentationen und Schlussfolgerungen aus Teil I in kurzer Form noch einmal aufgegriffen.

Die drei Phasen der in Kapitel 5 skizzierten Extraktionsstrategie spiegeln sich in der Gliederung dieses Teils wider. Kapitel 6 beschreibt die Kontextanalyse zu Beginn der Straßenextraktion. Anschließend wird in Kapitel 7 auf die Extraktion von Straßenobjekten, die sicher extrahierten Teilen des Straßennetzes entsprechen, eingegangen. Kapitel 8 erläutert schließlich die Fusion der bis dahin erzeugten Zwischenergebnisse aus überlappenden Bildern und die darauf aufbauende Straßennetzbildung. Die verwendeten Luftbilder haben eine Auflösung von ca. 15cm, die „hohe Auflösung“. Außerdem kommen zwei auflösungsreduzierte Versionen der Bilder zum Einsatz, die „mittlere Auflösung“ (ca. 1m) und die „niedrige Auflösung“ (ca. 4m). Zur Verkürzung der Rechenzeiten werden die Bilder nach Reduktion der Auflösung mit 75cm bzw. 2m unterabgetastet. Die verbleibende Reduktion der Auflösung auf 1m bzw. 4m ist in Form von Glättungsoperationen in die jeweiligen Bildverarbeitungsroutinen integriert¹⁴. In verschiedenen Tests hat sich nämlich herausgestellt, dass die bei der Unterabtastung hervorgerufenen Diskretisierungsfehler durch die restliche Glättung praktisch vollständig eliminiert werden.

6 Kontextanalyse

Die in Kapitel 5 entwickelte Extraktionsstrategie sieht als ersten Schritt des Verfahrens die Nutzung von Kontextwissen vor, weil hierdurch Hintergrundinformation gewonnen wird, die die eigentliche Straßenextraktion unterstützen soll. Dies setzt allerdings voraus, dass die Kontextobjekte und -relationen mit ausreichender Zuverlässigkeit extrahiert werden können. Bei größeren Unsicherheiten sollten die daraus gezogenen Schlussfolgerungen hauptsächlich zum Evidenzgewinn dienen, aber keine negative Evidenz verursachen.

Gemäß der Unterteilung des Kontextmodells in Abschnitt 3.3 in eine globale und eine lokale Komponente besteht dieses Kapitel aus zwei Hauptteilen: Zunächst wird in Abschnitt 6.1 die Bestimmung von globalen Kontextgebieten beschrieben. Sie dienen in erster Linie dazu, die Straßenextraktion auf solche Gebiete zu fokussieren, die eine gute Übereinstimmung zwischen dem Straßenmodell und den tatsächlichen Straßen erwarten lassen und somit für die Extraktion am erfolgversprechendsten sind. Abschnitt 6.2 erläutert dann die Gewinnung und Analyse lokaler Kontextinformation. Diese beinhaltet Kontextobjekte wie Fahrzeuge und Gebäude sowie deren Relationen, die signifikanten Einfluss auf das Erscheinungsbild von Straßen ausüben. Fahrzeugen kommt aufgrund ihres häufigen Auftretens in Innenstadtbereichen eine sehr große Bedeutung zu. Daher spielt ihre zuverlässige Detektion eine zentrale Rolle innerhalb der „Kontextanalyse“. Abbildung 19 fasst diesen Teil der Extraktionsstrategie zusammen.

¹⁴Vielen Bildverarbeitungsoperationen, z.B. Linien- und Kantenextraktion, ist eine Glättung sowieso inhärent.

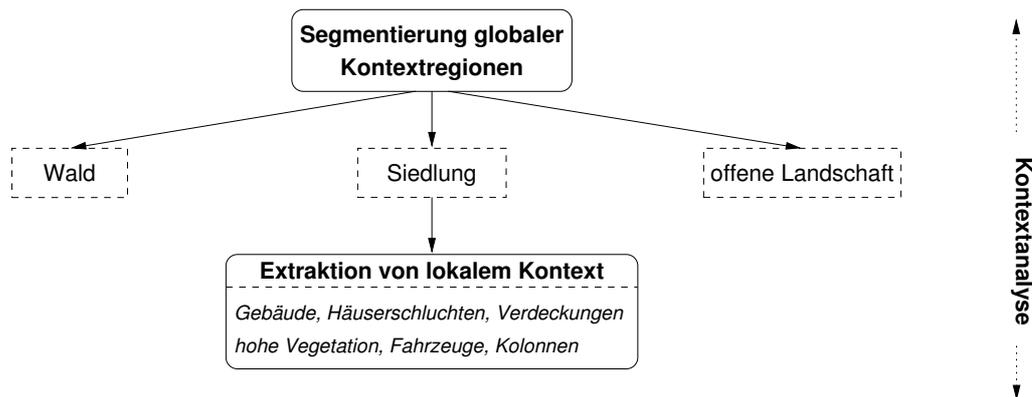


Abbildung 19. Extraktionsstrategie: Teil „Kontextanalyse“ (vgl. Abbildung 18).

6.1 Globaler Kontext

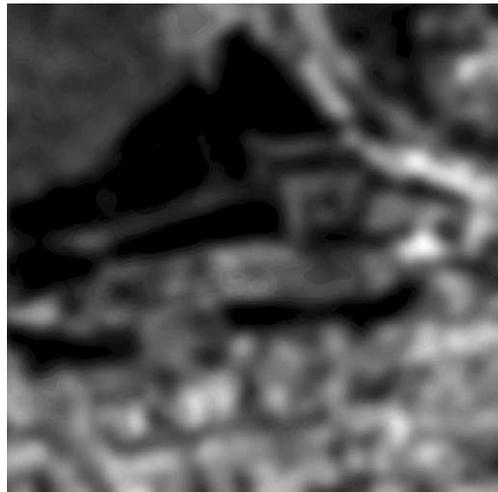
Die Segmentierung der globalen Kontextgebiete *Siedlung*, *offene Landschaft* und *Wald* wird mit Hilfe von Textur- und Strukturmerkmalen aus Bild- und Höhendaten mittlerer Auflösung (ca. 1m) durchgeführt. Offene Landschaft kann anhand ihrer Frequenzeigenschaften relativ einfach segmentiert werden, da sie sowohl in Luftbildern als auch im DOM wesentlich texturärmer als die beiden anderen Kontextgebiete erscheint. Um alle Regionen, die fast nur niederfrequente Anteile besitzen, zu segmentieren und gleichzeitig hochfrequentes Rauschen zu unterdrücken, werden Bild und DOM mit einem Gabor-Filter Bandpass-gefiltert, segmentiert und anschließend die Schnittmengen der segmentierten Regionen gebildet. Die resultierenden Regionen stellen mit sehr hoher Wahrscheinlichkeit offene Landschaft dar und dienen als Ausgangspunkt für eine Verfeinerung der Segmentierung mittels Regiongrowing und Morphologie. Abbildung 20 zeigt die Gabor-Filterung eines Bildes, das Ergebnis der vorläufigen Segmentierung sowie die anschließende Verbesserung durch Regiongrowing.

Um die restliche Region, die nicht der „offenen Landschaft“ zugewiesen wurde, in die Kontextgebiete Wald und Siedlung zu trennen, kommen nun strukturelle Merkmale zum Einsatz. Ein wesentliches Unterscheidungsmerkmal zwischen Siedlungs- und Waldgebieten ist die Anzahl an parallelen Kanten, die homogene Flächen mittlerer bis großer Helligkeit einschließen. Während solche Flächen in Waldgebieten nur sehr selten auftreten, kommen sie in Siedlungsgebieten aufgrund der Geometrie- und Materialeigenschaften von Dächern, Straßen und Parkplätzen relativ häufig vor. Für Städte mit größeren Gebäudekomplexen ist es darüber hinaus sinnvoll, die Modellierung von Strukturmerkmalen auch auf das DOM auszudehnen. In dieser Arbeit werden parallele Bruchkanten aus dem DOM in die Segmentierung miteingeführt. Eine weitere Möglichkeit wäre die Ableitung von Strukturmerkmalen aus lokalen Richtungshistogrammen z.B. von Grauwert- und DOM-Kanten wie in (Faber und Förstner, 2000; Hinz et al., 2001a). Tests auf unterschiedlichen Bilddaten haben allerdings ergeben, dass solche Merkmale nur zur Segmentierung von sehr regelmäßig bebauten Städten wie z.B. Washington DC oder Purdue dienlich sind.

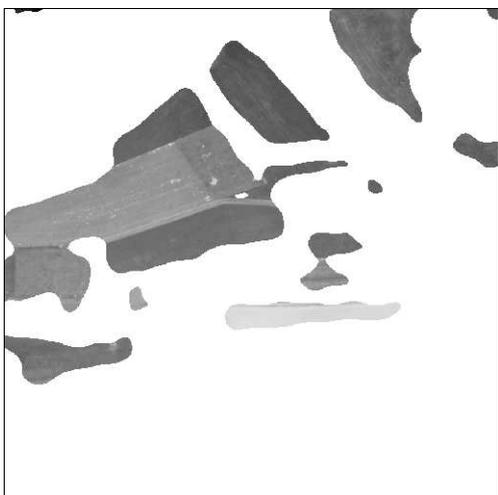
Abbildung 21 illustriert einige Zwischenergebnisse der Segmentierung von Wald- und Siedlungsgebieten. Zunächst werden homogene Flächen mit parallelen Kanten aus dem Bild und parallele Kanten aus dem DOM extrahiert. Aus dieser Eingangsinformation können nun Matrizen erzeugt werden, die die relativen Häufigkeiten dieser Merkmale in der lokalen Umgebung einer bestimmten Bildposition wiedergeben. Nach Kombination dieser Häufigkeiten entscheidet schließlich ein histogrammbasierter Schwellwert über die Zugehörigkeit eines jeden Pixels zum jeweiligen Kontextgebiet. Abschließend werden die segmentierten Regionen noch morphologisch nachbearbeitet, weil globale Kontextgebiete nur ab einer bestimmten Größenordnung – z.B. mehrere Häuserblöcke – sinnvoll für die Extraktion einsetzbar sind.



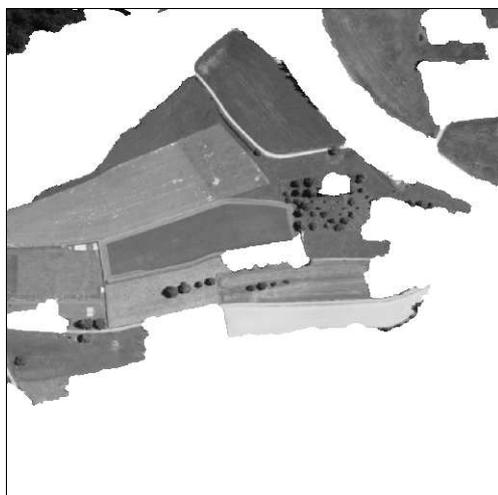
(a) Bildausschnitt mittlerer Auflösung



(b) Gabor-gefilterter Bildausschnitt



(c) Vorläufige Segmentierung offener Landschaft



(d) Kontextgebiet „offene Landschaft“: Verfeinerung von (c) durch Regiongrowing und Morphologie

Abbildung 20. Schritte zur Segmentierung des Kontextgebietes „offene Landschaft“

6.2 Lokaler Kontext

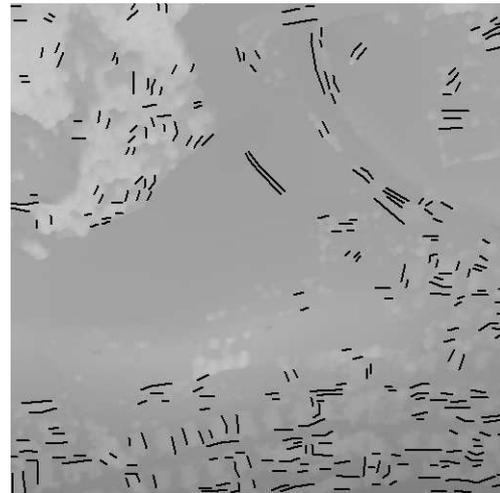
Das lokale Kontextmodell aus Abschnitt 3.3.1 beschreibt Beziehungen zwischen Straßenobjekten und den Kontextobjekten „Fahrzeuge“, „Gebäude“ und „hohe Vegetation“ und leitet daraus bestimmte Schlussfolgerungen für die Extraktion ab (siehe Tabelle 1). Während der Extraktion der Kontextobjekte wird jedoch keine Unterscheidung zwischen Gebäuden und Vegetation getroffen. Dies ist gerechtfertigt und bleibt ohne Auswirkung auf das Ergebnis, wenn der Algorithmus in der Lage ist, die von Gebäuden und hoher Vegetation ausgehenden Kontextrelationen unabhängig vom semantischen Unterschied dieser Objekte zu bestimmen. Abschnitt 6.2.1 skizziert diesen Algorithmus, bevor Abschnitt 6.2.2 ausführlicher auf die Fahrzeugextraktion eingeht.

6.2.1 Gebäude und hohe Vegetation

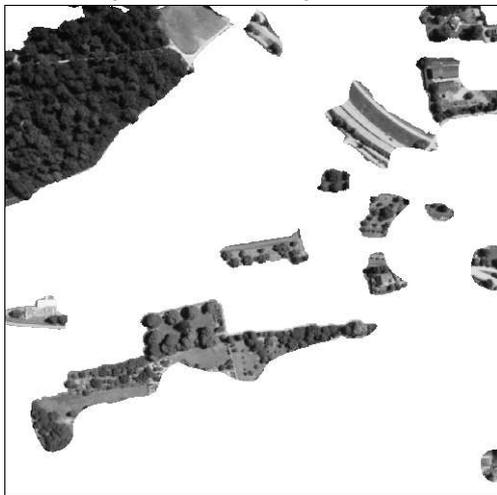
Der lokale Straßenkontext, der Gebäude und hohe Vegetation betrifft, umfasst die Relationen „begrenzt parallel“, „verdeckt“ und „wirft Schatten“. Zur Analyse dieser Relationen muss keine exakte Rekonstruktion der Gebäude oder detaillierte Extraktion der Vegetation durchgeführt werden. Es genügt, die Objekte im DOM, ggf. unter Zuhilfenahme der Bildinformation, zu detektieren. In der Darstellung eines DOM als höhencodiertes



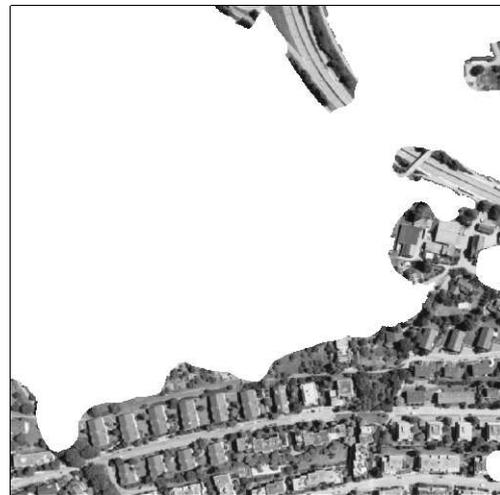
(a) Parallele Bildkanten mit eingeschlossener homogener Fläche



(b) Parallele DOM-Kanten



(c) Resultat nach Segmentierung: Kontextgebiet „Wald“



(d) Resultat nach Segmentierung: Kontextgebiet „Siedlung“

Abbildung 21. Verwendung von Strukturmerkmalen (parallele Kanten) zur Segmentierung der Kontextgebiete „Wald“ und „Siedlung“.

Grauwertbild sind isoliert stehende Gebäude und hohe Vegetation durch helle, kompakte Strukturen charakterisiert („Blobs“; siehe hierzu auch Abbildungen 7 und 8 in Abschnitt 3.1). Die Extraktion solcher Blobs geschieht wie in (Huertas et al., 2000) mittels Laplace-Of-Gaussian Filterung mit anschließender Bestimmung der Nulldurchgänge. Diese Vorgehensweise verursacht zwar eine etwas geringere Genauigkeit in der Lokalisierung der Blobs, sie hat aber den Vorteil, dass nur sehr wenige Parameter einzustellen sind und dass eine Verkettung der Nulldurchgänge direkt zu geschlossenen Umringspolygonen der Blobs führt. Das Innere der so erhaltenen Umrisse dient fortan als Ausschlussregion.

Im Falle von größeren, zusammenhängenden Gebäudekomplexen ist eine Bestimmung der Umrisse mittels Blob-Detektion wesentlich schwieriger. Um dennoch Kontextinformation zu gewinnen, wird das DOM nach Häuserschluchten und damit nach positiver Evidenz für Straßen analysiert (Kontextrelation „begrenzt parallel“). Häuserschluchten werden im höhecodierten Grauwertbild als dunkle, lineare Strukturen abgebildet, die mittels Linienextraktion leicht bestimmbar sind. In dicht bebauten Innenstädten ist sogar eine weitere Gruppierung dieser Linien zu Zyklen möglich und sinnvoll. Abbildungen 22 a) und b) visualisieren die im DOM

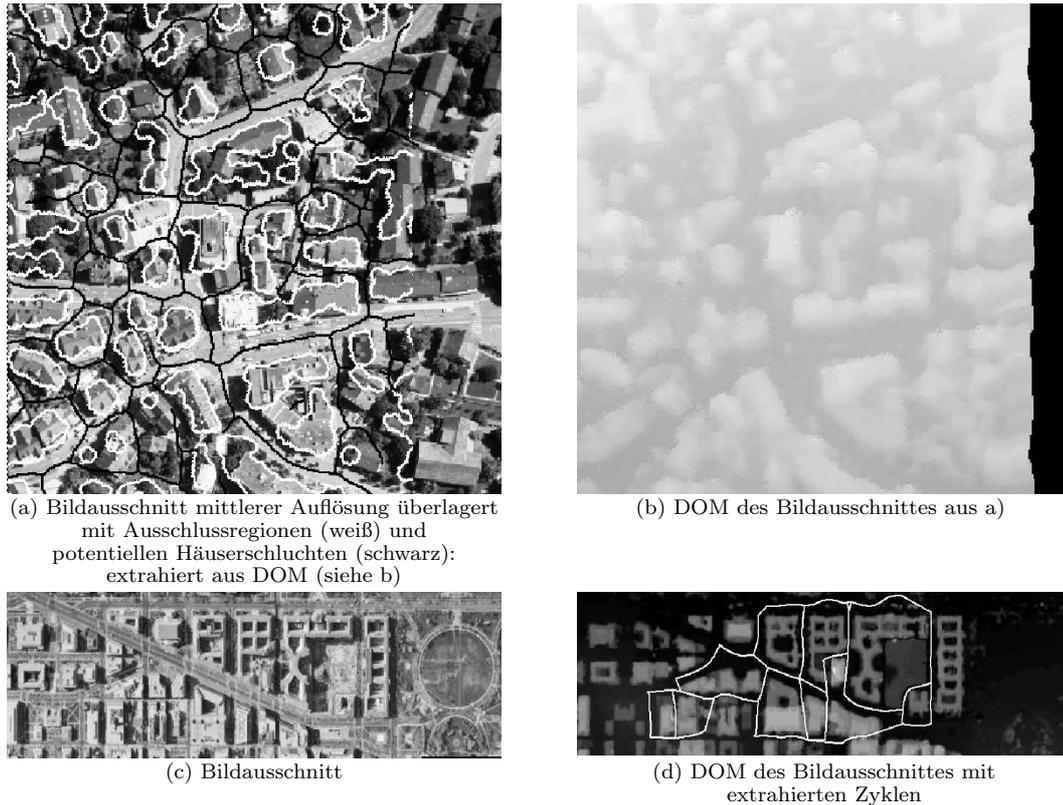


Abbildung 22. Bestimmung von Ausschlussregionen und potentiellen Häuserschluchten

detektierten Blobs und dunklen Linien sowie deren Projektion in den korrespondierenden Bildausschnitt. Die Erweiterung der Linienextraktion um eine Gruppierungskomponente zur Bestimmung von Zyklen ist in Abbildungen 22 c) und d) am Beispiel eines Datensatzes von Washington DC gezeigt.

Im nächsten Schritt der lokalen Kontextanalyse erfolgt die Untersuchung des Höhenmodells nach sichttoten Räumen (Kontextrelation „verdeckt“). Hierzu wird jedes Rasterelement des DOM mittels Ray-Tracing dahingehend überprüft, ob es im aktuellen Luftbild sichtbar ist oder nicht. Wie schon Abbildung 8 in Abschnitt 3.1 zeigt, gibt das DOM das „Stadtreief“ meist nur in unscharfer (geglätteter) Form wider. Daher wird in der Implementierung des Ray-Tracing-Verfahrens der Winkel des Sehstrahls zwischen DOM-Pixel und Projektionszentrum des betreffenden Luftbildes um einen vorgegebenen Wert (z.B. 10°) verringert. Diese Sichtbarkeitsanalyse wird für alle zur Verfügung stehenden Luftbilder durchgeführt, und das Ergebnis wird in einem mit dem DOM korrespondierenden Label-Bild gespeichert.

Abbildung 23 a) zeigt ein DOM-Mosaik, das aus jeweils zwei Stereo-Modellen abgeleitet wurde. Die Punkte im Mosaik markieren die Lage der Projektionszentren der Bilder 1 – 4. In Abbildung 23 b) und c) ist die Sichtbarkeitsanalyse für einen Szenenausschnitt (siehe schwarzes Quadrat in a) auf Grundlage von Bild 2 und Bild 4, die verschiedenen Stereo-Modellen zugeordnet sind, durchgeführt. Die eingezeichneten Regionen entsprechen sichttoten Räumen im jeweils anderen Bild. Wie eingangs erwähnt, hat aufgrund möglicher DOM-Fehler eine Kontextinformation, die alleine aus dem DOM abgeleitet wurde, eine geringere Zuverlässigkeit. Daher werden die detektierten Ausschluss- und Verdeckungsregionen nur bei der Extraktion von sicheren Teilen des Straßennetzes in vollem Umfang berücksichtigt. Im Gegensatz dazu können bei der Netzvervollständigung kleinere Ausschluss- und Verdeckungsregionen überbrückt werden, sofern genügend sichere Straßenteile in ihrer direkten Nachbarschaft extrahiert wurden.

Als letzte Kontextrelation, die von Gebäuden und hoher Vegetation ausgeht, wird die Relation „wirft Schatten“ analysiert. Die Vorgehensweise zur Bestimmung vorläufiger Schattenregionen gestaltet sich ähnlich wie die

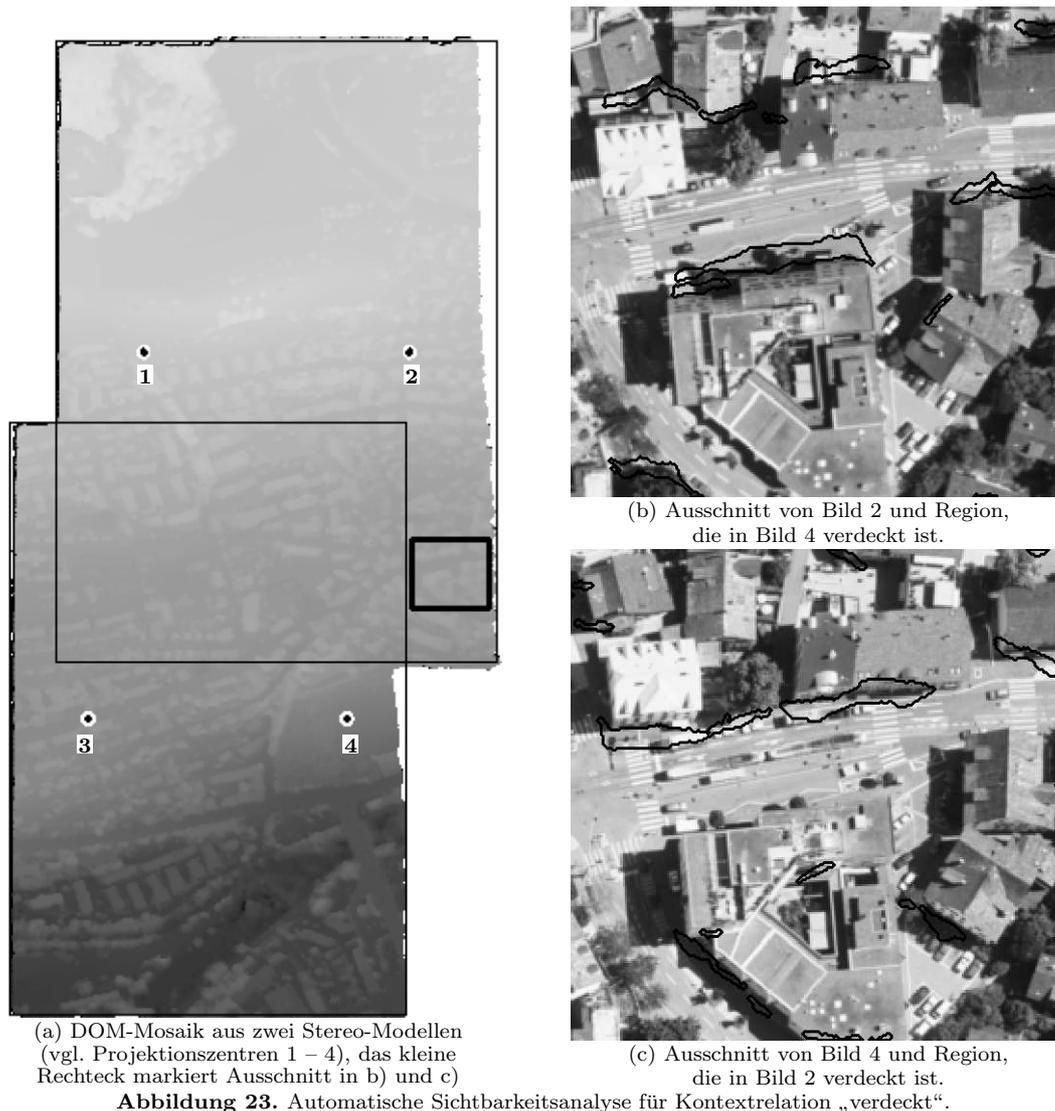


Abbildung 23. Automatische Sichtbarkeitsanalyse für Kontextrelation „verdeckt“.

Sichtbarkeitsanalyse, nur dass anstelle des Sehstrahls zum Projektionszentrum nun ein Sonnenstrahl angesetzt wird. Wegen der erwarteten Ungenauigkeiten des DOM werden die so erhaltenen Regionen nur als vorläufige Schattenbereiche angesehen. Ihre Verfeinerung geschieht nach der Projektion in das hochaufgelöste Bild durch Regiongrowing, das zusätzlich auf dunkle homogene Regionen eingeschränkt ist (Eckstein und Steger, 1996). Die Fokussierung auf homogene Regionen ist sinnvoll, da nur Schattenbereiche auf Straßen von Interesse sind. Abbildung 24 zeigt vorläufig und endgültig segmentierte Schattenregionen.

6.2.2 Fahrzeuge und Fahrzeugkolonnen

Die automatische Erkennung von Fahrzeugen in Luftbildern wird in dieser Arbeit in ähnlicher Weise wie in (Ruskoné, 1996) und (Quint, 1997) zur Unterstützung der Straßenextraktion herangezogen. Die Fahrzeugdetektion hat sich jedoch in den vergangenen Jahren auch als selbständiges Forschungsthema etabliert. Motiviert wird sie einerseits von militärischen und sicherheitsrelevanten Anwendungen, z.B. der automatischen Erkennung gegnerischer Ziele und der Überwachung von Transportaktivitäten. Auf der anderen – zivilen – Seite wird es durch die wachsende Verkehrsdichte zunehmend wichtiger, den Verkehrsfluss automatisch zu überwachen und durch Verkehrsleitsysteme oder so genannte ITS (Intelligent Transportation Systems) lenkend in den Verkehr einzugreifen. Die automatische Erkennung von Fahrzeugen aus Bildern und Bildsequenzen ist eine entscheidende



(a) Vorläufige Schattenregion abgeleitet aus DOM-Neigung und Sonnenstand.

(b) Endgültige Schattenregion nach Regiongrowing

Abbildung 24. Analyse der Kontextrelation „wirft Schatten“.

Grundlage zum Aufbau und Betrieb solcher Systeme. (Bogenberger et al., 1999) zeigen, dass luftbildgestützte Verfahren aufgrund der flächenhaften Erfassung verkehrsbezogener Daten hierfür eine innovative Lösung bieten und gleichzeitig wichtige Basis-Informationen für die Analyse und Planung von Straßennetzen liefern.

Entsprechend diesen vielfältigen Beweggründe für die automatische Fahrzeugerkennung sind die bislang entwickelten Verfahren in ihren Modellen und Strategien sehr heterogen. Ein ausführlicher Literaturüberblick hierzu ist in (Hinz, 2002, 2003) zu finden, wo auch mehr Details zur Implementierung des an der TU München entwickelten Verfahrens zur Fahrzeugdetektion nachgeschlagen werden können. Im Folgenden werden nur die wesentlichen Erkenntnisse zusammengestellt.

Fahrzeugmodell:

Aufgrund der häufigen Präsenz von Fahrzeugen in Stadtgebieten und ihrer daraus resultierenden Relevanz für die Straßenextraktion wird an die Fahrzeugdetektion ein weit höherer Anspruch hinsichtlich Genauigkeit, Vollständigkeit und Korrektheit gestellt als z.B. an die Detektion von Gebäuden und hoher Vegetation. Daher ist ein hoher Detaillierungsgrad des zugrunde liegenden Fahrzeugmodells zwingend erforderlich. Ein einfaches Modell – z.B. ein Rechteck-Modell – hätte aufgrund vieler ähnlich aussehender Objekte eine hohe Fehlerrate zur Folge.

Geometrisch werden Fahrzeuge wie in (Tan et al., 1998; Haag und Nagel, 1999; Zhao und Nevatia, 2001) als 3D Objekte mittels einer Drahtgitter-Repräsentation beschrieben und ihre Substrukturen wie Frontscheibe, Motorhaube, Dach etc. in das Modell miteinbezogen. Der Detailreichtum einer solchen Beschreibung hat in der Regel zur Folge, dass eine große Anzahl an Modellen nötig ist, um die unterschiedlichen Fahrzeugtypen und -größen zu modellieren. (Sullivan et al., 1995) zeigen aber durch eine eigenwertbasierte Untersuchung der geometrischen Variationen von PKW-Modellen, dass im Falle der für Luftbilder typischen Detailerkennbarkeit eine große Mehrheit an Personenkraftwagen mit relativ wenigen Modellen beschrieben werden kann (z.B. vier in (Zhao und Nevatia, 2001)). Flexible oder hierarchische Modelle wie in (Dubuisson-Jolly et al., 1996; Olson et al., 1996; Hinz und Baumgartner, 2001) kommen daher in dieser Arbeit nicht zum Einsatz. Stattdessen wird der Suchraum für Fahrzeuge durch die Segmentierung der globalen Kontextgebiete und durch die Bestimmung der Häuserschluchten und Ausschlussregionen in den Abschnitten 6.1 und 6.2.1 eingeschränkt.

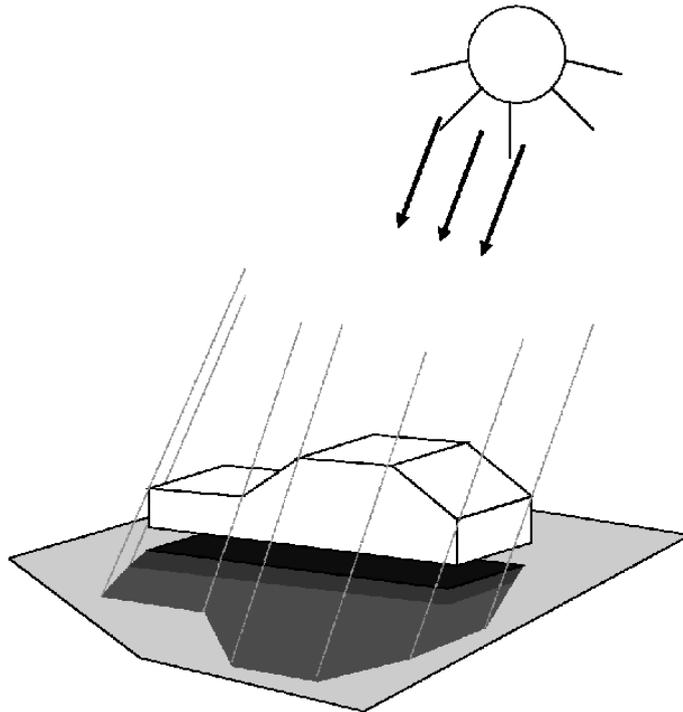


Abbildung 25. Skizze zum geometrischen Teil des Fahrzeugmodells.

Wie aus Abbildung 25 zu entnehmen ist, wird das geometrische Modell noch um seine **Schattensilhouette** ergänzt. In mehreren Tests hat sich herausgestellt, dass der Schattenwurf bei so kleinen Objekten wie Fahrzeugen einen wesentlichen Beitrag liefert, um auf die Präsenz eines 3D-Objektes zu schließen. Dies ist allerdings nur möglich, wenn bereits eine 3D-Fahrzeughypothese zugrunde gelegt wird. Versuche, aus einer 2D-Hypothese über den Fahrzeugschatten und die Kenntnis des Sonnenstandes auf die 3D-Form von Fahrzeugen zurückzuschließen, stellten sich als wenig robust heraus (Hinz et al., 2001b).

Bestandteile des **radiometrischen Modells** sind die Annahmen, dass Motorhaube und Dach gleiche Farbe bzw. Helligkeit aufweisen und dass sich der Fahrzeugschatten auf der Straßenfläche als homogen dunkle Region abzeichnet. Außerdem werden in Anlehnung an (Zhao und Nevatia, 2001) alle modellierten Kanten mit einem Wert für ihre erwartete Signifikanz belegt. Dieser Wert hängt von Helligkeit, Sonnenstand, Blickwinkel und Orientierung des Fahrzeugs ab. Die Notwendigkeit dieser Modellierung wird offensichtlich, wenn man die Stärken der Bildkanten verschiedener Fahrzeuge und ihrer Schatten betrachtet. Wie in den Abbildungen 26 und 27 zu sehen ist, sind bei hellen Fahrzeugen vor allem die Kanten der Front- und Heckscheibe sowie die Schattenkante auffällig. Bei dunklen Fahrzeugen hingegen sind innerhalb des Fahrzeugs kaum Kanten zu erkennen. Dafür ergeben sich sehr deutliche Kanten zwischen Fahrzeug und Straßenbelag. Diese Art der radiometrischen Modellierung wurde ursprünglich von (Hoogs und Bajcsy, 1995) für die Segmentierung von Gebäudekanten vorgeschlagen. Die Erweiterung in der vorliegenden Arbeit – auch gegenüber dem in (Zhao und Nevatia, 2001) verwendeten radiometrischen Modell – bezieht sich auf die explizite Einbeziehung der Radiometrie der Fahrzeugflächen.

Oftmals reicht diese zwar detaillierte, aber sehr lokal ausgerichtete Beschreibung für eine zuverlässige Fahrzeugdetektion nicht aus. Insbesondere kontrastarme Fahrzeuge, die – wie das in Abbildung 27 gezeigte Fahrzeug – bei lokaler Betrachtung lediglich als homogener „Blob“ erscheinen, können nur unter Verwendung von Umgebungsinformation mit größerer Zuverlässigkeit detektiert werden. Daher wird das strenge parametrische Modell durch ein generisches **Modell für Fahrzeugkolonnen** ergänzt. Dies erlaubt eine globalere Sichtweise auf Fahrzeuge als Bestandteile des Verkehrs. Da sich Fahrzeuge durch große Symmetrie bezüglich ihrer Mit-

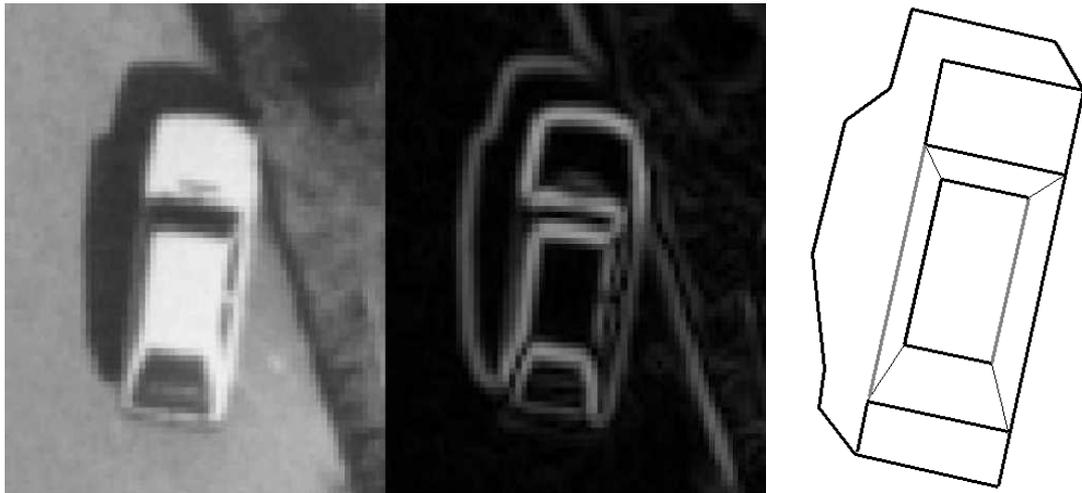


Abbildung 26. Helles Fahrzeug (links), Gradientenbild (Mitte), Signifikanz der Kanten (rechts): Schwarz / fett = auffällig. Grau / fett = optional. Schwarz / dünn = nicht benutzt, da zu kurz.

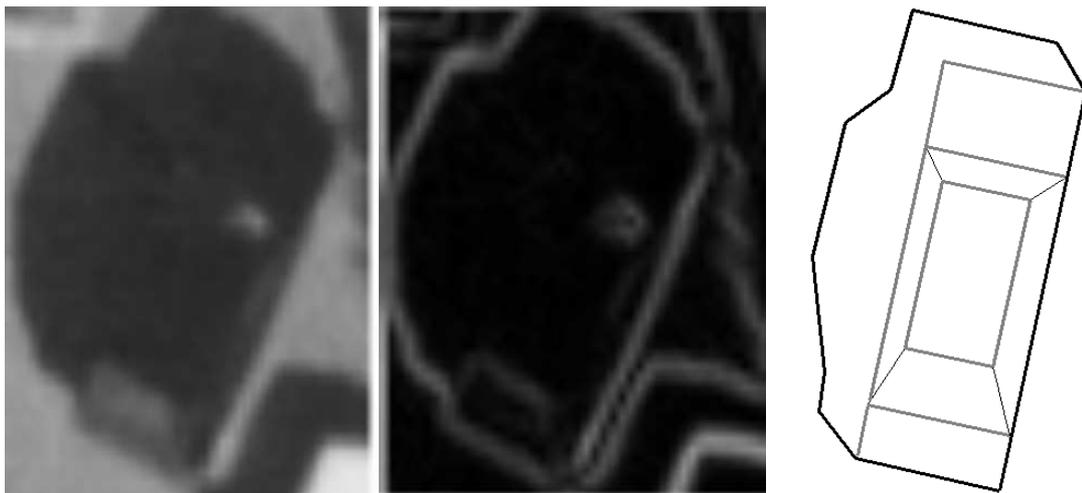


Abbildung 27. Dunkles Fahrzeug (links), Gradientenbild (Mitte), Signifikanz der Kanten (rechts): Schwarz / fett = auffällig. Grau / fett = optional. Schwarz / dünn = nicht benutzt, da zu kurz.

telachse auszeichnen, wird eine Kolonne als Band einheitlicher Breite (etwa eine Fahrzeugbreite) mit großer Grauwertsymmetrie in orthogonaler Richtung zur Mittelachse modelliert. Zusätzlich muss die Bandmittelachse zumindest abschnittsweise von parallelen Kanten flankiert werden und in bestimmten Abständen von kurzen Kanten senkrecht geschnitten werden. Die parallelen Kanten korrespondieren dabei mit den Fahrzeugseiten und die senkrechten Kanten mit Fahrzeugkanten an Front, Heck und Windschutzscheiben (siehe auch Abbildung 28). Dieses Kolonnenmodell ähnelt in gewisser Weise der Beschreibung regelmäßig angeordneter Objekte durch globale Merkmale im Frequenzraum (Burlina et al., 1997). Gegenüber (Burlina et al., 1997) hat das hier vorgeschlagene Kolonnenmodell jedoch den Vorteil, dass Regelmäßigkeiten der Objekte und ihrer Anordnung nicht ausschließlich auf Grauwertinformation basieren, sondern dass über die explizite Modellierung von Bandmittelachse und Kanten höherwertige Information miteinfließt.

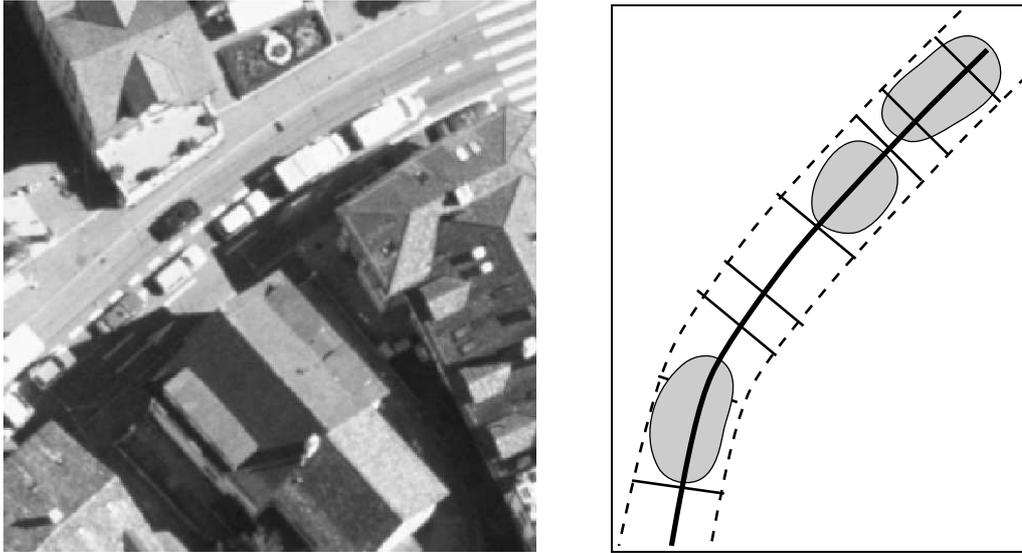


Abbildung 28. Bild einer typischen Fahrzeugkolonne (links) und Skizze des generischen Kolonnenmodells (rechts).

Extraktionsstrategie für Fahrzeuge:

Die Strategie zur Extraktion von Fahrzeugen gliedert sich in drei Phasen: (1) Die Fahrzeugdetektion basierend auf dem strengen, parametrischen Fahrzeugmodell, (2) die Detektion von Fahrzeuggruppen mittels des generischen Kolonnenmodells und (3) die Detektion weiterer Fahrzeuge nach der Fusion der unabhängig voneinander erzielten Ergebnisse aus (1) und (2).

Zur **Fahrzeugdetektion** wird eine überwiegend modellgetriebene Strategie herangezogen, weil diese eine effiziente Vorgehensweise für die Zuordnung von relativ einfach zu beschreibenden Modellen zu komplexen Bilddaten darstellt (siehe Abschnitt 2.1.4). Da gemäß dem Modell jedes Fahrzeug mindestens eine deutliche Kante besitzen muss, kann zunächst eine Suchraumeinschränkung durchgeführt werden. Hierzu wird ein Kantenstärkebild erzeugt, aus dem Teilstücke von potentiellen Fahrzeugkanten durch Schwellwertbildung segmentiert werden. Anhand von Position x, y und Gradientenrichtung ϕ jedes beteiligten Kantenpixels wird im nächsten Schritt eine Ansicht des 3D-Modells in entsprechender Orientierung generiert und als Fahrzeughypothese inklusive der Schattenregion an die betreffende Stelle des Bildes projiziert. Um eine Vielzahl an redundanten Projektionen zu vermeiden, wird analog zu (Michaelsen und Stilla, 2000) im Voraus eine Datenbank mit Ansichten erzeugt, auf die während der Extraktion zugegriffen wird¹⁵. Im Folgeschritt wird aus der Dachregion der Hypothese die Referenzhelligkeit des Fahrzeugs bestimmt, die es nun ermöglicht, die erwartete Signifikanz jeder Fahrzeugkante festzulegen. Schließlich erfolgt die Bewertung der Hypothese mittels geometrischer und radiometrischer Merkmale wie z.B. Kantenunterstützung und Farbkonstanz (Hinz, 2003). Übersteigt die Bewertung einen vorgegebenen Schwellwert, wird die Hypothese als Fahrzeug akzeptiert.

Das Prinzip der Fahrzeugdetektion ist in Abbildung 29 anhand des dunklen Fahrzeugs aus Abbildung 27 illustriert. Die fett eingezeichneten Pixel stehen beispielhaft für jene Kantenpixel, die zur Suchraumeinschränkung herangezogen werden. Die restlichen Kantenpixel finden nur im Rahmen der Hypothesenbewertung Verwendung. Ergebnisse der Fahrzeugdetektion sind in Abbildung 30 dargestellt. Wie zu sehen ist, konnte die Mehrheit der Fahrzeuge detektiert werden. Die geringe Anzahl an fehlerhaften Detektionen bestätigt die Qualität von Modell und Strategie. Allerdings bedingt das detaillierte Modell auch mehrere fehlende Detektionen, z.B. bei stärkeren geometrischen Abweichungen vom Modell oder bei Störungen der Radiometrie durch Reflexionen.

¹⁵Die Diskretisierung des x, y, ϕ -Raumes wird aus Bildmaßstab und mittlerer Fahrzeuggröße abgeleitet.

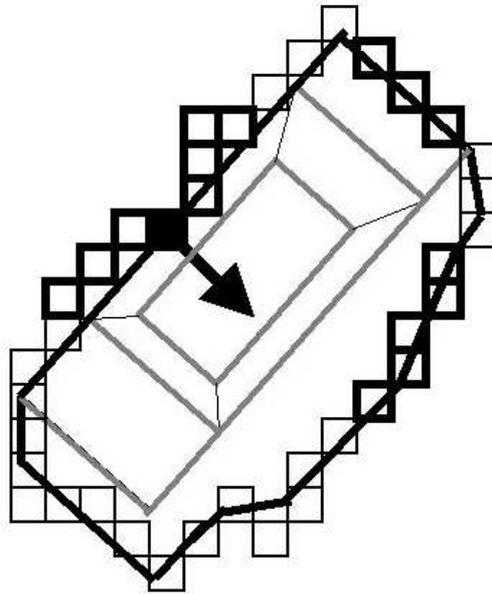


Abbildung 29. Prinzipskizze der Zuordnung von projiziertem Fahrzeugmodell zur Bildinformation (der Übersichtlichkeit wegen nur für Kanteninformation illustriert): Pfeil = Referenzrichtung des Fahrzeuges (entsprechend der Gradientenrichtung im betreffenden Pixel). Fett umrandete Pixel = auffälligen Kantenpixel, dünn umrandete Pixel = restliche Kantenpixel. Strichcodierung des Modells wie in Abbildung 27.

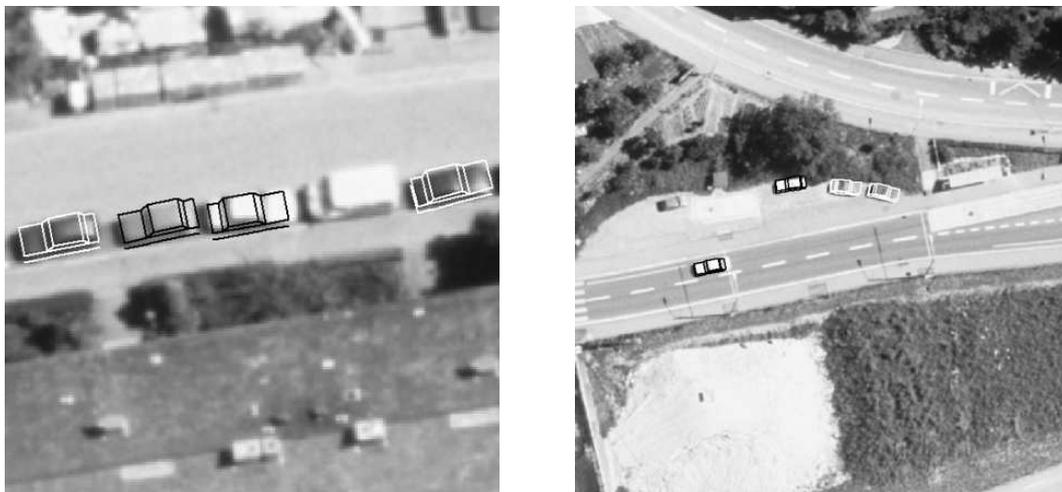
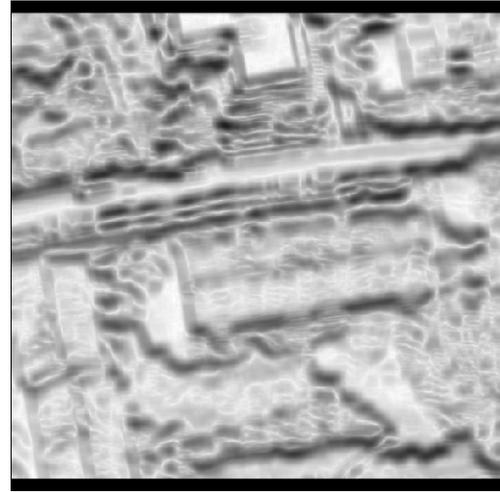


Abbildung 30. Ergebnisse der Fahrzeugdetektion: Als dunkel deklarierte Fahrzeuge haben weiße Drahtgitterdarstellung und als hell deklarierte Fahrzeuge haben schwarze Drahtgitterdarstellung.

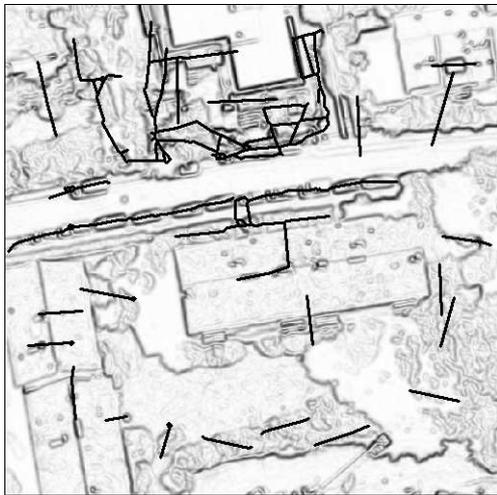
Die **Detektion von Fahrzeugkolonnen** startet mit der Extraktion von Bändern mit den oben angesprochenen Symmetrieeigenschaften. Zu diesem Zweck wird eine so genannte „richtungsabhängige Symmetrie-Karte“ erzeugt, die Stärke und Richtung der Grauwertsymmetrie für jedes Pixel angibt (Hinz, 2003). Um erste Hypothesen für potentielle Mittelachsen von Fahrzeugkolonnen zu erhalten, werden in dieser Symmetrie-Karte lokale Symmetrie-Maxima mit ähnlicher Symmetrie-Richtung zu Konturen verbunden. Die endgültige Aus-



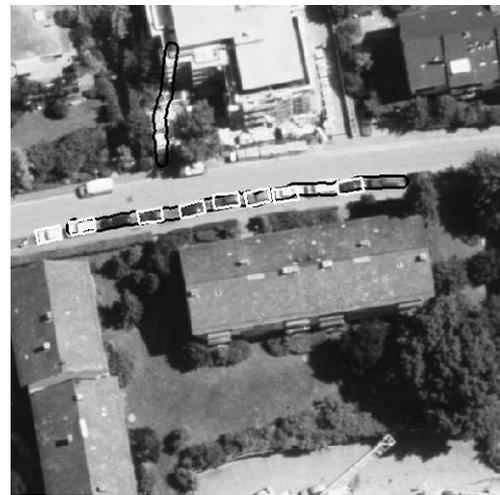
(a) Bildausschnitt



(b) Symmetrie-Karte für vertikale Richtung $\alpha = 90^\circ$
große Symmetrie = hell.



(c) Symmetrie-Achsen (Mittelachsen potentieller
Kolonnen) dem Gradientenbild überlagert



(d) Anhand von Kanteninformation ausgewählte
Kolonnen (schwarz), überlagert mit Ergebnis
aus Fahrzeugdetektion (weiß).

Abbildung 31. Schritte zur Detektion von Kolonnen.

wahl an Hypothesen geschieht durch die Bewertung einer Mittelachse aufgrund ihrer Länge sowie aufgrund von Kanteninformation in ihrer unmittelbarer Umgebung: Wie im Kolonnenmodell spezifiziert, müssen kurze, senkrecht schneidende Kanten entlang der Mittelachse fahrzeugtypische Anzahl und räumliche Verteilung haben. Außerdem erfordert eine gute Kolonnenhypothese ausreichend viele flankierende Kanten im Abstand etwa einer halben Fahrzeugbreite. Abbildung 31 zeigt die Symmetrie-Karte für eine festgelegte Richtung, sowie die Bildung von Mittelachsen und die weiteren Schritte zur endgültigen Auswahl der Fahrzeugkolonnen. Im Gegensatz zur Fahrzeugdetektion tendiert die Kolonnendetektion aufgrund des weniger streng formulierten Modells zu einer höheren Vollständigkeit. Wie an den Fehldetektionen in Abbildung 31 zu sehen ist, geht dies aber deutlich zu Lasten der Zuverlässigkeit.

Deshalb zielt die letzte Phase der Fahrzeugextraktion darauf ab, die Vorteile beider Detektionen synergetisch zu nutzen. Zu diesem Zweck findet eine **Fusion** der unabhängig voneinander erhaltenen Ergebnisse statt. Fahrzeuge und Kolonnen werden auf gegenseitige Überlappung und Parallelität geprüft, und nur solche Kolonnen werden weiter analysiert, die durch ausreichend viele detektierte Fahrzeuge gestützt werden. Im Gegenzug wird das Wissen über die Kolonnenbildung ausgenutzt, um unterbliebene Detektionen zu entdecken. Weil diese Art



Abbildung 32. Ergebnisse der gesamten Fahrzeugextraktion: Weiß: Ergebnisse aus Fahrzeugdetektion. Schwarz: Durch Fusion mit Kolonnendetektion zusätzlich detektierte Fahrzeuge.

von Fehler sehr häufig durch kontrastarme Fahrzeuge zustande kommt, werden nun homogene Blobs innerhalb der Kolonnen extrahiert und durch Rechtecke approximiert. Ein Rechteck wird als neu gefundenes Fahrzeug deklariert, sofern es kollinear zur Kolonne ausgerichtet ist und seine Ausmaße mit den Dimensionen eines Fahrzeugs übereinstimmen. Abschließend kommen noch jene detektierten Fahrzeuge zum Ergebnis hinzu, die nicht Teil einer Kolonne sind. Dies erscheint aufgrund des strengen Modells der Fahrzeugdetektion und der daraus resultierenden niedrigen Fehlerrate als gerechtfertigt.

Abbildung 32 zeigt das Endergebnis der Fahrzeugextraktion für zwei Bildausschnitte. Das Ergebnis der Fahrzeugdetektion ist mit weißen Rechtecken gekennzeichnet. Schwarze Rechtecke markieren Fahrzeuge, die durch die Fusion mit der Kolonnendetektion zusätzlich gefunden werden konnten. Wie zu sehen ist, konnte durch die Fusion von Fahrzeug- und Kolonnendetektion die Vollständigkeit des Ergebnisses deutlich gesteigert werden. Lediglich solche Fahrzeuge wurden nicht extrahiert, die aufgrund radiometrischer Störungen oder ihrer gänzlich anderen Geometrie zu sehr vom Fahrzeugmodell abweichen und gleichzeitig kein Bestandteil einer Kolonne sind. In Abbildung 32 b) ist im oberen Bereich eine fehlerhafte Detektion zu finden. Sie ist darauf zurückzuführen, dass ein Teil des dunklen Baumes fälschlicherweise als Abschnitt einer Kolonne detektiert wurde. Außerdem zeichnet sich die Region durch ausreichende Homogenität aus, was zur Detektion eines Blobs führte, dessen approximierendes Rechteck die Kriterien für ein Fahrzeug erfüllt.

Die Detektion von Fahrzeugen schließt die Phase der Kontextanalyse ab. Der Großteil der verfügbaren Kontextinformation steht damit der gesamten Straßenextraktion zur Verfügung. Im Zuge der Bestimmung sicherer Straßenobjekte im folgenden Kapitel wird in geringerem Umfang bei der Verifikation von Fahrspur- und Fahrbahnverknüpfungen durch Fahrzeuge noch einmal Kontextinformation extrahiert.

7 Straßenobjekte

Die Zielsetzung der zweiten Phase der Extraktionsstrategie ist die Bestimmung von sicheren Straßenobjekten. Um eine hohe Zuverlässigkeit zu erreichen, wird in dieser Phase in besonderem Ausmaß von detailliertem Modellwissen Gebrauch gemacht. Das Ergebnis sind Hypothesen für die Teilobjekte „Fahrspuren“ und „Fahrbahnen“, die sehr gut mit dem Modell übereinstimmen und damit eine relativ zuverlässige Ausgangsbasis für die in Kapitel 8 folgende Netzvervollständigung bieten. Desweiteren kommt hier das in Kapitel 4 entwickelte Konzept zur internen Bewertung von Extraktionsergebnissen zum Einsatz. Die Kriterien K I und K II werden für die jeweiligen Straßenobjekte tabellarisch zusammengefasst. Hierbei ist zu beachten, dass die Kriterien K II die *autonome Bewertung* eines Straßenobjektes definieren. Bottom-up und top-down Bewertung ergeben sich aus dem Ablauf der Extraktion.

Wie auch in Abbildung 33 zu sehen ist, werden als Eingangsinformation zur Fahrspurextraktion einerseits vorläufige Straßenabschnitte und andererseits Fahrspurbegrenzungen benötigt. Die Straßenabschnitte bestehen aus Straßenachsen und Fahrbahnen, wobei die Bestimmung von Straßenachsen in der niedrigen Auflösung und die Extraktion von Fahrbahnen über ihre Teilobjekte „Fahrbahnrand“ und „Fahrbahnbelag“ in der mittleren Auflösung geschieht (Abschnitt 7.1). Um den Detailreichtum der hohen Auflösung zu nutzen, kommen als Spezialisierungen von Fahrspurbegrenzungen die „durchgezogenen“ bzw. „strichlierten Markierungen“ in Abschnitt 7.2 hinzu. Neben Markierungen werden hier weitere Straßenränder ermittelt, die nicht Teil der Fahrbahnen aus Abschnitt 7.1 sind. Somit können im weiteren Verlauf der Extraktion auch Fahrbahnen und Straßenabschnitte erzeugt werden, die alleine in mittlerer und niedriger Auflösung zu wenig Evidenz haben. In Abschnitt 7.3.1 folgt die Konstruktion von Fahrspuren in drei verschiedenen Varianten: (a) aus Markierungen mit parallelen Fahrbahnrandern, (b) nur aus Markierungen und (c) nur aus vorläufigen Straßenabschnitten. In Variante (c) werden die Fahrspuren aus der Form der zugrunde liegenden Fahrbahnen geschätzt. Alle Fahrspuren inklusive derer aus Schattenbereichen und der detektierten Fahrzeuge werden anschließend fusioniert und gruppiert (Abschnitt 7.3.2). Der abschließende Abschnitt 7.4 erläutert die Konstruktion und Gruppierung von Fahrbahnen basierend auf den zuvor erhaltenen Fahrspuren.

7.1 Vorläufige Straßenabschnitte

Die Bestimmung vorläufiger Straßenabschnitte zu Beginn der eigentlichen Straßenextraktion verfolgt das Ziel, erste Hypothesen für Teilstücke des Straßennetzes zu erhalten. Der „vorläufige“ Charakter dieser Straßenabschnitte rührt daher, dass an dieser Stelle nur Wissen der hohen und mittleren Abstraktionsstufe zum Einsatz kommt. Die Fusionierung mit der Detailinformation aus der niedrigen Abstraktionsstufe bzw. der hohen Auflösung geschieht erst später im Rahmen der Konstruktion von Fahrspuren.

Das in Abbildung 13 skizzierte Straßenmodell beschreibt in der hohen Abstraktionsstufe einen Straßenabschnitt in Form seiner Straßenachse als limitiert gekrümmte Linie mit relativ konstantem Höhenverlauf. In der mittleren Abstraktionsstufe besteht ein Straßenabschnitt aus mindestens einer Fahrbahn, die als homogenes Band limitierter Neigung modelliert wird. Die seitliche Begrenzung des Bandes entspricht den Fahrbahnrandern und seine Fläche dem Fahrbahnbelag. Die Extraktion der vorläufigen Straßenabschnitte profitiert somit von den Maßstabsraumeigenschaften einer Straße. Während die niedrige Auflösung den Verlauf der Straßenachse in relativ stark generalisierter Form wiedergibt, kann in der mittleren Auflösung die präzise Lage und Orientierung der Straßenränder ermittelt werden. Flächenhafte Information über Homogenität und Neigung des Fahrbahnbelags steuert ebenfalls die mittlere Auflösung bei (siehe z.B. auch Abbildung 10 in Abschnitt 3.2).

Die Extraktion vorläufiger Straßenabschnitte startet mit der Generierung initialer Hypothesen für Straßenränder durch Kantenextraktion in der mittleren Auflösung bzw. für Straßenachsen durch Linienextraktion in der niedrigen Auflösung. Im Gegensatz zu den meisten anderen Ansätzen wird die Linienextraktion jedoch

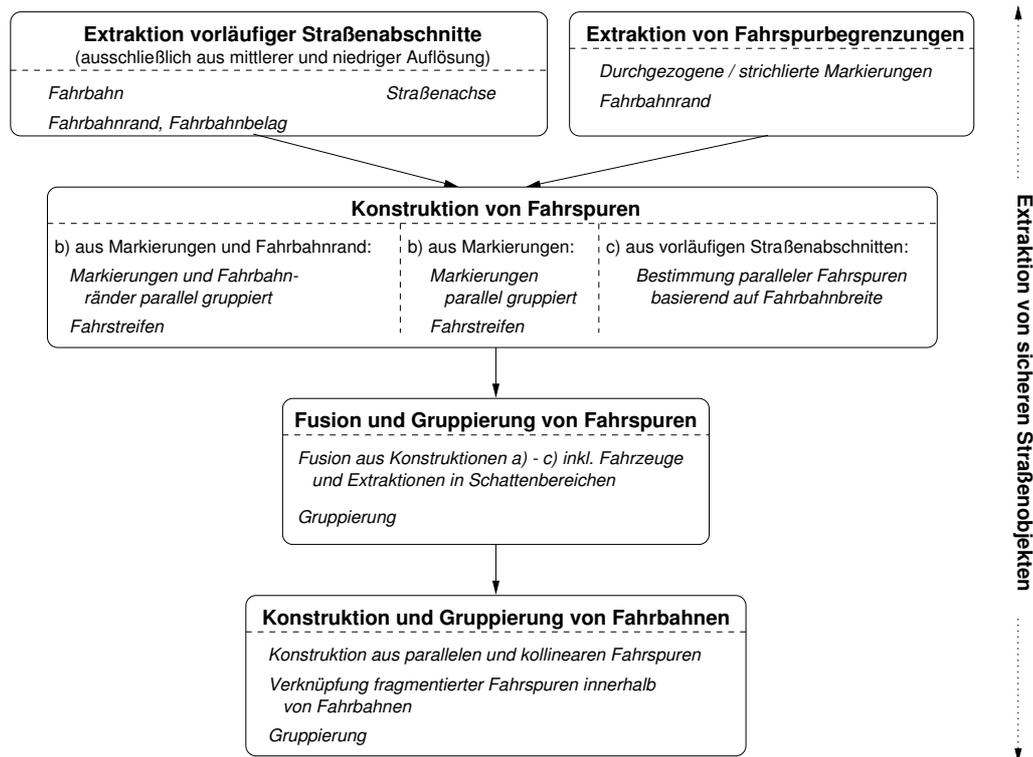


Abbildung 33. Extraktionsstrategie: Teil „Extraktion von sicheren Straßenobjekten“ (vgl. Abbildung 18).

nicht auf helle Linien eingeschränkt. Vielmehr können durch die Extraktion dunkler Linien im Bild der Gradientenstärken simultan Linien bestimmt werden, die im Originalbild ein helles, dunkles oder treppenförmiges Grauwertprofil besitzen. Nach Transformation der initialen Straßenachsenhypothesen in die mittlere Auflösung erfolgt die Zuordnung von Linien- und Kanteninformation (Abbildung 34). Hierzu wird entlang einem orthogonal zur lokalen Linienrichtung erzeugten Suchprofil jenes Kantenpixel lokalisiert, das die kürzeste Entfernung zur Linie besitzt. Die Entfernung definiert somit die Breite des Bandes an dieser Position und wird als Attribut pro Linienpixel gespeichert. Wird kein Kantenpixel bis zu einer maximalen Entfernung gefunden, erfolgt die Breitenbestimmung durch nachträgliche Interpolation aus den Breitenwerten benachbarter Linienpixel.

Die Repräsentation der vorläufigen Straßenabschnitte in Form von verketteten Linienpixel mit Breitenattribut ist bei großen Bildern unter Umständen sehr speicheraufwändig. Zur Datenreduktion werden daher geradlinige Pixelketten durch Polygonseiten durch den Algorithmus von (Ramer, 1972) vereinfacht – mit der zusätzlichen Bedingung, dass auch bei starken Breitenänderungen Polygonpunkte erzeugt werden, um größere Fehler aufgrund der Mittelung der Breitenwerte entlang jeder Polygonseite zu vermeiden. Abbildung 35 a) zeigt das Ergebnis für einen Bildausschnitt, wobei zur leichteren Interpretation das Bild in originaler Auflösung den extrahierten Objekten hinterlegt wurde.

Zur Extraktion der vorläufigen Straßenabschnitte wurde in erster Linie strukturelle Information verwendet. Die flächenhaften Merkmale des Straßenbelags „Homogenität“ und „Neigung“ kamen bislang nicht zum Einsatz, so dass sie als Kriterien zur internen Bewertung der Hypothesen von großem Nutzen sind. Als weitere Bewertungskriterien kommen noch die durchschnittliche absolute Krümmung der Straßenachse und die Breitenvarianz der zugrunde liegenden Fahrbahn(en) hinzu. Da die Berechnung jeder einzelnen Bewertung auf Grundlage der *gesamten* Hypothese erfolgt, stützt sich die Bewertung nicht nur auf bislang ungenutztes, sondern auch auf höherwertiges Wissen. Die einzelnen zur Extraktion und Bewertung von vorläufigen Straßenabschnitten heran-

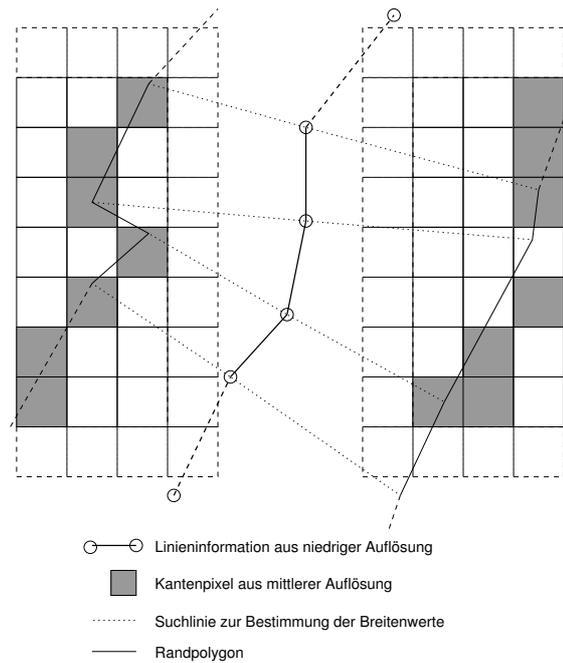


Abbildung 34. Konstruktionsprinzip von homogenen Bändern (siehe Text).

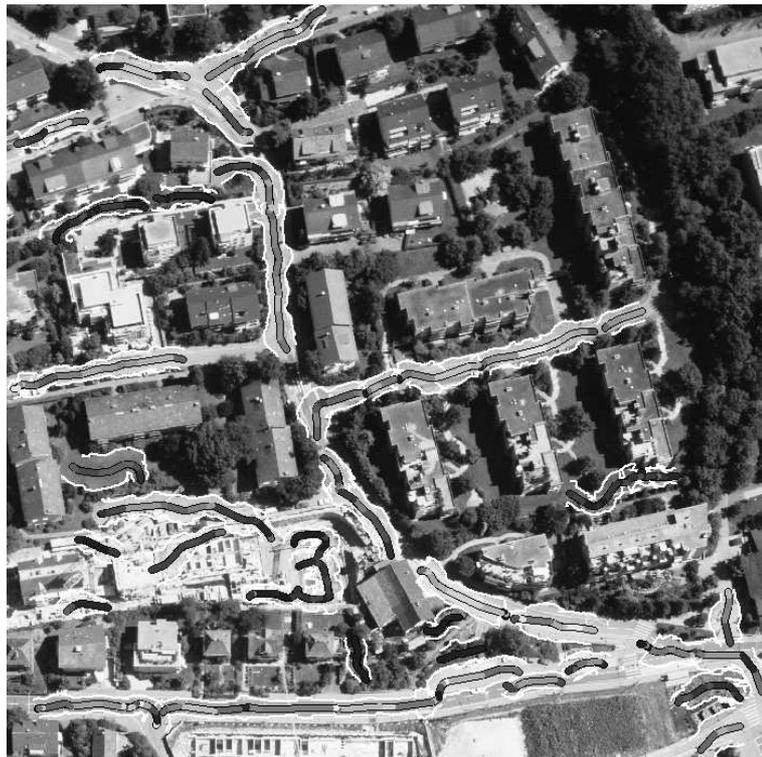
gezogenen Kriterien stellt Tabelle 3 zusammenfassend dar. Während die Kriterien K I durch feste Schranken parametrisiert sind („minimal“, „maximal“, ...), werden Begriffe wie „gering“ oder „konstant“ der Kriterien K II auf Fuzzy-Funktionen abgebildet. Die autonome Bewertung der vorläufigen Straßenabschnitte visualisiert Abbildung 35 b). Je heller die Straßenachse dargestellt wird, um so besser ist die Qualität eines Straßenabschnitts. Man sieht z.B., dass die fehlerhaften Extraktionen in den Vorgärten im linken oberen Quadranten oder im Baustellenbereich im linken unteren Quadranten des Bildes eine deutlich schlechtere Bewertung erhalten als die meisten korrekt extrahierten Straßenabschnitte. Ausnahmen, die fälschlicherweise schlecht bewertet wurden, sind vor allem Abschnitte mit großer Krümmung oder Breitenvarianz (z.B. im Kreuzungsbereich in der Bildmitte).

Kriterien K I	Kriterien K II
Dunkle lineare Struktur im Gradientenbild niedriger Auflösung	Große Grauwerthomogenität in mittlerer Auflösung
Beidseitig begrenzende Kante im Bild mittlerer Auflösung	Geringe Längsneigung
Minimale / maximale Breite	Sehr geringe Querneigung
Minimale Länge	Geringe mittlere Krümmung
	Geringe Breitenvarianz

Tabelle 3. Kriterien zur Extraktion und Bewertung von vorläufigen Straßenabschnitten.



(a) Vorläufige Straßenabschnitte:
Straßenachse (schwarz), Straßenrand (weiß).



(b) Bewertung der vorläufigen Straßenabschnitte:
Je heller die Straßenachse, desto besser die autonome Bewertung.

Abbildung 35. Extraktion und Bewertung von homogenen Bändern als vorläufige Straßenabschnitte.

7.2 Fahrspurbegrenzungen

Durch die Bestimmung von Fahrspurbegrenzungen wird nun der Extraktion detaillierteres Wissen hinzugefügt. Im Straßenmodell ist eine Fahrspurbegrenzung in drei verschiedene Konzepte spezialisiert: „Durchgezogene Markierung“ (Abschnitt 7.2.1), „strichlierte Markierung“ (Abschnitt 7.2.2) und „Straßenrand“ (Abschnitt 7.2.3). Die Extraktion der Straßenränder ergänzt dabei die im vorangegangenen Abschnitt bestimmten Straßenränder um solche, die nicht als Bestandteil vorläufiger Straßenabschnitte herangezogen werden konnten.

7.2.1 Durchgezogene Markierungen

Durchgezogene Markierungen sind Detailstrukturen, die nur in der hohen Auflösung modelliert werden können. Trotzdem sind sie ein sehr typisches und auffälliges Merkmal urbaner Straßen. Ihre charakteristischen Kriterien für die Extraktion sind z.B. ihre lineare Struktur, ihre große Helligkeit und Länge sowie ihr symmetrischer Kontrast beiderseits der Markierung. Initiale Hypothesen für durchgezogene Markierungen lassen sich daher über eine Extraktion dünner, heller Linien mit dem Steger-Algorithmus ermitteln (Steger, 1998a). Die initialen Hypothesen werden an Punkten mit starker Krümmung oder starker Grauwertänderung getrennt (Abbildung 36 a), da durchgezogene Markierungen einerseits durch Geradlinigkeit und andererseits in aller Regel auch durch einheitliche Helligkeit gekennzeichnet sind. Von den resultierenden Linien dienen fortan nur ausreichend lange Linien als Hypothesen. Deren Bewertung erfolgt über die Kriterien „Grauwerthomogenität“, „mittlere absolute Krümmung“ und „Kontrastsymmetrie“. Der Kontrast eines Linienpixels ist hierbei als Grauwertdifferenz zwischen Linienpixel und zugehöriger Kantenposition definiert (Abbildung 36 b). Das Verhältnis zwischen den gemittelten Kontrasten beiderseits der Linie gibt schließlich Auskunft über die Kontrastsymmetrie. Damit dient hier – wie zuvor auch bei den vorläufigen Straßenabschnitten – wieder die gesamte Hypothese als Grundlage für die Bewertung.

Unter dem Gesichtspunkt der Datenreduktion erfolgt abschließend eine Polygonapproximation und Mittelung der Attributwerte pro Polygonseite. Tabelle 4 fasst die einzelnen zur Extraktion und Bewertung involvierten Kriterien zusammen, und Abbildung 37 zeigt Ergebnisse der Extraktion durchgezogener Markierungen.

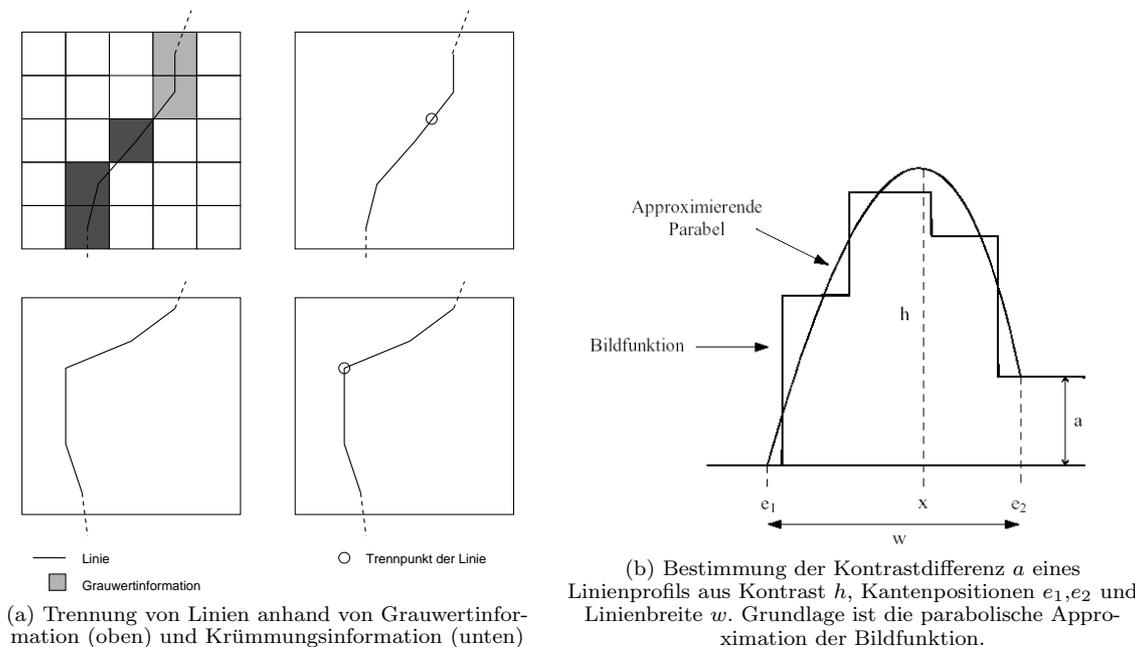


Abbildung 36. Prozessierung von Linien.



Abbildung 37. (a), (b) Ergebnisse für durchgezogene Markierungen.

Kriterien K I	Kriterien K II
Helle, lineare Struktur im Grauwertgebirge	Große Grauwerthomogenität entlang der Linie
Minimale Linienlänge	Große Kontrastsymmetrie entlang der Linie
Minimale Helligkeit	Geringe mittlere Krümmung

Tabelle 4. Kriterien zur Extraktion und Bewertung von durchgezogenen Markierungen.

7.2.2 Strichlierte Markierungen

Im Gegensatz zu einer durchgezogenen Markierung, die lediglich aus einer einzigen Linie besteht, sind strichlierte Markierung aus mehreren, topologisch verknüpften Linienteilstücken mit dazwischen liegenden Lücken aufgebaut. Linienteilstücke werden auf die gleiche Weise extrahiert und bewertet wie die durchgezogenen Markierungen in Abschnitt 7.2.1 – mit dem einzigen Unterschied, dass eine verkleinerte minimale und eine zusätzliche maximale Grenze für die Linienlänge eingeführt werden. Wie aus Tabelle 5 ersichtlich, kommen zur Extraktion der gesamten Markierung nun noch Kriterien hinzu, die Relationen zwischen Linienteilstücken beinhalten: Dies sind die maximale Lückenlänge zwischen zwei Linienteilstücken, deren Kollinearität und die Gesamtlänge der strichlierten Markierung.

Kriterien K I	Kriterien K II
Kamm-ähnliche Struktur im Grauwertgebirge	Große Grauwerthomogenität entlang der Linien
Minimale und maximale Linienlänge	Große Kontrastsymmetrie der Linien
Minimale Kollinearität von Linienpaaren	Konstante Linien- und Lückenlängen
Maximale Lückenlänge	Geringe mittlere Krümmung
Minimale Gesamtlänge	

Tabelle 5. Kriterien zur Extraktion und Bewertung von strichlierten Markierungen.

Der Algorithmus zur Extraktion von strichlierten Markierungen baut zunächst einen gewichteten Graphen auf, dessen Kanten die Linienteilstücke und dessen Knoten ihre Endpunkte sind. Die Gewichte werden aus den Kriterien K I ermittelt. Lücken, die gemäß den Kriterien K I potentiell überbrückt werden können, kommen als weitere Kanten hinzu. Eine strichlierte Markierung ergibt sich schließlich durch die Verkettung aller Linienteilstücke und Verknüpfungen zwischen überdurchschnittlich stark gewichteten Linienteilstücken. Die Verkettung erfolgt unter der Nebenbedingung, dass Linienteilstücke und Verknüpfungen immer alternieren. Abschließend werden für die endgültige Repräsentationsform Polygonverbindungen für überbrückte Lücken erzeugt und Linienteilstücke und Polygonverbindungen mit einheitlicher „Laufriichtung“ versehen. Abbildung 38 zeigt das Grundprinzip dieses Algorithmus.

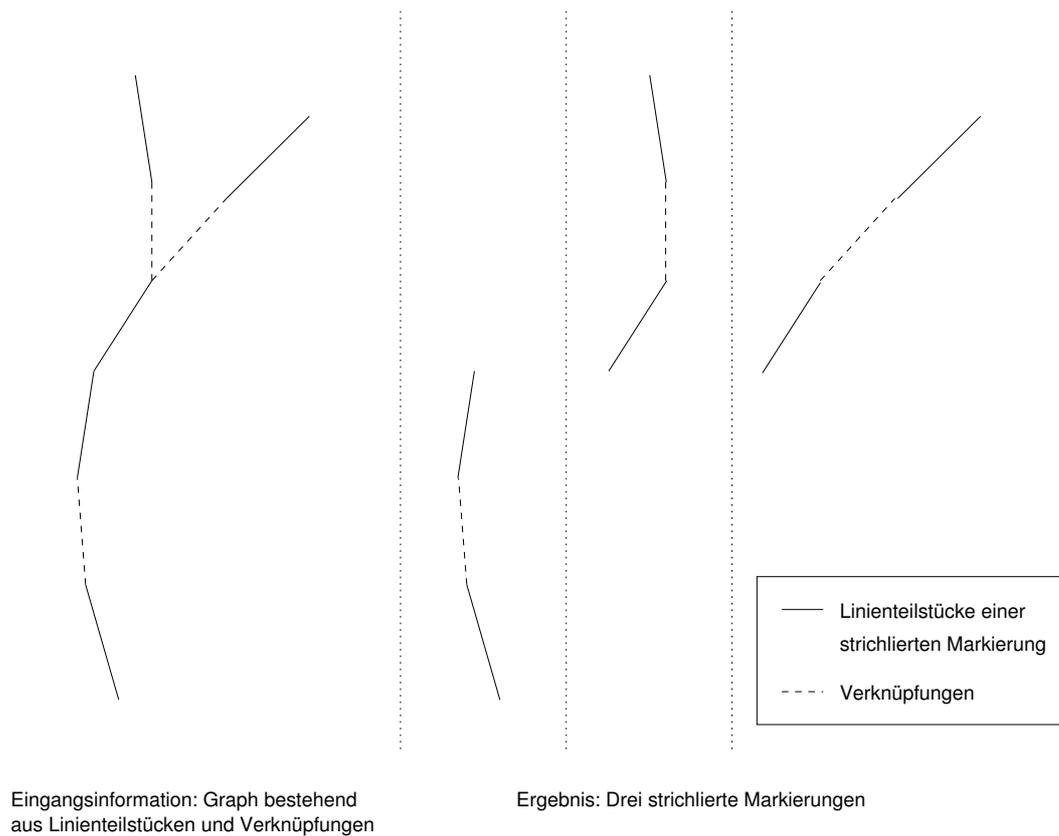
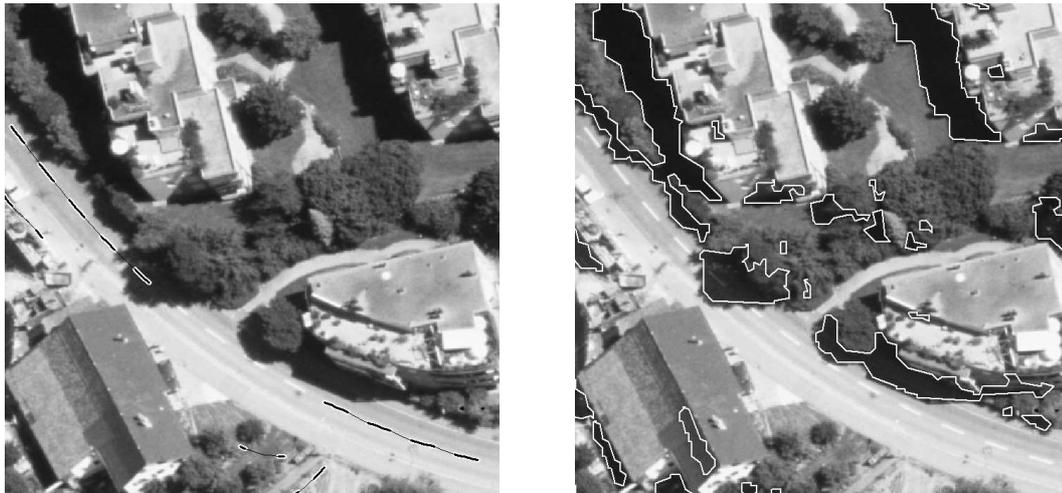


Abbildung 38. Grundprinzip der Extraktion strichlielter Markierungen.

Zur Bewertung von strichlierten Markierungen werden die gleichen Kriterien wie für durchgezogenen Markierungen herangezogen, ergänzt durch das Kriterium, dass die Längen der Linienteilstücke und der Lücken (jeweils) weitgehend konstant sein müssen. Für den Fall, dass eine extrem lange strichlierte Markierung (100 m und mehr) extrahiert wurde, so dass innerhalb der Markierung verschiedene Längen auftreten könnten, wird die Berechnung abschnittsweise durchgeführt.

Die Extraktion von strichlierten Markierungen ist in Abbildung 39 a) für einen detaillierten Ausschnitt dargestellt. Abbildung 39 b) zeigt die in diesem Ausschnitt detektierten Schattenregionen der Kontextanalyse. Wie in der Bildmitte zu sehen ist, konnte auch ein Linienteilstück in der Schattenregion, die über die Straßenmitte hinausreicht, erfolgreich extrahiert werden. Dadurch konnte immerhin die strichlierte Markierung von einer Seite her in die Schattenregion fortgesetzt werden. Abbildung 40 zeigt das Ergebnis der Extraktion durchgezogener und strichlielter Markierungen für eine größere Szene.



(a) Strichlierte Markierungen: Linienteilstücke (fett, weiß umrandet), Verknüpfungen (dünn, schwarz).

(b) Schattenregionen aus Kontextanalyse vgl. dort extrahiertes Linienteilstück in (a)

Abbildung 39. Extraktion strichlierter Markierungen in beschienenen und abgeschatteten Regionen.



Abbildung 40. Ergebnisse für durchgezogene und strichlierte Markierungen.

7.2.3 Straßenränder

Das Objektmodell beschreibt einen Straßenrand als Materialwechsel zwischen Fahrbahnbelag und unmittelbarer Straßenumgebung, der sich im Bild als deutliche Kante kennzeichnet. Folglich geschieht die Bestimmung von Hypothesen für Straßenränder lediglich durch eine Kantenextraktion. Verglichen mit Straßenrändern in der offenen Landschaft geben in Stadtgebieten Kanten-Kontrast und -Geradlinigkeit kaum weitere Evidenz



Abbildung 41. Extraktion von Straßenrändern: Man beachte auch die Fehlextraktion an der Dachkante. Sie kommt zustande, weil die Dachfläche homogen ist und im Zuge der Kontextanalyse nicht die gesamte Hausregion eliminiert werden konnte.

für Straßenränder. Beide finden daher keinen Eingang in die Bewertung (vgl. Tabelle 6). Allerdings impliziert die Modellierung von Straßenrändern, dass sich zumindest auf der straßenzugewandten Seite einer Kante eine homogene und limitiert geneigte Region – die Fahrbahn – anschließt. Deshalb wird für die Bewertung einer Hypothese beiderseits der Kante eine Region mit etwa der Breite einer Fahrspur erzeugt und auf Homogenität und Neigung getestet. In Abbildung 41 a) ist dies schematisiert dargestellt, und Abbildung 41 b) zeigt Ergebnisse der Extraktion.

Kriterien K I	Kriterien K II
Kante im Grauwertgebirge	Große Grauwerthomogenität einer angrenzenden Region
Minimale Kantenlänge	Geringe mittlere Neigung der gleichen angrenzenden Region

Tabelle 6. Kriterien zur Extraktion und Bewertung von Straßenrändern.

7.3 Fahrspuren

Die Bestimmung von vorläufigen Straßenabschnitten und Fahrspurbegrenzungen machte bereits von Wissen aus allen drei Abstraktionsstufen Gebrauch. Allerdings geschah die Extraktion dieser Straßenobjekte ohne gegenseitige Beeinflussung, so dass die Objekte im aktuellen Stadium der Straßenextraktion noch keine hohe Semantik besitzen. Würde z.B. anstelle eines Luftbildes ein gescannter Leitungsplan als Inputbild vorgelegt werden, hätte dies zur Folge, dass eine große Anzahl an fehlinterpretierten durchgezogenen und strichlierten Markierungen extrahiert würde. Erst die Kopplung der verschiedenen Straßenobjekte untereinander erzeugt eine höherwertige Semantik, die eine letztlich korrekte Interpretation von Globalstrukturen wie Straßenabschnitten und Details wie Markierungen erlaubt.

Eine erste Kopplung der bisher extrahierten Straßenobjekte geschieht durch die Extraktion von Fahrspuren. Diese wird in zwei grundlegenden Schritten vollzogen: die *Konstruktion* von Fahrspuren (Abschnitt 7.3.1) und deren *Fusion und Gruppierung* (Abschnitt 7.3.2). Ein weiterer Aspekt der Kopplung ist, dass von nun an nicht nur die autonome Bewertung zur Bemessung der Qualität von Hypothesen herangezogen werden kann, sondern auch die bottom-up und top-down Bewertungen (siehe Abschnitt 4.3.2).

7.3.1 Konstruktion von Fahrspuren

Unabhängig von den weiter unten beschriebenen Konstruktionsmethoden haben alle Fahrspuren die Eigenschaft gemeinsam, weitgehend parallele Randpolygone zu besitzen. Daher ist ein wesentlicher Bestandteil der Konstruktion von Fahrspuren ein Algorithmus, der aus einer Menge an Polygonen jene, die abschnittsweise parallel zueinander verlaufen, extrahiert und ihre Mittelachse bestimmt (siehe Abbildung 42).

Die Bestimmung paralleler Polygone beginnt mit der Überprüfung der individuellen Polygonseiten auf Überlappung und Parallelität bezüglich einer vorgegebenen Winkeltoleranz $\Delta\alpha$. Die Mittellinie zwischen zwei als ausreichend parallel gefundenen Polygonseiten ist durch jenen Abschnitt ihrer Winkelhalbierenden definiert, der den minimalen Überlappungsbereich der Polygonseiten einschließt (Abbildung 42 a). Um anschließend parallel verlaufende Polygonabschnitte leichter konstruieren zu können, werden an Mittellinien und Polygonseiten Informationen über Indizes, Reihenfolge der Polygonseiten und Überlappungsprozensatz als Attribut gespeichert. Die Konstruktion geschieht durch Verkettung von aneinander angrenzenden Seiten $P_{i,n}$, $P_{i,n+1}$, etc. des Polygons P_i mit der Bedingung, dass alle zugehörigen Parallelen *nur* dem Polygon P_j entstammen und zusätzlich die Seitenreihenfolge der Parallelen der Seitenreihenfolge von P_j entspricht (d.h. $P_{j,m}$, $P_{j,m\pm 1}$, $P_{j,m\pm 2}$, ...; siehe Abbildung 42 b). Abbildung 42 c) zeigt einen Fall, in dem diese Bedingung nicht erfüllt ist und daher nur parallele Teilstücke extrahiert werden. Der letzte Schritt des Algorithmus beinhaltet einerseits die Verkettung der Mittellinien und andererseits das Entfernen nicht-überlappender Endstücke der parallelen

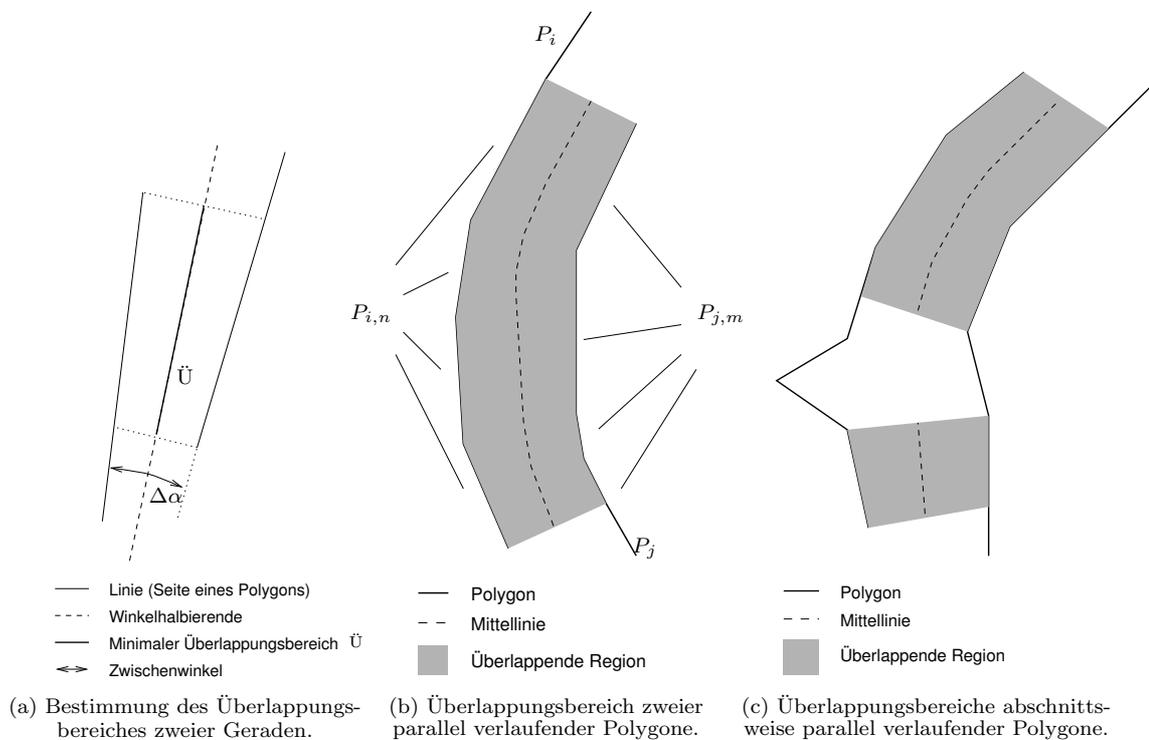


Abbildung 42. Bestimmung parallel verlaufender Polygone.

Polygone, so dass die resultierenden parallelen Polygone exakt mit der überlappenden Region enden (siehe Abbildung 42 a) und b).

Basierend auf diesem Algorithmus werden Fahrspuren in drei verschiedenen Varianten konstruiert: (a) aus parallelen Markierungen und Straßenrändern, (b) ausschließlich aus Markierungen und (c) aus vorläufigen Straßenabschnitten. Zusätzlich zur rein geometrischen Repräsentation durch Mittelpolygon und parallele Randpolygone erhalten die resultierenden Fahrspuren noch Attribute, die Auskunft über ihre Art der Konstruktion (a — c), ihren Kontext und ihre Bewertung geben.

a) Fahrspuren aus parallelen Markierungen und Straßenrändern:

Zunächst werden Markierungen und Straßenränder in eine einheitliche Auflösung transformiert, in Polygone konvertiert und parallele Polygonabschnitte extrahiert. Da insbesondere Straßenränder keinen geradlinigen Verlauf haben müssen, werden keine großen Ansprüche an die Winkeltoleranz zwischen parallelen Polygonseiten gestellt. Minimale und maximale Distanz zwischen parallelen Seiten dürfen aber die vom Modell vorgegebenen Breiten für Fahrspuren nicht über- oder unterschreiten (vgl. Kriterien K I in Tabelle 7). Während sich die Extraktion von Fahrspuren auf strukturelle Kriterien stützt, dienen die flächenhaften Kriterien des eingeschlossenen Fahrstreifens zu deren autonomer Bewertung. Der Fahrstreifen muss sich im DOM durch geringe Neigung und im hochaufgelösten Bild durch große Homogenität zumindest in Fahrtrichtung auszeichnen. Die Wirkungsweise der Bewertung verdeutlichen die extrahierten Fahrspuren in Abbildung 43. Je heller die Mittelachse einer Fahrspur in Abbildung 43 b) eingezeichnet ist, um besser ist ihre Bewertung.

Kriterien K I	Kriterien K II
Parallele Polygone (ggf. abschnittsweise)	Große Grauwerthomogenität in Fahrspurrichtung
Minimale und maximale Breite	Geringe Längsneigung
Minimale Länge	Sehr geringe Querneigung
	Geringe mittlere Krümmung
	Geringe Breitenvarianz

Tabelle 7. Kriterien zur Extraktion und Bewertung von Fahrspuren.

b) Fahrspuren ausschließlich aus Markierungen:

Die Konstruktion von Fahrspuren aus Markierungen beginnt mit der Bestimmung paralleler Markierungen unabhängig davon, ob es sich um strichlierte oder durchgezogene Markierungen handelt. Methodisch geschieht die Konstruktion in analoger Weise wie unter (a), lediglich die Winkeltoleranz zwischen parallelen Polygonseiten kann nun wesentlich stärker eingeschränkt werden, da Markierungen relativ gerade verlaufen müssen. Zur internen Bewertung kommen ebenfalls wie unter (a) die Kriterien aus Tabelle 7 zum Einsatz. Die Ergebnisse in Abbildung 44 zeigen einige auf diese Weise extrahierte Fahrspuren.

Wie im Straßenmodell definiert, dürfen strichlierte Markierungen nicht als Fahrbahnrand auftreten (Relation „strichlierte Markierung mittig“ des Konzeptes „Fahrspurbegrenzung“ in Abbildung 13). Deshalb wird die im momentanen Stadium der Extraktion erzeugte Menge an Fahrspuren (inkl. derer aus a) nun nach Fahrspurabschnitten analysiert, die mindestens auf einer Seite durch eine strichlierte Markierung begrenzt sind. Für diese



(a) Extrahierte Fahrspuren



(b) Autonome Bewertung:
Je heller die Fahrspurachse, desto besser die Bewertung

Abbildung 43. Extraktion und Bewertung von Fahrspuren aus Markierungen und Straßenrändern.



Abbildung 44. Extraktion von Fahrspuren aus parallelen Markierungen.

Abschnitte wird die Annahme getroffen, dass eine weitere Fahrspur vorhanden sein sollte, deren zweite Begrenzung weder in Form einer durchgezogenen Markierung noch in Form eines Straßenrandes extrahiert werden konnte. Als Konsequenz dieser Annahme werden für diese Abschnitte hypothetische Fahrspurbegrenzungen im Abstand der durchschnittlichen Fahrspurbreite erzeugt. Aus diesen werden weitere Fahrspuren konstruiert und schließlich nach obigen Kriterien bewertet. Abbildung 45 zeigt einige Beispiele für diese Fälle.

c) Fahrspuren aus vorläufigen Straßenabschnitten:

Zur abschließenden und zugleich unsichersten Variante der Konstruktion von Fahrspuren kommen die vorläufigen Straßenabschnitte aus Abschnitt 7.1 zum Einsatz. Ihre Verwendung an dieser Stelle rührt daher, dass nicht alle Straßen in Stadtgebieten mit Markierungen versehen sind oder die Farbe der Markierungen schon so abgetragen ist, dass eine zuverlässige Extraktion schwer möglich ist. Um dennoch Hypothesen für solche Teile des Straßennetzes generieren zu können, werden Anzahl und Verlauf der Fahrspuren aus den vorläufigen Straßenabschnitten geschätzt. Die Bestimmung dieser Fahrspuren erfolgt durch den Fahrbahnrand und durch die Erzeugung von hypothetischen Fahrspurbegrenzungen (Polygone parallel zur Achse des betreffenden Straßenabschnitts).

Wie aus Abbildung 46 zu sehen ist, können hierdurch Hypothesen für eine Reihe von bislang nicht extrahierten Teilen des Straßennetzes gewonnen werden. Allerdings ist die Fehlerrate auch erwartungsgemäß hoch. Neben einer entsprechenden Attributierung dieser Fahrspuren als „Variante (c)“ werden daher zur Bewertung der geometrischen Kriterien strengere Parametereinstellungen als bei den Varianten (a) und (b) herangezogen. Dies ist dadurch begründet, dass für die mit Variante (c) erzeugten Fahrspuren keine strukturelle Evidenz aus der hohen Auflösung vorliegt und dies mit der strengeren Parametereinstellung kompensiert werden soll. Weiterhin fällt in Abbildung 46 auf, dass etliche Fahrspuren auch schon durch die Varianten (a) und (b) extrahiert wurden.



Abbildung 45. Extraktion von Fahrspuren aus Markierungen mit hypothetischen Fahrspurbegrenzungen.

Die Behandlung solcher möglicherweise konkurrierender oder sich gegenseitig unterstützender Fahrspuren wird im nächsten Abschnitt erläutert.

7.3.2 Fusion und Gruppierung von Fahrspuren

Mit den Fahrspuren liegen nun semantisch höherwertige Objekte vor. Neben der rein geometrischen Repräsentation über ihre Mittelachse und Randpolygone verfügen sie über eine Reihe von Attributen, die Auskunft über zugrunde liegende Auflösung, Bewertungen, Kontexte und Konstruktionsvarianten geben und außerdem Verweise auf die zur Konstruktion verwendeten Straßenbegrenzungen beinhalten. Weiterhin können die Fahrspuren in Raster-Regionen gewandelt werden, so dass Zugriff auf die korrespondierende Bildregion bzw. nach entsprechender Transformation auf die Höhenwerte des DOM möglich ist. Diese reichhaltigen Informationen unterstützen die folgende Fusion und Gruppierung der Fahrspuren in hohem Maße.

Fusion von Fahrspuren:

Ziel der Fusion von Fahrspuren ist die Generierung eines einheitlichen Datensatzes ohne redundante Fahrspurabschnitte. In diesem Zusammenhang gilt ein Abschnitt als „redundant“, wenn er sich innerhalb eines Puffers der Breite b mit einem anderen Fahrspurabschnitt überlappt und die Winkeldifferenz der Fahrspurmittellinien im Überlappungsbereich eine gewisse Winkeltoleranz $\Delta\alpha$ nicht überschreitet (siehe Abbildung 47).

Der Algorithmus zur Fusion von Fahrspuren basiert auf jenem zur Bestimmung von parallelen Polygonabschnitten, erweitert ihn aber um einige Komponenten. Zunächst werden zwei zu fusionierende Fahrspuren F_i



Abbildung 46. Extraktion von Fahrspuren auf Basis vorläufiger Straßenabschnitte.

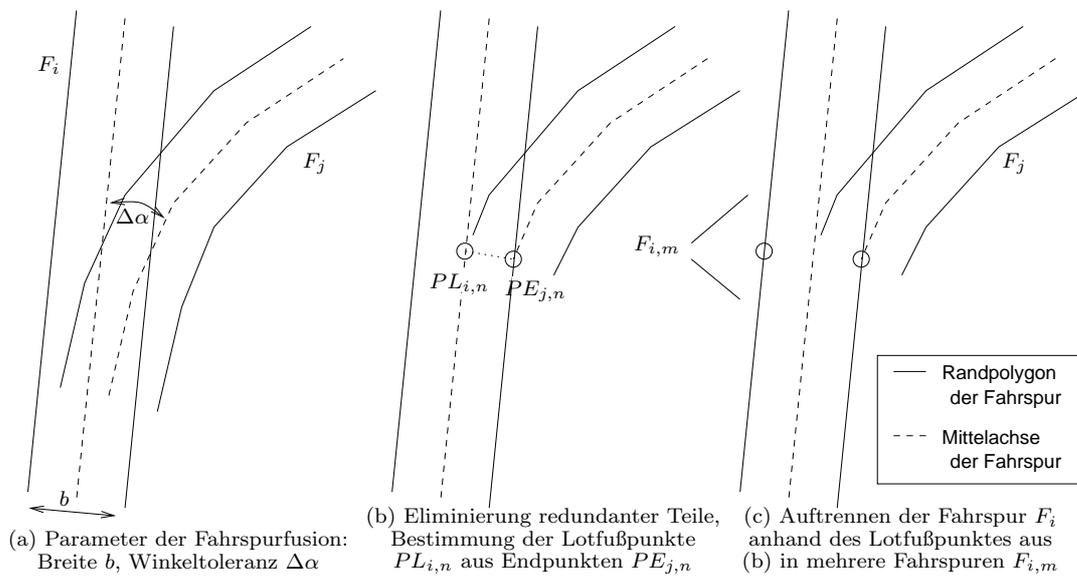


Abbildung 47. Fusion von Fahrspuren.

und F_j ausgewählt, ihre Achsen in Polygone konvertiert und unter Vorgabe der Parameter $\Delta\alpha$ und b (üblicherweise die mittlere Breite von F_i) parallele Polygonabschnitte detektiert. Diese Abschnitte erhalten bei F_i das Attribut „überlappt“ und bei F_j das Attribut „redundant“. Ausgehend von den Endpunkten $PE_{j,n}$ der nicht-redundanten Abschnitte von F_j ermöglicht der Algorithmus (optional), die spätere Generierung von Kreuzungspunkten vorzubereiten. Hierzu wird die Fahrspur F_i an den Fußpunkten $PL_{i,n}$ der Lote von $PE_{j,n}$ auf F_i getrennt, so dass aus F_i zwei oder mehr Fahrspuren $F_{i,m}$ entstehen (siehe Abbildung 47). Abschließend

werden aus den nicht-redundanten Abschnitten von F_j neue Fahrspuren erzeugt und zusammen mit den Fahrspuren $F_{i,m}$ als Ergebnis der Fusion zurückgeliefert. Redundant gekennzeichnete Fahrspurabschnitte kommen nicht in die Ergebnismenge.

Da bei der Fusionierung jedoch nicht nur zwei, sondern eine größere Menge an Fahrspuren bearbeitet wird, muss ein iteratives Schema zum Einsatz kommen: Ausgangspunkt ist eine zunächst leere Ergebnismenge und die alle Fahrspuren umfassende Ausgangsmenge. Die erste Fahrspur wird unverändert in die Ergebnismenge kopiert. Anschließend wird eine beliebige Fahrspur aus der Ausgangsmenge herausgenommen, auf Redundanz mit der Fahrspur in der Ergebnismenge geprüft und die nicht-redundanten Abschnitte der Ergebnismenge hinzugefügt. Es folgt die gleiche Operation mit der nächsten Fahrspur der Ausgangsmenge, nur mit dem Unterschied, dass diese Fahrspur nun mit *allen* Fahrspuren der Ergebnismenge verglichen wird. Dieser Prozess wird so lange wiederholt, bis alle Fahrspuren der Ausgangsmenge bearbeitet wurden. Da durch das Entfernen redundanter Abschnitte die Fahrspuren unter Umständen stark fragmentiert werden, übt die Reihenfolge, nach der die paarweise Überprüfung auf Redundanz erfolgt, großen Einfluss auf die Topologie der fusionierten Fahrspuren und dadurch auch auf die nachfolgende Gruppierung aus. Denn je geringer die Fragmentierung der Fahrspuren aufgrund der Fusion ist, um so weniger aufwändig gestaltet sich die Gruppierung der Fahrspuren in längere Einheiten.

Von großem Nutzen zur Erstellung einer günstigen Bearbeitungsreihenfolge ist die interne Bewertung. Sie ermöglicht eine Sortierung der Fahrspuren, so dass beim paarweisen Vergleich der Fahrspuren immer die besser bewertete unverändert bleibt. Dies ist hier auch unter dem Gesichtspunkt zu sehen, dass überlappende Fahrspuren je nach Konstruktionsvariante nicht nur unterstützende, sondern auch konkurrierende Hypothesen sein können. Die Bewertung ermöglicht es daher, auf die augenblicklich erfolgversprechendere Hypothese zu fokussieren. Ein weiterer Aspekt bei der Erstellung der Bearbeitungsreihenfolge ist die Behandlung der Kontextobjekte „Fahrzeuge“ und jener Fahrspuren, die aus niedrigeren Auflösungsstufen extrahiert wurden (Variante (c) in Abschnitt 7.3.1). Da zur Bestimmung redundanter Fahrspurabschnitte ausschließlich geometrische Kriterien herangezogen werden, wird den Fahrspuren aus der hohen Auflösung eine höhere Priorität zuerkannt. Anschließend folgen die ebenfalls in der hohen Auflösung extrahierten Fahrzeuge. Ihre geringere Priorität wird mit der Tatsache begründet, dass die Fahrzeugposition und -orientierung auf der Fahrspur nicht völlig festgelegt ist und daher das umschreibende Rechteck eines Fahrzeugs (als Ersatz für eine Fahrspur) nur näherungsweise mit der tatsächlichen Fahrspur übereinstimmt. Zuletzt werden jene Fahrspuren hinzugefügt, die auf Basis der vorläufigen Straßenabschnitte erstellt wurden. Ihre Geometrie ist am schlechtesten definiert, so dass sie für die Bearbeitungsreihenfolge die geringste Priorität besitzen.

Gruppierung von Fahrspuren:

An die Fusion der Fahrspuren schließt sich deren Gruppierung an. Sie ist aus zweierlei Gründen wichtig: Zum einen kann bei der Fusion nicht gewährleistet werden, dass die resultierenden Fahrspuren „bündig“ aneinander anschließen. Eine solche Situation entsteht z.B., wenn eine Fahrspur aus einem Straßenrand mit paralleler durchgezogener Markierungen konstruiert wurde und ihre Fortsetzung aus einem Straßenrand mit paralleler strichliertem Markierung. Die Bedingung, dass eine strichlierte Markierung immer mit einem Linienteilstück beginnen und enden muss, hat zur Folge, dass am Übergang von der strichlierten Markierung zur durchgezogenen Markierung zwangsläufig eine Lücke entsteht, die durch einen Gruppierungsschritt zu überbrücken ist. Ein ähnlicher Fall tritt z.B. auch auf, wenn Fahrspuren aus beschienenen und abgeschatteten Bereichen fusioniert wurden. Neben dieser algorithmischen Motivation kommt als weiterer Grund hinzu, dass bei der Konstruktion der Fahrspuren mit Beteiligung von Markierungen (Varianten (a) und (b)) hauptsächlich die detaillierte Modellierung der hohen Auflösung genutzt wurde. Da in dieser Auflösung auch kleinere Störobjekte die Extraktion stark beeinflussen, muss mit einem fragmentierten Ergebnis der Fahrspurextraktion gerechnet werden.

Als Konsequenz daraus bietet sich eine Gruppierung an, die nicht nur geometrische Kriterien wie z.B. die Kontinuität einer Fahrspur nutzt, sondern auch Information aus den niedrigeren Auflösungen, wo Störobjekte besser unterdrückt werden und die Topologie des Straßennetzes deutlicher zum Ausdruck kommt, miteinbindet. Als Basis der Gruppierung dient eine modifizierte Version des graphbasierten, bereits zur Extraktion von strichlierten Markierungen verwendeten Gruppierungsalgorithmus (siehe Abschnitt 7.2.2). Die Modifikation bezieht sich darauf, dass die Bestimmung von Verknüpfungen zwischen den Fahrspuren in drei Teilschritten geschieht:

- ◊ Im ersten Teilschritt der Gruppierung werden nur dann zwei Fahrspuren miteinander verbunden, wenn nicht nur geometrische Kriterien für eine Verknüpfung sprechen, sondern die beiden Fahrspuren auch in den niedrigeren Auflösungen durch eine Fahrbahn bzw. einen Straßenabschnitt verbunden sind. Die Überprüfung dieses Kriteriums geschieht mit Hilfe der bereits extrahierten vorläufigen Straßenabschnitte, indem getestet wird, ob die einzufügende Verknüpfung in den Bereich eines vorläufigen Straßenabschnittes fällt und nur geringe Richtungsabweichungen zwischen den Achsen der Verknüpfung und des vorläufigen Straßenabschnittes auftreten (Abbildung 48).
- ◊ Können diese Kriterien nicht erfüllt werden, wird im zweiten Teilschritt der Fahrspurgruppierung untersucht, ob die Verknüpfung in der mittleren Auflösung durch große Grauwerthomogenität gekennzeichnet ist und damit zumindest Evidenz für den Fahrbahnbelag gegeben ist (Abbildung 49).
- ◊ Im letzten Teilschritt der Gruppierung können noch Fahrspuren miteinander verknüpft werden, die von *verschiedenen* lokalen Kontexten beeinflusst sind (beschienene Bereiche¹⁶ / Schattenbereiche / Fahrzeuge, Abbildung 50). In diesen Fällen können keine radiometrischen Kriterien für die Bewertung der Fahrspurverknüpfung herangezogen werden. Da die lokalen Kontexte ebenfalls nah aneinander angrenzen, darf die zu überbrückende Distanz zwischen zwei Fahrspuren in solchen Fällen nur sehr gering sein (z.B. < 2m). Eine Beschränkung auf rein geometrische Kriterien ist hier also unproblematisch.

Abbildungen 48, 49 und 50 illustrieren Beispiele für die drei Teilschritte der Gruppierung. Oben ist jeweils die Eingangsinformation dargestellt und unten das Ergebnis der Gruppierung.

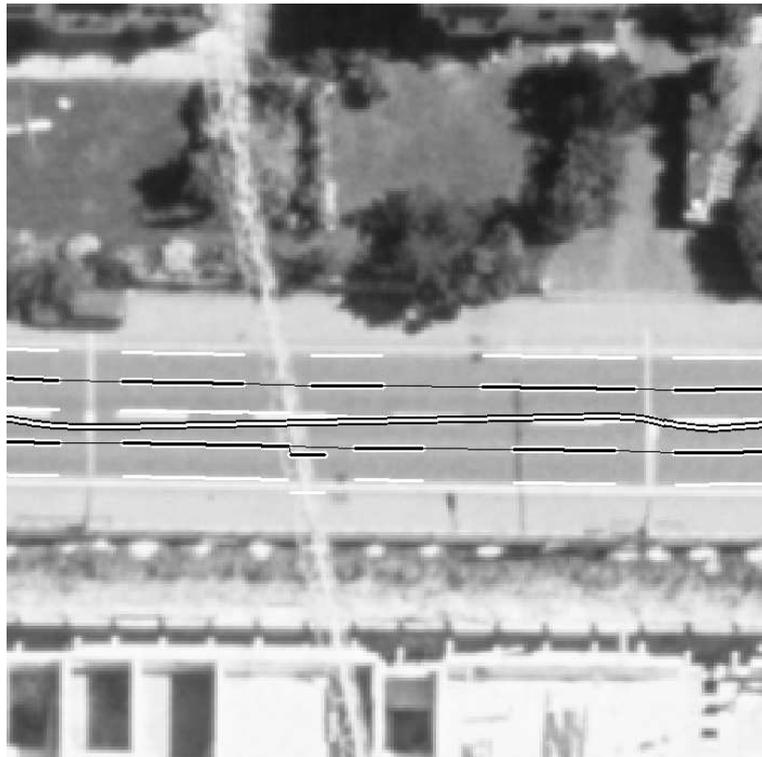
Fahrspuren werden nach der Gruppierung in analoger Weise wie bei ihrer Konstruktion durch Mittelpolygone und parallele Randpolygone repräsentiert. Ein wesentlicher Unterschied ist nun aber die zusätzliche Anzahl an Attributen, die während der Fusion und Gruppierung angefügt wurden. Die Attribute stellen Verweise auf die Fahrspuren vor der Gruppierung her, so dass z.B. Informationen über die Konstruktion und Bewertung der Fahrspuren auch in späteren Schritten der Extraktion zugänglich bleiben.

7.4 Fahrbahnen

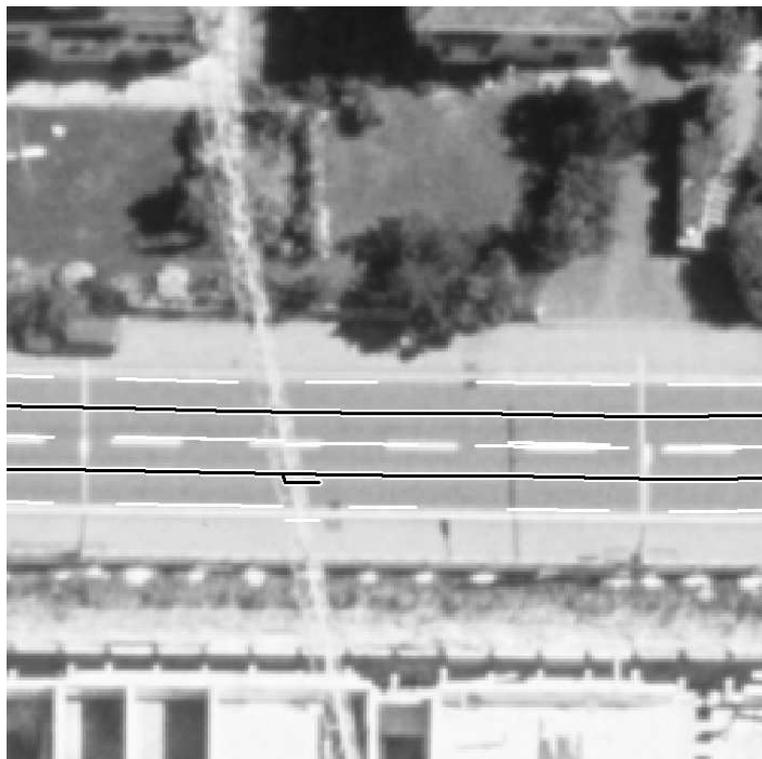
Mit der Konstruktion und Gruppierung von Fahrspuren wurde das in der niedrigsten Abstraktionsstufe modellierte Wissen vollständig genutzt. Vom Objektwissen der mittleren und hohen Abstraktionsstufe wurde dagegen nur in geringem Umfang Gebrauch gemacht. Der nächste Schritt zur Aggregation semantisch höherwertiger Information ist nun die Extraktion von Fahrbahnen, die auf der mittleren Abstraktionsstufe modelliert sind (vgl. Abbildung 13).

Fahrbahnen sind im Bild und DOM mittlerer Auflösung als homogenes Band limitierter Neigung beschrieben, dessen eingeschlossene Region dem Fahrbahnbelag und dessen seitliche Kanten dem Fahrbahnrand entsprechen. Die Koppelung dieser Beschreibung mit der Information aus der hohen Auflösung geschieht über die im vorangegangenen Abschnitt extrahierten Fahrspuren, da parallel und kollinear gruppierte Fahrspuren Bestandteile einer Fahrbahn sind. Eine der vorrangigen Aufgaben bei der Generierung von Fahrbahnen ist also eine geeignete Integration von Informationen, die in der hohen und mittleren Auflösung extrahiert wurden.

¹⁶im Modell kein lokaler Kontext, sondern der „Normalfall“



(a) Fahrspuren: Achse (schwarz), Begrenzung (weiß), Verknüpfung (schwarz, dünn).
Achse der vorläufigen Straßenabschnitte (dicke Linie), vgl. auch Abb. 35).



(b) Ergebnis nach Gruppierung.

Abbildung 48. Gruppierung von Fahrspuren: Verifikation durch vorläufige Straßenabschnitte in niedriger Auflösung.



(a) Fahrspuren: Achse (schwarz), Begrenzung (weiß), Verknüpfungen (schwarz, dünn).



(b) Ergebnis nach Gruppierung.

Abbildung 49. Gruppierung von Fahrspuren: Verifikation durch homogenen Straßenbelag.



(a) Fahrspuren: Achse (schwarz), Begrenzung (weiß), Fahrzeuge (weiße Rechtecke), Verknüpfung (schwarz, fett).



(b) Ergebnis nach Gruppierung.

Abbildung 50. Gruppierung von Fahrspuren: Verifikation durch lokalen Kontext (hier Fahrzeug).

Wie schon zuvor im Falle der Fahrspuren ist die Bestimmung von Fahrbahnen in zwei Phasen unterteilt: die *Konstruktion* von Fahrbahnen (Abschnitt 7.4.1) und die *Gruppierung* von Fahrbahnen (Abschnitt 7.4.2).

7.4.1 Konstruktion von Fahrbahnen

Die wichtigste Grundlage zur Konstruktion von Fahrbahnen sind parallele Fahrspuren. Sie geben einen eindeutigen Hinweis auf die Präsenz einer mehrspurigen Straße. Da bei der Fahrspurextraktion in den vorangegangenen Abschnitten lokale Straßeneigenschaften im Vordergrund standen, kann nicht davon ausgegangen werden, dass bereits in diesem Stadium der Prozessierung über längere Strecken des Straßennetzes parallele Fahrspuren vorliegen. Daher ist neben der Detektion von – unter Umständen mehrfach – parallelen Fahrspuren ein weiteres Ziel bei der Konstruktion von Fahrbahnen, dass fehlende Fahrspurabschnitte detektiert und verifiziert werden, um damit die Fragmentierung einer Fahrbahn zu reduzieren oder ganz zu eliminieren. Bei der Bewertung und Verifikation von fehlenden Fahrspurabschnitten kommen vor allem die Fahrbahneigenschaften der mittleren Abstraktionsstufe zum Tragen.

Die Konstruktion von Fahrbahnen ist in drei Teilschritte gegliedert. Im ersten Teilschritt (a) werden Parallelen-Beziehungen zwischen Fahrspuren genutzt, um so genannte „Fahrspurcluster“ aufzubauen. Diese Cluster werden dann im zweiten Teilschritt (b) sukzessive vereinfacht, indem auf Basis vorläufiger „Teil-Fahrbahnen“ fehlende Fahrspurabschnitte detektiert und im Falle einer erfolgreichen Verifikation eingefügt werden. Der letzte Teilschritt (c) umfasst schließlich die Bestimmung der endgültigen Fahrbahnen aus den vereinfachten Fahrspurclustern.

a) Generierung von Fahrspurclustern:

Die Generierung von Fahrspurcluster startet mit der Detektion von abschnittsweise parallelen Fahrspuren. Dies sind Fahrspurabschnitte, deren gegenseitig zugewandte Randpolygone geringen Abstand haben und die innerhalb einer vorgegebenen Winkeltoleranz parallel zueinander verlaufen. Für die Detektion solcher Abschnitte kommt wiederum der Algorithmus zur Bestimmung paralleler Polygone zum Einsatz. Eine entsprechende Attributierung der Randpolygone ermöglicht die Speicherung der Information, welche Abschnitte einer Fahrspur parallel zu einer anderen Fahrspur verlaufen.

Mit Hilfe dieser Information wird anschließend ein Graph aufgebaut, dessen Knoten die Fahrspuren und dessen Kanten die Parallelen-Relationen zwischen den Fahrspuren darstellen (siehe Abbildung 51 a). Der Vorteil dieser Graphrepräsentation liegt darin begründet, dass jedes Fahrspurcluster genau einer Zusammenhangskomponente im Graph entspricht und die Bestimmung solcher Subgraphen ohne großen Rechenaufwand möglich ist. Die Topologie rechtsseitiger und linksseitiger Parallelen-Beziehungen einer Fahrspur wird dadurch repräsentiert, dass pro Knoten die Kanten zwei unterschiedlichen Mengen für das rechte bzw. das linke Randpolygon zugewiesen werden. Um auch die Topologie der Parallelen-Beziehungen *entlang* eines Randpolygons zu erfassen, wird jede Kantenmenge entsprechend der Fahrspur-„Laufriichtung“ sortiert. Abbildung 51 zeigt schematisch diese Repräsentationsform.

b) Vereinfachung der Fahrspurcluster auf Basis von Teil-Fahrbahnen:

Das Ziel der Vereinfachung von Fahrspurclustern ist die iterative Detektion und Verifikation von Verknüpfungen zwischen Fahrspurpaaren, die die gleiche parallele Fahrspur besitzen. Im Gegensatz zur Fahrspurgruppierung in Abschnitt 7.3.2 wird jetzt Evidenz aufgrund paralleler Fahrspuren genutzt, um fehlende Abschnitte zu ex-

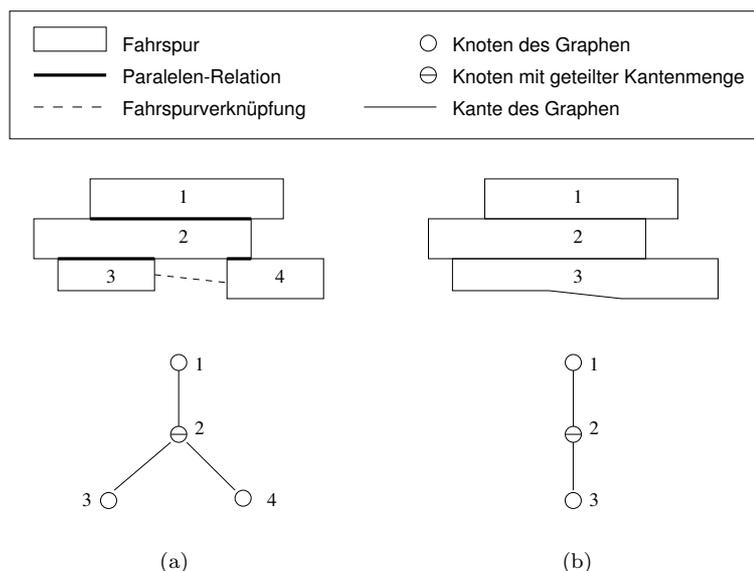


Abbildung 51. Fahrspurcluster (jeweils oben) und Graph-Repräsentation (jeweils unten): (a) Generierung einer Fahrspurverknüpfung, (b) Resultat nach erfolgreicher Verifikation der Verknüpfung.

trahieren. Große Unterstützung leistet hierfür die Graphrepräsentation der Fahrspurcluster, weil sich fehlende Fahrspurabschnitte darin ohne Suchaufwand detektieren lassen. Jede Fahrspur ist lediglich daraufhin zu überprüfen, ob in der rechten oder linken Kantenmenge ihres korrespondierenden Knotens mehr als eine Kante vorhanden ist. Da die Fahrspuren, die durch die Zielknoten der Kanten repräsentiert werden, parallel zur aktuellen Fahrspur verlaufen, kann auch die Verknüpfung als Parallele der aktuellen Fahrspur erzeugt werden. Weiterhin ist durch die Sortierung der Kanten sichergestellt, dass Verknüpfungen immer zwischen Fahrspuren erzeugt werden, die bezüglich der Laufrichtung der aktuellen Fahrspur direkt aufeinander folgen. Im Falle einer erfolgreichen Verifikation können die beiden Fahrspuren verbunden werden und dadurch die beiden korrespondierenden Graph-Kanten durch eine einzige ersetzt werden. Diese Vereinfachung der Fahrspurcluster iteriert solange, bis keine neuen Verknüpfungen mehr generiert oder verifiziert werden können. Für den in Abbildung 51 b) gezeigten Idealfall ist der resultierende Subgraph ein Pfad, der die beiden äußeren Fahrspuren einer Fahrbahn miteinander verbindet.

Infolge dessen, dass bisher überwiegend lokale Eigenschaften zur Extraktion von Fahrspuren herangezogen wurden und deshalb auch mit einer gewissen Anzahl an Fehlextraktionen zu rechnen ist, wäre eine Verknüpfung zweier Fahrspuren *allein* aufgrund der Evidenz durch eine parallele Fahrspur relativ unzuverlässig. Deshalb geschieht die Generierung und Bewertung der Verknüpfungen auf Basis von so genannten „Teil-Fahrbahnen“. Eine Teil-Fahrbahn besteht aus den zwei zu verknüpfenden Fahrspuren und der korrespondierenden parallelen Fahrspur, z.B. die Fahrspuren 2, 3 und 4 sowie die Fahrspurverknüpfung zwischen 3 und 4 in Abbildung 51 a). Da im Straßenmodell kein Unterschied zwischen Fahrbahnen mit zwei oder mit mehr Fahrspuren gemacht wird, kann eine Teil-Fahrbahn nach den gleichen Kriterien extrahiert und bewertet werden wie später die endgültigen Fahrbahnen. Wie auch Tabelle 8 zu entnehmen ist, muss zur Extraktion neben dem Parallelen-Kriterium noch die Bedingung erfüllt sein, dass alle drei Fahrspuren einer Teil-Fahrbahn ähnliche mittlere Helligkeit besitzen¹⁷.

Für die Entscheidung, ob eine Fahrspurverknüpfung akzeptiert werden kann, stehen drei Alternativen zur Verfügung, die sequenziell bearbeitet werden:

- (1) Verifikation aufgrund der autonomen Bewertung:

¹⁷ Als Ersatz für die nicht vorhandenen Fahrspurbegrenzungen werden in diesem Fall Korbbögen zwischen den zu verknüpfenden Fahrspuren generiert.

Kriterien K I	Kriterien K II
Parallele Fahrspuren	Große Länge
Ähnliche mittlere Helligkeit der Fahrspuren	Niedrige Fragmentierung der Fahrspuren
	Ähnliche Fahrspurbreiten
	Geringe Längsneigung
	Sehr geringe Querneigung

Tabelle 8. Kriterien zur Extraktion und Bewertung von Fahrbahnen.

Hierzu wird die Verknüpfung wie im Falle der Fahrspurkonstruktion und -gruppierung mit den Kriterien K II aus Tabelle 7 autonom evaluiert. Ist die Überprüfung der Krümmungs-, Neigungs-, Breiten- und Grauwertkriterien erfolgreich, gilt die Verknüpfung als verifiziert und alle weiteren Alternativen werden übersprungen.

(2) Verifikation durch Fahrzeuge:

Mit der nächsten Alternative wird versucht, die Lücke durch die Detektion eines Fahrzeuges zu erklären. Da aufgrund der parallelen Fahrspur nun schon große Evidenz für eine Straße vorliegt, erfolgt die Fahrzeugdetektion jetzt nicht mehr mit dem detaillierten, parametrischen Modell aus der Kontextanalyse, sondern mit einem weniger stringenten, auf Rechteckstrukturen basierenden Modell aus (Hinz und Baumgartner, 2001). Abbildung 52 zeigt einen Fall, in dem das rechte Fahrzeug während der Kontextanalyse detektiert wurde und somit keine Unterbrechung der Fahrspur auftritt. Das linke Fahrzeug konnte wegen der inhomogenen Dachregion zunächst nicht detektiert werden. Die erste Alternative zur Verifikation der Fahrspurverknüpfung schlug aufgrund der großen Inhomogenität innerhalb der Verknüpfung fehl, wohingegen die Fahrzeugdetektion mit dem schwächeren Modell erfolgreich war.

(3) Verifikation aufgrund der top-down Bewertung:

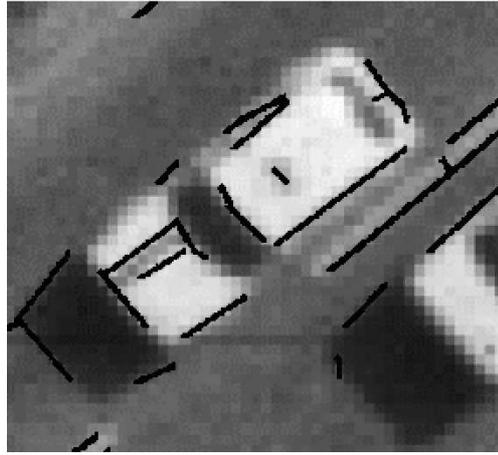
Abbildung 53 zeigt eine Situation, in der die obere der beiden Fahrspuren über eine längere Distanz extrahiert werden konnte. Die untere Fahrspur weist jedoch aufgrund der Inhomogenitäten im Straßenbelag Lücken auf, die im Zuge der Fahrspurgruppierung nicht überbrückt werden konnten. Dementsprechend besitzt die Verknüpfung auch eine schlechte autonome Bewertung. Deutet aber die Bewertung der drei angrenzenden Fahrspuren durch einen hohen Fuzzy-Wert (im Mittel z.B. > 0.5) auf die Präsenz einer Straße hin, darf auch im Falle einer schlechten autonomen Bewertung (Fuzzy-Wert z.B. < 0.1) die top-down Bewertung ausschlaggebend für eine erfolgreiche Verifikation sein. Als Voraussetzung hierfür muss die Fahrbahn zum einen durch große Länge, ähnliche Fahrspurbreiten und geringe Neigung gekennzeichnet sein (siehe Tabelle 8); zum anderen sollte die Fragmentierung durch das Einfügen der Verknüpfung erheblich reduziert werden. Die Fragmentierung errechnet sich aus der Gesamtlänge der parallel verlaufenden Fahrspurabschnitte im Verhältnis zur Länge der Fahrspurverknüpfung. Das bedeutet, je länger Fahrspuren parallel zueinander verlaufen oder je kürzer eine Verknüpfung ist, um so größer ist der Wert für die Fragmentierung, die eine Lücke verursacht, und um so höhere Evidenz liegt vor, die Lücke zu schließen¹⁸. Fällt die top-down Bewertung wie in Abbildung 53 positiv aus, wird die Verknüpfung in jedem Fall als Fahrspur eingefügt; d.h. die autonome Bewertung einer Fahrspurverknüpfung kann „überstimmt“ werden, wenn einerseits große Unterstützung durch Detailinformation in Form von Fahrspuren vorhanden ist, und wenn gleichzeitig genügend Evidenz vorliegt, dass die einzufügende Fahrspur aus globaler Sicht korrekt sein muss.

Abschließend wird aus der Verknüpfung und den beiden angrenzenden Fahrspuren eine neue Fahrspur erzeugt,

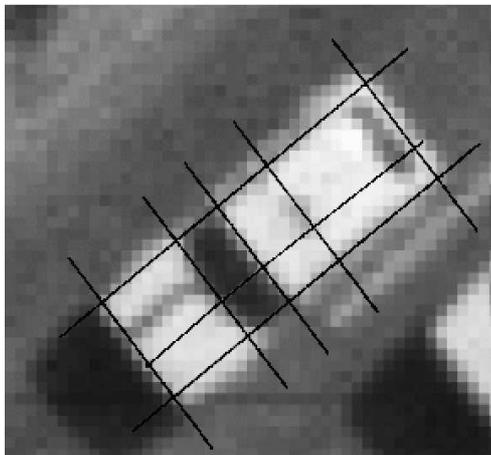
¹⁸Die Fragmentierung hat somit ähnliche Bedeutung wie der so genannte „Umwegfaktor“ in (Wiedemann, 2001) – wenn auch in einem viel kleinräumigeren Bereich.



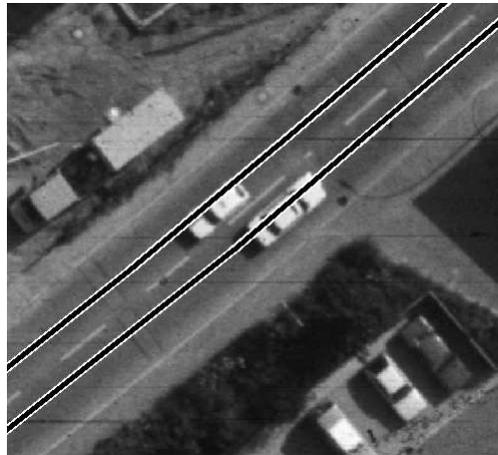
(a) Extrahierte Fahrspuren (schwarz fett, weiß umrandet), Verknüpfung (schwarz, dünn)



(b) Extrahierte Bildkanten in Region der Verknüpfung



(c) Gruppierung der Bildkanten aus (b) zu Rechteckstrukturen



(d) Verifizierte Verknüpfung.

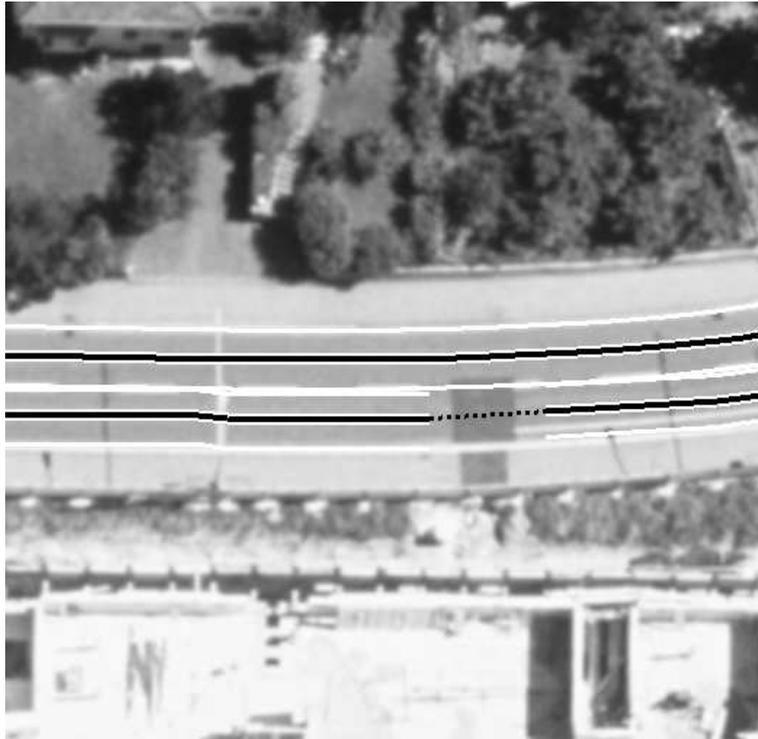
Abbildung 52. Fahrzeugdetektion mittels Gruppierung von Bildkanten in Rechteckstrukturen (Details in (Hinz und Baumgartner, 2001)): Hier eingesetzt zur Verifikation von Fahrspurverknüpfungen.

das Fahrspurcluster bzw. der Subgraph wie in Abbildung 51 skizziert aktualisiert und zur nächsten Iteration der Vereinfachung des Clusters übergegangen.

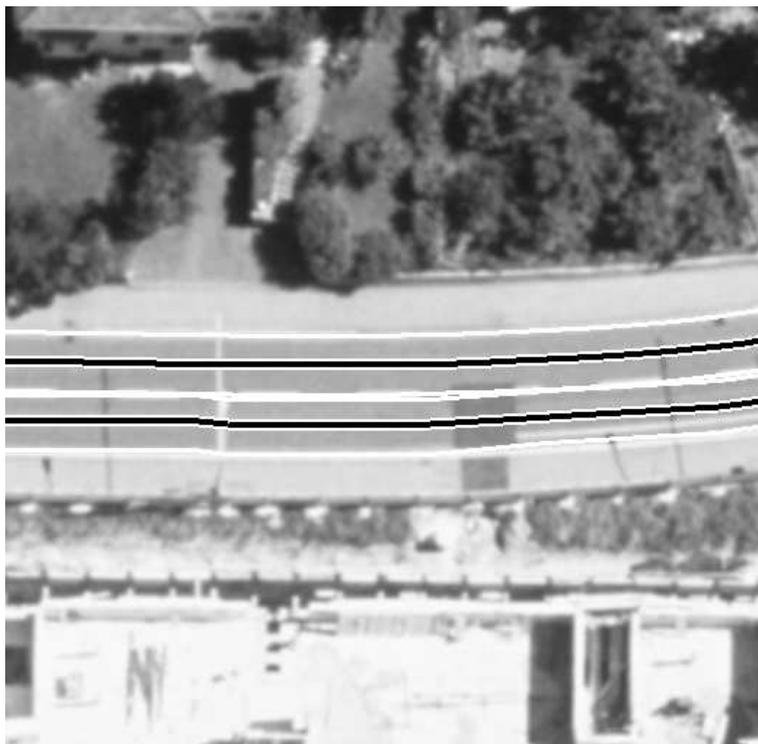
Abbildung 54 zeigt ein weiteres Beispiel der vereinfachten Fahrspurcluster. Der Vergleich mit Abbildung 49 – dem Ergebnis nach der Fahrspurgruppierung – verdeutlicht den Nutzen dieses Teilschritts auf dem Weg zur Konstruktion von Fahrbahnen: Die rechte äußere Fahrspur der Hauptstraße führt nun auch durch die Schattenregion im mittleren Bereich des Bildes. Außerdem konnten zwei weitere Unterbrechungen im Bereich der beiden Zebrastreifen überbrückt werden.

c) Bestimmung endgültiger Fahrbahnen:

Im letzten Teilschritt der Fahrbahnkonstruktion werden aus den vereinfachten Fahrspurclustern endgültige Fahrbahnen bestimmt. Die Konstruktionsvarianten der ursprünglichen Fahrspuren haben hier keinen Einfluss mehr; ein Zugriff auf sie bleibt aber über eine entsprechende Indizierung möglich. Eine Fahrbahn wird hier als der minimale Überlappungsbereich von bis zu fünf parallelen Fahrspuren definiert. Dies impliziert, dass alle



(a) Fahrbahnen (durchgezogene Linien) und Verknüpfungen (gepunktet). Aufgrund der Inhomogenität im Fahrbahnbelag ist weder die Verifikation aufgrund der autonomen Bewertung noch eine Fahrzeugdetektion möglich.



(b) Verifizierte Verknüpfung durch erfolgreiche top-down Bewertung: Die Verknüpfung reduziert die Fragmentierung der Fahrbahn erheblich.

Abbildung 53. Verknüpfung von Fahrspuren bei schlechter autonomer, aber guter top-down Bewertung.



Abbildung 54. Vereinfachte Fahrspurcluster: Im Vergleich zur Fahrspurgruppierung (Abbildung 49) konnten nun Unterbrechungen in der rechten äußeren Spur der Hauptstraße (durch Schatten) und im Bereich der beiden Zebrastreifen überbrückt werden.

Fahrspuren einer Fahrbahn an den Enden des minimalen Überlappungsbereiches abgeschnitten werden müssen. Die Repräsentation einer Fahrbahn erfolgt einerseits über ihre Fahrspuren. Andererseits werden darüber hinaus ihre Mittelachse und die beiden Randpolygone – die „Fahrbahnränder“ in Abbildung 13 – bestimmt. Die von den Randpolygonen eingeschlossene Fläche entspricht dem Fahrbahnbelag.

Um Fahrbahnen mit mehr als zwei Fahrspuren zu konstruieren, muss die Suche nach parallelen Fahrspuren im Vergleich zur Generierung von Fahrspurclustern hinsichtlich multipler Parallelen-Relationen erweitert werden. Dies geschieht durch einen Algorithmus, der zunächst die Parallelen-Relationen iterativ detektiert und anschließend daraus Fahrbahnen konstruiert (siehe auch Abbildung 55):

In der ersten Iteration zur Detektion von Parallelen-Beziehungen werden unter Berücksichtigung der Kriterien K I für Fahrbahnen (Tabelle 8) parallele Fahrspurpaare aus den Fahrspurclustern entnommen und die Mittelachse jeder Parallelen-Relation bestimmt. Um Dreier-Relationen zwischen Fahrspuren zu finden, werden in der zweiten Iteration Parallelen-Relationen zwischen den Mittelachsen gesucht und daraus neue Mittelachsen erzeugt. Diese wiederum sind die Eingangsinformation für die dritte Iteration, in der die nächsthöheren Relationen bestimmt werden, etc. Die Mittelachse der höchsten Parallelen-Relation stellt somit den minimalen Überlappungsbereich der Fahrspuren eines Clusters dar. Die obere Hälfte von Abbildung 55 illustriert die iterative Detektion von multiplen Parallelen.

Um Fahrbahnen aus den Relationen zu konstruieren, werden von der höchsten Stufe ausgehend alle Parallelen-Relationen bis zur untersten Stufe zurückverfolgt. Wie in der unteren Hälfte von Abbildung 55 zu sehen ist, werden im ersten Schritt alle Fahrspurabschnitte, die in den Bereich der höchsten Parallelen-Relation fallen, aus den ursprünglichen Fahrspuren entnommen und der neu erzeugten Fahrbahn zugewiesen (Abbildung 55 untere Hälfte). Auf der nächstniedrigeren Stufe wird mit allen dort verbliebenen Parallelen-Relationen in gleicher Weise verfahren, usw. Der Prozess endet, wenn die unterste Stufe erreicht ist.

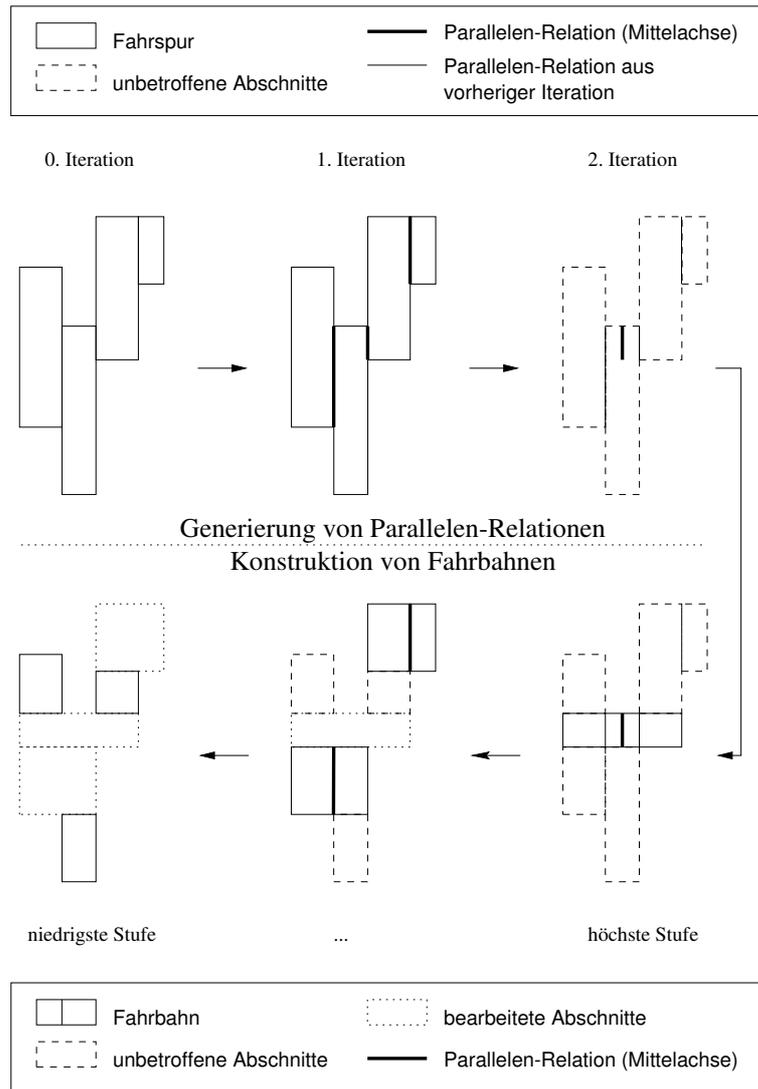


Abbildung 55. Schema zur Detektion von multiplen Parallelen-Relationen zwischen Fahrspuren und darauf aufbauende Konstruktion von Fahrbahnen.

Es folgt die abschließende autonome Bewertung der Fahrbahnen gemäß den Kriterien K II „große Länge“, „ähnliche Fahrspurbreiten“, „geringe Längsneigung“ und „sehr geringe Querneigung“ (vgl. Tabelle 8). Da die Fahrspuren einer endgültigen Fahrbahn keine Unterbrechungen besitzen, hat das Kriterium „niedrige Fragmentierung“ keinen Einfluss mehr. Ein Ergebnis der Konstruktion und Bewertung von Fahrbahnen ist in Abbildung 56 dargestellt. Die dicken Linien in Abbildung 56 a) entsprechen den Mittelachsen der Fahrbahnen, wobei die Helligkeit über die Güte der autonomen Bewertung Auskunft gibt. Am exakten Verlauf entlang von Markierungen erkennt man, dass Fahrbahnen, deren Fahrspuren aus Markierungen und Straßenrändern aufgebaut sind, durchweg gute Qualität besitzen. Im Gegensatz dazu fallen solche Fahrbahnen stärker ab, die entweder sehr gekrümmt sind – z.B. vor dem Schattenbereich im oberen Drittel der Hauptstraße – oder die Markierungen kreuzen, wie die Fahrbahn im unteren Drittel der Hauptstraße. Eine sehr schlechte Bewertung erhält die Fehlextraktion am linken Bildrand. Dies ist vor allem auf die große Querneigung entlang des Gebäudes zurückzuführen.



(a) Bewertete Fahrbahnen:
Je heller der Grauton um so höher die Qualität.



(b) Detail: Linke obere Ecke von (a).
Dargestellt sind die zugrunde liegenden Fahrspuren der Fahrbahnen.

Abbildung 56. Ergebnis der Fahrbahnkonstruktion und -Bewertung.

Die Tatsache, dass viele Mittelachsen in Abbildung 56 a) nicht kollinear verlaufen, liegt an der unterschiedlichen Anzahl an Fahrspuren und der dadurch bedingten Verschiebung der Mittelachsen. Dies wird auch aus der Detailansicht von Abbildung 56 b) ersichtlich, wo die zugehörigen Fahrspuren der Fahrbahnen dargestellt sind. Weiterhin ist dort der Einfluss der Forderung nach parallelen Fahrspuren zu erkennen. Beim Vergleich mit dem Ergebnis der vereinfachten Fahrspurcluster in Abbildung 54 sieht man, dass insbesondere im Kreuzungsbereich nun einige Fahrspuren eliminiert wurden, deren Verlauf nicht ausreichend parallel ist.

7.4.2 Gruppierung von Fahrbahnen

Wie im Fall der Fahrspuren schließt sich an die Konstruktion der Fahrbahnen ein Gruppierungsschritt an. Grundlage des Algorithmus ist wiederum ein Graph, dessen Kanten die Fahrbahnen bzw. Fahrbahnverknüpfungen und dessen Knoten deren Endpunkte repräsentieren. Die Gewichtung der Kanten leitet sich aus der Bewertung der Fahrbahnen bzw. der Verknüpfungen ab. Die Auswahl der endgültigen Verknüpfungen, mit denen die Fahrbahnen zu verbinden sind, geschieht über die Berechnung optimaler Pfade zwischen überdurchschnittlich bewerteten Fahrbahnen.

Entsprechend der höheren Semantik von Fahrbahnen stützt sich der Algorithmus auf relativ spezielle Kriterien für die Generierung und Bewertung von Verknüpfungen zwischen Fahrbahnen. Als Basis für die Generierung einer Verknüpfung dienen *Fahrspur*verknüpfungen zwischen beiden Fahrbahnen. Daher können für den Regelfall, dass eine Fahrbahn aus mehreren Fahrspuren besteht, mehrere Verknüpfungen zwischen zwei Fahrbahnen erzeugt und dementsprechend auch mehr Kanten in den Graph eingetragen werden. Abbildung 57 zeigt ein Beispiel für zweispurige Fahrbahnen. Hier können eine zweispurige Verknüpfung und zwei einspurige Verknüpfungen generiert werden. Begründet werden die drei Verknüpfungen durch die Bedingung, dass möglichst alle freien Fahrspurenden miteinander verbunden werden sollen, aber ohne dass sich Fahrspuren kreuzen. Als weitere Bedingung wird eingeführt, dass alle Fahrspurverknüpfungen den Kriterien K I für Fahrspuren und die resultierende Fahrbahnverknüpfung den Kriterien K I für Fahrbahnen genügen müssen (Tabellen 7 und 8). Wie zuvor kommen als Ersatz für die fehlenden Fahrspurbegrenzungen Korbbögen zwischen den freien Fahrspurenden zum Einsatz, aus denen die Fahrspuren konstruiert werden.

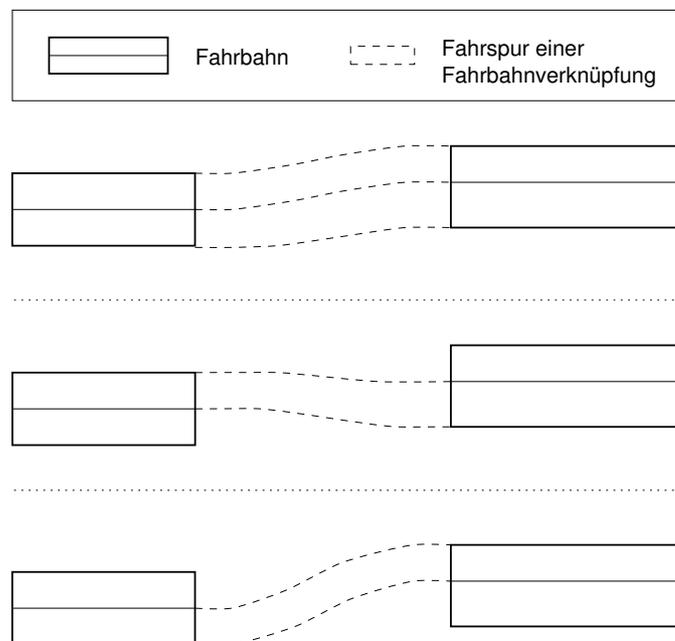
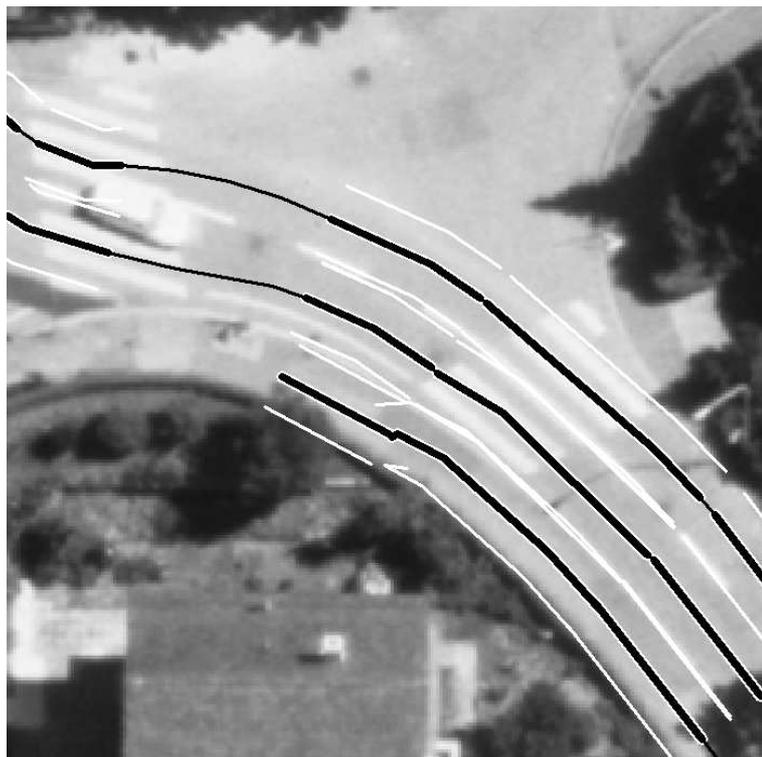


Abbildung 57. Verknüpfung von Fahrbahnen.



(a) Mittellinien der gruppierten Fahrbahnen (schwarz, fett),
Randpolygone (weiß), akzeptierte Verknüpfungen (schwarz, dünn).



(b) Fahrspuren der Fahrbahnen und Fahrbahnverknüpfungen:
Strichcodierung wie für Fahrbahnen in (a).

Abbildung 58. Ergebnis der Fahrbahngruppierung.

Die Bewertung einer Fahrbahnverknüpfung erfolgt in analoger Weise wie für Fahrspuren und Fahrbahnen. Zunächst werden die Fahrspuren der Verknüpfung mittels der Kriterien K II der Fahrspurbewertung evaluiert (Tabelle 7) und daraus die bottom-up Bewertung ermittelt. Im Falle einer schlechten Bewertung wird die Verknüpfung hinsichtlich Fahrzeugen untersucht. War die Fahrspurbewertung erfolgreich, dienen anschließend die Kriterien K II der Fahrbahnbewertung aus Tabelle 8 zur autonomen Bewertung der gesamten Verknüpfung. Als weiteres Kriterium zur autonomen Bewertung kommt noch der maximale Prozentsatz an Fahrspurenden, die eine Fahrbahnverknüpfung miteinander verbindet, hinzu. Durch dieses Kriterium wird z.B. die zweispurige Verknüpfung in Abbildung 57 gegenüber den beiden einspurigen bevorzugt.

Im letzten Verfahrensschritt wird die Entscheidung getroffen, welche Verknüpfungen letztendlich akzeptiert werden sollen. Alle Fahrbahnen und Verknüpfungen werden deshalb mit ihren Bewertungen in den Graph eingetragen, wobei das Gewicht der Kanten als Mittel aus bottom-up und autonomer Bewertung definiert ist. Es folgt die Auswahl überdurchschnittlich bewerteter Fahrbahnen als sichere Straßenobjekte, zwischen denen optimale Pfade berechnet werden. Alle Fahrbahnen und Verknüpfungen, die auf mindestens einem Pfad liegen, werden als verifiziert angesehen.

Ähnlich wie bei der Bestimmung von Fahrspurverknüpfungen auf Basis von Teil-Fahrbahnen in Abschnitt 7.4.1 kann hier eine schlechtere autonome Bewertung durch eine gute bottom-up und top-down Bewertung ausgeglichen werden. Denn trotz einer schlechten autonomen Bewertung bleibt aufgrund der Mittelbildung mit der bottom-up Bewertung die Chance erhalten, dass die Verknüpfung als Teil eines optimalen Pfades extrahiert wird. Auf der anderen Seite ist die Pfadberechnung gleichbedeutend mit einer top-down Bewertung für Fahrbahnen und Fahrbahnverknüpfungen. Denn wenn ein nur mäßig bewertetes Teilstück dennoch auf einem optimalen Pfad liegt, muss es aus globaler Sicht wichtig für den Zusammenhang der Fahrbahnen sein.

In Abbildung 58 sind Detailansichten auf die gruppierten Fahrbahnen dargestellt, wobei Abbildung 58 a) die Mittelachse der Fahrbahnen inklusive ihrer Randpolygone zeigt und Abbildung 58 b) die zugrunde liegenden Fahrspuren und Fahrspurverknüpfungen im Kreuzungsbereich visualisiert. Trotz des komplexen Bildinhaltes



Abbildung 59. Extrahierte Fahrbahnen eines größeren Bildausschnittes.

mit Kreuzung, Zebrastreifen und einem (nicht detektierten) Fahrzeug stimmt der Verlauf der Fahrspuren und Fahrspurverknüpfungen relativ gut mit der Realität überein. An Abbildung 58 a) fällt auf, dass im Zuge der Pfadberechnung die Überbrückung einer größeren Lücke Vorzug gegenüber mehreren kurzen Fahrbahnen erhalten hat. Dies ist eine eindeutige Konsequenz der schlechten Bewertung der stark gekrümmten Fahrbahn im Bereich der Schattenregion (vgl. Abbildung 56 a).

Mit der Gruppierung von Fahrbahnen ist die Extraktion sicherer Straßenobjekte abgeschlossen. Wie aus Abbildung 59 zu sehen ist, entsprechen die meisten der extrahierten Fahrbahnen auch tatsächlich Teilen des Straßennetzes. Allerdings konnten insbesondere in der Umgebung von Kreuzungen und bei teilweise abgeschatteten oder verdeckten Bereichen die Fahrbahnen nur in eingeschränktem Maße extrahiert werden. Dies liegt unter anderem auch daran, dass die Extraktion bisher nur auf Grundlage eines Bildes geschah und die zusätzliche Information aus mehreren überlappenden Bildern noch nicht genutzt wurde. Ein weiterer Grund für einige fehlende Straßenstücke ist die Tatsache, dass sich die Extraktion bis zur Bestimmung von Fahrbahnen auf eine relativ kleinräumige Umgebung beschränkte (einige wenige hundert Meter). Die Netzwerkeigenschaften von Straßen wurden noch nicht in die Extraktion eingebunden. Aus diesen Gründen behandelt das folgende Kapitel die Fusion von Extraktionen aus überlappenden Bildern und die Netzbildung durch Nutzung der Netzwerkeigenschaften von Straßen.

8 Fusion und Netzbildung

Die letzte Phase der Extraktionsstrategie hat die „Straßennetzbildung“ zum Inhalt (siehe Abbildung 60). Um einen einheitlichen Datensatz ohne Redundanzen zu erhalten, werden vor dem eigentlichen Aufbau des Straßennetzes die Ergebnisse, die in allen überlappenden Bildausschnitten separat erzielt wurden, fusioniert (Abschnitt 8.1). Die Fusion geschieht auf Basis von Fahrbahnen und wird im Objektraum durch Überlagerung aller Fahrbahnen auf dem DOM durchgeführt. Die resultierenden Fahrbahnen dienen dann als Eingangsinformation für den nachfolgenden Aufbau des Straßennetzes (Abschnitt 8.2.1). Dieser vollzieht sich durch iteratives Generieren und Verifizieren von Fahrbahnverknüpfungen und Kreuzungen unter Nutzung von Wissen aller Abstraktionsstufen inklusive von Wissen über Kontext. Im Unterschied zu allen vorangegangenen Gruppierungsschritten ist diese Gruppierung stärker auf die globale Netzwerkcharakteristik ausgelegt. Nach Beendigung des Netzaufbaus folgt die abschließende Analyse des Straßennetzes (Abschnitt 8.2.2). Kürzere Fahrbahnen, die nicht an das restliche Netz angeschlossen werden konnten, werden eliminiert und aus den verbleibenden Fahrbahnen werden Straßenabschnitte mit konstanter Anzahl an Fahrspuren gebildet. Knoten, an denen mehr als zwei Straßenabschnitte einmünden, definieren Kreuzungen, während Pfade, die Kreuzungen oder Endstücke verbinden, die endgültigen Straßen ergeben.

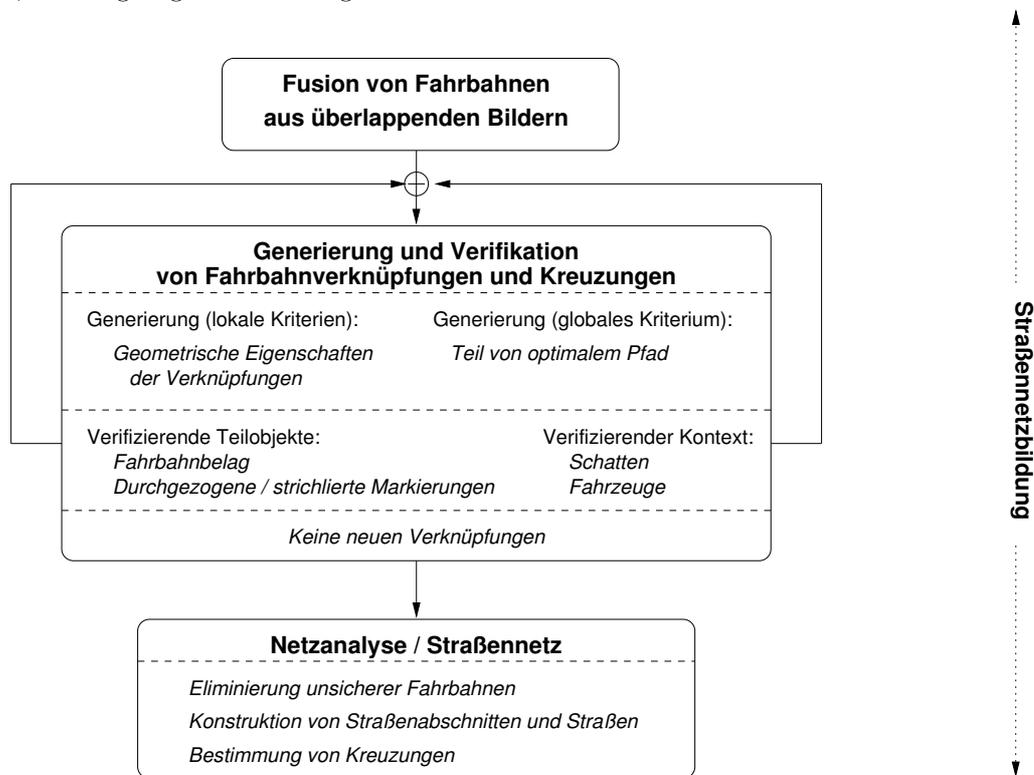


Abbildung 60. Extraktionsstrategie: Teil „Straßennetzbildung“ (vgl. Abbildung 18).

8.1 Fusion

Im eigentlichen Sinn ist bei der Fusion von Extraktionsergebnissen aus mehreren Bildern ein Korrespondenzproblem zu lösen. Denn Ziel der Fusion ist die Zuordnung von Objekten aus verschiedenen Bildern, um mehrfach („redundant“) extrahierte Objektteile von einmalig extrahierten zu unterscheiden. Redundant extrahierte Abschnitte des Straßennetzes stützen sich gegenseitig, so dass ihnen eine höhere Zuverlässigkeit beigemessen werden kann. Abschnitte, die nur in einem Bild extrahiert wurden, haben geringere Zuverlässigkeit. Allerdings

tragen sie zu einer größeren Vollständigkeit der gesamten Extraktion bei. In Abbildung 61 wird dies für das Extraktionsergebnis der Fahrbahnen aus Abbildung 59 illustriert. Beim Vergleich der Ergebnisse ist unschwer zu erkennen, dass viele Abschnitte des Straßennetzes in allen drei Bildern extrahiert wurden, einige jedoch nur in zwei oder gar nur einem.

Eine grundlegende Fragestellung bei der Zuordnung von Objekten in mehreren Bildern ist der Umgang mit Unsicherheiten, die durch die vorangegangene Extraktion oder durch die Erfassung der Datengrundlage entstehen. Wie in Abschnitt 3.1 beschrieben, sind z.B. die Höhenwerte eines DOM von Stadtgebieten nur mäßig genau und können vor allem in kleinräumiger Umgebung stark von der korrekten Höhe abweichen. Darüber hinaus treten auch Unsicherheiten in der Lage der extrahierten Objekte auf. Zum Beispiel kann es vorkommen, dass die in Abschnitt 7.4 bestimmten Fahrbahnen in Teilbereichen nur in der mittleren Auflösung extrahiert wurden und deshalb eine geringere geometrische Genauigkeit besitzen als Fahrbahnen, die aus Fahrspuren der hohen Auflösung konstruiert wurden.

Daraus wird deutlich, dass die exakte Zuordnung von struktureller Information unter Berücksichtigung von Unsicherheiten eine sehr komplexe Thematik ist. Sie steht jedoch nicht im Fokus dieser Arbeit. Stattdessen wird hier ein vereinfachtes, aber für die Zwecke der Straßenextraktion ausreichendes Vorgehen realisiert. Ziel ist, die Fusion auf Basis möglichst großer Objekte durchzuführen, so dass grobe Höhenfehler leichter eliminiert werden können und die geschätzten Objekthöhen eine für die Fusion ausreichende Genauigkeit besitzen. Zusätzlich werden all jene Regionen, die im Zuge der Kontextanalyse als Gebäude oder Vegetation detektiert wurden, von der Berechnung ausgeschlossen. Unsicherheiten in der Geometrie der extrahierten Objekte werden in Form von Toleranzintervallen für die Parameter der Zuordnung beschrieben. Durch diese Intervalle können auch die meist recht geringen Ungenauigkeiten in den Orientierungs- und Kameraparametern abgefangen werden.

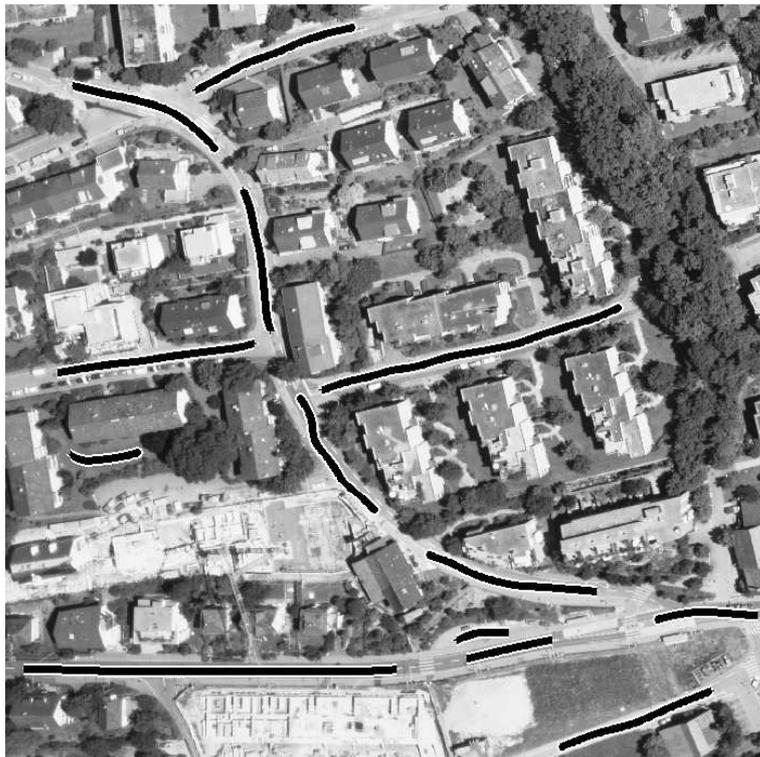
In (Hinz und Baumgartner, 2003) wird die Fusion auf Basis von Fahrspuren vollzogen. Abbildung 62 zeigt ein Beispiel von Fahrspuren, die in zwei verschiedenen Bildern extrahiert wurden und deren Höhe durch Medianbildung aus dem DOM ermittelt wurde. Die Transformation der Fahrspuren des mittleren Ausschnitts anhand ihrer Höhenwerte in den linken Ausschnitt zeigt eine sehr gute geometrische Übereinstimmung mit den dortigen Fahrspuren. Diese Genauigkeit kann jedoch nicht in allen Fällen erwartet werden, insbesondere wenn mehrere Häuser und dichte Vegetation relativ nahe an der Straße stehen und dadurch die Höhengenaugigkeit des DOM stark beeinflusst wird. Solch ein Fall ist in Abbildung 63 illustriert. Hier sind zwei Fahrbahnmittelachsen, die in verschiedenen Bildausschnitten extrahiert wurden, überlagert. Wie insbesondere am Verlauf der Mittelachsen in der Bildmitte zu sehen ist, kommt es zu geometrischen Unsicherheiten, die größer als die Breite einer halben Fahrspur sind. Nur wenn diese Bedingung eingehalten wird, können auf rein geometrischem Wege einzelne Fahrspuren eindeutig zugeordnet werden. Für die Zuordnung von Fahrbahnen reicht die gezeigte Genauigkeit jedoch aus. Daher erfolgt im Gegensatz zu (Hinz und Baumgartner, 2003) in dieser Arbeit die Fusion auf Basis von Fahrbahnen.

Der Algorithmus zur Fusion startet mit der Sortierung *aller* extrahierten Fahrbahnen gemäß ihrer Bewertung. Daran schließt sich die Projektion auf das Höhenmodell und die Ermittlung der Höhenwerte durch Medianbildung an. Um großräumigen Höhenänderungen Rechnung zu tragen, werden zur Höhenberechnung lange Fahrbahnen in Teilstücke zerlegt (z.B. 20m-Abschnitte). Die nachfolgende Fusion von Fahrbahnen entspricht exakt dem Verfahren zur Fahrspurfusion in Abschnitt 7.3.2. Unter Vorgabe eines Puffers und einer Winkeltoleranz werden die Fahrbahnen zugeordnet, redundante Teile eliminiert und alle restlichen Teile der Ergebnismenge hinzugefügt. Mögliche Höhendifferenzen zwischen zugeordneten Fahrbahnen und dadurch verursachte kleine Verschiebungen in ihrer Lage werden in Anbetracht der Objektgröße als vernachlässigbar erachtet. Neben einer Attributierung der resultierenden Fahrbahnen gemäß ihrem Ursprungbild wird eine Gewichtung in Abhängigkeit des Überlappungsprozentsatzes mit Fahrbahnen aus anderen Bildern durchgeführt.

Ein für die weitere Verarbeitung der fusionierten Fahrbahnen wesentlicher Punkt ist deren Bearbeitung im Hinblick auf potentielle Kreuzungen. Analog zu dem in Abbildung 47 für Fahrspuren illustrierten Fall, werden Fahrbahnen auch an solchen Positionen getrennt, an denen eine zugeordnete Fahrbahn den Puffer verlässt. Dies



(a)



(b)

Abbildung 61. (a), (b) Ergebnisse der Extraktion von Fahrbahnen in anderen (überlappenden) Bildausschnitten, die die gleiche Szene wie Abbildung 59 darstellen.

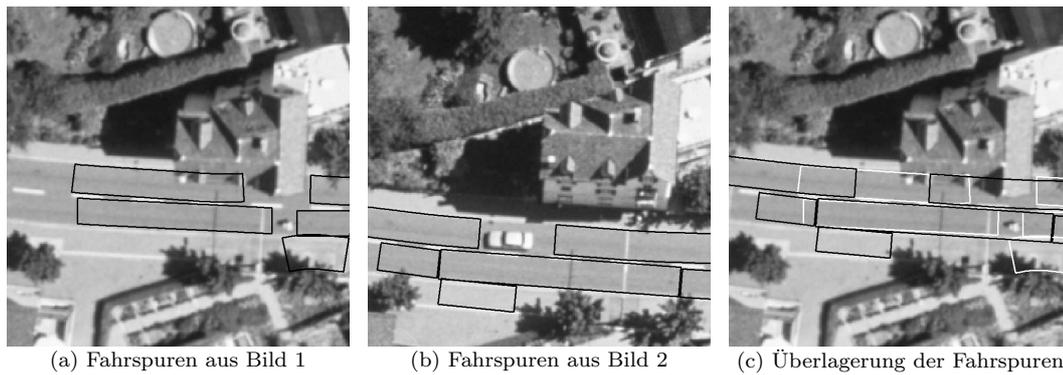


Abbildung 62. Beispiel zur korrekten Fusion von Fahrschulden (aus (Hinz und Baumgartner, 2003)): Fahrschulden aus Bild 2 (b) über die Höhenwerte des DOM in Bild 1 (a) transformiert, Überlagerung des Ergebnisses in (c).



Abbildung 63. Überlagerung der Achsen von Fahrschulden aus verschiedenen Bildern. Lageabweichungen der Fahrschulden können größer als eine halbe Fahrschuldenbreite sein, so dass eine eindeutige Zuordnung von Fahrschulden nicht unbedingt gewährleistet ist.

ermöglicht im Rahmen der Netzbildung das Einfügen eines Verbindungsstückes und somit die Konstruktion einer Kreuzung.

Ein gewisser Nachteil bei der gewählten Form der Fusion ist die Tatsache, dass die Fahrschuldenzahl bei der Fusion nicht berücksichtigt wird. Dies hat zur Folge, dass unter Umständen auch Fahrschulden mit unterschiedlicher Anzahl an Fahrschulden fusioniert werden und damit die Information über die Fahrschuldenzahl nicht mehr nützlich ist. Am Ende der Extraktion wird diese Information im Zuge der Netzanalyse (Abschnitt 8.2.2) jedoch wieder zurückgewonnen. Abbildung 64 zeigt das Ergebnis der Fusion aus den drei in Abbildung 59 und Abbildung 61 a) und b) dargestellten Einzelergebnissen. Wie beim Vergleich mit diesen Abbildungen auffällt, hat sich allein schon durch die Fusion (ohne weitere Gruppierung) die Vollständigkeit der Extraktion um einiges erhöht. Dank der großen Zuverlässigkeit der vorangegangenen Extraktionen hält sich die Anzahl an fehlerhaften Fahrschulden in Grenzen. Allerdings sind einige geometrische Unsicherheiten in den Achsen der fusionierten Fahrschulden zu beobachten. Besonders auffällig sind diese an manchen Übergängen zwischen Fahrschulden, die in verschiedenen

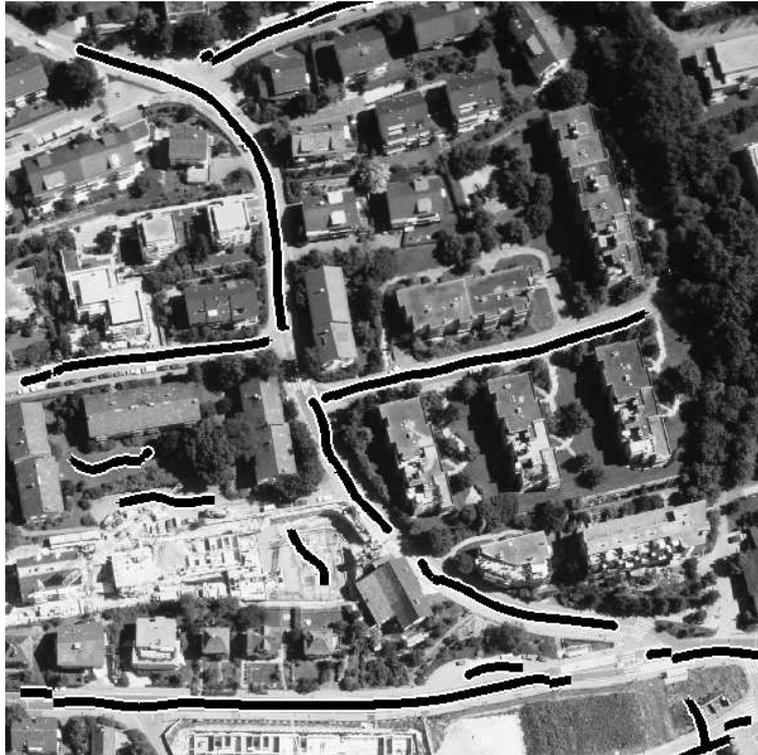


Abbildung 64. Ergebnis der Fusion von Fahrbahnen (vgl. Abbildungen 61 und 59).

Bildern extrahiert wurden (siehe z.B. die Fahrbahnen entlang des unteren Bildrandes). Das bedeutet, dass die folgenden Algorithmen an die Art von Unsicherheit angepasst sein müssen.

8.2 Netzbildung

Der abschließende Schritt der Extraktion ist die Straßennetzbildung. Ihre Zielsetzung ist der Aufbau eines topologisch zusammenhängenden Straßennetzes, das eine Verbindung von einem beliebigen Punkt „A“ zu einem beliebigen Punkt „B“ gewährleistet. Die entsprechende Modellkomponente für den Aufbau des Straßennetzes aus den fusionierten Fahrbahnen ist die *Netzwerkeigenschaft*. Sie ist im Modell auf der hohen Abstraktionsstufe durch die Relation „verknüpft“ der Konzepte „Straße“ und „Kreuzung“ charakterisiert (siehe Abbildung 13). Die Koppelung mit der mittleren Abstraktionsstufe wird über die Straßen-Teilobjekte „Straßenabschnitte“ hergestellt, die dort als kollinear und parallel angeordnete Fahrbahnen modelliert sind.

Um die Netzwerkeigenschaft für die Extraktion zu nutzen, wird in Anlehnung an (Steger et al., 1997; Baumgartner und Hinz, 2000; Wiedemann, 2001) in Abschnitt 8.2.1 eine so genannte „globale Gruppierung“ eingesetzt. Ihr Ziel ist die Bestimmung von Verknüpfungen, die aus globaler Sicht wichtig für die Netztopologie sind. Dies wird dadurch realisiert, dass optimale Verbindungen zwischen weit entfernten Punkten auf dem Straßennetz berechnet und all jene Verknüpfungen selektiert werden, die Bestandteil solcher Verbindungen sind. Eine Verknüpfung wird als Fahrbahn akzeptiert, wenn sie durch Evidenz im Bild gestützt wird. Darüber hinaus ist die globale Gruppierung in einen iterativen Prozess eingebettet, der mit relativ restriktiv eingestellten Parametern beginnt, um diese in folgenden Iterationen schrittweise zu lockern. Dadurch wird auch im Rahmen der globalen Gruppierung die Maßgabe verfolgt, zunächst auf zuverlässig extrahierbare Teile zu fokussieren und anschlie-

ßend von diesen ausgehend das Netz zu vervollständigen. Der iterative Prozess setzt sich solange fort, bis keine weiteren Verknüpfungen mehr generiert oder verifiziert werden können.

Die Konstruktion und Bewertung der extrahierten Straßen erfolgt im Rahmen einer Netzanalyse (Abschnitt 8.2.2). Hierzu werden die Fahrbahnen zu Straßenabschnitten mit konstanter Anzahl an Fahrspuren und weiter zu Straßen zusammengefasst. Die Straßenextraktion endet mit der abschließenden Bewertung von Straßen und Kreuzungen.

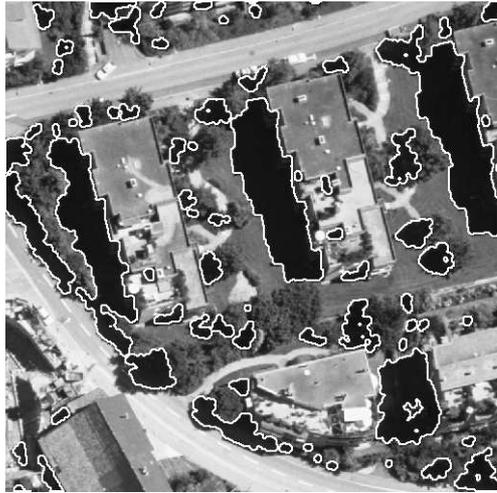
8.2.1 Generierung und Verifikation von Fahrbahnverknüpfungen und Kreuzungen

Als Basis für die Nutzung der Netzwerkeigenschaften wird in den folgenden Schritten wiederum eine Graphrepräsentation herangezogen. Die Fahrbahnen bilden die Kanten des Graphen und die Endpunkte seine Knoten. Um eine erste Auswahl an potentiellen Fahrbahnverknüpfungen zu erhalten, werden Verknüpfungen zunächst anhand *lokaler Kriterien* aufgestellt und bewertet. Hierzu dienen in erster Linie die geometrische Eigenschaften „Neigung“, „Länge“, „relative Länge“ und „Kollinearität/Orthogonalität“. Die relative Länge bewertet die Verknüpfungsdistanz bezüglich der kürzeren der beiden zu verbindenden Fahrbahnen. Sie wird dadurch begründet, dass kurze Fahrbahnen auch nur durch kurze Verknüpfungen ans Netz angeschlossen werden sollten. Durch das Kollinearität-/Orthogonalität-Kriterium werden sowohl geradlinige Verknüpfungen als auch Verknüpfungen mit Richtungsabweichungen um 90° , die auf die Präsenz einer Kreuzung schließen lassen, bevorzugt gewichtet. Im Falle von Verknüpfungen, die kürzer als die Pufferbreite der Fahrbahnfusion sind, bleibt dieses Kriterium jedoch unberücksichtigt, weil dort ein kleiner Versatz zwischen zwei Fahrbahnachsen als tolerierbar angesehen wurde.

Die Entscheidung, welche der potentiellen Verknüpfungen tatsächlich einer Verifikation unterzogen werden soll, stützt sich auf das *globale Kriterium*, dass eine Verknüpfung Teil eines optimalen Pfades zwischen zwei weit entfernten Punkten auf dem Netz sein muss. Zu diesem Zweck werden Kanten des Graphen, die mit Fahrbahnen korrespondieren, anhand ihrer Bewertung aus Abschnitt 7.4 gewichtet. Die Evaluierung von Kanten, die Verknüpfungen repräsentieren, erfolgt anhand von Fuzzy-Funktionen für die oben erwähnten Kriterien. Anschließend werden Paare von weit entfernten und überdurchschnittlich gut bewerteten Fahrbahnen gebildet, optimale Pfade zwischen ihnen berechnet und Verknüpfungen selektiert, die mindestens auf einem dieser Pfade liegen. Die Bedingung, dass die Fahrbahnpaare weit entfernt und überdurchschnittlich gut bewertet sein müssen, soll zum einen sicherstellen, dass die globale Verbindungseigenschaft von Straßen genutzt wird, und zum anderen, dass die berechneten Pfade mit hoher Wahrscheinlichkeit tatsächliche Fahrbahnen miteinander verbinden.

Zur Verifikation einer Verknüpfung wird zunächst anhand der Sichtbarkeitsanalyse aus Abschnitt 6.2.1 überprüft, aus welchen Bildern die Verknüpfung eingesehen werden kann und ob lokale Kontextinformation in den korrespondierenden Bildbereichen extrahiert wurde. Die anschließenden Varianten zur Verifikation werden in jedem Bild separat durchgeführt, wobei zur Annahme einer Verknüpfung die erfolgreiche Verifikation in einem einzigen Bild ausreicht:

- (1) Liegt keine Kontextinformation vor, erfolgt die Verifikation über die Teilobjekte „Fahrbahnbelag“ (die Grauwerthomogenität in der mittleren Auflösung) und „durchgezogene oder strichlierte Markierung“ aus der hohen Auflösung. Bildregionen, die extrahierte Fahrzeuge beinhalten, werden von der Berechnung der Grauwerthomogenität ausgeschlossen. Im Gegensatz zur Grauwerthomogenität werden Markierungen nur zum zusätzlichen Evidenzgewinn herangezogen und stellen keine notwendige Bedingung dar.
- (2) Im Falle einer Schattenregion wird die gleiche Verifikationsmethode wie oben angewendet, jedoch mit angepassten Parametersätzen.
- (3) Ist die Bildregion zu einem hohen Prozentsatz mit extrahierten Fahrzeugen belegt ($> 50\%$), wird die Verknüpfung ohne weitere Analyse akzeptiert.



(a) Detektierte Schattenregionen in der Umgebung einer Verknüpfung (siehe c)



(b) Detektierte sichttote Räume in der Umgebung einer Verknüpfung: Die hier gezeigten Regionen können in Bild (d) nicht eingesehen werden.



(c) Verknüpfung (schwarz, dünn) kann trotz des Schattens verifiziert werden (siehe a).



(d) Verknüpfung (schwarz, dünn) wird in diesem Bild wegen der Verdeckung nicht überprüft (siehe b).

Abbildung 65. Nutzung von Kontextinformation bei der Verifikation von Verknüpfungen: Verknüpfung = dünne schwarze Linie, Fahrbahnachse = dicke Linie.

Abbildung 65 zeigt eine Situation, in der ein Straßenstück in einem Bild abgeschattet ist und in einem anderen Bild durch einen Baum zur Hälfte verdeckt ist. Dank der im Voraus extrahierten Kontextinformation ist das System in der Lage, die Verknüpfung zu akzeptieren. Die Verifikation wird in diesem Fall lediglich im Bildausschnitt mit der Schattenregion durchgeführt, da die Information über die Verdeckungssituation eine Überprüfung in dem anderen Bildausschnitt korrekterweise verhindert.

Kann eine Fahrbahn nicht verifiziert werden, bleibt die entsprechende Kante zwar im Graph, sie erhält aber das Gewicht „Null“. Dies verhindert einerseits, dass die Kante in der nächsten Iteration Teil eines optimalen Pfades wird. Andererseits bleibt dadurch die Information auch in folgenden Iterationen erhalten, dass die betreffende Verknüpfung bereits geprüft und als falsch eingestuft wurde.

Sind alle Verknüpfungen bearbeitet worden, startet die nächste Iteration mit schwächer eingestellten Parametern für die lokalen Kriterien zur Generierung von Verknüpfungen. Dies betrifft insbesondere die Lockerung der zulässigen minimalen Kollinearität und Orthogonalität. Als Ausgangspunkte für die Pfadberechnungen dienen fortan all jene Knoten im Graph, die in der vorherigen Iteration Startpunkt eines vollständig verifi-

zierten Pfades waren. Diese Bedingung stellt ein gewisses Gegengewicht zu den abgeschwächten Parametern dar, weil dadurch Startpunkte ausgewählt werden, deren Zuverlässigkeit eingehend überprüft wurde. Im Falle, dass keine neue Verknüpfungen mehr generiert werden können, bricht das Verfahren ab. Am Ende der globalen Gruppierung enthält der Graph keine Knoten mehr mit Grad 2, d.h. eine Fahrbahn endet jetzt immer an einer Kreuzung oder an einem Endpunkt (Sackgasse). In Abbildung 66 a) ist ein Zwischenergebnis der globalen Gruppierung dargestellt. Beim Vergleich mit Abbildung 64 wird deutlich, dass in den ersten beiden Iterationen vor allem relativ geradlinige Verknüpfungen eingefügt wurden. Im weiteren Verlauf bis zum Endergebnis nach vier Iterationen (Abbildung 66 b) wurden dann auch die Kreuzungsbereiche überbrückt.

Eine wesentliche Charakteristik des vorgestellten Algorithmus zur globalen Gruppierung ist, dass Wissen aus allen drei Abstraktionsstufen genutzt wird: Potentielle Verknüpfungen werden auf Basis von Fahrbahnen generiert und stützen sich somit in erster Linie auf Wissen der mittleren Abstraktionsstufe. Die Selektion der Verknüpfungen mit Hilfe optimaler Pfade geschieht durch Verwendung der Netzwerkeigenschaften der hohen Abstraktionsstufe. Im Rahmen der Verifikation schließlich wird nach Evidenz für Straßenbelag, Markierungen, Fahrzeuge und Schatten gesucht und somit vorwiegend Wissen aus der niedrigen Abstraktionsstufe und Wissen über lokalen Kontext integriert.

Das Resultat am Ende der globalen Gruppierung besteht in der Regel aus einer großen Anzahl an verknüpften und aus einer geringen Anzahl an isolierten Fahrbahnen, die nicht an das restliche Netz angeschlossen werden konnten. Um dieses Resultat geeignet weiter zu verarbeiten, folgt die Netzanalyse.

8.2.2 Netzanalyse

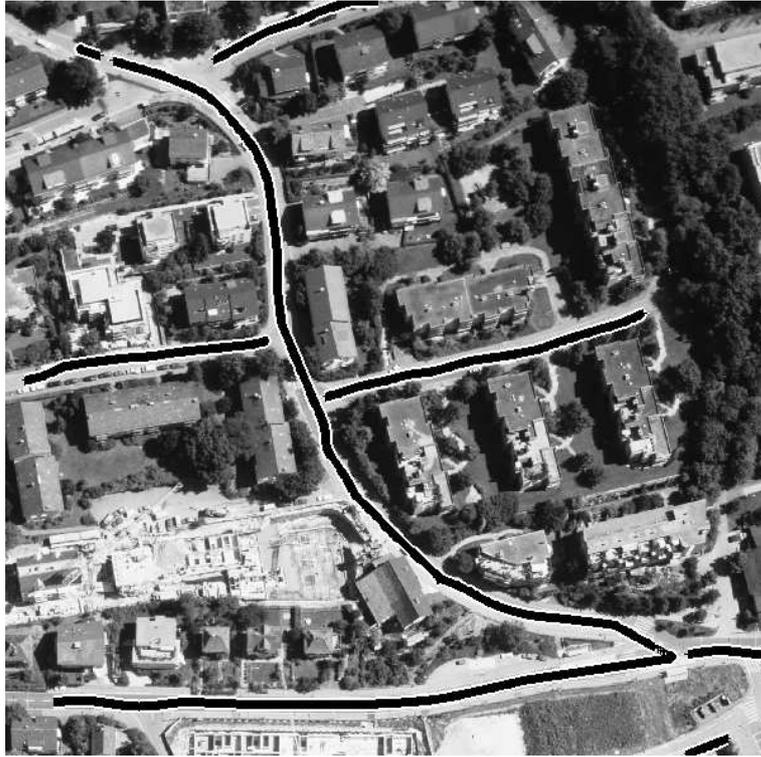
Die Netzanalyse stellt den letzten Schritt der Straßenextraktion dar. Mit ihr werden zweierlei Ziele verfolgt: Zum einen soll anhand der Eigenschaften des extrahierten Fahrbahn-Netzes entschieden werden, welche Teile des Netzes als fehlerhafte Extraktion angesehen werden und zu eliminieren sind. Zum anderen sind auf Basis der verbliebenen Teile die endgültigen Straßen und Kreuzungen zu extrahieren und zu bewerten. Straßenabschnitte dienen hierfür als Zwischenkonzept (vgl. Abbildung 13).

Abgesehen von sehr seltenen Ausnahmen bilden Straßen in der realen Welt ein vollständig zusammenhängendes Netz. Eine direkte Übertragung dieser Eigenschaft auf die Straßenextraktion wäre jedoch nicht zielführend. Einerseits kann ein Bild nur einen endlichen Ausschnitt der Szene erfassen, so dass der abgebildete Ausschnitt des Straßennetzes nicht zwangsläufig zusammenhängend sein muss. Andererseits ist es aus heutiger Sicht nicht möglich, alle Erscheinungsformen von Straßen und alle umgebenden Einflüsse auf ihr Erscheinungsbild zu modellieren, weshalb auch bei einer qualitativ sehr hochwertigen Extraktion einige Lücken im extrahierten Netz zu erwarten sind. Unter diesen Voraussetzungen wäre weder eine Netzanalyse, die eine einzige Zusammenhangskomponente voraussetzt, noch eine Analyse, die Verbindungen zwischen Kreuzungs- oder Endpunkten des Netzes als isolierte Einheiten untersucht, sinnvoll. Daher verbindet die folgende Netzanalyse Eigenschaften, die sich direkt auf Straßen beziehen, mit Eigenschaften, die die jeweilige Zusammenhangskomponente, in der sich eine Straße befindet, betreffen.

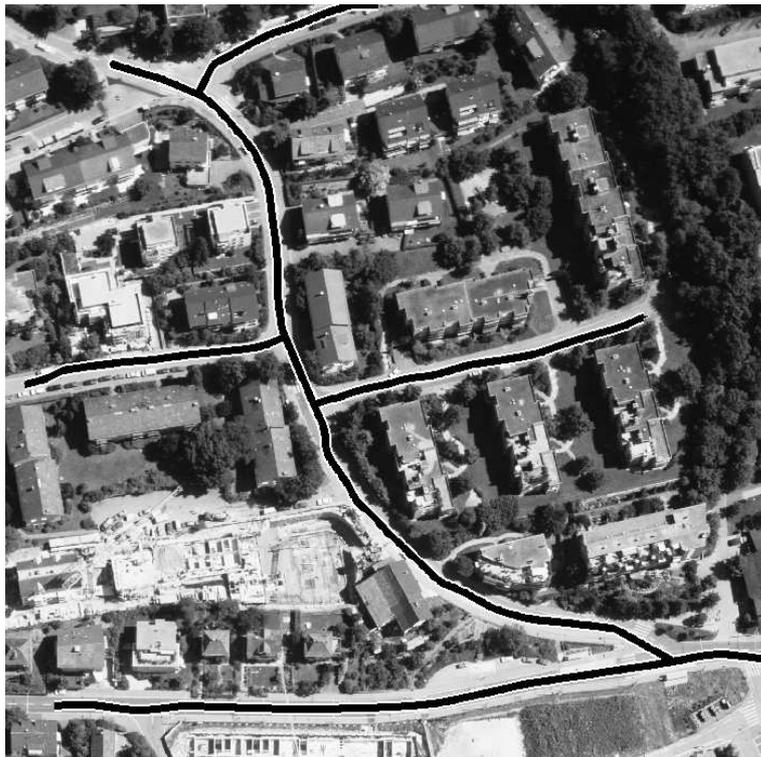
Die Analyse des Fahrbahn-Netzes und die Bestimmung der endgültigen Straßen und Kreuzungen umfasst folgende drei Schritte (siehe auch Abbildung 60): (a) Die Elimination unsicherer Fahrbahnen, (b) die Konstruktion von Straßen und (c) die Bestimmung von Kreuzungen.

a) Elimination unsicherer Fahrbahnen:

Unter unsicheren Fahrbahnen werden solche Fahrbahnen verstanden, deren Funktion im extrahierten Netz als zu gering eingeschätzt wird, um sie zur endgültigen Bestimmung der Straßen heranzuziehen. Unsicher sind in



(a)



(b)

Abbildung 66. (a) Zwischenergebnis nach 2 Iterationen und (b) Endergebnis der globalen Gruppierung nach 4 Iterationen.

diesem Sinne z.B. Fahrbahnen, die im Netz kurze Sackgassen darstellen. In der Realität sind solche Straßen nur in Verbindung mit der Funktion sinnvoll, dass sie ein spezielles Gebäude oder ein spezielles Gebiet an das restliche Straßennetz anschließen. Da diese Funktion aber nicht im Modell enthalten ist, fehlt eine zuverlässige Entscheidungsgrundlage, ob es sich bei einer derartigen Fahrbahn um eine Fehlextraktion oder eine tatsächliche Zufahrt handelt. Aus diesem Grund werden kurze Fahrbahnen, die mindestens an einem Ende keinen Anschluss an eine andere Fahrbahn haben, eliminiert (vgl. auch Tabelle 9). Die Einschränkung auf kurze Fahrbahnen trägt der Tatsache Rechnung, dass das Straßennetz nur in Teilen extrahiert werden kann.

Darüber hinaus werden unter der verbleibenden Menge an Fahrbahnen diejenigen als unsicher eingeschätzt, die sich in einer Zusammenhangskomponente mit *ausschließlich* kurzen Fahrbahnen befinden. Dieses Kriterium spiegelt die Eigenschaft wider, dass die Struktur des Straßennetzes nicht beliebig engmaschig sein kann, sondern dass der mittlere Abstand zwischen Kreuzungen eine gewisse Mindestlänge aufweisen muss. Fällt also die mittlere Fahrbahnlänge einer Zusammenhangskomponente unter einen bestimmten Schwellwert – etwa die Seitenlänge eines minimalen Häuserblocks –, werden alle Fahrbahnen der betreffenden Zusammenhangskomponente eliminiert.

b) Konstruktion von Straßenabschnitten und Straßen:

Straßen sind im Modell als Verbindungen zwischen Kreuzungen definiert. Sie sind aus Teilobjekten, den Straßenabschnitten, aufgebaut, deren Fahrbahn eine konstante Anzahl an Fahrspuren besitzt. Dies steht in Analogie zur Konstruktion von Fahrspuren aus vorläufigen Straßenabschnitten.

Wie in Abschnitt 8.1 erläutert, wurde die Fahrspuranzahl bei der Fusion von Fahrbahnen nicht berücksichtigt. Dies hatte zur Folge, dass unter Umständen Fahrbahnen mit unterschiedlicher Fahrspuranzahl fusioniert wurden. Um solche Konflikte zu lösen und um Straßenabschnitte mit einheitlicher Fahrspuranzahl zu bestimmen, werden das Fahrbahnnetz und sein Graph in kleine Teilstücke zerlegt, die Teilstücke in die einzelnen Bilder projiziert und dort den Fahrbahnen im Stadium vor der Fusion zugeordnet. Über die zugeordneten Fahrbahnen lässt sich nun auf die Anzahl der korrespondierenden Fahrspuren in jedem Bild zugreifen. Das längengewichtete, gerundete Mittel aus allen Bildern ergibt dabei die endgültige Anzahl an Fahrspuren eines jeden Teilstücks. Für Teilstücke, die in keinem Bild zugeordnet werden konnten, weil sie ausschließlich aus einer Fahrbahnverknüpfung bestehen, wird die Fahrspuranzahl aus den angrenzenden Teilstücken interpoliert. Abbildung 67 a) illustriert die so ermittelten Fahrspuren für das Straßennetz aus Abbildung 66 b).

Die endgültigen Straßenabschnitte ergeben sich dann aus der Verkettung von Teilstücken, die im Graphen durch Knoten mit Grad 2 verbunden sind und die gleiche Anzahl an Fahrspuren besitzen. Straßen erhält man schließlich über die Verkettung von Straßenabschnitten, die durch Knoten mit Grad 2 verbunden sind, ohne dass die Bedingung einer konstanten Anzahl an Fahrspuren berücksichtigt wird. Die resultierenden Straßen enden also entweder an Kreuzungspunkten (Knoten mit Grad > 2) oder an Endpunkten (Knoten mit Grad = 1).

Die abschließende Bewertung der extrahierten Straßen erfolgt anhand der Kriterien „mittlere Krümmung“, „mittlere Neigung“ und „konstante Anzahl an Fahrspuren“. Alle Bewertungsmaße werden für eine gesamte Straße berechnet (siehe Tabelle 9). Die Anzahl der Fahrspuren entlang einer Straße muss nicht völlig konstant sein, damit die Straße eine ideale Bewertung erhält. Eine große Varianz in der Fahrspuranzahl lässt jedoch auf eine Straße schließen, die zumindest mit geringerer Zuverlässigkeit extrahiert wurde – sofern sie überhaupt korrekt ist. Die Kriterien zur Extraktion und Bewertung von Straßen sind in Tabelle 9 zusammengestellt, und Abbildung 67 b) visualisiert das Ergebnis der Bewertung.

Kriterien K I	Kriterien K II
Minimale Länge unter Berücksichtigung von Kreuzungen und Sackgassen	Geringe mittlere Krümmung
Minimale mittlere Fahrbahnlänge einer Zusammenhangskomponente	Geringe mittlere Neigung
Verkettete Straßenabschnitte (an Knoten mit Grad 2)	Weitgehend konstante Anzahl an Fahrspuren

Tabelle 9. Kriterien zur Extraktion und Bewertung von Straßen.

c) Bestimmung von Kreuzungen:

Der abschließende Schritt der Extraktion besteht aus der Bestimmung von Kreuzungen. Während das Modell in Abbildung 13 zwischen einfachen und komplexen Kreuzungen unterscheidet, wurde in der bisherigen Extraktion auf diese Differenzierung keine Rücksicht genommen. Die Bestimmung von Kreuzungen erfolgte bisher nur implizit im Rahmen der globalen Gruppierung in Abschnitt 8.2.1. Sie ergeben sich dadurch, dass zwei optimale Pfade in einem Teilbereich identisch und in anderen Bereichen verschieden sein können. Endet der identische Teilbereich nicht mit Startpunkten der Pfadberechnung, muss er an Kreuzungspunkten enden.

Als Kandidaten für Kreuzungen werden zunächst alle Knoten mit Grad 3 und höher aus dem Graphen extrahiert und auf Plausibilität überprüft. Für diese einfache Form der Extraktion dient als Bewertungskriterium, dass Kreuzungen üblicherweise nicht mehr als fünf einmündende Straßen besitzen (Mackaness und Mackechnie, 1997). Der Vollständigkeit halber sind die beiden Kriterien in Tabelle 10 eingetragen.

Um auf Basis des extrahierten Straßennetzes eine Klassifikation in einfache und komplexe Kreuzungen vorzunehmen, wird jede als plausibel bewertete Kreuzung daraufhin überprüft, ob sie in der niedrigen Auflösung mit einem differential-geometrischen Verfahren zur Ermittlung von Linienknoten extrahiert werden kann (Steger, 1998b). Zu diesem Zweck wird eine Kreisregion um jede Kreuzung erzeugt und getestet, ob in diesen Bereich ein Linienknoten fällt, der sich durch die Extraktion von Straßenachsen in der niedrigen Auflösung bestimmen lässt (analog zur Generierung vorläufiger Straßenabschnitte in Abschnitt 7.1). Da sich das in (Steger, 1998b) verwendete Kreuzungsmodell stark auf lineare Strukturen stützt, ist das Verfahren nur zur Extraktion von Kreuzungen mit relativ kleinräumigem Kreuzungsbereich und somit für „einfache Kreuzungen“ geeignet. Im Umkehrschluss bedeutet dies, dass Kreuzungen, die nicht mit diesem Modell extrahiert werden können, nur dank der Konstruktion und Gruppierung von Fahrbahnen detektiert werden konnten und somit komplexerer Natur sein müssen. Eine detailliertere Extraktion komplexer Kreuzungen, z.B. mit Hilfe von Fahr- und Abbiegespuren im Kreuzungsbereich oder baulichen Trennelementen, ist in jedem Fall sinnvoll. Dies würde aber den Rahmen dieser Arbeit sprengen und wird daher nicht weiterverfolgt.

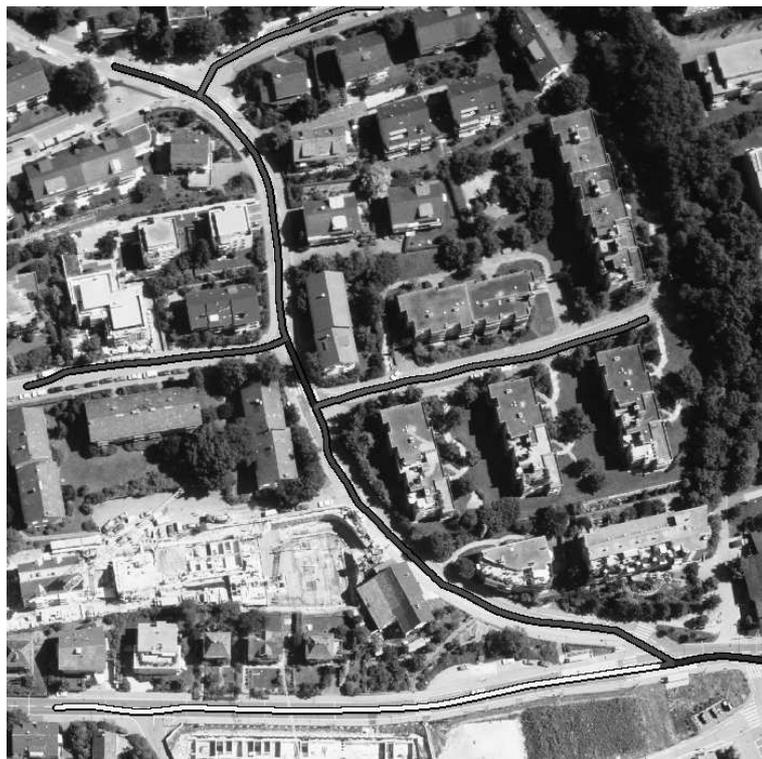
In Abbildung 68 sind die klassifizierten Kreuzungen des Straßennetzes dargestellt. Die hell umrandeten Kreuzungsbereiche entsprechen dabei einfachen Kreuzungen und der schwarz umrandete einer als komplex eingestuften Kreuzung. Abbildung 68 b) zeigt eine Detailansicht zweier einfacher Kreuzungen. Hier wurde im Kreuzungsbereich die Linienextraktion aus der niedrigen Auflösung den endgültigen Straßenachsen überlagert.

Kriterien K I	Kriterien K II
Kreuzungspunkt von Straßen (gleichbedeutend: Endpunkt einer Straße)	Anzahl der einmündenden Straßen

Tabelle 10. Kriterien zur Extraktion und Bewertung von Kreuzungen.

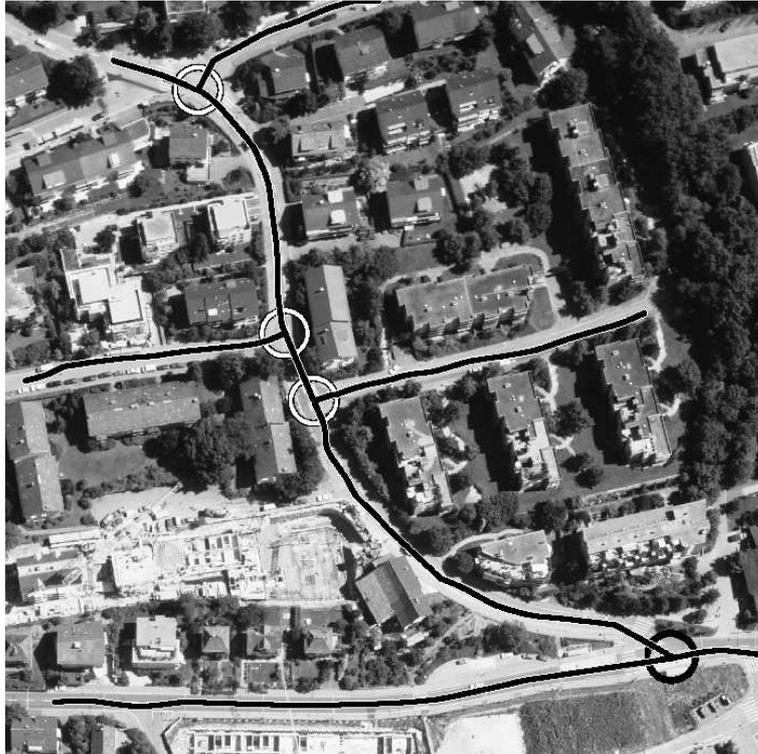


(a) Fahrspuren der endgültigen Straßen

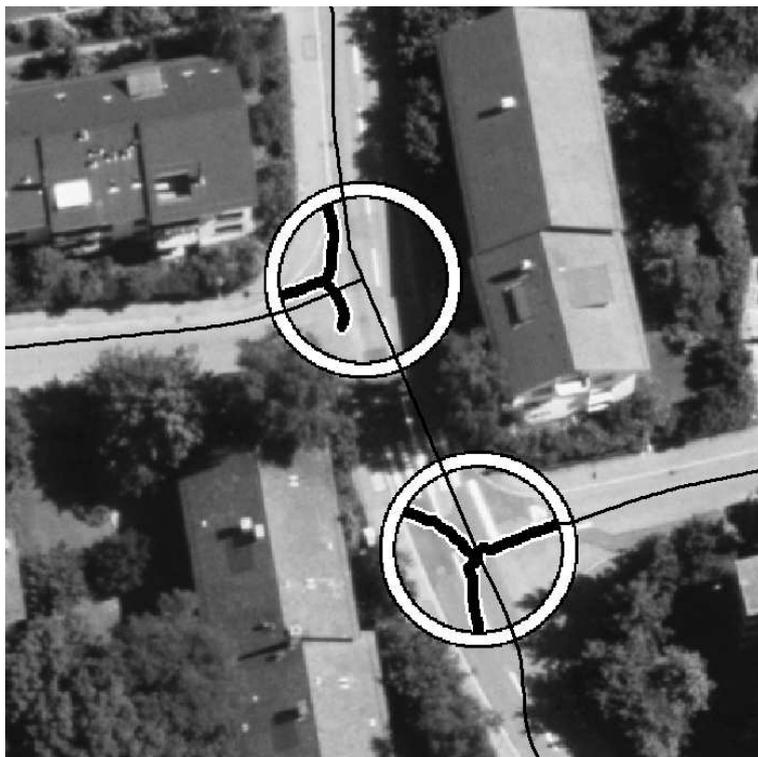


(b) Bewertung der endgültigen Straßen:
Je heller die Achse desto besser die Bewertung.

Abbildung 67. (a) Fahrspuren, die den Straßen zugrunde liegen. (b) Bewertung des Endergebnisses anhand der Kriterien K II aus Tabelle 9.



(a) Klassifizierte Kreuzungen:
Einfache Kreuzung (weißer Ring), komplexe Kreuzung (schwarzer Ring)



(b) Detail: Kreuzungsbereich (Ring), Linien- und Kreuzungsextraktion aus niedriger Auflösung (schwarz, fett), überlagerte Straßenachsen (schwarz, dünn).

Abbildung 68. Klassifikation von Kreuzungen in einfache und komplexe Kreuzungen; Grundlage ist eine Linien- und Kreuzungsextraktion in niedriger Auflösung nach (Steger, 1998b).

Abschluss Teil II

Die in den Abbildungen 67 und 68 dargestellten Resultate entsprechen dem Endergebnis der Straßenextraktion. Abgesehen von einigen Abschnitten an den Bildrändern wurden alle großen Straßen extrahiert und durch Kreuzungen miteinander verknüpft. Dies deutet darauf hin, dass das implementierte System grundsätzlich in der Lage ist, Straßennetze in komplexen städtischen Umgebungen zu extrahieren. Bei genauerer Betrachtung fällt jedoch auf, dass insbesondere die resultierenden Fahrspuren einige Mängel aufweisen. Dies betrifft zum Teil die Vollständigkeit und zum Teil die Lagegenauigkeit der extrahierten Fahrspuren. Um die Gründe dieser und anderer Defizite der Extraktion aufzudecken folgt in Teil III dieser Arbeit eine detaillierte Untersuchung der Qualität der Extraktion anhand von Referenzdaten. Darüber hinaus werden dort weitere Ergebnisse vorgestellt.

Zum Abschluss von Teil II „Realisierung der Straßenextraktion“ werden die elementaren Schritte der Extraktion und ihre grundlegenden Ziele noch einmal stichpunktartig zusammengestellt. Die Implementierung der Straßenextraktion umfasst die drei Phasen (1) Kontextanalyse, (2) Extraktion sicherer Straßenobjekte und (3) Straßennetzbildung:

- (1) Ziel der **Kontextanalyse** ist die Gewinnung globaler und lokaler Hintergrundinformation, die für den weiteren Verlauf der Straßenextraktion von Bedeutung ist. Die Kontextanalyse umfasst folgende Teilschritte:
 - ◇ Segmentierung globaler Kontextgebiete: „Wald“, „Siedlung“ und „offene Landschaft“.
 - ◇ Analyse von lokalem Kontext:
 - DOM-basierte Detektion von Gebäuden und hoher Vegetation als Ausschlussregionen.
 - Bestimmung von sichttoten Räumen und Schattenbereichen mittels DOM- oder Bildinformation,
 - Detektion von Fahrzeugen und Fahrzeugkolonnen.
- (2) Mit der **Extraktion von sicheren Straßenobjekten** wird das Ziel verfolgt, durch die integrierte Nutzung von detailliertem und abstrahiertem Modell-Wissen sowie durch die Einbindung von Kontextinformation möglichst zuverlässige Teile des Straßennetzes zu extrahieren. Dies geschieht mit folgenden Teilschritten:
 - ◇ Detektion vorläufiger Straßenabschnitte aus der mittleren und niedrigen Auflösung.
 - ◇ Bestimmung von Fahrspuren durch:
 - Extraktion von Fahrspurbegrenzungen in Form von Markierungen und Straßenrändern aus hoher und mittlerer Auflösung,
 - Konstruktion von Fahrspuren aus Fahrspurbegrenzungen und aus vorläufigen Straßenabschnitten,
 - Fusion der so erhaltenen Fahrspuren mit anschließender Gruppierung.
 - ◇ Konstruktion von Fahrbahnen aus parallelen Fahrspuren mit integriertem Schließen von Lücken auf Basis von Fahrspurclustern.
 - ◇ Gruppierung der Fahrbahnen.
- (3) Eingangsinformation in die **Straßennetzbildung** sind Fahrbahnen, die in mehreren überlappenden Bildern extrahiert wurden. Ziel ist die Fusion dieser Daten und die darauf aufbauende Generierung eines möglichst zusammenhängenden Netzes. Im einzelnen sind folgende Teilschritte hierfür implementiert:
 - ◇ Fusion der Fahrbahnen aus überlappenden Bildern durch Projektion auf das DOM.
 - ◇ Iterative Netzbildung durch:
 - Generierung von Fahrbahnverknüpfungen mittels lokaler geometrischer Kriterien und globaler Pfadsuche,

- Verifikation der Verknüpfungen über gezielte Suche nach Straßenevidenz oder Kontext im Bild.
- ◇ Abschließende Netzanalyse zur Bestimmung von Straßenabschnitten und Straßen sowie zur Unterscheidung von einfachen und komplexen Kreuzungen.

Anhand eines Beispieldatensatzes wurde der Funktionsnachweis des implementierten Systems erbracht. In Teil III folgt nun eine genauere Untersuchung der Qualität der Extraktion.

Teil III:

Diskussion der Ergebnisse und Ausblick

Dieser Teil befasst sich mit der Bewertung und Diskussion der erzielten Ergebnisse. Grundlage hierfür ist ein etabliertes Schema zur externen Bewertung von Ergebnissen einer Straßenextraktion. An die kurze Darstellung dieses Schemas in Kapitel 9 schließt sich in Kapitel 10 die Diskussion und Bewertung der erzielten Extraktionsergebnisse an. Die Evaluierung von Zwischen- und Endergebnissen bezüglich Referenzdaten ermöglicht die Analyse von Stärken und Schwächen des Gesamtsystems und seiner einzelnen Komponenten. Das letzte Kapitel dieser Arbeit (Kapitel 11) fasst die wesentlichen Aspekte der Arbeit zusammen und zeigt in einem Ausblick potentielle Erweiterungen des hier verfolgten Ansatzes auf.

9 Externe Bewertung

Die externe Bewertung dient zur quantitativen Beschreibung der Qualität von Extraktionsergebnissen. Die Qualität wird durch den Vergleich der Extraktion mit einem Referenzdatensatz – also einem Datensatz, der dem gewünschten Ergebnis entspricht – ermittelt. Auch wenn die Qualitätsaussagen im strengen Sinn nur für das bei der Bewertung verwendete Datenmaterial gelten, kann mit ihnen dennoch abgeschätzt werden, wie geeignet ein Extraktionsalgorithmus für anderes Datenmaterial ist, in dem die zu extrahierenden Objekte durch ähnliche Eigenschaften wie im bewerteten Datensatz gekennzeichnet sind. Ein weiterer Vorteil einer externen Bewertung ist die Möglichkeit der systematischen Analyse von Stärken und Schwächen eines Ansatzes.

Neben den bei der Bewertung involvierten Kriterien spielen die Eigenschaften der Referenzdaten eine zentrale Rolle für die Bewertung. Daher wird in Abschnitt 9.1 zunächst auf ihre Bedeutung eingegangen, bevor die Erläuterung des in dieser Arbeit eingesetzten Bewertungsschemas folgt (Abschnitt 9.2).

9.1 Bedeutung der Referenzdaten

Referenzdaten werden häufig manuell erstellt. Deshalb zeichnen sie sich üblicherweise durch eine weit höhere Qualität als die Extraktion aus, so dass beim Vergleich von Extraktion mit Referenz oft von einer „objektiven“ Bewertung gesprochen wird. In diesem Zusammenhang sind jedoch zwei Punkte von großer Bedeutung:

- (1) Auch wenn die Referenzdaten höhere Qualität besitzen, sind sie in aller Regel *nicht* vollständig fehlerfrei. Fehler können einerseits Geometrie und Topologie von Teilstücken der Referenz betreffen, z.B. die richtige Position einer Straßenachse oder ihr korrekter Anschluss ans Netz. Andererseits kann bei der Erfassung der Referenz die Semantik einzelner Teilstücke nicht immer eindeutig geklärt werden, z.B. können Fahrspuren nahe dem Straßenrand auch Parkstreifen darstellen. (Heller et al., 1998) handhaben solche semantischen Unsicherheiten durch die Definition von so genannten „Don't-evaluate“ Regionen. Dies sind Regionen, zu deren Bewertung einige (oder alle) Kriterien ausgeschlossen werden.
- (2) Wenn die Referenzdaten unter Zuhilfenahme anderer (höherwertiger) Datenquellen erstellt wurden, kann sich ihre Erfassung auf Informationen stützen, die nicht in der Datengrundlage der Extraktion inbegriffen

sind und somit auch dem Extraktionsverfahren nicht zur Verfügung stehen. Daher muss von vornherein die Zielsetzung der externen Bewertung geklärt werden: Soll die Extraktion mit einer Referenz verglichen werden, die der Realität möglichst nahe kommt ? Oder soll die Extraktion einer Referenz gegenübergestellt werden, die auf Basis der selben Datengrundlage, aber mit einem „optimalen Verfahren“ erfasst wurde ? Zur Vermeidung dieser Problematik unterscheidet (Sörgel, 2003) „reale“ und „erfassbare“ Referenz, wobei die erfassbare Referenz von einem geübten menschlichen Operateur erstellt wurde. Dies impliziert, dass die menschliche Fähigkeit der Bildinterpretation eine obere Schranke für die erzielbare Qualität der Extraktion darstellt.

Zu einer korrekten Einschätzung der Bewertungsergebnisse ist es also zwingend erforderlich, die Zielstellung der Bewertung und die Vorgehensweise bei der Referenzdatenerfassung mit in Betracht zu ziehen. Der Schwerpunkt der in dieser Arbeit durchgeführten externen Bewertung liegt auf einem Vergleich zwischen dem implementierten Extraktionsverfahren und einem hypothetischen, optimalen Verfahren, das durch die menschliche Fähigkeit der Bildinterpretation repräsentiert wird. Die Referenzdaten sollen somit der erfassbaren Referenz entsprechen. Alle im Folgenden verwendeten Referenzdaten wurden daher auf Basis derjenigen Luftbilder erstellt, die auch der Extraktion zur Verfügung stehen. Da die Digitalisierung der Straßendaten von einem versierten, aber ortsfremden Operateur durchgeführt wurde, ist es eine realistische Annahme, dass außer der Bildinformation keine zusätzliche Information über die Szene in die Datenerfassung eingeflossen ist. Ein Indiz dafür ist auch, dass bei der Datenerfassung vermehrt semantische Unsicherheiten aufgetreten sind. Während die Referenz für Straßenachsen davon nahezu unbetroffen ist, trat die Unsicherheit bei der detaillierten Erfassung der Fahrspur-Referenz in größerem Ausmaß auf. Um diesem Umstand Rechnung zu tragen, wurden zwei Typen von Referenzdaten erstellt:

- ◊ Die **minimale Referenz**: Sie umfasst Objekte, deren Semantik sicher bestimmt werden konnte, und die deshalb auf jeden Fall extrahiert werden sollten. Das Augenmerk liegt also auf der *Vollständigkeit* der Extraktion.
- ◊ Die **maximale Referenz**: Sie beinhaltet zusätzlich all jene Objekte, deren Semantik mit geringerer Wahrscheinlichkeit festgelegt werden konnte. Hierunter fallen auch mehrdeutige Objekte wie z.B. Fahrspuren, die gleichzeitig auch als Parkraum dienen. Die maximale Referenz bildet eine Obermenge der minimalen Referenz. Sie vermindert daher das Risiko einer Unterschätzung der *Extraktionszuverlässigkeit*.

Zusammenfassend stehen für jede Extraktion vier verschiedene Referenzdaten zur Verfügung: Eine minimale und eine maximale Referenz, und für beide die digitalisierten Straßenachsen sowie ihre korrespondierenden Fahrspuren. Analog zur Bewertung der Straßenachsen erfolgt die Bewertung der Fahrspuren auf Basis der Fahrspurmittelachsen. Im folgenden Abschnitt wird in Kurzform ein etabliertes Schema zur externen Bewertung von Straßennetzen beschrieben. Details hierzu finden sich in (Wiedemann, 2001, 2003).

9.2 Bewertungsschema

Das für die externe Bewertung verwendete Verfahren evaluiert das extrahierte Straßennetz einerseits durch die Kriterien „Vollständigkeit“, „Zuverlässigkeit“ und „geometrische Genauigkeit“. Diese können durch die *lokale* Zuordnung von Teilstücken des Extraktions- und Referenznetzes berechnet werden. Andererseits werden durch die Maße „topologische Vollständigkeit“, „topologische Korrektheit“, „mittlerer Umwegfaktor“ und „mittlerer Abkürzungsfaktor“ auch die *globalen* topologischen Eigenschaften zwischen und innerhalb von Zusammenhangskomponenten des Straßennetzes bewertet.

Die Ermittlung der Bewertungsergebnisse gliedert sich in zwei Schritte: (1) Die Zuordnung von Extraktion und Referenz und (2) die Berechnung der Qualitätsmaße. Abgesehen von algorithmischen Unterschieden erfolgt die Zuordnung der beiden Straßennetze mit einem konzeptionell identischen Ansatz, wie er bei der Fusion von Fahrspuren und Fahrbahnen beschrieben wurde. Mit Hilfe eines Puffers und einer vorgegebenen Winkeltole-

ranz werden redundante – in diesem Fall besser *homologe* – Teilstücke der Netze bestimmt. Die Berechnung der Bewertungsmaße geschieht auf Basis der zugeordneten Teilstücke oder ersatzweise darauf gleichabständig definierten Punkten. Zu beachten ist jedoch, dass zwischen Zuordnung und Bewertungskriterien eine Kopplung besteht. Beispielsweise hat eine tolerantere Zuordnung eine geringere geometrische Genauigkeit der Extraktion zur Folge. Umgekehrt steigen aber durch die größere Toleranz Vollständigkeit und Zuverlässigkeit der Extraktion. Aufgrund dieser gegenseitigen Beeinflussung muss die Zuordnung als fester Bestandteil der Bewertung angesehen werden.

Die Bewertungskriterien sind wie folgt definiert:

- ◇ **Vollständigkeit:** Sie gibt an, welcher Prozentsatz der in den Referenzdaten enthaltenen Straßen auch im Extraktionsergebnis enthalten ist. Der Optimalwert beträgt 100%.

$$\text{Vollständigkeit} = \frac{\text{Länge der zugeordneten Referenz}}{\text{Länge der Referenz}} \cdot 100 \text{ [\%]}$$

- ◇ **Zuverlässigkeit:** Sie spiegelt den Prozentsatz des Extraktionsergebnisses wider, der in den Referenzdaten enthalten ist. Dieser drückt die Wahrscheinlichkeit aus, mit der ein beliebig herausgegriffenes Teilstück der Extraktion in der Referenz vorhanden ist. Der Optimalwert liegt ebenfalls bei 100%.

$$\text{Zuverlässigkeit} = \frac{\text{Länge der zugeordneten Extraktion}}{\text{Länge der Extraktion}} \cdot 100 \text{ [\%]}$$

- ◇ **Geometrische Genauigkeit:** Die geometrische Genauigkeit der extrahierten Straßenachsen wird durch den Root-Mean-Square-(RMS-)Fehler definiert. Hierzu werden auf beiden Netzen in kurzen Abständen homologe Hilfsknoten eingeführt und die Distanzen zwischen ihnen berechnet.

$$\text{RMS-Fehler} = \sqrt{\frac{\text{Quadratsumme der Distanzen zwischen homologen Knoten}}{\text{Anzahl der homologen Knoten}}} \text{ [m]}$$

- ◇ **Mittlerer Umwegfaktor:** Grundlage seiner Berechnung ist die Streckenlänge des optimalen Pfades zwischen einem Punktepaar auf dem Referenznetz (Referenzlänge) und zwischen dem homologen Punktepaar auf dem extrahierten Netz (Extraktionslänge). Sofern die Extraktionslänge *größer* als die Referenzlänge ist, definiert das Verhältnis zwischen Extraktionslänge und Referenzlänge einen Umwegfaktor. Die Berechnung ist aber nur möglich, wenn sich beide Punktepaare in gleichen Zusammenhangskomponenten befinden. Eine Mittelung der Umwegfaktoren zwischen allen paarweisen Punktombinationen ergibt den mittleren Umwegfaktor:

$$\text{Mittlerer Umwegfaktor} = \sum_{n=1}^N \frac{\text{Extraktionslänge zwischen Punktepaar } PP_n}{\text{Referenzlänge zwischen Punktepaar } PP_n} \text{ []}$$

mit N Anzahl an Punktepaaren PP . Der optimale Wert für den mittleren Umwegfaktor beträgt 1. Er vergrößert sich mit wachsender Anzahl an Lücken in der Extraktion oder mit längeren Umwegen, die die Lücken verursachen.

- ◇ **Mittlerer Abkürzungsfaktor:** Der mittlere Abkürzungsfaktor berechnet sich analog zum mittleren Umwegfaktor. Im Gegensatz zum mittleren Umwegfaktor werden jedoch jene Verhältnisse an Streckenlängen in Betracht gezogen, die eine *kürzere* Extraktionslänge als die Referenzlänge aufweisen. Der mittlere Abkürzungsfaktor erreicht sein Optimum bei 1 und fällt gegen 0 je mehr Punktepaare in der Extraktion durch eine kürzere Strecke als in der Referenz verbunden sind. Kürzere Strecken treten entweder im Fall von fehlerhaften Teilstücken oder im Fall einer stark generalisierten Extraktion auf (z.B. bei unzureichender Approximation glatter Kurven durch Polygone).
- ◇ **Topologische Vollständigkeit:** Sie bemisst den Prozentsatz an Punktepaaren in Referenz und Extraktion, die innerhalb gleicher Zusammenhangskomponenten liegen. Die topologische Vollständigkeit erreicht ihr Optimum bei 100% und fällt mit zunehmender Fragmentierung der Extraktion bezüglich der Referenz ab.

$$\textit{Topologische Vollständigkeit} = \frac{\textit{Anzahl an Punktepaaren, die in beiden Netzwerken verbunden sind}}{\textit{Anzahl an Punktepaaren, die in der Referenz verbunden sind}} \cdot 100 [\%]$$

- ◇ **Topologische Zuverlässigkeit:** Sie berechnet sich auf die gleiche Weise wie die topologische Vollständigkeit. Vergleichsmaßstab ist hier aber das extrahierte Netzwerk, so dass die topologische Korrektheit von ihrem Optimalwert 100% absinkt, je mehr Zusammenhangskomponenten der Referenz durch fehlerhafte Teilstücke in der Extraktion verbunden sind.

$$\textit{Topologische Zuverlässigkeit} = \frac{\textit{Anzahl an Punktepaaren, die in beiden Netzwerken verbunden sind}}{\textit{Anzahl an Punktepaaren, die in der Extraktion verbunden sind}} \cdot 100 [\%]$$

Auf die Angabe der Redundanz, d.h. der mehrfachen Extraktion desselben Teilstücks, wird verzichtet. Solche Teilstücke kommen in den Ergebnissen praktisch nicht vor. In (Shufelt, 1999; Harvey, 1999) wird noch eine Reihe weiterer Bewertungsmaße eingeführt. Auch sie werden für die Diskussion der Ergebnisse im folgenden Kapitel nicht herangezogen. Diese Maße lassen sich alle aus den hier verwendeten Bewertungskriterien herleiten und enthalten somit keine zusätzliche Information.

Für die Bewertung der Ergebnisse wird die Position der Straßenachse herangezogen. Die Evaluierung der im Endergebnis enthaltenen Fahrspuren wird ebenfalls durchgeführt, allerdings nur hinsichtlich ihrer Vollständigkeit, Zuverlässigkeit und geometrischen Genauigkeit. Eine Bewertung der Fahrspurtopologie erscheint nicht sinnvoll, da das Extraktionsverfahren keine Rekonstruktion der Fahrspuren in Kreuzungsbereichen vorsieht. Deshalb haben die topologischen Bewertungskriterien angewandt auf Fahrspuren keine große Aussagekraft.

10 Ergebnisse und Diskussion

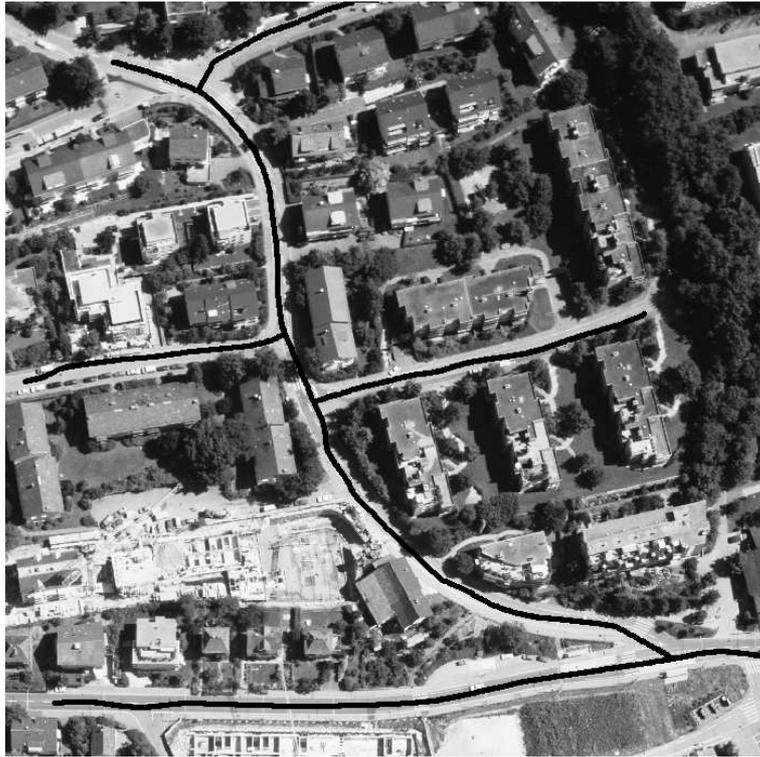
Das in dieser Arbeit entwickelte System zur Straßenextraktion wird anhand der im vorangegangenen Kapitel vorgestellten externen Bewertungskriterien evaluiert. Grundlage hierfür sind zwei Beispieldatensätze komplexer Stadtszenen. In Abschnitt 10.1 erfolgt zunächst die Bewertung und Analyse der Extraktionsergebnisse auf Basis dieser Datensätze. Darauf aufbauend werden in Abschnitt 10.2 die Charakteristiken des hier verfolgten Ansatzes anderen verwandten Arbeiten internationaler Forschergruppen gegenübergestellt und diskutiert. Die Diskussion belegt die große Leistungsfähigkeit des Systems, sie deutet jedoch auch auf einige grundlegende Einschränkungen hin, die Anlass zu Verbesserungen geben.

10.1 Beispieldatensätze und Analyse der Extraktionsergebnisse

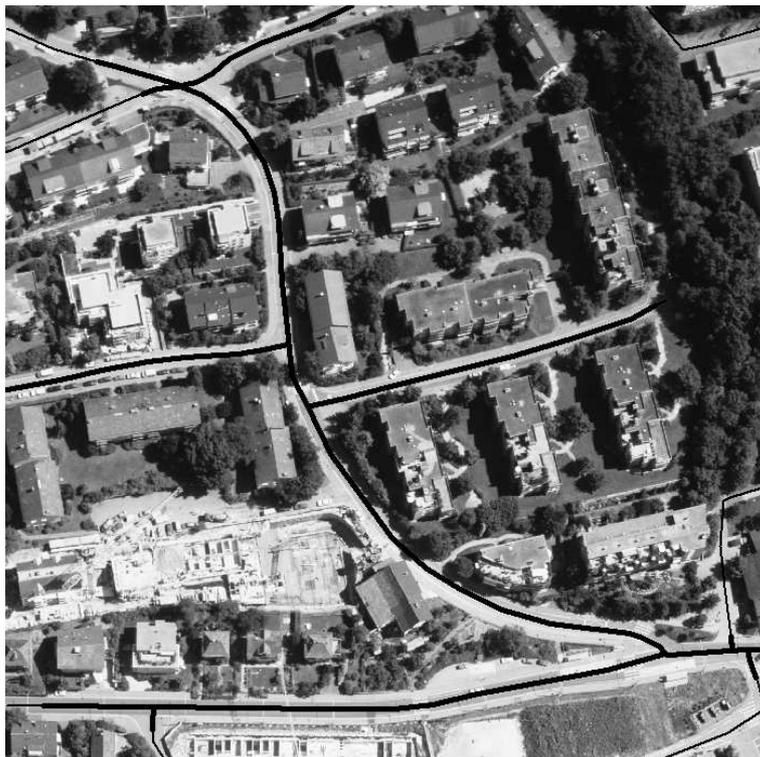
Die externe Bewertung wird für zwei verschiedene Datensätze durchgeführt. Bei *Datensatz I* handelt es sich um eine Stadtszene, die die Komplexität ländlicher und suburbaner Szenen deutlich übersteigt, die aber nicht die enorme Komplexität eines Innenstadtbereichs einer Großstadt besitzt. Die Bewertung dieses Datensatzes erfolgt nicht nur für das Endergebnis sondern auch für ausgewählte Zwischenstände der Extraktion. Anhand der Resultate lassen sich Vorteile und Defizite einzelner Verfahrensschritte charakterisieren. Anschließend wird das Extraktionsergebnis von *Datensatz II* analysiert. Verglichen mit Datensatz I weist Datensatz II eine höhere Komplexität auf. Einerseits dient dieser Datensatz zur Validierung des Verfahrens für sehr komplexe Innenstadtbereiche. Andererseits lassen sich daran aber auch Grenzen des Verfahrens aufzeigen, da die Extraktion hier generell größere Schwierigkeiten hat als in Datensatz I. Die Referenzdaten wurden durch manuelle Digitalisierung mit einem Monoplotting-vergleichbaren Verfahren erzeugt. Abgesehen von den Straßenachsen in Datensatz I, deren Semantik bei der Referenzdatenerfassung zweifelsfrei bestimmt werden konnte, wurden für Straßenachsen und Fahrspuren sowohl minimale als auch maximale Referenz erfasst.

Datensatz I: Abbildung 69 zeigt eines von drei überlappenden, hochaufgelösten Luftbilder aus Datensatz I und dort hineinprojiziert das Ergebnis von Extraktion und externer Bewertung. Abgesehen vom Baustellenbereich im linken unteren Bildquadranten wird die Szene durch eine Reihe von Einzelgebäuden und einige mittelgroße Gebäudeblöcke, die alle durch Vorgärten voneinander getrennt stehen, charakterisiert. Die Gebäude werden einerseits durch die beiden Hauptstraßen, wovon eine am unteren Bildrand und eine diagonal von links oben nach rechts unten verläuft, und durch einige davon abzweigende Wohnstraßen erschlossen. Während die Hauptstraßen helle Markierungen und Zebrastreifen aufweisen, sind bei den Wohnstraßen keine deutlichen Begrenzungen von Fahrspuren oder Parkstreifen zu erkennen. Neben dieser Variabilität des Objekts „Straße“ stellen die zahlreichen Schattenbereiche auf der diagonalen Hauptstraße, die von rechts bis in die Straßenmitte hineinreichen eine Herausforderung für die Extraktion dar. Partielle Verdeckungen und Fahrzeuge kommen dagegen weniger häufig vor.

Datensatz I, Bewertung der Straßenachsen: In Abbildung 69 a) sind die Ergebnisse der *Extraktion* und ihrer Bewertung visualisiert. Die dicken schwarzen Linien stellen die extrahierten Straßenachsen dar, die der Referenz bei einer Pufferbreite von 5m zugeordnet werden konnten. Eine nicht-zugeordnete Extraktion kommt praktisch nicht vor. Lediglich ein durch einen Randeffect bedingtes kurzes Stück am äußersten oberen Bildrand – gekennzeichnet durch eine dünne schwarze Linie – wird von der externen Bewertung als Fehlextraktion eingestuft. Abbildung 69 b) illustriert in dicker Strichstärke jene Teile der *Referenz*, die der Extraktion zugeordnet werden konnten, und in dünner Strichstärke die Referenz, die in der Extraktion fehlt. Hierbei fällt auf, dass nur solche Teile des Straßennetzes nicht extrahiert werden konnten, die zum Bildrand hinführen. Diese Erscheinung liegt in der Methodik der Straßennetzbildung begründet. Denn im Zuge der globalen Gruppierung können nur dann Straßenstücke am Bildrand extrahiert werden, wenn sie Startstück eines optimalen Pfades sind. Eine Behebung dieses Effektes stellt kein grundsätzliches Problem dar, weil die Extraktionsergebnisse benachbarter Luftbilder mit integriert werden können. Alternativ könnten Tracking-Verfahren angewendet werden, um die



(a) Extraktion, die der Referenz zugeordnet werden konnte (fett), fehlerhafte Extraktion (dünn, tritt nur am äußersten oberen Bildrand auf)



(b) Referenz, die der Extraktion zugeordnet werden konnte (fett), Referenz, die nicht extrahiert werden konnte (dünn)

Abbildung 69. Externe Bewertung der Straßenachsen aus Datensatz I: (a) Extraktion, (b) Referenz

Endpunkte der Extraktion in Richtung Bildrand fortzusetzen. Die Tatsache, dass sowohl die Hauptstraßen als auch die Wohnstraßen zum Großteil extrahiert wurden, kann als klares Indiz für die ausreichende Flexibilität des Modells gewertet werden.

Der gute visuelle Eindruck der Ergebnisse zeigt sich auch in der quantitativen externen Bewertung (Tabelle 11). Die Extraktion erreicht eine Vollständigkeit von knapp 77% und eine Zuverlässigkeit von nahe 99%. Bei Elimination der Referenz in den Randbereichen des Bildes würde sich die Vollständigkeit noch um einige Prozentpunkte erhöhen. Ein klares Defizit zeigt die Extraktion jedoch in ihrer geometrischen Genauigkeit. Ein RMS-Fehler von 1.3m bedeutet einen mittleren Lagefehler von mehr als dem Achtfachen der hohen Auflösung. Eine Hauptursache dieser vergleichsweise schlechten Lagegenauigkeit der Straßenachsen ist die Fahrbahnfusion auf Basis des nur mäßig genauen DOM. Dieser Sachverhalt wurde bei der Beschreibung der Implementierung bereits diskutiert und bewusst in Kauf genommen. Denn das Hauptziel der Fahrbahnfusion liegt auf einer korrekten semantischen Zuordnung, wobei ein Verlust in der geometrischen Genauigkeit toleriert wird. Ein weiterer Grund für die resultierende Ungenauigkeit ist das Schließen von Lücken durch Verknüpfungen, die sich nur auf geometrische Annahmen über den tatsächlichen Straßenverlauf stützen. Eine bessere Anpassung der Verknüpfungen an die Bildinformation könnte z.B. durch Ziplock-Snakes erreicht werden. Gemäß den Untersuchungen von (Baumgartner, 2003) ist jedoch zu beachten, dass eine snake-basierte Überbrückung von Lücken im Falle von sehr dunklen Schatten, wie sie z.B. von Gebäuden geworfen werden, Schwierigkeiten verursacht. Im Gegensatz zur Geometrie konnte die Topologie des Straßennetzes wesentlich besser bestimmt werden. Die guten Werte für die Netzwerk-Kriterien belegen, dass das Verfahren ein zusammenhängendes Netz extrahiert und nicht nur einen hohen Prozentsatz an Teilstücken. Lediglich das kurze isolierte Straßenstück in der rechten oberen Bildecke konnte aufgrund der bereits erwähnten Randeffekte nicht extrahiert werden (siehe Abbildung 69 b).

Externe Bewertungskriterien	Datensatz I: Endergebnis Straßenachsen
Vollständigkeit [%]	76.6
Zuverlässigkeit [%]	98.8
RMS-Fehler [m]	1.3
Umwegfaktor []	1.04
Abkürzungsfaktor []	0.95
Topologische Vollständigkeit [%]	100.0
Topologische Zuverlässigkeit [%]	96.2

Tabelle 11. Externe Bewertung für Straßenachsen aus Datensatz I

Datensatz I, Bewertung der Fahrspuren: Neben den Straßenachsen wurden auch die resultierenden Fahrspuren evaluiert. Da bei der Referenzdatenerfassung eine eindeutige Bestimmung der Semantik nicht in allen Fällen möglich war, wurden minimale und maximale Referenzdaten für Fahrspuren erstellt. Die resultierenden unterschiedlichen Bewertungen der extrahierten Fahrspuren sind in Abbildung 70 gegenübergestellt – in der linken Spalte die Bewertung mit minimaler Referenz und in der rechten Spalte mit maximaler Referenz. Die Unterschiede in den Referenzdaten sind beim Vergleich der Abbildungen 70 c) und d) ersichtlich. Während bei der minimalen Referenz die Straßen höchstens zwei Fahrspuren besitzen, weisen die diagonale Hauptstraße und die Wohnstraße im linken zentralen Bildbereich in der maximalen Referenz drei Fahrspuren auf.

Wie anhand der dicken durchgezogenen Linien in den Abbildungen 70 a) und b) zu erkennen ist, konnte die Mehrheit der extrahierten Fahrspuren der Referenz bei einer Pufferbreite von 1.5m zugeordnet werden. Dies bestätigt auch die quantitative Bewertung der Vollständigkeit in Tabelle 12 mit Werten von 68% für

die minimale Referenz und 65.2% für die maximale Referenz. Die Zuverlässigkeit der Extraktion liegt auf sehr hohem Niveau bei 90% – 95%. Wie im Falle der Straßenachsen fällt die geometrische Genauigkeit mit einem RMS-Fehler von 1m deutlich ab. Beim einem visuellen Vergleich mit der großen Genauigkeit, die die Fahrspuren bei ihrer Konstruktion in den Einzelbildern erreichen (vgl. Abbildungen in Abschnitt 7.3) wird wiederum deutlich, dass die Fusion auf Basis des nur mäßig genauen DOM der limitierende Faktor für die geometrische Genauigkeit ist.

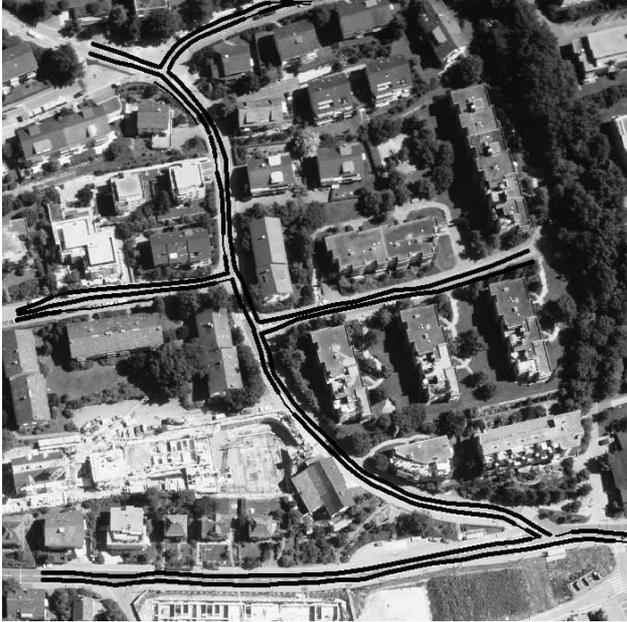
Externe Bewertungskriterien	Datensatz I: Ergebnis Fahrspuren (minimale Referenz)	Datensatz I: Ergebnis Fahrspuren (maximale Referenz)
Vollständigkeit [%]	68.0	65.2
Zuverlässigkeit [%]	89.8	95.5
RMS-Fehler [m]	1.0	1.0

Tabelle 12. Externe Bewertung für Fahrspuren aus Datensatz I

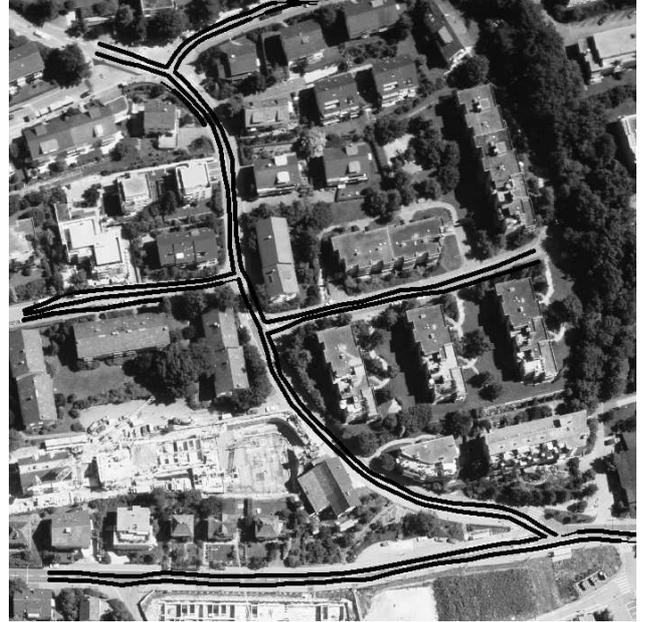
Obwohl die externen Bewertungen von Fahrspuren und Straßenachsen ein grundsätzlich konsistentes Verhalten zeigen, fällt dennoch auf, dass die Fahrspurbewertung für Vollständigkeit und Zuverlässigkeit bis zu 10% schlechter ausfällt. Die Erklärung hierfür wird aus den Abbildungen 70 a) und b) ersichtlich: Bei größeren geometrischen Abweichungen, wie sie z.B. durch das „Schlingern“ einiger Fahrspuren deutlich werden, können die extrahierten Fahrspuren aufgrund der beschränkten Pufferbreite nicht mehr der Referenz zugeordnet werden. Dies führt einerseits zu einer niedrigeren Vollständigkeit und Zuverlässigkeit und andererseits aufgrund der größeren Fragmentierung auch zu einer schlechteren topologischen Qualität. Die Alternative einer toleranteren Zuordnung durch eine größere Pufferbreite ist jedoch nicht zielführend. Aufgrund der gegenseitigen Parallelität von Fahrspuren würde eine größere Pufferbreite dazu führen, dass ein extrahierter Fahrspurabschnitt fälschlicherweise mehreren Fahrspuren in der Referenz zugeordnet würde. Trotz der Verwendung einer relativ engen Pufferbreite von 1.5m sind den Abbildungen 70 b) und d) Beispiele dieser Problematik zu entnehmen: Die diagonale Hauptstraße wurde abgesehen von einem kurzen Teilstück in der Nähe der oberen Kreuzung mit zwei parallelen Fahrspuren extrahiert. Diese wurden aber in Teilbereichen den drei parallelen Fahrspuren der maximalen Referenz zugeordnet. Die Weiterentwicklung der externen Bewertung in diese Richtung steht zwar nicht im Fokus der vorliegenden Arbeit, die Resultate der Fahrspurbewertung zeigen aber deutlich, dass bei komplexeren Modellen für die Extraktion in Zukunft auch komplexere Modelle zur externen Bewertung herangezogen werden müssen.

Datensatz I, Bewertung von Fusion und Netzbildung: Aus den obigen Analysen lässt sich schlussfolgern, dass das Verfahren zur Fusion von Fahrbahnen einen negativen Einfluss auf die geometrische Genauigkeit des Endergebnisses hat. Bei der Konzipierung des Verfahrens war jedoch der dominierende Beweggrund für die Fusion von Zwischenergebnissen aus verschiedenen Bildern, dass einerseits redundante Information aus mehreren Bildern zur gegenseitigen Unterstützung und andererseits supplementäre Information zur Erhöhung der Vollständigkeit genutzt werden sollte. Daher wird in der folgenden Untersuchung der Verlauf bestimmter Bewertungsmaße für drei ausgewählte (Zwischen-)Ergebnisse der Extraktion analysiert: für die Eingangsinformation der Fusion aus den einzelnen Bildern, für das Ergebnis nach der Fusion und schließlich für das Endergebnis nach der Straßennetzbildung. Abbildung 71 illustriert die externe Bewertung der drei Eingangsdatensätze in die Fusion sowie die Bewertung des Fusionsergebnisses. Zur besseren Vergleichbarkeit wurden die Ergebnisse aus den unterschiedlichen Bildern in das gleiche Luftbild projiziert. Die Überlagerung mit den Originaldaten zeigt Abbildung 61.

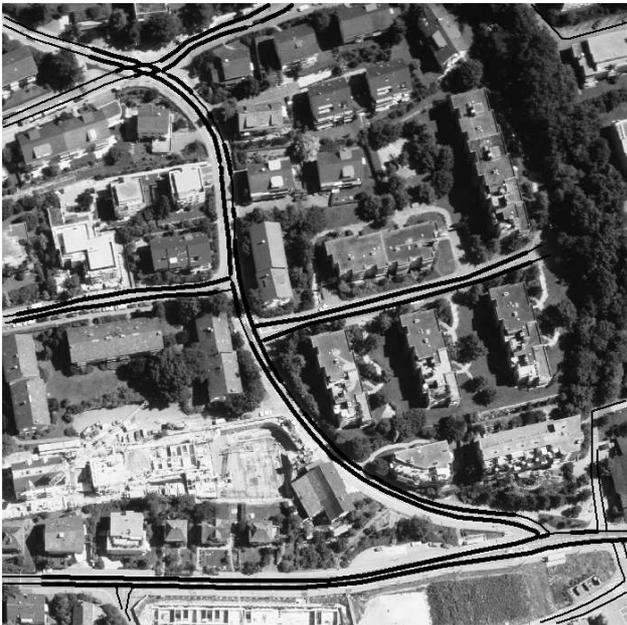
In Tabelle 13 sind die externen Bewertungen für die extrahierten Fahrbahnachsen vor und nach der Fusion zusammengestellt. Während die Vollständigkeit der Zwischenergebnisse für die einzelnen Bilder zwischen 45% und 70% liegt, erreicht der fusionierte Datensatz einen Wert von knapp 77%. Auf der einen Seite unterstreicht



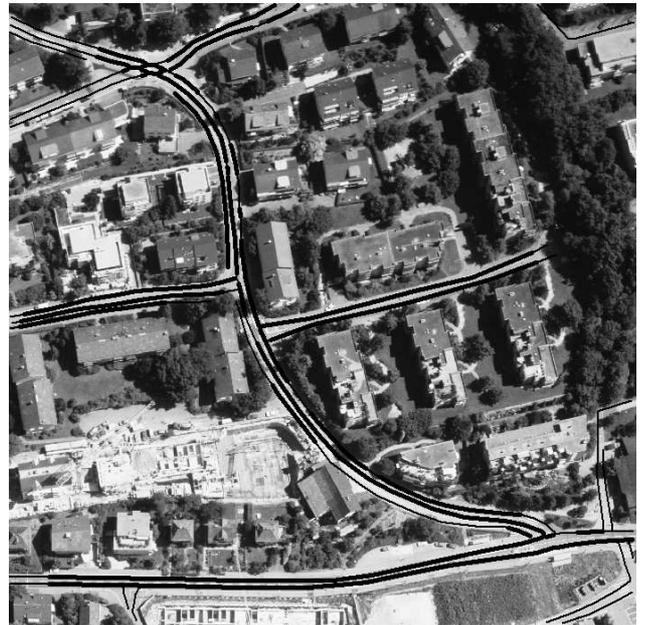
(a) Zuordnung der Extraktion zur minimalen Referenz



(b) Zuordnung der Extraktion zur maximalen Referenz

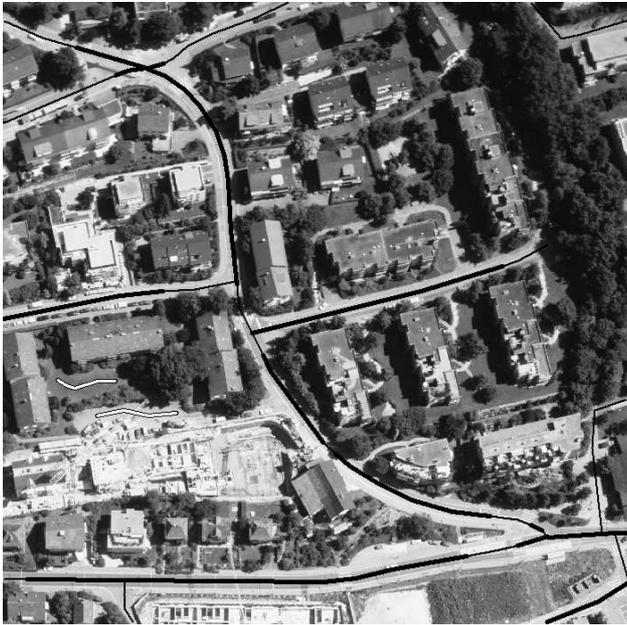


(c) Zuordnung der minimalen Referenz zur Extraktion

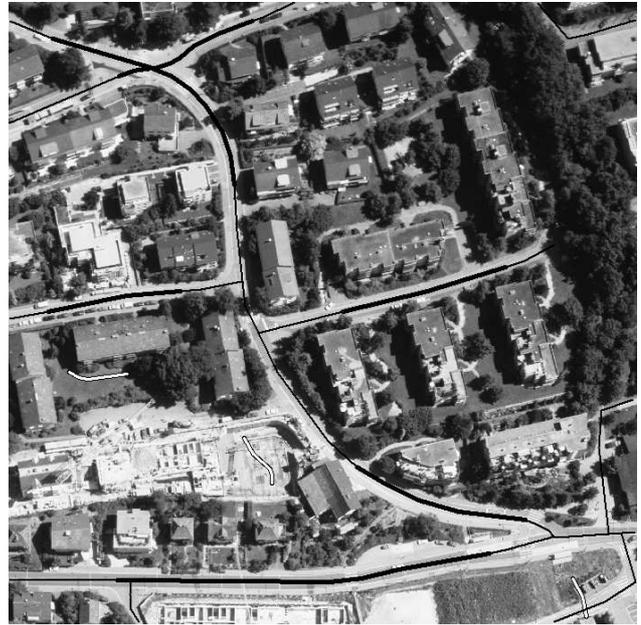


(d) Zuordnung der maximalen Referenz zur Extraktion

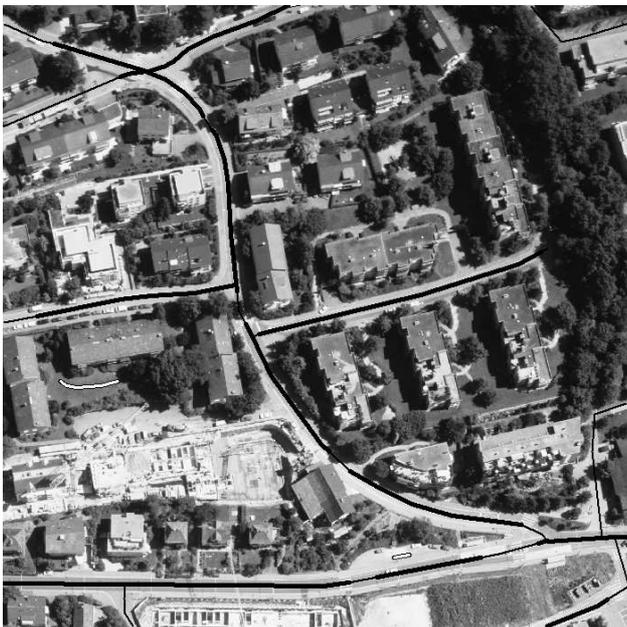
Abbildung 70. Externe Bewertung der Fahrspuren aus Datensatz I: (a),(b) Zuordnung von Extraktion zu Referenz, (c) Zuordnung von minimaler Referenz, (d) von maximaler Referenz zu Extraktion. Erfolgreiche Zuordnung (dick) und fehlende Zuordnung (dünn).



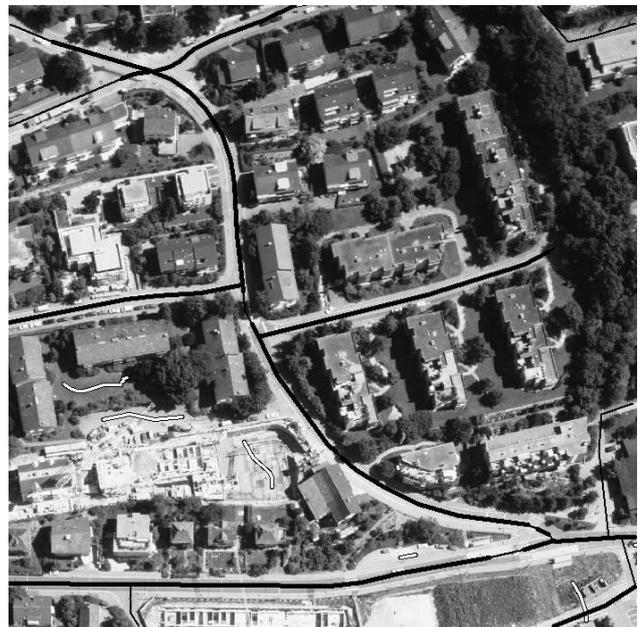
(a) Externe Bewertung der Fahrbahnextraktion aus Bild 1



(b) Externe Bewertung der Fahrbahnextraktion aus Bild 2 (dargestellt in Bild 1)



(c) Externe Bewertung der Fahrbahnextraktion aus Bild 3 (dargestellt in Bild 1)



(d) Externe Bewertung für Ergebnis der Fahrbahnfusion

Abbildung 71. Externe Bewertung für die Eingangsinformation zur Fahrbahnfusion (a – c) und für Ergebnis der Fusion (d): Referenz, die der Extraktion zugeordnet werden konnte (schwarz, dick), die nicht zugeordnet werden konnte (schwarz, dünn) und fehlerhafte Extraktion (weiß, dick, schwarz umrandet).

dies den Gewinn an Vollständigkeit durch die Fusion, weil sich die Zwischenergebnisse teilweise ergänzen. Auf der anderen Seite zeigt die große Schwankungsbreite zwischen 45% und 70% Vollständigkeit vor der Fusion, dass die Zwischenergebnisse z.B. aufgrund ungünstiger Parametereinstellungen stark variieren können. Der Einfluss kann durch die hier verwendete Methodik der Fusion – eine Vereinigungsoperation – gedämpft werden, weil diese im Sinne der Vollständigkeit das bestmögliche Ergebnis erzielt. Da aber durch die Vereinigung auch die Menge an fehlerhaften Extraktionsergebnissen größer werden kann, sollten die Eingangsdaten der Fusion eine hohe Zuverlässigkeit besitzen. Die hohen Werte von etwa 87% bis über 91% Zuverlässigkeit untermauern, dass das bis zur Fusion genutzte Modellwissen umfassend genug ist, um zuverlässige Eingangsdaten für die Fusion zu liefern. Dass dies bei der gewählten Fusionsmethode grundsätzlich auch nötig ist, verdeutlicht der Rückgang auf etwa 83% Zuverlässigkeit nach der Fusion. In Abbildung 71 sind inkorrekte Fahrbahnen als weiße Linien mit schwarzem Rand gekennzeichnet und finden sich hauptsächlich in der Umgebung des Baustellenbereiches im linken unteren Bildquadranten.

Externe Bewertungskriterien	Datensatz I: Bild 1	Datensatz I: Bild 2	Datensatz I: Bild 3	Datensatz I: Fusion
Vollständigkeit [%]	61.9	45.1	70.2	76.6
Zuverlässigkeit [%]	91.1	86.8	89.4	83.2
RMS-Fehler [m]	1.19	2.2	1.82	1.6
Umwegfaktor []	1.01	1.0	1.0	1.0
Abkürzungsfaktor []	0.96	0.85	0.97	0.96
Topologische Vollständigkeit [%]	19.2	23.8	12.4	12.8
Topologische Zuverlässigkeit [%]	95.1	100.0	97.3	98.8

Tabelle 13. Externe Bewertung für Zwischenschritte der Extraktion aus Datensatz I.

Vergleicht man die Werte für Vollständigkeit und Zuverlässigkeit zwischen dem Ergebnis nach der Fusion (Tabelle 13) und dem Endergebnis (Tabelle 11), so fällt auf, dass bei konstanter Vollständigkeit die Zuverlässigkeit von etwa 83% auf knapp 99% ansteigt. Das bedeutet, dass im Rahmen der Straßennetzbildung quasi der gesamte Anteil an fehlerhafter Eingangsinformation eliminiert und durch einen längenmäßig gleich großen Anteil an korrekten Straßenstücken ersetzt wird. Dies ist ein wichtiges Indiz dafür, dass die vor und während der Straßennetzbildung involvierten internen Bewertungen plausible Werte liefern. Denn sowohl die individuelle Gewichtung der einzelnen Fahrbahnen und Fahrbahnverknüpfungen als auch die Auswahl der Startstücke für die globale Gruppierung stützen sich auf die Resultate der internen Bewertung. Auch in diesem Zusammenhang liefert die Fusion einen bedeutsamen Beitrag, weil mehrfach extrahierte Fahrbahnen ein höheres Gewicht erhalten und dadurch mit höherer Wahrscheinlichkeit Teilstück eines optimalen Pfades bzw. des Endergebnisses sind.

Da im Zuge der Erstellung der Referenzdaten auch die Höhen aus dem DOM verwendet wurden, hat der RMS-Fehler vor der Fusion keine große Aussagekraft. Hervorzuheben ist aber, dass der RMS-Wert durch die Straßennetzbildung von 1.6m nach der Fusion auf 1.3m im Endergebnis sinkt. Die neu eingefügten Verknüpfungen des Straßennetzes haben also eine höhere Genauigkeit als die eliminierten Fahrbahnen. Der größte Nutzen der Straßennetzbildung zeigt sich in der topologischen Vollständigkeit. Vor der globalen Gruppierung bewegt sich ihre Bewertung zwischen 12% und 24%, am Ende wird jedoch ein komplett zusammenhängendes Straßennetz mit einer topologischen Vollständigkeit von 100% erreicht. Die Ergebnisse der Bewertungen für topologische Zuverlässigkeit, Umweg- und Abkürzungsfaktor liegen nahe an ihren Optimalwerten und entsprechen damit den erwarteten Werten. Denn aufgrund der großen Fragmentierung der Fahrbahnen vor der Straßennetzbildung können kaum Fehlverbindungen oder keine besonders langen Umwege und Abkürzungen auftreten. Die Tatsache, dass der Abkürzungsfaktor in der Regel etwas mehr vom Idealwert 1 abweicht als der Umwegfak-

tor, kann mit der Polygonapproximation nach (Ramer, 1972) begründet werden. Bei dieser Methode sind die approximierenden Polygone immer kürzer als die Eingangsdaten.

Datensatz II: Zur Validierung der Ergebnisse und Bewertungen von Datensatz I wurde ein zweiter Datensatz herangezogen. Das den Ergebnissen in Abbildung 72 hinterlegte Luftbild zeigt eine komplexe Innenstadtszene. Die Häuser stehen sehr eng beisammen und bilden oft zusammenhängende Gebäudekomplexe, die bis an den Straßenrand grenzen. Nur vereinzelt findet man wie am oberen Bildrand frei stehende Gebäude mit Vorgärten, deren Vegetation bis zur Straße reicht. Das Straßennetz ist gekennzeichnet durch die beiden Hauptstraßen, die oberen und unteren bzw. rechten und linken Bildrand verbinden und sich in einer komplexen Kreuzung mit ausgedehntem Kreuzungsbereich schneiden. Während die Fahrspuren der Hauptstraßen zur Straßenmitte hin mit Markierungen begrenzt sind, können die Straßenränder meist nur über die Präsenz von Gehsteigen, Parkstreifen oder überhaupt nicht eindeutig identifiziert werden. Von den Hauptstraßen zweigen insgesamt drei größere Nebenstraßen ab, von denen nur die horizontal zum linken Bildrand verlaufende mit Markierungen versehen ist. Die Nebenstraße im linken oberen Quadranten und jene entlang des rechten Bildrandes sind hauptsächlich aufgrund von Umgebungsinformation zu erkennen, aber kaum anhand straßentypischer Merkmale. Anhaltspunkte sind z.B. die Unterbrechung der Bebauung, der daraus resultierende Schattenwurf und der mit Markierungen und Zebrastrifen versehene Kreuzungsbereich der Hauptstraßen. Neben den genannten Straßen wurde noch eine kurze, vertikal verlaufende Stichstraße in der Bildmitte in die minimale Referenz aufgenommen (siehe auch Abbildung 72 c). Die maximale Referenz umfasst zusätzlich noch eine längere, vertikal in der Bildmitte verlaufende Gasse (Abbildung 72 d). Aus den Bilddaten ist es nicht eindeutig ersichtlich, ob diese Gasse eine sehr schmale Straße oder z.B. ein Fußweg ist, der von Anliegern zum Be- und Entladen benutzt werden darf.

Die beschriebene Szene ist sehr herausfordernd für die Straßenextraktion: Ein Hauptgrund ist, dass die Straßen eine sehr unterschiedliche Erscheinungsform haben (unterschiedliche Anzahl an Fahrspuren, mit / ohne Markierungen, oft keine deutlichen Straßenränder, etc.). Ein weiterer Grund ist das dichte Verkehrsaufkommen, hauptsächlich auf der horizontal verlaufenden Hauptstraße. Die Fahrzeug- und Kolonnenextraktion spielt deshalb eine wichtige Rolle. Schließlich erschweren die Extraktion noch große Schattenregionen im mittleren und unteren Bereich der vertikal verlaufenden Hauptstraße sowie einige Verdeckungen durch Gebäude z.B. links der komplexen Kreuzung.

Datensatz II, Bewertung der Straßenachsen: Die Visualisierung der Bewertung in Abbildung 72 und ihre quantitative Erfassung in Tabelle 14 verdeutlichen, dass die Straßenextraktion angesichts der skizzierten Herausforderungen mit knapp 82% Vollständigkeit bei Vergleich mit der minimalen Referenz und 71% bei maximaler Referenz ein passables Ergebnis für die resultierenden Straßenachsen liefert. Fehlende Abschnitte sind in den Abbildungen 72 c) und d) mit dünner Strichstärke gekennzeichnet. Die auffälligsten Lücken in der Extraktion sind die Nebenstraße am rechten Bildrand und – bezogen auf die maximale Referenz – die schmale Gasse in der Bildmitte. Während bei letzterer mehrere Umstände für das Scheitern der Extraktion verantwortlich sind (z.B. die geringe Breite, die starken Abschattungen, das unzureichende Abbild im DOM, etc.), liegt der Hauptgrund für die fehlende Extraktion am rechten Bildrand am Mangel eines zuverlässigen Startstückes für die Pfadberechnung bei der Straßennetzbildung. Bei Integration von Ergebnissen benachbarter Ausschnitte vor der Pfadberechnung bestünde aber die Chance, solch einen Straßenabschnitt zu extrahieren. Dass grundsätzlich auch ein längerer, problematischer Abschnitt extrahiert werden kann, zeigt z.B. der mittlere Abschnitt der korrekt extrahierten vertikalen Hauptstraße in Abbildung 72 a).

Der Nutzen der umfassenden Modellierung drückt sich insbesondere durch die hohe Zuverlässigkeit von 95% aus. Allerdings fällt die geometrische Genauigkeit mit 2.5m wieder deutlich ab. Wie aus Abbildung 72 ersichtlich, ist der Hauptgrund hierfür die geometrisch unzureichende Bestimmung von Straßenachsen in Kreuzungsbereichen und Schattenregionen. Die beiden oberen Graphiken 72 a) und b) zeigen in dicker Strichstärke die extrahierten Straßenachsen, die als korrekt bewertet wurden, und die beiden unteren Graphiken 72 c) und d) zeigen die entsprechende Referenz. Beim Vergleich der Kreuzungs- und Schattenbereiche in der jeweils oberen und unteren Graphik ist zu erkennen, dass die Straßenachsen und Kreuzungspunkte teilweise um mehr als eine

Fahrspurbreite versetzt sind, was letztendlich zu dem großen RMS-Fehler führt.

Positiv hervorzuheben sind wie schon bei Datensatz I die Netzwerkbewertungen, die für die minimale Referenz nahe an die Optimalwerte heranreichen. Da die maximale Referenz eine zusätzliche Verbindung innerhalb einer Zusammenhangskomponente beinhaltet, die nicht extrahiert werden konnte, verschlechtern sich Umwegfaktor und topologische Vollständigkeit verglichen zur Bewertung mit minimaler Referenz.

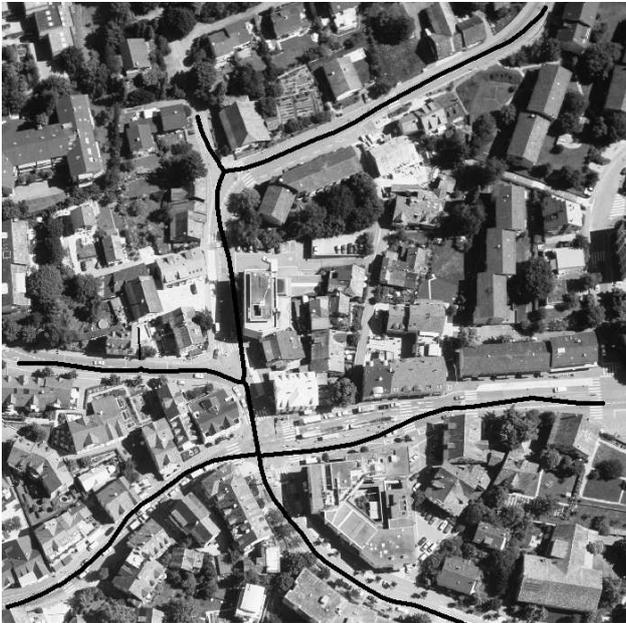
Externe Bewertungskriterien	Datensatz II: Ergebnis Straßenachsen (minimale Referenz)	Datensatz II: Ergebnis Straßenachsen (maximale Referenz)
Vollständigkeit [%]	81.6	71.0
Zuverlässigkeit [%]	95.0	95.0
RMS-Fehler [m]	2.5	2.5
Umwegfaktor []	1.05	1.22
Abkürzungsfaktor []	0.95	0.95
Topologische Vollständigkeit [%]	84.0	76.2
Topologische Zuverlässigkeit [%]	100.0	98.2

Tabelle 14. Externe Bewertung für Straßenachsen aus Datensatz II

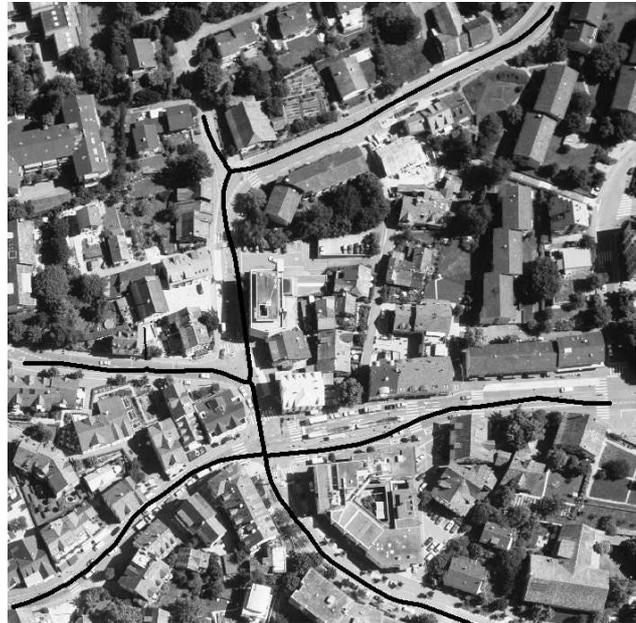
Externe Bewertungskriterien	Datensatz II: Ergebnis Fahrspuren (minimale Referenz)	Datensatz II: Ergebnis Fahrspuren (maximale Referenz)
Vollständigkeit [%]	61.3	55.9
Zuverlässigkeit [%]	75.2	76.9
RMS-Fehler [m]	1.0	1.0

Tabelle 15. Externe Bewertung für Fahrspuren aus Datensatz II

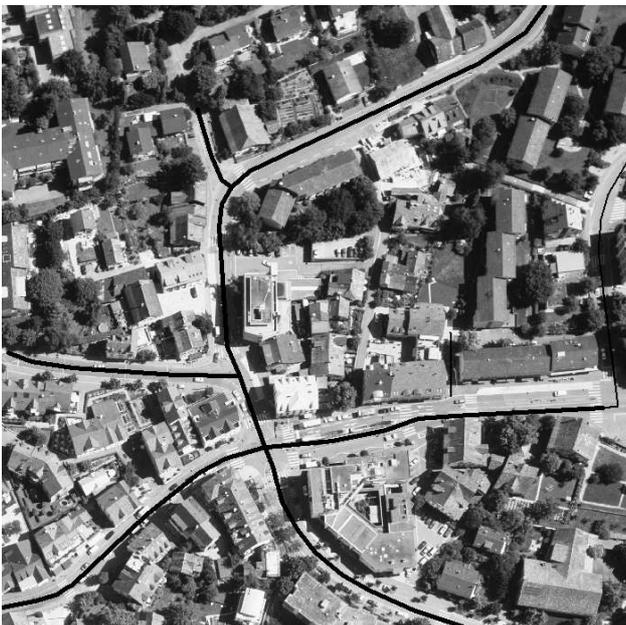
Datensatz II, Bewertung der Fahrspuren: Auch für Datensatz II wurde eine externe Bewertung der resultierenden Fahrspuren durchgeführt. Tabelle 15 ist zu entnehmen, dass die Bewertungen mit ca. 56% – 61% für Vollständigkeit und ca. 75% – 77% für Zuverlässigkeit nicht die Güte der Straßenachsen erreichen und sich damit der gleiche Trend wie bei Datensatz I zeigt. Als grundlegendes Manko erweist sich hier wiederum die geometrische Genauigkeit der Fahrspuren. Analog zur Analyse der Fahrspurbewertung für Datensatz I gelten hier die gleichen Aussagen bezüglich der Fahrbahnfusion auf dem DOM und der erweiterungsbedürftigen externen Bewertung. Allerdings fällt auf, dass die Differenz zwischen den Bewertungen von Straßenachsen und Fahrspuren bei Datensatz I etwa 10 Prozentpunkte und bei Datensatz II 15 Prozentpunkte und mehr beträgt. Wie man Abbildung 73 entnehmen kann, ist die komplexe Kreuzung im zentralen Bildbereich hauptverantwortlich für diese hohe Differenz. Während abseits dieser Kreuzung das Ergebnis qualitativ etwa dem von Datensatz I entspricht, können vor allem die Fahrspuren des von rechts in die Kreuzung einmündenden Straßenabschnitts nur teilweise rekonstruiert werden. Der Grund für das Scheitern der Fahrspurextraktion liegt in der großen Anzahl an Lastkraftwagen, die vor dem eigentlichen Kreuzungsbereich stehen (siehe z.B. Abbildung 73 a). Das Erscheinungsbild der Lastkraftwagen weicht zu sehr vom Modell der Fahrzeugdetektion ab, so dass eine automatische Erkennung nicht möglich ist. Wie nutzvoll eine erfolgreiche Fahrzeugdetektion für die Straßenextraktion ist, zeigt die von links in den Kreuzungsbereich einmündende Straße. Dank der Detektion der Fahrzeugkolonne vor der Ampel können die Fahrspuren dieses Straßenabschnitts korrekt extrahiert werden



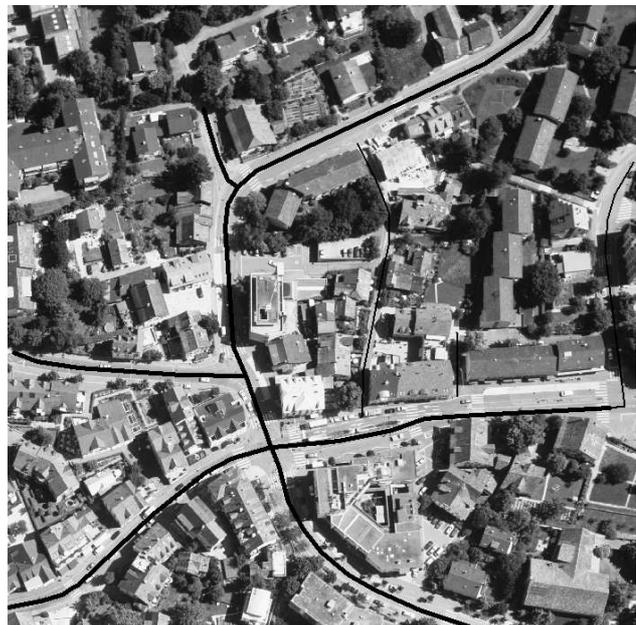
(a) Zuordnung der Extraktion zur minimalen Referenz



(b) Zuordnung der Extraktion zur maximalen Referenz



(c) Zuordnung der minimalen Referenz zur Extraktion



(d) Zuordnung der maximalen Referenz zur Extraktion

Abbildung 72. Externe Bewertung der Straßenachsen aus Datensatz II: (a),(b) Zuordnung von Extraktion zu Referenz, (c) Zuordnung von minimaler Referenz, (d) von maximaler Referenz zu Extraktion. Erfolgreiche Zuordnung (dick) und fehlende Zuordnung (dünn).

(siehe auch Abbildung 32 b für eine detaillierte Ansicht). Auch wenn in Stadtgebieten eine derartige Häufung von Lastkraftwagen wie in diesem Beispiel eher selten sein dürfte, sollte die Fahrzeugdetektion in Zukunft dennoch auf Lastkraftwagen ausgeweitet werden.

Die Analyse der Ergebnisse **abschließend** bleiben folgende Punkte festzuhalten:

- ◇ Das Extraktionssystem ist auch bei sehr komplexen Szenen in der Lage, das urbane Straßennetz auf Basis der Straßenachsen mit hoher Zuverlässigkeit, zufriedenstellender Vollständigkeit und mit sehr gut rekonstruierter Topologie zu extrahieren. Die Straßenextraktion ist sowohl bei Straßen mit sehr unterschiedlichen Merkmalen erfolgreich als auch bei Straßenabschnitten, deren Erscheinungsbild stark von Kontextobjekten beeinflusst ist.
- ◇ Die Lagegenauigkeit der resultierenden Straßenachsen und Fahrspuren weist Mängel auf. Diese werden größtenteils von der DOM-basierten Fahrbahnfusion verursacht. Weiterhin haben geometrische Ungenauigkeiten den Nebeneffekt, dass insbesondere die Ergebnisse der Fahrspurextraktion durch die externe Bewertung tendenziell unterschätzt werden, weil hier im Vergleich zum Extraktionssystem ein zu einfaches Straßenmodell zugrunde liegt.
- ◇ Komplexe Kreuzungen können auf Basis der Straßenachsen extrahiert werden. Die Rekonstruktion der Fahrspuren in Kreuzungsbereichen ist jedoch mangelhaft. Da vor Kreuzungen häufig dichter Verkehr herrscht, muss neben einer detaillierteren Fahrspurrekonstruktion im Kreuzungsbereich auch die Fahrzeug- und Kolonnenextraktion erweitert werden.

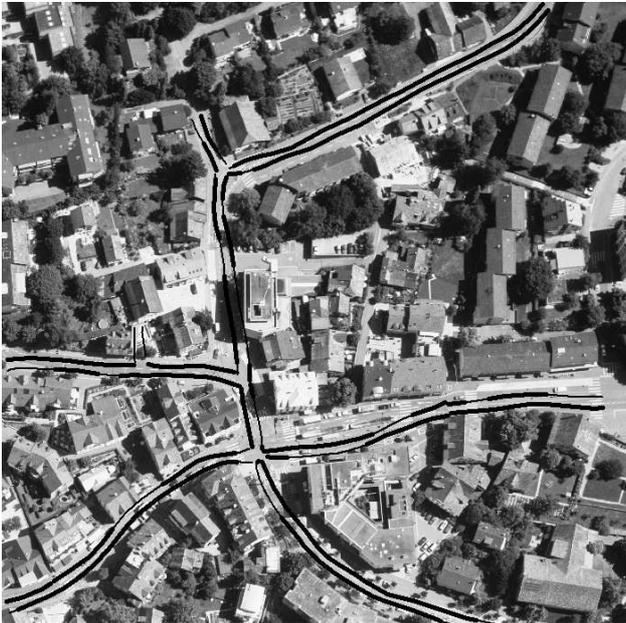
Auf Basis der durchgeführten externen Bewertungen werden im folgenden Abschnitt allgemeine Stärken und Schwächen dieses Ansatzes diskutiert und Lösungsvorschläge für die angesprochenen Mängel skizziert. Darüber hinaus werden andere Forschungsarbeiten auf dem Gebiet der Straßenextraktion dem entwickelten System gegenübergestellt und die jeweiligen Charakteristiken diskutiert.

10.2 Allgemeine Diskussion und Vergleich mit anderen internationalen Arbeiten

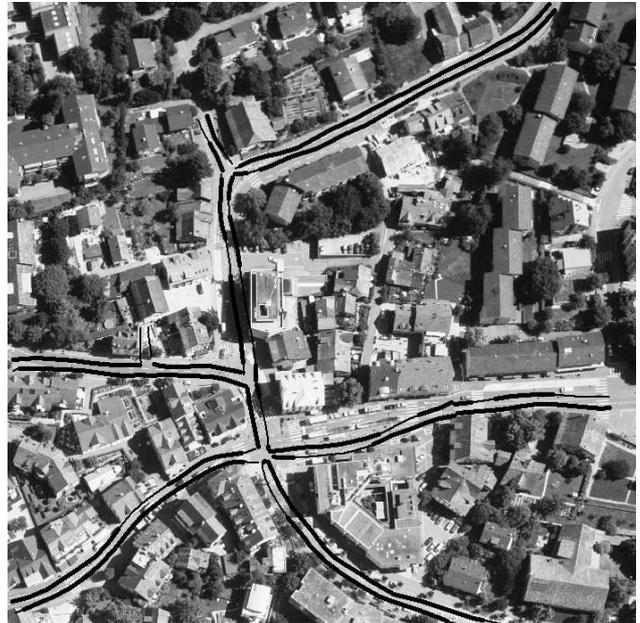
Aus der Analyse der Ergebnisse des vorangegangenen Abschnitts lässt sich generell schlussfolgern, dass Modell und Strategie des in dieser Arbeit entwickelten Extraktionssystems ausreichend an die Erfordernisse komplexer Umgebungen angepasst sind. Je nach Bewertungskriterium werden zufriedenstellende bis sehr gute Ergebnisse erzielt. Die Grundlage für die hohe Zuverlässigkeit des Endergebnisses wird durch die detaillierte Modellierung von Teilobjekten wie Markierungen und Fahrspuren in der hohen Auflösung gelegt, weil hierdurch von Beginn der Extraktion an die Anzahl an fehlerhaften Hypothesen gering gehalten wird. Belegt wird dies durch die relativ hoch bewertete Zuverlässigkeit von Zwischenergebnissen.

Die Modellierung von Fahrbahnen als zentrales Objekt der mittleren Auflösung dient einerseits zur Aggregation detaillierter Information aus der hohen Auflösung. Andererseits können Fahrbahnen in Form von homogenen Bändern auch ohne die Information aus der hohen Auflösung extrahiert werden. Dies ist von entscheidender Bedeutung für die Vollständigkeit des Ergebnisses, weil die Details der hohen Auflösung nicht immer extrahierbar bzw. in etlichen Fällen gar nicht vorhanden sind. Nur wegen der dadurch ermöglichten Flexibilität konnten in den Beispieldatensätzen nicht nur Hauptstraßen, sondern auch Wohnstraßen extrahiert werden. Die in solchen Fällen nachträglich durchgeführte Schätzung der Fahrspuren dient der Homogenisierung des Ergebnisses unter der Vorgabe, dass jede Straße und Fahrbahn aus einer gewissen Anzahl an parallelen Fahrspuren aufgebaut sein soll. Die tatsächliche Position der Fahrspuren ist dann nur ungenau bestimmt. In einigen Fällen bleibt es offen, ob in der Realität die Unterteilung einer Fahrbahn in einzelne Fahrspuren überhaupt vorliegt.

Der Hauptnutzen der Modellierung der Netzwerkeigenschaften von Straßen wird vor allem aus der erheblichen Verbesserung der Topologie des Straßennetzes ersichtlich. Wie anhand der guten Netzwerkbewertung zu sehen



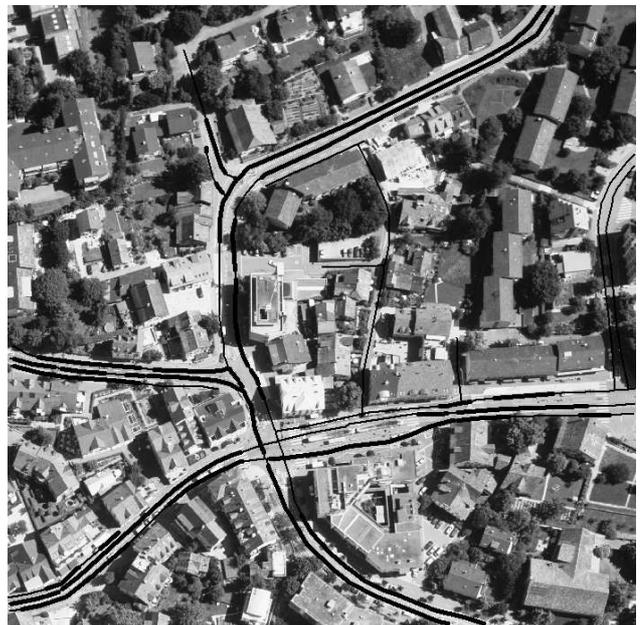
(a) Zuordnung der Extraktion zur minimalen Referenz



(b) Zuordnung der Extraktion zur maximalen Referenz



(c) Zuordnung der minimalen Referenz zur Extraktion



(d) Zuordnung der maximalen Referenz zur Extraktion

Abbildung 73. Externe Bewertung der Fahrspuren aus Datensatz II: (a),(b) Zuordnung von Extraktion zu Referenz, (c) Zuordnung von minimaler Referenz, (d) von maximaler Referenz zu Extraktion. Erfolgreiche Zuordnung (dick) und fehlende Zuordnung (dünn).

ist, ergibt sich erst durch die Integration dieses Wissens ein Straßennetz, das durchgängige Verbindungen zwischen weit entfernten Punkten auf dem Netz ermöglicht.

Abgesehen von sehr schmalen Gassen, die für die meisten Anwendungen eine eher untergeordnete Rolle spielen, sind bei der Straßenextraktion kaum modellbedingte Schwierigkeiten aufgetreten. Dies unterstreicht die Tauglichkeit des entwickelten Straßenmodells zur Extraktion der überwiegenden Mehrheit an urbanen Straßen¹⁹. Dagegen zeigt die Modellierung von Kreuzungen – insbesondere von komplexen Kreuzungen – einige Defizite. Zwar werden die meisten Kreuzungen im Rahmen der Netzbildung durch die Gruppierung von Fahrbahnen mit korrekter Topologie extrahiert, für eine detaillierte Rekonstruktion komplexer Kreuzungen reicht das Modell jedoch nicht aus. Hierfür müsste die Modellierung des Kreuzungsbereiches z.B. hinsichtlich Markierungen und baulichen Trennelementen ausgedehnt werden. Das Kreuzungsmodell sollte auch die in den Kreuzungsbereich einmündenden Straßen miteinbeziehen, da in vielen Fällen typische Straßeneigenschaften wie Anzahl der Fahrspuren oder maximal erlaubte Krümmung einer Straße in enger Beziehung zur Kreuzung stehen. Dies verdeutlicht, dass die Rekonstruktion von Kreuzungen eine umfassende Thematik ist, die in Zukunft näher untersucht werden sollte. Das Straßennetz, wie es bei dem in dieser Arbeit verfolgten Ansatz extrahiert wird, kann dabei als wertvolle Eingangsinformation für die Kreuzungsrekonstruktion dienen.

Die Extraktionsstrategie mit ihrer Dreiteilung in eine vorgeschaltete Kontextanalyse, eine Phase der Extraktion sicherer Straßenobjekte und eine Phase der Straßennetzbildung hat sich einerseits als robust gegenüber Störungen und andererseits als flexibel hinsichtlich der jeweiligen kontextuellen Situation erwiesen. Große Unterstützung für einen effizienten Ablauf leistet die systematische Verwendung interner Bewertungen von Zwischenergebnissen. Belegt wird die Leistungsfähigkeit vor allem durch die guten Ergebnisse, die für die hochkomplexe Szene von Datensatz II erzielt wurden. Aus Sicht einer geometrisch exakten Rekonstruktion stellt die resultierende geometrische Genauigkeit des Straßennetzes jedoch einen Schwachpunkt des Verfahrens dar. Der Hauptgrund dafür liegt in der mangelnden Genauigkeit des zur Fusion von Fahrbahnen herangezogenen Höhenmodells. Dieses Manko wurde bewusst in Kauf genommen, weil der Schwerpunkt der Fusion zunächst auf die semantisch korrekte Zuordnung der Fahrbahnen gelegt wurde, um die Zuverlässigkeit der Zwischenergebnisse nicht wesentlich zu beeinträchtigen. Der Nutzen dieser Vorgehensweise konnte durch die externe Bewertung von Zwischenergebnissen vor und nach der Fusion verdeutlicht werden.

Zur Verbesserung der geometrischen Genauigkeit gibt es verschiedene Alternativen. Eine davon wäre die Verwendung eines sehr genauen Höhenmodells, wie es aus Laserscanning-Daten oder hochredundanten Bilddaten abgeleitet werden kann. Aus Kostengründen dürfte vor allem die Verwendung von hochredundanten Bilddaten in Zukunft sehr erfolgversprechend sein (Gruber et al., 2003). Eine zweite Alternative zur Verbesserung der Genauigkeit wäre die Erweiterung der Extraktionsstrategie im Hinblick auf eine 3D-Rekonstruktion durch ausgefeilte Zuordnungsverfahren für Fahrbahnen oder Fahrspuren. Aufgrund der Komplexität von Stadtgebieten erscheint es jedoch schwierig, die Rate an Fehlzuordnungen gering und damit die Zuverlässigkeit der Extraktion auf hohem Niveau zu halten. Erfolgversprechender dürfte die dritte Alternative sein, die Achsen des resultierenden Straßennetzes als Näherungswerte für eine nachträgliche, verfeinerte Bestimmung der Geometrie in den Originalbildern heranzuziehen. Dies hätte darüber hinaus den praktischen Vorteil, dass die Rekonstruktion ohne Anpassungsmaßnahmen direkt auf das implementierte Extraktionssystem aufsetzen könnte.

Eine gewisse Problematik, die aus heutiger Sicht allen umfangreichen Bildanalyse-Systemen inhärent ist, stellt die korrekte Wahl von Parametereinstellungen für verschiedene Algorithmen und Bewertungsfunktionen dar. Dies gilt auch für das hier entwickelte System zur Straßenextraktion. In der Erprobungsphase des Systems hat sich herausgestellt, dass Parameter, die Geometrie und Homogenität von Straßen beschreiben oder bewerten, recht einfach eingestellt werden können und kaum einer Anpassung für verschiedene Bilddaten bedürfen. Kritischer sind dagegen Parameter, die in enger Beziehung zu den Kontrasteigenschaften des Bildes stehen, wie z.B. Parameter zur Linien- und Kantenextraktion. Ihre Einstellung muss gewährleisten, dass eine ausreichende Anzahl an anfänglichen Hypothesen für Fahrspuren und Fahrbahnen generiert werden kann und keine

¹⁹ Der Sonderfall von Stadtautobahnen mit baulich getrennten Fahrbahnen wurde nicht in Betracht gezogen.

Untersegmentierung zu Beginn der Extraktion stattfindet. Eine Übersegmentierung – also die Generierung tendenziell zu vieler anfängliche Hypothesen – wirkt sich weniger kritisch auf das Endergebnis aus, da die meisten Gewichtungsfunktionen der internen Bewertung unabhängig von solchen kritischen Parametern sind und deshalb die Ergebnisse der internen Bewertung eine zuverlässige Grundlage für die Auswahl an korrekten Hypothesen ermöglichen.

Um zum Abschluss dieses Kapitels die Stärken und Schwächen des vorgestellten Ansatzes im internationalen Vergleich herauszuarbeiten, wird er ausgewählten, nahestehenden Arbeiten gegenübergestellt und qualitativ bewertet. Nahestehende Arbeiten sind in diesem Fall die Ansätze (Ruskoné, 1996; Price, 2000; Zhang et al., 2001), die alle den Anspruch der Handhabung urbaner Szenen erheben:

- ◊ Der Ansatz von (Ruskoné, 1996) verfolgt eine teilweise ähnliche Strategie wie diese Arbeit, indem zunächst sicher bestimmbare Teile des Straßennetzes extrahiert werden, und anschließend das Netz durch Gruppierungsalgorithmen vervollständigt wird. Ein grundlegender Unterschied besteht in der Nutzung von Kontextinformation. Zwar macht der Ansatz von (Ruskoné, 1996) in großem Umfang Gebrauch von lokalem Kontext, aber hauptsächlich erst nach der Netzbildung zur abschließenden Verifikation einzelner Teilstücke des Netzes. Dies setzt voraus, dass das Straßennetz auch ohne intensive Nutzung von Kontext extrahiert werden kann. Der vorgestellte Ansatz erscheint in diesem Punkt überlegen, weil detaillierte lokale Kontextinformation zu Beginn der Straßenextraktion extrahiert wird und somit eine flexiblere Extraktion und letztlich auch die Handhabung komplexerer Szenen möglich wird. Neben der Verwendung von Mehrbild- und Höhendaten besteht ein weiteres vorteilhaftes Merkmal des vorgestellten Ansatzes in der Verwendung eines umfassenden und maßstabsabhängigen Modells, was im Gegensatz zu (Ruskoné, 1996) eine integrierte Nutzung von Detailinformation und abstrahierter Information ermöglicht.
- ◊ Der Ansatz von (Price, 2000) ist grundsätzlich in der Lage, ähnlich komplexe Szenen wie der vorgestellte Ansatz zu bearbeiten. Auch die Nutzung von Höhen- und Mehrbildinformation im Rahmen der Extraktion ist vergleichbar. Das zugrunde liegende Straßenmodell ist jedoch wesentlich einfacher, da keine Detailinformation wie Markierungen und Fahrspuren und, abgesehen von der Modellierung von Häuserschluchten, auch keine Kontextinformation verwendet wird. Infolge der Einschränkung auf gitterförmige Straßennetze und der dadurch realisierbaren, stark modellgetriebenen Vorgehensweise ist mit dem Ansatz von (Price, 2000) dennoch eine Straßenextraktion in Innenstadtbereichen möglich. Der Ansatz scheitert jedoch bei einer weniger regulären Globalstruktur des Straßennetzes, wie sie nicht nur in den Beispieldaten dieser Arbeit vorkommt, sondern in vielen Europäischen Städten die Regel ist. Die grundlegende Idee, globale geometrische Eigenschaften für die Extraktion zu nutzen, ist aber sehr wertvoll. Auch der hier entwickelte Ansatz könnte von stärkeren geometrischen Restriktionen für den Straßenverlauf profitieren.
- ◊ Mit der Verbesserung von GIS-Daten verfolgt der Ansatz von (Zhang et al., 2001) zwar eine grundlegend andere Zielstellung als der vorgestellte Ansatz. Da der Ansatz von (Zhang et al., 2001) jedoch auch für urbane Szenen entwickelt wurde, stützt er sich auf ein umfassendes Straßenmodell. Im Unterschied zu dem hier verfolgten Ansatz werden jedoch die einzelnen Modellkomponenten nicht zur Extraktion von Straßenteilobjekten zusammengefasst, sondern sie dienen als einzelne Evidenzen zur exakten Bestimmung der Straßengeometrie. Die Nutzung von Kontextinformation ist weniger stark mit der Straßenextraktion verflochten als in diesem Ansatz. Vorteilhaft ist aber, dass (Zhang et al., 2001) neben der Verwendung eines Höhenmodells auch Verfahren zur Zuordnung von Bildprimitiven über mehrere Bilder hinweg einsetzt. Die überwiegende Mehrheit der Straßenextraktion wird daher in 3D durchgeführt und resultiert in einer wesentlich höheren geometrischen Genauigkeit, als sie der vorgestellte Ansatz zu erbringen vermag. Diese Genauigkeit könnte hier erst mit den oben erwähnten Erweiterungen erreicht werden. Die Vorgehensweise, erst nach der Straßennetzbildung eine verfeinerte Zuordnung von Fahrspuren oder Fahrbahnen in den Bildern durchzuführen, erscheint dabei effizienter als das in (Zhang et al., 2001) gewählte Verfahren der sehr frühen Bestimmung von 3D-Primitiven.

Ein wesentlicher Fortschritt des hier entwickelten Systems zur Straßenextraktion gegenüber allen drei angesprochenen Ansätzen besteht in der systematischen Nutzung von internen Bewertungskriterien.

Die Bewertung abschließend bleibt festzuhalten, dass abgesehen von vereinzelten Defiziten gegenüber anderen Ansätzen das System sowohl in seiner konzeptionellen Grundlage als auch in seiner praktischen Leistungsfähigkeit eine internationale Spitzenstellung einnimmt. Dies soll aber nicht darüber hinwegtäuschen, dass noch eine Vielzahl von interessanten wissenschaftlichen Fragestellungen für die automatische Straßenextraktion in Stadtgebieten unbeantwortet bleiben. Einige elementare Fragen werden im Ausblick des abschließenden Kapitels aufgegriffen.

11 Zusammenfassung und Ausblick

Die Zielsetzung der vorliegenden Arbeit war die Entwicklung eines Verfahrens zur vollautomatischen Extraktion von urbanen Straßennetzen aus Luftbildern einer Auflösung von ca. 15cm. Dominierender Gesichtspunkt bei der Konzipierung der Straßenextraktion ist die Handhabung der hohen Komplexität urbaner Szenen. Um Straßen mit ausreichender Vollständigkeit und Zuverlässigkeit automatisch extrahieren zu können, stützt sich das Extraktionssystem auf eine umfassende, maßstabsabhängige und explizite Modellierung des Objektes „Straße“ sowie des Kontexts von Straßen. Das Straßenmodell integriert Wissen über Detailstrukturen wie Markierungen, Straßenränder und Fahrspuren mit stärker abstrahiertem Wissen über die globale Verbindungsfunktion von Straßen. Während durch die Modellierung des lokalen Straßenkontexts Beziehungen zwischen einer Straße und den unmittelbar benachbarten Objekten erfasst werden, ermöglicht die Modellierung des globalen Kontexts, die Abhängigkeit der Gestalt und Erscheinungsform einer Straße von der großräumigen Umgebung zu berücksichtigen.

Ein weiteres Merkmal des vorgestellten Ansatzes ist die Entwicklung und systematische Anwendung eines Schemas zur internen Bewertung von Extraktionsergebnissen. Das Schema stützt sich auf das Objektmodell und teilt die Kriterien, die ein (Teil-)Objekt charakterisieren in zwei Kategorien ein: einerseits in Kriterien, die zur Extraktion des Objektes herangezogen werden, und andererseits in Kriterien, die zu seiner internen Bewertung dienen. Unter der Voraussetzung, dass die Kriterien der beiden Kategorien unabhängiges Wissen über das Objekt repräsentieren, erlaubt die interne Bewertung die unabhängige Bemessung der Konfidenz von Objekthypothesen. Insbesondere bei der Behandlung komplexer Szenen ist die Bearbeitung von Hypothesen entsprechend ihres Konfidenzwertes von großer Bedeutung, weil auf vielversprechende Hypothesen fokussiert werden kann und damit die Extraktion effizienter und robuster wird.

Die Implementierung der Straßenextraktion umfasst die drei Phasen (1) Kontextanalyse, (2) Extraktion von Straßenobjekten und (3) Straßennetzbildung:

- (1) Mit der Kontextanalyse wird globale und lokale Hintergrundinformation gewonnen, die für den weiteren Verlauf der Straßenextraktion von Bedeutung ist. Die Kontextanalyse umfasst einerseits die Segmentierung der globalen Kontextgebiete „Wald“, „Siedlung“ und „offene Landschaft“ und andererseits die Extraktion von lokalem Kontext. Neben der groben Bestimmung von Gebäuden und hoher Vegetation als Ausschlussregionen spielen vor allem die Detektion von Verdeckungen und Schattenbereichen sowie die Extraktion von Fahrzeugen und Fahrzeugkolonnen eine zentrale Rolle.
- (2) Bei der Extraktion von Straßenobjekten wird dem Ziel nachgegangen, durch die integrierte Nutzung von detailliertem und abstrahiertem Modell-Wissen sowie durch die Einbindung von Kontextinformation möglichst zuverlässige Teile des Straßennetzes zu extrahieren. Hierzu dient zum einen die Extraktion von Markierungen und Straßenrändern, die zur Konstruktion von Fahrspuren in der hohen Auflösung herangezogen werden; zum anderen werden vorläufige Straßenabschnitte aus der mittleren und niedrigen Auflösung bestimmt. Aus dieser Eingangsinformation werden durch mehrere Gruppierungsoperationen größere und semantisch höherwertige Straßenobjekte – die Fahrbahnen – aufgebaut.
- (3) Da die bis hierhin extrahierten Fahrbahnen in aller Regel noch kein zusammenhängendes Netz formen, folgt als abschließende Phase der Extraktion die Bildung des Straßennetzes. Nach der DOM-basierten Fusion von Fahrbahnen, die in überlappenden Bildern extrahiert wurden, folgt der Aufbau eines möglichst zusammenhängenden Straßennetzes durch einen iterativen Gruppierungsprozess: Die Generierung von Verknüpfungshypothesen zwischen Fahrbahnen nutzt Kriterien der lokalen Straßengeometrie und integriert sie mit dem globalen Kriterium, dass potentielle Verknüpfungen wichtig für Verbindungen zwischen weit entfernten Punkten auf dem Netzwerk sein müssen. Die Verifikation von Verknüpfungshypothesen erfolgt über gezielte Suche nach Straßenevidenz oder Kontext im Bild. Im Rahmen einer abschließenden Netzanalyse werden die endgültigen Straßen und Kreuzungen definiert.

Die erzielten Ergebnisse wurden anhand von manuell erstellten Referenzdaten evaluiert. Die Tatsache, dass für die extrahierten Straßenachsen auch im Falle sehr komplexer Innenstadtbereiche zufriedenstellende bis sehr gute Bewertungen erreicht wurden, belegt die Leistungsfähigkeit des Systems. Als Schwachstelle wurde vor allem die mangelhafte geometrische Genauigkeit der resultierenden Fahrspuren und Straßenachsen identifiziert. Ein möglicher Weg zu deren Behebung ist die exakte Zuordnung von Fahrspuren in den Originaldaten. Das extrahierte Straßennetz könnte hierfür als wertvolle Näherungsinformation dienen.

Als Konsequenz der detaillierten Modellierung von Straßen wäre als nächster Entwicklungsschritt des vorgestellten Ansatzes eine umfassendere Modellierung komplexer Kreuzungen zu sehen. Einerseits sollte die Modellierung eine grobe Unterteilung in verschiedene Kreuzungsklassen wie Einmündung, Kreuzung, Kreisverkehr, etc. vorsehen und Fahrspuren integrieren. Andererseits muss die Extraktionsstrategie für Kreuzungen in der Lage sein, viele unscheinbare Evidenzen wie kurze Markierungen, Stoppllinien, kleine Verkehrsinseln usw. für eine exakte Rekonstruktion der Fahrspuren innerhalb des Kreuzungsbereiches zu kombinieren. Dies verdeutlicht, dass die Bestimmung der Relationen zwischen den in den Kreuzungsbereich einmündenden Fahrspuren selbst bei einer Einschränkung auf höhengleiche Kreuzungen eine große Herausforderung darstellt.

Neben den direkt aus der Analyse der Ergebnisse abgeleiteten Möglichkeiten zur Weiterentwicklung des vorgestellten Ansatzes gibt es noch eine große Anzahl an weiteren, eher allgemeinen Fragestellungen zur tieferen wissenschaftlichen Durchdringung dieses Themas. Im Folgenden werden einige ausgewählte Fragen aufgegriffen:

- ◇ Wie in Kapitel 10.2 angesprochen, stellt die korrekte Einstellung der Parameter eine grundlegende Schwierigkeit in der automatischen Bildanalyse dar. Als eine aus der vorliegenden Arbeit erwachsende Fragestellung ergibt sich, ob und wie das Schema zur internen Bewertung verwendet werden kann, um eine automatisierte Justierung der Extraktionsparameter zu erreichen. Ziel wäre es, anhand der internen Bewertungen schon relativ frühzeitig zu erkennen, dass wichtige Teile des Straßennetzes nicht oder gewisse Teile inkorrekt extrahiert wurden.
- ◇ In diesem Zuge ist es interessant der Frage nachzugehen, inwieweit die Modellierung zusätzlichen Wissens für eine vollständigere und zuverlässigere Extraktion möglich bzw. nötig ist. Das in der vorliegenden Arbeit modellierte Wissen reicht in vielen Fällen aus, um Straßen mit zwei und mehr Fahrspuren zu extrahieren. Schmalere Gassen in sehr dicht bebauten Gebieten werfen aber Schwierigkeiten auf, da oft schon die Bestimmung von initialen Hypothesen scheitert. Zu klären wäre z.B., ob eine umfassendere Modellierung der Interaktion zwischen Gebäuden, Straßen, Parkplätzen und Fahrzeugen hierfür eine Lösungsmöglichkeit bietet.
- ◇ Eng verbunden damit ist die Fragestellung, inwieweit die Interpretation einer gesamten Szene hilfreich für die Extraktion einzelner Objektklassen sein kann. In den letzten Jahren wurden einige vielversprechende Ansätze zur Extraktion einzelner Objektklassen – allen voran Gebäude, Straßen und Vegetation – entwickelt. Aus einer Kombination der am weitesten entwickelten Ansätze für eine jeweilige Objektklasse ergeben sich zwei interessante Fragestellungen: Zum einen, ob eine Fusion solcher hochentwickelter Ansätze eine bessere Qualität der Szeneninterpretation mit sich bringt, als sie Ansätze erreichen, deren Konzeption von vorneherein auf die Interpretation der gesamten Szene ausgelegt ist; zum anderen wäre zu klären, wie ein optimaler Rückfluss des Ergebnisses der Szeneninterpretation zu gestalten ist, um eine Verbesserung der Extraktion der einzelnen Objektklassen zu erzielen.
- ◇ Neue Möglichkeiten ergeben sich auch durch die Fortschritte in der Sensorik. Die neuen kommerziellen Satellitendaten erreichen mit einer Auflösung von unter 1m im panchromatischen Kanal und von unter 4 m in multispektralen Kanälen eine Auflösungsklasse, in der sie für die Auswertung von Stadtszenen prinzipiell in Frage kommen. Zwar ist aufgrund der im Vergleich zu Luftbildern gröberen Auflösung eine Extraktion von Detailstrukturen kaum möglich. Interessant wäre es aber der Frage nachzugehen, ob die zusätzliche multispektrale Information den Mangel an Detailinformation kompensieren kann, so dass zumindest bezüglich der extrahierten Straßenachsen vergleichbare Ergebnisse erzielt werden können.
- ◇ Anderes Potential erschließt die verstärkte Nutzung von Laserdaten. Die Verwendung solcher Daten hat nicht nur den Vorteil, dass ein genaues Höhenmodell abgeleitet werden kann. Eine Reihe von Ansätzen zeigt, dass

über die Analyse von multiplen Echos der ausgesendeten Pulse nicht nur Regionen mit Vegetation extrahiert werden können, sondern teilweise auch die Grundfläche unter der Vegetation bestimmbar ist. Laserdaten würden die Möglichkeit eröffnen, z.B. das Kontextmodell hinsichtlich Alleen zu erweitern. Diese könnten für die Straßenextraktion eine ähnliche Funktion wie Fahrzeugkolonnen übernehmen. Prinzipiell liefern auch Radar-Verfahren, insbesondere interferometrisches und polarimetrisches SAR, komplementäre Informationen zu Bilddaten. Während Radar-Verfahren für die Straßenextraktion in der offenen Landschaft von großem Nutzen sein können, erweisen sie sich aus heutiger Sicht für die Auswertung von Stadtgebieten aufgrund der inhärenten Schrägsichtgeometrie und der schlechteren radiometrischen Bildqualität eher als problematisch.

Die Ergebnisse der Straßenextraktion in urbanen Gebieten zeigen eine hohe Zuverlässigkeit, eine zufriedenstellende Vollständigkeit und eine gute topologische Qualität. Trotz der weitaus umfassenderen Straßenmodellierung als für die Extraktion in der offenen Landschaft ist aber festzuhalten, dass die Ergebnisse in Stadtgebieten tendenziell weniger gut ausfallen als in der offenen Landschaft. Mit Blick auf einen möglichen operationellen Einsatz der Straßenextraktion in urbanen Gebieten muss konstatiert werden, dass trotz der erfreulichen Ergebnisse momentan keine Praxistauglichkeit eines vollautomatischen Systems in Sicht ist. Auch die Richtung, die zur Entwicklung eines nutzergeführten Systems für die Straßenextraktion in Stadtgebieten eingeschlagen werden sollte, ist noch nicht klar abzusehen. Aufgrund der unvermeidbaren Komplexität des Straßenmodells erscheint eine Real-time Überwachung der Extraktion durch einen menschlichen Operateur wie bei Tracking-Verfahren schwer erreichbar. Einen Ausweg würde eine Kombination von unüberwachter und überwachter Extraktion bieten. Gewisse Teile der Extraktion, so z.B. die Kontextanalyse und die Extraktion von Teilstücken des Straßennetzes, könnte ohne Überwachung vonstattengehen. In der überwachten Phase der Extraktion müssten dann mit Hilfe geeigneter Tools Bereiche, in denen die automatische Extraktion Schwierigkeiten hatte, editiert werden. Wie bereits in Kapitel 4 herausgearbeitet, ist die interne Bewertung der automatisch erzielten Ergebnisse eine entscheidende Grundlage für eine geeignete Nutzerführung bei der Editierung. Eine weitere ungelöste Fragestellung ist die Konzipierung eines Verfahrensablaufs, der aus Sicht der Extraktion an sinnvollen Stellen Interaktion mit dem Nutzer vorsieht und der gleichzeitig aus Sicht des Operateurs eine effiziente Arbeitsweise erlaubt, entweder mit dauerhafter Interaktion oder mit langen Zeiträumen ohne Interaktion.

Auch wenn am Ende dieser Arbeit noch viele Fragen offen bleiben, so zeigt insbesondere die Stellung des entwickelten Systems im internationalen Vergleich, dass die Straßenextraktion in komplexen Szenen ein nicht unwesentliches Stück vorangebracht werden konnte. Zu berücksichtigen ist, dass die Arbeit auf grundlagenorientierte Forschungsfragen fokussiert. Eine Integration in die Praxis liegt noch in weiter Ferne. Als erfreuliche Aussicht für die Zukunft sei aber angemerkt, dass die Straßenextraktion in der offenen Landschaft noch vor einigen Jahren einen ähnlichen Status hatte, inzwischen aber durch Kooperationen von Institutionen aus Forschung und Praxis näher an einen operationellen Einsatz herangerückt ist (Willrich, 2002; Bacher und Mayer, 2003). Die künftigen Entwicklungen werden zeigen, ob dies auch für die Straßenextraktion in Stadtgebieten möglich ist.

Literatur

- Airault, S., Ruskoné, R., Jamet, O., 1994. Road detection from aerial images: a cooperation between local and global methods. In: Desachy, J. (Hrsg.), *Image and Signal Processing for Remote Sensing*, Proc. SPIE 2315, 508–518.
- Aviad, Z., Carnine, P. D., 1988. Road Finding for Road-Network Extraction. In: *Computer Vision and Pattern Recognition*, 814–819.
- Bacher, U., Mayer, H., 2003. Fuzzy-Classification and Ziplock-Snakes for Road Extraction from IKONOS Images. In: *Joint Workshop on High Resolution Mapping from Space*, International Society for Photogrammetry and Remote Sensing, Hanover, Germany.
- Bajcsy, R., Tavakoli, M., 1976. Computer Recognition of Roads from Satellite Pictures. *IEEE Transactions on Systems, Man, and Cybernetics* 6 (9): 623–637.
- Baltsavias, E., Grün, A., van Gool, L. (Hrsg.), 2001. *Automatic Extraction of Man-Made Objects from Aerial and Space Images (III)*. Balkema Publishers, Lisse, The Netherlands.
- Barzohar, M., Cohen, M., Ziskind, I., Cooper, D., 1997. Fast Robust Tracking of Curvy Partially Occluded Roads in Clutter in Aerial Images. In: Grün et al. (1997), 277–286.
- Barzohar, M., Cooper, D., 1996. Automatic Finding of Main Roads in Aerial Images by Using Geometric-Stochastic Models and Estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 18 (7): 707–721.
- Baumgartner, A., 2003. *Automatische Extraktion von Straßen aus digitalen Luftbildern*. Dissertation, Deutsche Geodätische Kommission (C) 564, München.
- Baumgartner, A., Eckstein, W., Mayer, H., Heipke, C., Ebner, H., 1997. Context-Supported Road Extraction. In: Grün et al. (1997), 299–308.
- Baumgartner, A., Hinz, S., 2000. Multi-Scale Road Extraction Using Local and Global Grouping Criteria. In: *International Archives of Photogrammetry and Remote Sensing*, Vol. 33(B3/1), 58–65.
- Baumgartner, A., Hinz, S., Wiedemann, C., 2002. Efficient Methods and Interfaces for Road Tracking. In: *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. 34(3B), 28–31.
- Baumgartner, A., Steger, C., Mayer, H., Eckstein, W., Ebner, H., 1999. Automatic Road Extraction Based on Multi-Scale, Grouping, and Context. *Photogrammetric Engineering and Remote Sensing* 65 (7): 777–785.
- Bogenberger, K., Ernhofer, O., Schütte, C., 1999. Effects of Telematic Applications for a High Capacity Ring Road in Munich. In: *Proceedings of the 6th World Congress on Intelligent Transportation Systems*, Toronto.
- Bordes, G., Giraudon, G., Jamet, O., 1995. Road Extraction Guided by a Cartographic Database: Creating a Strategy. In: *IEEE Workshop on Context-based Vision*, 30–37.
- Bordes, G., Giraudon, G., Jamet, O., 1997. Road Modeling Based on a Cartographic Database for Aerial Image Interpretation. In: Grün et al. (1997), 123–139.
- Brachman, R., 1977. What's in a Concept: Foundations for Semantic Networks. *International Journal of Man-Machine Studies* 9: 127–152.
- Brachman, R., 1979. *On the Epistemological Status of Semantic Networks*. Academic Press, Inc., New York, USA.
- Briese, C., Belada, P., Pfeifer, N., 2001. Digitale Geländemodelle im Stadtgebiet aus Laser-Scanner-Daten. *Österreichische Zeitschrift für Vermessung und Geoinformation* 2: 83–91.
- Brunn, A., 2001. Statistical interpretation of DEM and image data for building extraction. In: Baltsavias et al. (2001), 171–180.
- Burlina, P., Chellappa, R., Lin, C., Zhang, X., 1995. Context-Based Exploitation of Aerial Imagery. In: *IEEE Workshop on Context-based Vision*, 38–49.
- Burlina, P., Parameswaran, V., Chellappa, R., 1997. Sensitivity Analysis and Learning Strategies for Context-Based Vehicle Detection. In: *Image Understanding Workshop '97*, Morgan Kaufmann Publishers, San Francisco, CA.
- Couloigner, I., Ranchin, T., 2000. Mapping of Urban Areas: A Multiresolution Modeling Approach for Semi-Automatic Extraction of Streets. *Photogrammetric Engineering and Remote Sensing* 66(7): 867–874.
- de Gunst, M., Vosselman, G., 1997. A Semantic Road Model for Aerial Image Interpretation. In: Förstner,

- W., Plümer, L. (Hrsg.), *Semantic Modeling for the Acquisition of Topographic Information from Images and Maps*, 108–122.
- de Gunst, M. E., 1996. *Knowledge-based Interpretation of Aerial Images for Updating of Road Maps*. Ph.D. thesis, Delft University of Technology, Delft.
- Doucette, P., Agouris, P., Musavi, M., Stefanidis, A., 1999. Automated Extraction of Linear Features from Aerial Imagery Using Kohonen Learning and GIS Data. In: *Lecture Notes in Computer Science*, 1737, Springer-Verlag, 20–33.
- Dubois, D., Prade, H., 1988. *Possibility Theory - An Approach to Computerized Processing of Unvertainty*. Plenum Press, New York, NY.
- Dubuisson-Jolly, M.-P., Lakshmanan, S., Jain, A., 1996. Vehicle Segmentation and Classification Using Deformable Templates. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 18 (3): 293–308.
- Ebner, H., Eckstein, W., Heipke, C., Mayer, H. (Hrsg.), 1999. Automatic Extraction of GIS Objects from Digital Imagery. *International Archives of Photogrammetry and Remote Sensing*, 32(3-2W5).
- Eckstein, W., Steger, C., 1996. Fusion of digital terrain models and texture for object extraction. In: *Proceedings of the Second International Airborne Remote Sensing Conference and Exhibition*, Vol. III, Environmental Research Institute of Michigan, 1–10.
- Eidenbenz, C., Käser, C., Baltsavias, E., 2000. ATOMI - Automated reconstruction of topographic objects from aerial images using vectorised map information. In: *International Archives of Photogrammetry and Remote Sensing*, Vol. 33(B3/1), 462–471.
- Englisch, A., Heipke, C., 1998. Erfassung und Aktualisierung topographischer Geo-daten mit Hilfe analoger und digitaler Luftbilder. *Photogrammetrie - Fernerkundung - Geoinformation (PFG)* 3: 133–149.
- Faber, A., Förstner, W., 2000. Detection of Dominant Orthogonal Structures in Small Scale Imagery. In: *International Archives of Photogrammetry and Remote Sensing*, Vol. 33(B3/1), 274–281.
- Fischer, A., Kolbe, T., Lang, F., Cremers, A., Förstner, W., Plümer, L., Steinhage, V., 1998. Extracting Buildings from Aerial Images Using Hierarchical Aggregation in 2D and 3D. *Computer Vision and Image Understanding* 72 (2): 185–203.
- Fischler, M., Heller, A., 1998. Automated Techniques for Road Network Modeling. In: *Image Understanding Workshop '98*, Morgan Kaufmann Publishers, San Francisco, CA, 501–516.
- Fischler, M. A., 1994. The Perception of Linear Structure: A Generic Linker. In: *Image Understanding Workshop '94*, Morgan Kaufmann Publishers, San Francisco, CA, 1565–1579.
- Fischler, M. A., Tenenbaum, J. M., Wolf, H. C., 1981. Detection of Roads and Linear Structures in Low-Resolution Aerial Imagery Using a Multisource Knowledge Integration Technique. *Computer Graphics and Image Processing* 15: 201–223.
- Förstner, W., 1996. 10 Pros and Cons Against Performance Characterization of Vision Algorithms. In: *Christensen, H. I., Förstner, W., Madsen, C. B. (Hrsg.), Workshop on Performance Characteristics of Vision Algorithms*, 13–29.
- Fuchs, C., Gülch, E., Förstner, W., 1998. OEEPE Survey on 3D-City Models. *OEEPE Official Publication* 35: 9–123.
- Geman, D., Jedynak, B., 1996. An Active Testing Model for Tracking Roads in Satellite Images. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 18 (1): 1–14.
- Geman, S., Geman, D., 1984. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 6 (6): 721–741.
- Gruber, M., Leberl, F., Petko, R., 2003. Paradigmenwechsel in der Photogrammetrie durch digitale Luftbildaufnahme ? *Photogrammetrie – Fernerkundung – Geoinformation (PFG)* 4: 285–297.
- Grün, A., Baltsavias, E., Henricsson, O. (Hrsg.), 1997. *Automatic Extraction of Man-Made Objects from Aerial and Space Images (II)*. Birkhäuser Verlag, Basel.
- Grün, A., Kuebler, O., Agouris, P. (Hrsg.), 1995. *Automatic Extraction of Man-Made Objects from Aerial and Space Images*. Birkhäuser Verlag, Basel.
- Grün, A., Li, H., 1997. Linear Feature Extraction with LSB-Snakes. In: *Grün et al. (1997)*, 287–298.
- Grün, A., Nevatia, R. (Hrsg.), 1998. Special Issue on Automatic Building Extraction from Aerial Images. *Computer Vision and Image Understanding*, 72(2).
- Haag, M., Nagel, H.-H., 1999. Combination of Edge Element and Optical Flow Estimates for 3D-Model-Based

- Vehicle Tracking in Traffic Sequences. *International Journal of Computer Vision* 35 (3): 295–319.
- Harvey, W. A., 1999. Performance Evaluation for Road Extraction. In: *International Workshop on 3D Geospatial Data Production: Meeting Application Requirements*, Paris, 79–87.
- Heipke, C., Mayer, H., Wiedemann, C., Jamet, O., 1998. External Evaluation of Automatically Extracted Road Axes. *Photogrammetrie - Fernerkundung - Geoinformation (PGF)* 2: 81–94.
- Heller, A., Fischler, M., Bolles, R., Connolly, C., 1998. An Integrated Feasibility Demonstration for Automatic Population of Spatial Databases. In: *Image Understanding Workshop '98*, Morgan Kaufmann Publishers, San Francisco, CA, 501–516.
- Hellwich, O., Günzl, M., Wiedemann, C., 2001. Fusion of SAR/INSAR Data and Optical Imagery for Landuse Classification. *Frequenz - Zeitschrift für Telekommunikation* 55(3/4): 129–136.
- Heuel, S., Förstner, W., 2001. Topological and geometrical models for building reconstruction from multiple images. In: *Baltsavias et al. (2001)*, 13–24.
- Hinz, S., 2002. Automatische Fahrzeugerkennung in optischen Bilddaten: Ein Überblick mit Konzeption eines neuen Verfahrens. *Technischer Bericht, PF-2002-02*, Lehrstuhl für Photogrammetrie und Fernerkundung, Technische Universität München.
- Hinz, S., 2003. Integrating Local and Global Features for Vehicle Detection in High Resolution Aerial Imagery. In: *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. 34(3W/8), 119–124.
- Hinz, S., Baumgartner, A., 2000. Road Extraction in Urban Areas supported by Context Objects. In: *International Archives of Photogrammetry and Remote Sensing*, Vol. 33(B3/1), 405–412.
- Hinz, S., Baumgartner, A., 2001. Vehicle Detection in Aerial Images Using Generic Features, Grouping, and Context. In: *Pattern Recognition (DAGM 2001)*, Lecture Notes on Computer Science 2191, Springer-Verlag, 45–52.
- Hinz, S., Baumgartner, A., 2003. Automatic Extraction of Urban Road Nets from Multi-View Aerial Imagery. *ISPRS Journal of Photogrammetry and Remote Sensing* (im Druck).
- Hinz, S., Baumgartner, A., Ebner, H., 2001a. Modelling Contextual Knowledge for Controlling Road Extraction in Urban Areas. In: *IEEE/ISPRS joint Workshop on Remote Sensing and Data Fusion over Urban Areas*.
- Hinz, S., Kurz, F., Baumgartner, A., Wasmeier, P., 2001b. The Role of Shadow for 3D-Object Reconstruction from Monocular Images. In: *Kahmen, H., Grün, A. (Hrsg.), 3D-Optical Measurement Techniques (V)*, 354–363.
- Hinz, S., Wiedemann, C., 2003. Increasing Efficiency of Road Extraction by Self-Diagnosis. *Photogrammetric Engineering and Remote Sensing* (im Druck).
- Hinz, S., Wiedemann, C., Baumgartner, A., 2000. A Scheme for Road Extraction in Rural Areas and its Evaluation. In: *5th IEEE Workshop on Applications of Computer Vision*, 134–139.
- Hoogs, A., Bajcsy, R., 1995. Using Scene Context to Model Segmentations. In: *IEEE Workshop on Context-based Vision*, 50–61.
- Huertas, A., Kim, Z., Nevatia, R., 2000. Multisensor integration for building modeling. In: *CVPR*, 203–210.
- Illert, A., 1993. Automationsgestützte Fortführung des ATKIS. *Technischer Bericht*, Institut für Photogrammetrie, Universität Bonn.
- Kass, M., Witkin, A., Terzopoulos, D., 1988. Snakes: Active contour models. In: *International Journal of Computer Vision*, 321–331.
- Kim, Z., Huertas, A., Nevatia, R., 2001. A model-based approach for multi-view complex building description. In: *Baltsavias et al. (2001)*, 181–193.
- Konecny, G., 1995. Satelliten-Fernerkundung und Kartographie. *GIS* 8 (2): 3–12.
- Köster, M., 1995. Kontextsensitive Bildinterpretation mit Markoff-Zufallsfeldern. *Dissertation*, Deutsche Geodätische Kommission (C) 444, München.
- Kulschewski, K., 1999. Modellierung von Unsicherheiten in dynamischen Bayes-Netzen zur qualitativen Gebäudeerkennung. *Shaker-Verlag*, Aachen.
- Laptev, I., Mayer, H., Lindeberg, T., Eckstein, W., Steger, C., Baumgartner, A., 2000. Automatic Extraction of Roads from Aerial Images Based on Scale Space and Snakes. *Machine Vision and Applications* 12(1): 22–31.
- Leberl, F., Kalliany, R. (Hrsg.), 2002. *Photogrammetric Computer Vision*. *International Archives of Photo-*

- grammetry, Remote Sensing and Spatial Information Sciences, 34.
- Liedke, C.-E., Bückner, J., Pahl, M., Stahlhut, O., 2001. Knowledge based system for the interpretation of complex scenes. In: Baltsavias et al. (2001), 3–12.
- Lindeberg, T., 1998. Edge detection and ridge detection with automatic scale selection. *International Journal of Computer Vision* 30 (2): 117–156.
- Mackaness, W. A., Mackechnie, G. A., 1997. Detection and Simplification of Road Junctions in Automated Map Generalisation. In: *International Cartographic Conference*, 1013–1021.
- Mayer, H., 1998. Automatische Objektextraktion aus digitalen Luftbildern. Habilitation, Deutsche Geodätische Kommission (C) 494, München.
- Mayer, H., Steger, C., 1998. Scale-Space Events and Their Link to Abstraction for Road Extraction. *ISPRS Journal of Photogrammetry and Remote Sensing* 53 (2): 62–75.
- Mayr, W., 2002. Bemerkungen zum Thema „True Orthoimage“. *Photogrammetrie – Fernerkundung – Geoinformation (PGF)* 4: 237–244.
- McKeown, D., 1990. Toward Automatic Cartographic Feature Extraction. In: *Mapping and Spatial Modelling for Navigation*, Vol. F65 of Nato ASI Series, Springer-Verlag, 149–180.
- McKeown, D., Bowling, M., Bulwinkle, G., Cochran, S., Ford, S., Harvey, W., Kalp, D., McGlone, C., McMahill, J., Polis, M., Shufelt, J., Yocum, D., 1997. Research in the Automated Analysis of Remotely Sensed Imagery: 1995 - 1996. In: *Image Understanding Workshop '97*, Morgan Kaufmann Publishers, San Francisco, CA, 779–812.
- McKeown, D., Denlinger, J., 1988. Cooperative Methods For Road Tracking In Aerial Imagery. In: *Computer Vision and Pattern Recognition*, 662–672.
- McKeown, D., Harvey, W., Wixson, L., 1989. Automating Knowledge Acquisition for Aerial Image Interpretation. *Computer Vision, Graphics, and Image Processing* 46: 37–81.
- Mendel, J.-M., 1995. Fuzzy Logic Systems for Engineering: A Tutorial. *Proceedings of the IEEE* 83 (3): 345–377.
- Merlet, N., Zerubia, J., Apr. 1996. New Prospects in Line Detection by Dynamic Programming. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 18 (4): 426–431.
- Michaelsen, E., Stilla, U., 2000. Ansichtenbasierte Erkennung von Fahrzeugen. In: Sommer, G., Krüger, N., Perwass, C. (Hrsg.), *Mustererkennung, Informatik aktuell*, Springer-Verlag, Berlin, 245–252.
- Neuenschwander, W., Fua, P., Székely, G., Kübler, O., 1995. From ziplock snakes to velcro surfaces. In: Grün et al. (1995), 105–114.
- Niemann, H., Sagerer, G., Schröder, S., Kummert, F., 1990. ERNEST: A semantic network system for pattern understanding. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 12 (9): 883–905.
- Olson, C., Huttenlocher, D., Doria, D., 1996. Recognition by Matching With Edge Location and Orientation. In: *Image Understanding Workshop '96*, Morgan Kaufmann Publishers, San Francisco, CA.
- Price, K., 1999. Road Grid Extraction and Verification. In: *International Archives of Photogrammetry and Remote Sensing*, Vol. 32, part 3-2W5, 101–106.
- Price, K., 2000. Urban Street Grid Description and Verification. In: *5th IEEE Workshop on Applications of Computer Vision*, 148–154.
- Quint, F., 1997. MOSES: A Structural Approach to Aerial Image Understanding. In: Grün et al. (1997), 323–332.
- Ramer, U., 1972. An iterative procedure for the polygonal approximation of plane curves. *Computer Graphics and Image Processing* 1: 244–256.
- Reigber, A., 2002. Airborne Polarimetric SAR Tomography. Dissertation, Universität Stuttgart.
- Rosenfeld, A., 1982. Computer image analysis: an emerging technology in the service of society. *Computer Science Technical Reports TR-1177*, University of Maryland.
- Ruskoné, R., 1996. Road Network Automatic Extraction by Local Context Interpretation: application to the production of cartographic data. Ph.D. thesis, Université Marne-La-Valleé.
- Ruskoné, R., Airault, S., Jamet, O., 1994. Road Network Interpretation: A Topological Hypothesis Driven System. In: *International Archives of Photogrammetry and Remote Sensing*, Vol. 30(3/2), 711–717.
- Ruskoné, R., Guiges, L., Airault, S., Jamet, O., 1996. Vehicle Detection on Aerial Images: A Structural Approach. In: *13th International Conference on Pattern Recognition*, Vol. 3, 900–903.
- Sagerer, G., 1982. Darstellung und Nutzung von Expertenwissen für ein Bildanalyzesystem. Springer-Verlag, Berlin.

- Schistad-Solberg, A., 1999. Contextual Data Fusion Applied to Forest Map Revision. *IEEE Transactions on Geoscience and Remote Sensing* 37 (3): 1234–1243.
- Scholze, S., Moons, T., van Gool, L., 2001. A probabilistic approach to roof patch extraction and reconstruction. In: Baltsavias et al. (2001), 195–215.
- Shafer, G., Pearl, J. (Hrsg.), 1990. *Readings in Uncertain Reasoning*. Morgan Kaufmann Publishers.
- Shufelt, J., 1999. *Geometric Constraints for Object Detection and Delineation*. Kluwer Academic Publishers, Boston.
- Sörgel, U., 2003. *Iterative Verfahren zur Detektion und Rekonstruktion von Gebäuden in SAR- und InSAR-Daten*. Ph.D. thesis, Universität Hannover, Hannover.
- Stark, L., Bowyer, K., 1991. Achieving Generalized Object Recognition through Reasoning and Association of Function to Structure. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 13 (10): 1097–1104.
- Steger, C., 1998a. An Unbiased Detector of Curvilinear Structures. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 20 (2): 113–125.
- Steger, C., 1998b. *Unbiased Extraction of Curvilinear Structures from 2D and 3D Images*. Dissertation, Fakultät für Informatik, Technische Universität München, Herbert Utz Verlag, München.
- Steger, C., 2000. Subpixel-precise extraction of lines and edges. In: *International Archives of Photogrammetry and Remote Sensing*, Vol. 33(B3/1), 141–156.
- Steger, C., Mayer, H., Radig, B., 1997. The Role of Grouping for Road Extraction. In: Grün et al. (1997), 245–256.
- Strat, T., 1992. *Natural Object Recognition*. Springer-Verlag, New York.
- Strat, T. M., 1995. Using Context to Control Computer Vision Algorithms. In: Grün et al. (1995), 3–12.
- Straub, B. M., Heipke, C., 2001. Automatic extraction of trees for 3D city models from images and height data. In: Baltsavias et al. (2001), 267–277.
- Suetens, P., Fua, P., Hanson, A., 1992. Computational strategies for object recognition. *ACM Computing Surveys* 24 (1): 5–60.
- Sullivan, G., Worrall, A., Ferryman, J., 1995. Visual Object Recognition Using Deformable Models of Vehicles. In: *IEEE Workshop on Context-based Vision*, 75–86.
- Tan, T., Sullivan, G., Baker, K., 1998. Model-Based Localisation and Recognition of Road Vehicles. *International Journal of Computer Vision* 27 (1): 5–25.
- Tönjes, R., Growe, S., 1998. Knowledge Based Road Extraction from Multisensor Imagery. In: *International Archives of Photogrammetry and Remote Sensing*, Vol. 32(3/1), 387–393.
- Tönjes, R., Growe, S., Bückner, J., Liedke, C.-E., 1999. Knowledge-Based Interpretation of Remote Sensing Images Using Semantic Nets. *Photogrammetric Engineering and Remote Sensing* 65(7): 811–821.
- Tupin, F., Bloch, I., Maitre, H., 1999. A First Step Toward Automatic Interpretation of SAR Images Using Evidential Fusion of Several Structure Detectors. *IEEE Transactions on Geoscience and Remote Sensing* 37(3): 1327–1343.
- van Cleynebreugel, J., Fierens, F., Suetens, P., Oosterlinck, A., 1990. Delineating Road Structures on Satellite Imagery by a GIS-Guided Technique. *Photogrammetric Engineering and Remote Sensing* 56(6): 893–898.
- Vosselman, G., 1992. *Relational Matching*. Springer-Verlag, Berlin.
- Vosselman, G., de Gunst, M., 1997. Updating Road Maps by Contextual Reasoning. In: Grün et al. (1997), 267–276.
- Vosselman, G., de Knecht, J., 1995. Road tracing by profile matching and kalman filtering. In: Grün et al. (1995), 265–274.
- Wiedemann, C., 2001. *Extraktion von Straßennetzen aus optischen Satellitenbilddaten*. Dissertation, Deutsche Geodätische Kommission (C) 551, München.
- Wiedemann, C., 2003. External Evaluation of Road Networks. In: *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. 34(3W/8), 93–98.
- Wiedemann, C., Ebner, H., 2000. Automatic Completion and Evaluation of Road Networks. In: *International Archives of Photogrammetry and Remote Sensing*, Vol. 33(B3/2), 979–986.
- Wiedemann, C., Hinz, S., 1999. Automatic Extraction and Evaluation of Road Networks from Satellite Imagery. In: *International Archives of Photogrammetry and Remote Sensing*, Vol. 32(3-2W5), 95–100.
- Willrich, F., 2002. Quality control and updating of road data by gis-driven road extraction from imagery. In:

- International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. 34(2), 761–767.
- Woods, W., 1975. What's in a Link: Foundations for Semantic Networks. In: Representation and Understanding: Studies in Cognitive Science, Academic Press, Inc., New York, USA, 35–85.
- Zadeh, L., 1989. Knowledge Representation in Fuzzy Logic. IEEE Transactions on Knowledge and Data Engineering 1 (1): 89–100.
- Zhang, C., Baltsavias, E., Grün, A., 2001. Updating of cartographic road databases by image analysis. In: Baltsavias et al. (2001), 243–253.
- Zhao, T., Nevatia, R., 2001. Car detection in low resolution aerial image. In: International Conference on Computer Vision.
- Zimmermann, P., 2001. Description of the Dataset of Zurich, Hoengg. Technical report, Institute of Geodesy and Photogrammetry, ETH Zurich, CH.
- Zlotnick, A., Carnine, P., 1993. Finding Road Seeds in Aerial Images. Computer Vision, Graphics, and Image Processing: Image Understanding 57 (2): 243–260.

Danksagung

An erster Stelle möchte ich meinem Doktorvater, Herrn Prof. Dr.-Ing. Heinrich Ebner, danken. Er verstand es, in einem freundschaftlichen und motivierenden Umfeld mir optimale Rahmenbedingungen zur Anfertigung dieser Arbeit zu bereiten und den Fortgang der Arbeit mit weitsichtigen und gleichzeitig präzisen Anregungen zu begleiten. Seine konsequente Einstellung, dass sich Engagement, Leistung und Übernahme von Verantwortung lohnen müssen, sind mir stets eine Ermunterung zu intensivem, eigenständigem Arbeiten gewesen. Für die Führung an der „langen Leine“ und das damit entgegengebrachte Vertrauen möchte ich mich bei ihm von ganzem Herzen bedanken.

Dem Zweitgutachter, Herrn Prof. Dr.-Ing. Helmut Mayer, möchte ich meinen Dank dafür aussprechen, dass er trotz seiner hohen Arbeitsbelastung ausgiebig Zeit zur Betreuung fand. Durch seine profunde und immer konstruktive Kritik hat die Arbeit ganz wesentlich an Profil gewonnen. Darüber hinaus soll nicht unerwähnt bleiben, dass Prof. Mayer durch seine sprühende Fachkenntnis mitverantwortlich dafür ist, dass bereits im Studium meine Begeisterung am Thema „Bildanalyse“ geweckt und während der Diplom- und Doktorarbeit intensiviert wurde.

Außerdem sei allen (Ex-)Kollegen und Kolleginnen vom Lehrstuhl für die heitere, freundschaftliche und hilfsbereite Atmosphäre gedankt – allen voran meinem Gruppenleiter Herrn Dr.-Ing. Albert Baumgartner, der sich den Anliegen seines Schützlings mit enormer Einsatzbereitschaft anvertraute. Aber auch aus dem gesamten Lehrstuhl-Team war in jeder Situation ein Ansprechpartner zur Stelle, egal ob gerade ein fachkundiger Rat nötig war oder die Versorgung mit Kaffee sicher gestellt werden musste.

Zum Schluss möchte ich noch meiner Frau Maria für die mir und meiner Begeisterung für die universitäre Arbeit entgegengebrachte Toleranz danken. Ihr Verständnis ist bei Weitem keine Selbstverständlichkeit.

Lebenslauf

Name: Stefan Janos Hinz

Geburtsort: Bamberg

Geburtsdatum: 21. Oktober 1972

Schulausbildung:

1979 — 1983 Grundschule: Kaulbergschule Bamberg
1983 — 1992 Kaiser-Heinrich-Gymnasium Bamberg
Allgemeine Hochschulreife

Zivildienst:

1992 — 1993 Bildungshaus „Marianum“, Bamberg

Studium:

1993 — 1998 Vermessungswesen, TU München
Dipl.-Ing. (univ.) Vermessungswesen

Beruf:

1998 — 2003 wissenschaftlicher Angestellter am
Lehrstuhl für Photogrammetrie und Fernerkundung

2003 — wissenschaftlicher Assistent am
Lehrstuhl für Methodik der Fernerkundung

1999 Gastwissenschaftler im Institute for Robotics and Intelligent Systems,
University of Southern California, Los Angeles, USA.