

Ein wissensbasiertes System zur automatischen Extraktion von semantischen Informationen aus digitalen Fernerkundungsdaten

Vom Fachbereich Elektrotechnik und Informationstechnik

der Universität Hannover

zur Erlangung des akademischen Grades

Doktor-Ingenieur

genehmigte

Dissertation

von

Dipl.-Inform. Jürgen Bückner

geb. am 29. Oktober 1964 in Bad Oeynhausen

2002

1.Referent: Prof. Dr.-Ing. C.-E. Liedtke

2.Referent: Prof. Dr.-Ing. C. Heipke

Tag der Promotion: 18.11.2002

Kurzfassung

Die vorliegende Arbeit befasst sich mit der automatischen Gewinnung semantischer Informationen aus Fernerkundungsdaten. Der Ansatz behandelt sowohl die Aspekte der Extraktion von speziellen Objekttypen durch die Verwendung von Bildverarbeitungsoperatoren als auch die Weiterverarbeitung dieser Ergebnisse auf einer symbolischen Ebene. Das entstandene operative System integriert a priori Wissen und die semantischen Eigenschaften der Objekte zur Erzeugung einer konsistenten Gesamtinterpretation einer in digitaler Form vorliegenden Szene.

Die hierarchisch arbeitende Interpretation verwendet Vorwissen über die Objekte sowie deren geometrische, topographische und radiometrische Ausprägung in den verschiedenen Fernerkundungssensoren und die topologischen Beziehungen der Objekte zueinander. Der hierarchische Ansatz fokussiert eine Betrachtung von großflächigen Strukturen zu immer kleineren Details. Diese Hierarchie ermöglicht eine Effizienzsteigerung der Analyse, da für ein Objekt einer Detaillierungsebene nur die darin möglichen Sub-Hypothesen generiert werden müssen.

Jede Hypothese besitzt eine holistische und eine strukturelle Methode, deren Anwendung von einem problemunabhängigen Analyseprozess aufgerufen wird. Der holistische Teil beinhaltet einen für die Objektklasse der Hypothese spezifischen Bildverarbeitungsoperator mit dem aktuellen Kontext angepassten Parametern zur Extraktion von Bildobjekten. Die strukturelle Verifizierung bewertet und gruppiert die vorliegenden Sub-Hypothesen bezüglich eines abstrakteren Begriffes.

Das verwendete Wissen wird in Form eines Semantischen Netzes eingebracht. Diese explizite Form der Wissensrepräsentation bietet eine strukturierte Darstellung der Wissensbasis. Vorhandenes Wissen aus bereits vorliegenden Interpretationen oder geographischen Datenbanken kann in die Analyse integriert werden. Die Integration von multisensoriellen und multispektralen Daten auch mit unterschiedlichen Auflösungen werden im System konsequent behandelt.

Das entstandene wissensbasierte Interpretationssystem wird für die automatische Erzeugung von hierarchischen Kartendarstellungen mit einer korrespondierenden strukturierten Beschreibung von komplexen Szenen in Fernerkundungsdaten eingesetzt. Die Ergebnisse enthalten semantische Zusatzinformationen und können für die Erstellung oder Aktualisierung von geographischen Datenbanken und die Erstellung von thematischen Karten eingesetzt werden.

Stichwörter: Wissensbasierte Mustererkennung, Semantische Netze, GIS, Photogrammetrie, Fernerkundung, Bildinterpretation

Abstract

The aim of this thesis is to take a close look at the possibilities for automatic extraction of semantic information from remote sensing data. The approach deals both with the aspects of extraction of special types of objects by using image processing operators and further processing of these results on a symbolic level. The operative system that has been developed integrates a priori knowledge and semantic features of the objects to generate a consistent overall interpretation of a scene presented in digital form.

The hierarchical interpretation employs previous knowledge of the objects and their geometrical, topographical and radiometric form in the different remote sensors as well as the topological relations between the objects. The hierarchical approach focusses the viewing of large format structures down to smaller details. This hierarchy increases the efficiency of the analysis, since for one object of a certain detail level only associated sub-hypothesis have to be generated.

Each hypothesis consists of a holistic and a structural method. The related applications of these methods are called by a problem-independent analysis process. The holistic part comprises an image processing operator that is specific to the hypothesis' class of objects. The parameters of the operator are adapted to the actual context of the object extraction. The structural verification values and groups the sub-hypotheses with regard to a more abstract form.

The knowledge that is used, is introduced in form of a semantic net. This explicit form of represented knowledge offers a structured description of the knowledge base. Available knowledge from already existing interpretations or geographical databases can be integrated into the analysis. The integration of multi-sensorial and multi-spectral data, even with different resolutions, is consequently processed in the system.

The implemented knowledge based interpretation system is used for automatic generation of hierarchical map representations with a corresponding, structured description of complex scenes in remote sensing data. The results contain additional semantic informations. They can be either used for creation or updating of geographical databases or for the generation of thematic maps.

Keywords: knowledge based system, image interpretation,
remote sensing, semantic net,

Vorwort

Die vorliegende Arbeit ist ein Resultat meiner Tätigkeit am Institut für Theoretische Nachrichtentechnik und Informationsverarbeitung an der Universität Hannover als wissenschaftlicher Mitarbeiter im Zeitraum vom Januar 1995 bis zum Juli 2001.

Herrn Prof. Dr.-Ing. C.-E. Liedtke danke ich für die wissenschaftliche Betreuung der Arbeit, die Ausführung des Hauptreferates und die gute Mitarbeiterbetreuung. Besonders geschätzt habe ich seine wissenschaftliche Kompetenz sowie die vorgefundene freie und motivierende Arbeitsumgebung. Diese Bedingungen und die interessanten Diskussionen und Anregungen haben das Entstehen dieser Arbeit ermöglicht.

Bei Herrn Prof. Dr.-Ing. C. Heipke - Leiter des Instituts für Photogrammetrie und GeoInformation - möchte ich mich hier für die Übernahme des Koreferats und die gute Zusammenarbeit auch über den Rahmen dieser Arbeit hinaus bedanken. Durch mehrere gemeinsame, interdisziplinäre Projekte habe ich ihn beruflich und persönlich schätzen gelernt.

Für die positive und erfolgreiche Zeit am Institut möchte ich mich bei allen Kollegen für die kooperative und sehr kollegiale Zusammenarbeit bedanken. Diese schaffte nicht nur eine hervorragende Arbeitsatmosphäre, sondern auch eine technisch gut funktionierende und organisierte Arbeitsumgebung die von Martin Pahl, Oliver Stahlhut, Mathias Schuh, Niko Meine und vielen anderen gestaltet wurde.

Mein besonderer Dank gilt Martin Pahl, Oliver Stahlhut und Sönke Müller für die sehr intensiven und konstruktiven Diskussionen und die ausgesprochen gute Zusammenarbeit, die wesentlich zum Gelingen dieser Arbeit beigetragen haben, sowie in der Anfangszeit den Kollegen Stefan Growe, Ralf Tönjes und Wolfgang Niem. Die kleinen und großen Probleme wurden weiterhin von den Kollegen Sebastian Weik, Jochen Wingbermhühle, Daniel Schaper und Mathias Schuh unterstützend aufgegriffen.

Auch möchte ich die gute Kooperation mit anderen Instituten ansprechen und die Kollegen Kian Pakzad, Markus Gerke, Stefan Jung und Hans Koch erwähnen, als auch die Diplomanden und studentischen Hilfskräfte Sönke Müller, Tim Lorenz und Peter Kostelnik nennen. Der Firma TopoSys danke ich für die angenehme Zusammenarbeit und die problemlose Bereitstellung von Fernerkundungsdaten.

Zum Schluß möchte ich allen Freunden und Bekannten danken, die das Gelingen der Arbeit durch aktive Unterstützung oder durch Nachsicht gefördert haben.

Porta Westfalica, im April 2003

Eine elektronische Form dieser Arbeit mit farbigen Abbildungen ist unter:
<http://www.tnt.uni-hannover.de/~bueckner/diss.html>
zu finden.

Inhaltsverzeichnis

1	Einleitung	1
1.1	Einordnung der Arbeit	2
1.2	Problemstellung und Motivation	2
1.3	Stand der Technik	3
1.4	Ziele der Arbeit	6
1.5	Aufbau der Arbeit	7
2	Grundlagen der digitalen Bildinterpretation	11
2.1	Begriffe der Bildverarbeitung	12
2.2	Anwendungsfelder	18
2.3	Datenerfassung	18
2.4	Geographische Informationssysteme	20
2.5	Wissensbasierte Luftbildinterpretation	21
3	Operatoren der digitalen Bildverarbeitung	25
3.1	Extraktion von Straßenverläufen aus Fernerkundungsdaten . . .	26
3.1.1	Das Modell	28
3.1.2	Extraktion der Basisinformation	30
3.1.3	Einbringen von Vorinformationen	31
3.1.4	Bestimmung von Startbereichen	34
3.1.5	Erweitern und Verbinden von Straßensegmenten	36
3.1.5.1	Straßensegmenterweiterung	38
3.1.5.2	Bestimmung der Kosten	39
3.1.5.3	Modellverifikation der Einzelkosten	41
3.1.6	Vektorisieren	45

3.1.7	Ergebnisse der Straßenextraktion	46
3.1.8	Ergebnisbewertung	48
3.1.9	Zusammenfassung	50
3.2	Automatisierte Parameter-Adaption	52
3.2.1	Komponenten des Adaptionssystems	54
3.2.2	Parameteradaption für die Straßenextraktion	58
3.3	Schattendetektion unter Verwendung von Laserscandaten	61
3.3.1	Datenmaterial	63
3.3.2	Schattendetektion im visuellen Bild	63
3.3.3	Schattenprojektion	65
3.3.3.1	Bestimmung des Sonnenstandes	65
3.3.3.2	Projektion	66
3.3.4	Radiometrische Entzerrung	69
3.3.4.1	Neigungsbild	69
3.3.4.2	Algorithmische Umsetzung	71
3.3.5	Ergebnisbewertung der radiometrischen Entzerrung	73
4	Wissensrepräsentationsformen und ihre Nutzung	75
4.1	Wissensbasierte Systeme	79
4.2	Regelbasierte Systeme	80
4.3	Logik	82
4.4	Frames	83
4.5	Semantische Netze	84
4.6	Agenten	85
4.7	Unsicheres Wissen	87
4.8	Theorem von Bayes	89
4.8.1	Bayes Netze	90
4.9	Certainty Factors	91
4.10	Fuzzy Logik	91
4.11	Zusammenfassung	92
5	Wissensrepräsentation mit GeoAIDA	93
5.1	Systemanforderungen	94
5.2	Zielsetzung und Anwendungsgebiete	95
5.3	Analyseablauf	96

5.4	Semantische Netze	97
5.5	Strukturelle Komponenten von GeoAIDA	101
5.5.1	Konzeptdefinitionen	101
5.5.2	Instanzen	102
5.5.3	Kanten	104
5.5.4	Attribute	106
5.5.5	Relationen	107
5.5.6	Modalitäten	107
5.5.7	Bewertungen	108
5.6	Analysestrategie	110
5.6.1	Top-Down-Analyse	111
5.6.2	Bottom-Up-Analyse	113
5.6.2.1	Generischer Bottom-Up-Operator	115
5.6.3	Analysestruktur	115
5.6.4	Holistische Operatoren	118
5.7	Zusammenfassung der Eigenschaften von GeoAIDA	121
5.7.1	Systemvergleich	122
6	Interpretationsschritte	125
6.1	Analyseablauf	126
6.2	Generischer Bottom-Up-Operator	128
6.3	Behandlung von Alternativen	130
6.4	Datengetriebene Aufspaltung von Regionen	131
6.5	Topologische Beziehungen	134
6.6	Spezialisierung von Objekten	135
6.6.1	Spezialisierung durch neue Konzeptdefinition	138
6.6.2	Spezialisierung über Attribuierung	139
7	Zusammenfassung	141
	Literaturverzeichnis	145
	Index	157

Abkürzungsverzeichnis

AIDA	Automatic Image Data Analyzer
ALK	Automatisierte Liegenschaftskarte
ATKIS	Amtliche Topographisch-Kartographische Informationssystem
BKG	Bundesamt für Kartographie und Geodäsie
BVO	Bildverarbeitungsoperator
CORBA	Common Object Request Broker Architecture
DGM	Digitales Geländemodell = DTM
DHM	Digitales Höhenmodell = DSM
DKM	Digitales kartographisches Modell
DLM	Digitales Landschaftsmodell
DOM	Digitales Oberflächenmodell
DSM	Digital Slope Modell = DHM
DTM	Digital Terrain Modell = DGM
GeoAIDA	Automatic Image Data Analyzer für geographische Fragestellungen
GIS	Geographisches Informationssystem
GPS	Global Positioning System

IR	Infrarot
KI	Künstliche Intelligenz
MSS	Multispectral Scanner (Sensor)
NIR	nahes Infrarot
OMA	Object Managment Architecture
OMG	Object Managment Group
ORB	Object Request Broker
Pixel	Picture Element, Bildpunkt
RMK	Reihenmeßkamera
SAR	Synthetic Aperture Radar
SNR	Signal/Rausch Verhältnis
TM	Thematic Mapper (Sensor)
UML	Unified Modeling Language
UPN	Umgekehrt Polnische Notation
VIS	Bilddaten im visuellen Bereich
XML	Extened Make-Up Language

Definitionen:

$G = 0, 1, \dots, 255$ Grauwertmenge

$I = g(i, j)$ Bildmatrix eines Grauwertbildes

$i = 0, 1, \dots, L$ Bildzeilen

$j = 0, 1, \dots, M$ Bildspalten

(i, j) Ortskoordinaten eines Bildpunktes

$g(i, j) \in G$ Grauwert des Bildpunktes i, j

Ref_{match} Länge der übereinstimmenden Referenz

Ref_{all} Länge der Referenz

Ex_{match} Länge der übereinstimmenden, extrahierten Daten

Ex_{all} Länge der extrahierten Daten

1 Einleitung

Die Umgestaltung des menschlichen Lebensraumes erfolgt weltweit mit zunehmender Geschwindigkeit als Folge meist anthropogener Einflüsse (“Changing Earth“). In Zukunft muss daher erheblich schneller und stärker diversifiziert die Erfassung und Darstellung dieser Entwicklungen durchgeführt werden, um fundierte Vorhersagen und Entscheidungen für die Zukunft oder im Katastrophenfall treffen zu können. Es zeichnet sich z. B. ab, dass für eine Reihe von Fragestellungen – Landnutzung, Monitoring, Katastrophenschutz – die Landschaftsveränderungen in Zukunft quasi kontinuierlich zu erfassen sein werden.

Die Rechner- und Sensorentwicklung hat Geographische-Informationssysteme (GIS) und Satellitenfernerkundung an die Schwelle operationeller Nutzung gebracht (ATKIS, LANDSAT), ohne dass jedoch bisher die Möglichkeiten gegenseitiger Synergien wissenschaftlich ausreichend erforscht wären. Hierzu soll der vorliegende Ansatz einen methodischen Beitrag liefern, wobei er inhaltlich auf das aktuelle Thema “Erzeugung und Verifikation von Daten für GI-Systeme” bezogen ist.

Dieser Ansatz vereinigt grundlegende Untersuchungen der semantischen und geometrischen Objektmodellierung mit dem Ziel, bei der Interpretation von Fernerkundungsdaten einen möglichst hohen Automatisierungsgrad zu erreichen. Hierbei hat sich das Einbringen von Zusatzwissen sowohl in Form von Wissen über die Zusammenhänge als auch Hinweise aus vorangegangenen Interpretationen aus anderen Informationsquellen in den Interpretationsablauf bewährt.

1.1 Einordnung der Arbeit

Motiviert ist die Arbeit durch Fragestellungen aus dem Bereich der *Angewandten Informatik*, hier im Speziellen der Photogrammetrie, Kartographie oder auch Geo-Informatik, zur automatischen Gewinnung von speziellen Objektklassen aus Fernerkundungsdaten. Ziel ist es in den genannten Disziplinen, aus vorliegenden digitalen Bilddaten Objekte und semantische Kontextinformationen automatisch zu extrahieren, sodass diese in einer Datenbank (GIS) persistent gehalten werden können. Die zusätzlich gewonnenen Attribute und semantischen Zusatzinformationen können ebenfalls Eingang in die Datenbank finden. Diese Datenbestände bilden die Grundlage für Karten unterschiedlicher Anwendungsfelder (Thematische-Karten) in variablen Maßstäben, der elektronischen Präsentation oder als Kolateraldaten für spezielle Fragestellungen, wie etwa der Städteplanung, des Umweltschutzes oder des Katastrophenschutzes.

Zur Umsetzung wurde auf Methoden der *Praktischen Informatik* und *Theoretischen Informatik* zurückgegriffen. Es wurde auf einen modularen Aufbau, eine effiziente Umsetzung der Algorithmen (Parallelisierung von Prozessen) und eine formale Beschreibung von Teilaufgaben geachtet.

1.2 Problemstellung und Motivation

Die große Anzahl bereits existierender Satelliten mit bildgebenden Sensoren erzeugen teilweise kontinuierlich digitale Daten. Diese Datenmengen können bereits heute nicht mehr manuell ausgewertet werden. Hinzu kommen laufend neue Satelliten mit einer steigenden Anzahl von immer höher auflösenden Sensoren. Um eine Nutzung der Daten zu ermöglichen, ist eine Interpretation mit einem möglichst hohen Automatisierungsgrad von großer Bedeutung. Dem gegenüber steht das Problem, dass es sich um stark variierendes Datenmaterial mit sehr komplexem Inhalt handelt und eine Auswertung mit einfachen Mitteln nicht zu erreichen ist.

Grundlegend für den Prozess der Interpretation von Fernerkundungsdaten ist die Verwendung einer Reihe von unterschiedlichen Verfahren zur Bildauswer-

tung. Hierzu gehören Filterverfahren, Kanten-, Textur- und Regionendetektionsverfahren sowie modellbasierte Verfahren und spezielle Objektklassifikationssysteme. Es existieren bereits eine große Anzahl solcher Verfahren, die in der Lage sind, aus speziellen Daten einzelne Objekte oder Objektklassen zu extrahieren. Je nach vorliegendem Datenmaterial und Objekttyp bestehen diese Verfahren aus einfachen Schritten, wie etwa das Setzen eines Schwellwertes für den gesamten Datensatz, bis hin zu Verfahren, die mit sich adaptierenden Modellen versuchen, Objekte in den Daten zu verifizieren.

Existierende Bildverarbeitungssysteme wie *KHOROS*, *HALCON* oder *ERDAS* liefern eine Reihe von Verfahren, um Bilder zu manipulieren oder Bildeigenschaften zu extrahieren. Einige dieser Systeme sind an spezielle Anwendungsgebiete angepasst, wie beispielsweise *ERDAS*, das spezielle Verfahren für die Auswertung von Fernerkundungsdaten beinhaltet.

Die Nachteile der bestehenden Systeme liegen darin begründet, dass Wissen über Zusammenhänge der Objektbeziehungen nur selten genutzt und in die Analyse integriert wird. Weiterhin ist die Zusammenarbeit von symbolischen und subsymbolischen Verfahren für eine automatische Interpretation nicht ausreichend umgesetzt, da diese Verbindung meist nur unidirektional, von der subsymbolischen zur symbolischen Ebene ausgelegt ist. Einen letzten Punkt betrifft die Integration von Wissen aus vorangegangenen Interpretationen, das beispielsweise in geographischen Datenbanken oder als thematische Karte vorliegen kann und dessen Integration in den genannten Systemen nicht berücksichtigt ist.

1.3 Stand der Technik

Die wissensbasierte Bildinterpretation hat zum Ziel, Objekte in einem Bild zu erkennen und diese mit einer symbolischen Bedeutung zu versehen, die der Beschreibung, wie sie ein menschlicher Auswerter erstellen würde, ähnlich ist. Hierbei gliedert sich das Problem der wissensbasierten Bildinterpretation in die Bereiche der subsymbolischen und der symbolischen Interpretation. Der subsymbolische Bereich umfasst die Extraktion von Regionen oder Segmenten im Bildbereich, deren Prozess immer auf der Anwendung von Bildverarbeitungs-

operatoren basiert. Der symbolische Bereich betrifft die Bedeutungszuweisung zu den Regionen oder Gruppierungen dieser und die Gewinnung von semantischer Information zu den extrahierten Objekten.

Als erste Systeme zur wissensbasierten Bildinterpretation von Fernerkundungsdaten sind ECHO [56], VISION [43], ACRONYM [21], SPAM [65], BPI [62], und MESSIE [26] zu nennen. Im Vordergrund steht hier die Erkennung von Einzelobjekten wie Straßen, Gebäude oder von Flugzeugen in Einzelbildern. Die meisten dieser Systeme arbeiten mit Regeln, deren Regelbasen schnell ansteigen und unübersichtlich werden. Ein weiteres Problem bei der Anwendung von Regeln ist, dass die Systeme erst terminieren, falls keine Regel mehr anwendbar ist. Da hier nicht der optimale Weg zum Ziel gewählt, sondern eher jede Alternative getestet wird, sind die Laufzeiten solcher Systeme groß. Einige der Systeme behandeln nur die wissensbasierte Gruppierung von Objektteilen also den symbolischen Teil und gehen dabei von einer idealen Segmentierung aus.

Wharton [105] entwickelte ein wissensbasiertes Expertensystem zur Klassifikation von urbanem Gelände und zur Bestimmung weiterer Landnutzungstypen. Das von Matsuyama beschriebene System SIGMA [75] [76] wurde zur Analyse von Fernerkundungsdaten eingesetzt. Ton [103] kombiniert ein Expertensystem mit Kontextinformationen. Einen guten Überblick über die bis 1990 entstandenen Ansätze zur Bildinterpretation findet man bei Arigialis und Harlow [7], die ihre Ergebnisse folgendermaßen zusammenfassen: “Knowledge based image interpretation can upgrade the state of image analysis capabilities from brute force mathematical and statistical approaches to analysis techniques based on interpretation logic and heuristics.”

Das System ERNEST [88] [73] [74] [89] [58] [59] ist ein allgemeiner Ansatz zur Behandlung von Mustererkennungsaufgaben. Das zu verwendende Vorwissen wird übersichtlich in Form von Semantischen Netzen abgelegt, wobei die Strukturiertheit durch eine Beschränkung auf wenige Relationstypen erreicht wird. Die Analyse eines Datenbestandes wird als Suchproblem aufgefasst, wobei zu einem Zeitpunkt immer der am besten bewertete Suchbaumknoten weiterverfolgt wird. Ein Suchbaumknoten enthält dabei eine in sich konsistente (Teil-) Interpretation des zu analysierenden Datensatzes. Der als nächstes auf dieser

Interpretation auszuführende Analyseschritt wird durch eine feste Anzahl von Regeln ausgewählt.

AIDA [61][101][40] ist ein an die in ERNEST verwendeten Strategien angelehntes System zur Interpretation von Bilddaten. Um das System flexibler zu gestalten, wurden die in ERNEST statischen Regeln durch eine dynamische Regelmengende ersetzt, deren Auswahl über ein Prioritätensystem getroffen wurde. Weiterhin wurde das System in [40] und [80] zur multitemporalen Auswertung erweitert und verwendet.

Neuere Ansätze der wissensbasierten Interpretation [20][9][10] [32][66] von Fernerkundungsdaten sind auch in den kommerziellen Programmen der Firmen ERDAS und Definiens zu finden.

ERDAS bietet mit dem Tool *ExpertClassifier* einen regelbasierten, hierarchischen Klassifikationsansatz, der wissensbasierte Entscheidungsbäume verwalten und nutzen kann. Jede Objektklasse beinhaltet einen eigenen Entscheidungsbaum, der aus Hypothesen, Variablen, Bedingungen und Regeln aufgebaut ist. Die Regeln greifen auf die Variablen und Bedingungen zu, die wiederum aus den Eingangsdaten gewonnen werden. Eingangsdaten können beliebigen Sensoren entstammen oder auch synthetisch aus diesen bestimmt worden sein. Ebenfalls sind Vektordaten z. B. aus einem GIS integrierbar. Unterschiedliche Auflösungen des Rasters werden vom System umgesetzt. Dieser jedoch nur pixelbasierte Ansatz bezieht keine Nachbarschaftsrelationen zwischen Objekten in die Analyse mit ein.

Die Firma Definiens geht mit ihrem Produkt eCognition von den rein pixelorientierten Methoden hin zu segmentorientierten Ansätzen. Die Basis zum Bilden von Segmenten beruht auf der "Fractal Net Evolution-Technik", wobei Homogenitätskriterien zum Bilden der Segmente [19] Verwendung finden. Die komplexere Gruppierung erfolgt anschließend mittels eines Semantischen Netzes, das als wissensbasierter Entscheidungsbaum die eigentliche Klassifikation entweder anhand von "fuzzy logic" basierten Zugehörigkeitsfunktionen oder den Vergleich mit Testflächen (Nahester Nachbar Klassifikator) realisiert. Jedes Segment wird als Knoten im Semantischen Netz dargestellt. Innerhalb des aus Knoten entstehenden hierarchischen Ergebnisnetzes können Relationen zu benachbarten, unter- und übergeordneten Knoten für die Analyse genutzt werden. Nachteil des

Ansatzes ist die Festlegung auf eine Funktion zur Segmentierung der Bilddaten und die Bedingung, dass alle Eingangsdaten über dieselbe Auflösung verfügen. Weiterhin ist der Import von Vektor-Daten nicht möglich und die Segmentierungsergebnisse sind bereits bei kleinen Veränderungen des Bildausschnittes im Überschneidungsbereich nicht reproduzierbar [70]. Die Ursache hierfür liegt in dem Segmentierungsprozess, der weiterhin nicht veröffentlicht ist. Durch den Einsatz von nur einem Verfahren zur Aufteilung der betrachteten Szene wird die Behandlung von Alternativen umgangen, jedoch auf die Potentiale anderer Verfahren verzichtet.

Steinnocher und Bauer [15][97] verwenden die auf der Graphtheorie basierende Software SAMS [11], um die Struktur von klassifizierten Bilddaten zu analysieren. Die Ergebnisse benutzen sie anschließend zur Konstruktion des Netzwerkes von Regeln in eCognition.

Die genannten Systeme sind bisher nicht zur vollautomatisierten Auswertung von Fernerkundungsdaten geeignet, da die Parametrisierung der anzuwendenden Operatoren manuell vorgenommen werden muss und die Ergebnisse sehr stark vom verwendeten Bildmaterial abhängen. Dies liegt bei den Systemen von ERDAS und eCognition darin begründet, dass nur Regionen auf Basis eines Einheitlichkeits-Kriteriums gebildet werden, jedoch keine symbolische Analyse stattfindet.

1.4 Ziele der Arbeit

Die zuvor geschilderten Probleme bei der Integration von zusätzlichem Wissen in den Analyseprozess sollen behoben werden, wodurch nicht nur die Qualität, sondern auch die Plausibilität der Ergebnisse erhöht werden soll. Ein wichtiger Gesichtspunkt dieser Arbeit ist dabei der umfassende Ansatz zur Bildinterpretation, indem sowohl subsymbolische als auch symbolische Verfahren integrativ Verwendung finden, und die Möglichkeit der leichten Anbindung externer Verfahren durch eine modulare Gestaltung des Gesamtsystems.

Die mit diesem hybriden Ansatz einhergehende Erhöhung des Automatisierungsgrades bei der Interpretation von Fernerkundungsdaten stellt ein zentrales

Ziel dieser Arbeit dar. Weiterhin wird auch auf die Problematik der subsymbolischen Verfahren eingegangen, die die grundlegenden Daten für eine strukturelle Analyse liefern und ohne die ein wissensbasiertes System nur eine leere Hülle wäre.

Im Vordergrund steht die Realisierung eines nicht nur zu akademischen, sondern auch für praktische Belange einsetzbaren Systems. Es wurde auf klare Schnittstellen zwischen den Modulen, leichte Integrationsmöglichkeiten für externe Bildverarbeitungsoperatoren und die bereits angesprochene Parallelisierung von Analyseprozessen zur Beschleunigung der Analyse geachtet.

1.5 Aufbau der Arbeit

Das **Kapitel 2** bildet die Einführung in die Thematik der vorliegenden Arbeit. Hier wird der Themenbereich motiviert, grundlegende Begriffe definiert und ein Überblick über den behandelten Themenkomplex gegeben.

Im **Kapitel 3** werden beispielhaft zwei im Rahmen der Arbeit realisierte Bildverarbeitungsoperatoren vorgestellt. Diese und weitere Verfahren bilden die Grundlage der in Kapitel 5 folgenden symbolischen Analyse.

Um den Automatisierungsgrad einer Interpretation von Fernerkundungsdaten zu erhöhen und die Ergebnisqualität zu verbessern, müssen Ansätze gefunden werden, wann die Verfahren, wie sie in Kapitel 3 angesprochenen wurden, anzuwenden sind, wie diese problemspezifisch initialisiert und wie die gegebenenfalls konkurrierenden Ergebnisse verschiedener Verfahren zu bewerten sind. Das **Kapitel 4** gibt einen Überblick und eine Einordnung von Wissensrepräsentationsformen und ihrer Nutzung, die für die vorliegende Fragestellung in Betracht kommt.

Aus der Diskussion der in Kapitel 4 vorgestellten Formalismen resultiert das in **Kapitel 5** beschriebene System. Die verschiedenen Aspekte des realisierten Systems werden hier motiviert und diskutiert.

Die im Rahmen der Untersuchungen erzielten Ergebnisse sind im **Kapitel 6** dargestellt. Anhand konkreter Anwendungsbeispiele aus der Praxis der Luft-

bildauswertung wird die Leistungsfähigkeit des vorgeschlagenen Verfahrens zur Diskussion gestellt.

Mit der Zusammenfassung (**Kapitel 7**) erfolgt ein bewertender Überblick aller Komponenten in Form einer Gesamtdarstellung.

Aus dem Überblick über den Aufbau der Arbeit ist eine Fokussierung auf die zwei wesentlichen Problemstellungen bei der wissensbasierten Interpretation von Fernerkundungsdaten ersichtlich. Bei dem hier zur Diskussion stehenden hybriden Ansatz steht sowohl die Extraktion von symbolischer Information (Kapitel 3) als auch die weitere Verarbeitung der entstandenen Symbole (Kapitel 5) in einem wissensbasierten System im Vordergrund. Die Notwendigkeit des Einsatzes eines wissensbasierten Systems wird durch die Darstellung der Problematiken, wie sie bei der Straßenextraktion (Kapitel 3.1) auftreten, motiviert, wobei der vorgestellte Bildverarbeitungsoperator als repräsentatives Beispiel anzusehen ist.

Die Arbeit spiegelt somit die Hierarchie der in der Bildverarbeitung vorliegenden Ansätze wieder, die im einfachsten Fall durch Bildbearbeitungsoperatoren repräsentiert werden. Hierzu gehören Filteroperatoren oder das Ändern der Helligkeit des Bildes. Diese Operatoren zeichnen sich dadurch aus, dass sie jeden Punkt eines Bildes mit derselben Funktion bearbeiten und folglich keine Strategie besitzen. Sie eignen sich folglich für eine interne Parallelisierung der Bildbearbeitungsaufgabe und sind allgemein einsetzbar. Besitzen diese Operatoren Parameter, so ist die Auswirkung auf das Ergebnis, die durch eine Parameteränderung hervorgerufen wird, vorhersagbar.

Die modellbasierten Verfahren decken spezielle Aufgaben der Erkennung von Objekten aus Bilddaten ab. Beispielsweise werden hier rechteckige und rote oder längliche miteinander verbundenen Objekte extrahiert. Das erste Beispiel könnte Hinweise auf Hausdächer, das zweite Hinweise für Straßen in Fernerkundungsdaten liefern. Diese Verfahren zeichnen sich dadurch aus, dass sie aufwendiger sind als die erstgenannten Bildbearbeitungsoperatoren, jedoch bezüglich ihrer Aufgabenstellung auch erheblich leistungsstärker sind. Die Bilddaten werden nicht mehr monoton, sondern anhand einer internen – impliziten – Strategie abgearbeitet. Die Anwendbarkeit dieser Operatoren ist auf eine mehr oder weniger umfangreiche Klasse von Bildern beschränkt, in denen die Objekte die im

Modell beschriebenen Eigenschaften besitzen.

Die wissensbasierten Verfahren (Kapitel 4) arbeiten auf Symbolen, die in dem hier vorliegendem Kontext durch die oben genannten Verfahren generiert werden und somit die Basis für die auch als symbolische Verfahren bezeichneten Ansätze bilden. Zentral ist hier die Strategie, die unabhängig von den Bilddaten explizit formuliert wird und somit ein höheres Abstraktionsniveau (Metastruktur) darstellt.

Das in Kapitel 5 beschriebene hybride System integriert rein numerische Verfahren der Signalverarbeitung bis hin zu symbolischen Repräsentationen und ihre Verarbeitung (Gruppierung, Strukturierung und Konfliktlösung). Die Aufgaben werden systematisch über die verschiedenen Abstraktionsebenen strukturiert. Hierdurch wird die Basis für eine Lösungsstrategie von komplexen Problemstellungen aus dem Bereich der bildhaften Mustererkennung bereitgestellt.

2 Grundlagen der digitalen Bildinterpretation

Ein allgemein bekanntes und kompetentes Bildinterpretationssystem ist die menschliche visuelle Wahrnehmung. Es ist nicht nur enorm leistungsfähig, sondern auch konzeptionell ästhetisch, da es auf einem Systemmodell basiert. Die automatische Interpretation von Fernerkundungsdaten ist nun der Versuch, vergleichbare Kompetenzen unter Verwendung eines digitalen Systems nachzubilden.

Die Nachbildung des gesamten neurobiologischen Systems des Menschen ist jedoch bisher nicht möglich und bei der zu behandelnden Fragestellung auch nicht Thema, jedoch lassen sich die Funktionen einzelner Komponenten der menschlichen Wahrnehmung mehr oder weniger gut für spezielle Fragestellungen nachbilden. Diese Funktionen sind meist von einer Reihe von Parametern abhängig und bei der Anwendung auf anderes Datenmaterial (neuer Sensor, andere Auflösung) fehleranfällig, wodurch die mit diesen Methoden zuverlässig lösbaren Aufgaben eingeschränkt sind. Die automatische Bearbeitung von komplexen Bildinterpretationsaufgaben ist jedoch bei Themenstellungen wie der Fernerkundung aufgrund der großen und kontinuierlich anfallenden Datenmengen wünschenswert. Da es in näherer Zukunft keine allumfassenden und generellen Bildinterpretationssysteme geben wird, bietet es sich an, durch Einbringen und Akquirieren von Wissen der Beschränktheit, Parameterabhängigkeit und Unzulänglichkeit der existierenden Verfahren entgegenzuwirken.

2.1 Begriffe der Bildverarbeitung

Ziel der Luftbildinterpretation ist die Zuweisung einer Beschreibung oder Bedeutung zu einem Bild und dessen Teilen. Um die Bilddaten in einem Rechner bearbeiten zu können, müssen diese in digitaler Form vorliegen. Die Digitalisierung von analogen Daten geschieht in zwei Schritten, der Rasterung und der Quantisierung. Bei der Rasterung wird das Bild mit einem quadratischen (in seltenen Fällen einem nicht quadratischen, aber rechteckigen) Gitter in Rasterstücke, den Bildpunkten, unterteilt. Die Gitterweite hängt von der Auflösung der vorliegenden Bilddaten und der späteren Anwendung ab. Bei der Quantisierung wird jedem Rasterflächenstück ein Wert zugewiesen. Bei grauwertigen Bildern ist dieses meist ein Wert aus der Menge $G = 0, 1, \dots, 255$, wobei 0 den dunkelsten Wert (schwarz) und 255 den hellsten Wert (weiß) repräsentiert.

Da der Begriff *Bild* ein breites Spektrum von Bedeutungen hat, soll er und weitere, mit der Arbeit im Zusammenhang stehenden Begriffe definiert werden.

Bild: Ein Bild ist eine zweidimensionale in digitaler Form gespeicherte Matrix von Bildelementen p , auch Pixel genannt.

$$\mathbf{I} = g(i, j) \text{ mit: } 0 \leq i < L - 1, \quad 0 \leq j < M - 1 \quad L, M \in \mathbf{N}$$

Hierbei ist L eine ganze positive Zahl, die Größe – Anzahl der Bildpunkte – des Bildes in X-Richtung und M entsprechend für die Y-Richtung. Ein digitales Bild \mathbf{I} ist im einfachsten Fall eine zwei dimensionale Matrix von natürlichen Zahlen.

Bildelement: Kurzform des Bildelement ist Pixel für “Picture Element” als kleinstes Informationselement in einem zweidimensionalen Rastermodell. Ein Pixel enthält als Informationseinheit im einfachsten Fall eine binäre Information. Es können aber auch beliebige Bereiche des ganzzahligen Zahlenraumes, Fließkommazahlen – z. B. bei Höhendaten – oder ein Vektor von Zahlen sein. Im letzten Fall spricht man von mehrbändigen- oder mehrkanäligen Bildern. Die mathematische Darstellung der Bilddaten wird hierfür um eine Dimension erhöht $\mathbf{I} = g(i, j, b)$ mit $0 \leq b < B - 1$, wobei $B \in \mathbf{N}$ die Anzahl der Kanäle

bezeichnet. Bei Bildsequenzen kommt die Zeit als weitere Dimension t hinzu: $\mathbf{I} = g(i, j, n, t)$ mit $0 \leq t < T - 1$ und $T \in \mathbf{N}$.

Die aus der digitalen Bildverarbeitung stammende Bezeichnung Pixel wurde in die Geoinformatik übernommen. Bei Rastermodellen werden die Begriffe nicht immer einheitlich benutzt. Man sagt statt Rasterzelle häufig nur Raster, obwohl mit Raster nur die sich in den Gitterpunkten kreuzenden Gitterlinien gemeint sind. Von einem Gitter (grid, lattice) spricht man, wenn das System der Gitterpunkte gemeint ist. Grid-Strukturen gehören also zur Klasse der Vektor-Modelle.

Bildobjekt: Ein Bildobjekt, auch kurz Objekt, ist eine Menge von Pixeln, die aufgrund ihrer Charakteristika eine strukturierte Entität in einer Domäne darstellen. Die gewählten Charakteristika sind abhängig von der Anwendung.

Extraktion: Die Extraktion betrifft das *Erkennen* von Bildobjekten (hier: in digitalen Bilddaten), die als Grundlage einer Interpretation dienen können. Sie erfolgt anhand gegebener Kriterien und liefert Objekte, denen aber noch keine Bedeutung zugeordnet sein muss.

Fernerkundungsdaten im Speziellen bestehen aus einer Menge von Bildelementen, die Informationen über die reflektierte oder emittierte Strahlung der Erdoberfläche innerhalb eines mehr oder weniger engen Bereiches des elektromagnetischen Spektrums abbilden. Der abgebildete Spektralbereich entscheidet – neben der Auflösung – darüber, welche Objektklassen wie sicher erkannt werden können.

Semantikfunktion: Eine Semantikfunktion oder auch Klassifikation ist eine Abbildung, die jedem Objekt o_i in einem Bild I ein Symbol s_j aus einer Menge von Bedeutungen \mathcal{L} zuweist. \mathcal{L} wird auch als *Bildbeschreibungssprache* bezeichnet und deren Objekte als *Klassen*. Eine Semantikfunktion ist gegeben durch:

i) Eine Menge \mathcal{L} von Symbolen, die die möglichen Bedeutungen der Bildobjekte umfassen.

ii) Die Abbildung: $f_S : (o_1, \dots, o_n | o_i \subseteq I) \rightarrow (s_1, \dots, s_m | s_j \in \mathcal{L})$

Interpretation: Eine vorliegende Interpretation oder Bedeutung ist gegeben durch ein Bild, dessen Objekten eine Bedeutung zugewiesen wurde. Es gibt Definitionen, die die Interpretation als eine Bedeutungszuweisung zu dem gesamten Bildinhalt verstehen (Beispiel: "Dieses ist ein dadaistisches Bild").

Die vorliegende Interpretation eines Bildes wird in dieser Arbeit als die bestehende Zerlegung eines Bildes in Objekte und den Bedeutungen, die diesen Objekten zugewiesen wurden, verstanden. Da ein Bild unter Umständen auf mehrfache Weise zerlegt werden kann, ist auch die Interpretation nicht immer eindeutig bestimmt. Der Prozess der Interpretation wird als Interpretationsprozess oder auch kurz als Interpretation bezeichnet, wobei die Doppeldeutigkeit des Wortes über den Kontext aufgelöst wird in dem es Verwendung findet. Die dem Interpretationsprozess innewohnende Problematik liegt sowohl in der Bestimmung der Bildobjekte überhaupt als auch darin, die Mehrdeutigkeiten durch geeignete Einschränkungen aufzulösen.

Interpretationsprozess: Der Interpretationsprozess (kurz Interpretation) oder auch Bilddeutungsprozess liefert die Zuweisung von Semantik (Bedeutungsinhalt) zu Objekten in Bildern unter Verwendung einer Semantikfunktion.

Die Interpretation eines Bildes oder dessen Bildobjekte, also das Auffinden seiner Semantik, ergibt sich durch die Verwendung eines Modells. Im Modell werden die Beziehungen zwischen den Objekten und deren Geometrie sowie anderen Merkmalen, die für die Interpretation von Bedeutung sind, abgelegt.

Die Interpretation ist eine Abbildung zwischen zwei verschiedenen Repräsentationsformen, wobei der Ursprung konkreter als das Ergebnis ist. Mathematisch ist es eine surjektive Abbildung, die auf einer Zuordnung basiert.

Interpretationen, die auf unterschiedlichen Ansätzen beruhen, führen im Normalfall bei denselben Daten zu unterschiedlichen Ergebnissen. Diese Unterschiede können die Qualität, die Darstellungsform und den Inhalt betreffen. Will man die Ergebnisse qualitativ beschreiben, so müssen die zu Grunde liegenden Modelle und die der Interpretation zu Grunde liegenden Ziele mitberücksichtigt werden.

Bildbearbeitung: Ist \mathcal{B} eine Menge von Bildern, so ist die Bildbearbeitung eine Transformation von existierenden Bildern I in neue Bilder I' . Hierbei wird

jedoch keine Beschreibung der Bedeutungsinhalte erzeugt, also keine Interpretation durchgeführt.

$$f_B : (I_1, \dots, I_n | I_i \in \mathcal{B}) \rightarrow (I'_1, \dots, I'_m | I'_j \in \mathcal{B})$$

Bildverarbeitung: Die Bildverarbeitung ist eine mathematische Methode bzw. ein Algorithmus, der eine Transformation eines Bildes durchführt. Das Ergebnis der Transformation ist nicht zwingend in neues Bild. Es können auch Eigenschaften des Bildes in einer anderen Form repräsentiert werden. Die digitale Bildverarbeitung auf einem Rechner wird durch Bildverarbeitungsoperatoren realisiert, die eine programmtechnische Umsetzung eines Algorithmus darstellen.

Bildklassen: Die Ergebnisse eines Bildverarbeitungsoperators sind in vielen Fällen sehr stark von den verwendeten Eingangsbildern abhängig. Die Anwendung eines Bildverarbeitungsoperators ist daher meist auf eine Klasse von Bildern beschränkt. Werden Bilder einer anderen Klasse verarbeitet, so ist ein sinnvolles Ergebnis nicht garantiert. Die Klassenzugehörigkeit entscheidet sich anhand von zum Bild gehörenden Attributen. Im vorliegenden Themenkomplex z. B. durch den Sensortyp und die Auflösung eines Bildes.

Parameter: Die Parameter eines Bildverarbeitungsoperators umfassen die Daten, die für die Ausführung des ihm innewohnenden Algorithmus notwendig sind. Die verwendeten Bilder und die Ergebnisse gehören nicht zu den Parametern eines Bildverarbeitungsoperators.

Parameterabhängigkeit: (Interbildklassenabhängigkeit) Wird ein Bildverarbeitungsoperator der für eine Klasse von Bildern mit konstanten Parametern gute Ergebnisse liefert, auf eine andere Bildklasse angewendet, die generell für eine Anwendung mit dem vorliegenden Bildverarbeitungsoperator geeignet ist, so müssen die Parametereinstellungen im Normalfall angepasst werden.

Parameterempfindlichkeit: (Intrabildklassenabhängigkeit) Die Ergebnisqualität kann auch innerhalb einer Klasse von Bildern für unterschiedliche Parameterbelegungen unterschiedlich sein. So können etwa kleine Parameteränderungen für einige Bilder einer Klasse zu besseren Ergebnissen führen, wohingegen andere Bilder dieser Klasse hiermit schlechtere Ergebnisse liefern.

subsymbolische Verfahren: Die subsymbolischen Bildverarbeitungsopera-

toren umfassen alle Algorithmen, die Bilddaten verarbeiten. Diese Operatoren liefern gegebenenfalls semantische Informationen zu den vorliegenden Bilddaten, die die Grundlage (*Symbole*) der *symbolischen Verfahren* darstellen. Subsymbolische Verfahren sind implizit in dem Sinne, dass das *Wie* vom Anwender nicht verändert werden kann. Auch ist die intern vorliegenden Strategie nicht nach außen sichtbar – oft interessiert es den Anwender auch nicht, wie eine Aufgabe gelöst wird. Sie werden auch als implizit oder *Low-Level-Verfahren* bezeichnet.

symbolische Verfahren: Diese Verfahren arbeiten auf symbolischer Ebene. Der Ursprung der Symbole ist hier nicht wichtig. Ausschlaggebend sind ihre Attribute und die Relationen der Symbole untereinander. Symbolische Verfahren werden auch als explizite oder *High-Level-Verfahren* bezeichnet, da die Strategie der Verarbeitung der Daten dem Anwender zugänglich gemacht wird und diese auch von ihm modifiziert werden kann.

Hybride Systeme: Hybride Systeme bezeichnen Ansätze, in denen nicht nur ein einziger Formalismus zur Lösung einer Aufgabe verwendet wird, sondern eine Kombination verschiedener Ansätze. Für die unterschiedlichen Aufgaben und die differenten vorliegenden Abstraktionsebenen, in die ein Problem zerlegt werden kann liegen hier einzelne Schritte in einer modularen Form vor. Jede Ebene verfügt über spezielle Funktionalitäten und sie ermöglichen verschiedene Ansichten auf das vorliegende Problem. Diese Schichten sollten hierarchisch sein und von den speziellen Problemen bis zu einer abstrakten Betrachtung reichen. Die Bearbeitung und Lösung von komplexen Problemen wird durch den Einsatz von hybriden Systemen ermöglicht.

Geometrische Auflösung: Die Geometrische Auflösung beschreibt die Anzahl der Pixel pro Längeneinheit (Rasterung, Abtastung).

Radiometrische Auflösung: Die Radiometrische Auflösung gibt die Anzahl der Werte oder die Anzahl Bits an, die zu jedem Pixel gespeichert werden können (Quantisierung).

Vektordaten: Ein Kurvenverlauf wird in Form von Vektordaten als eine Folge von Punktkoordinaten dargestellt. Zusätzlich muss es eine Festlegung geben, wie die Folge von Punktkoordinaten zu interpretieren, also die Punkte zu verbinden sind (Gerade, Polynome, Spline, u.s.w.).

Kurve: Eine Kurve ist ein eindimensionales Primitiv, welches eine Linie stetig abbildet und topologisch zusammenhängt. Weiterhin ist sie durch einen ersten und einen letzten Punkt abgeschlossen.

Modell: Der Begriff *Modell* spielt nicht nur in der hier behandelten Fragestellung eine große Rolle, sondern auch im Forschungsgebiet der Künstlichen Intelligenz. Das Modell steht im Kontext der wissensbasierten Systeme für eine Abstraktion der realen Welt, niedergelegt in Form von Wissen (semantisch strukturierter Information) über einen Themenbereich und wird häufig als (Wissens-) Domäne bezeichnet. Eine Wissensbasis kann in diesem Sinne auch als Modell betrachtet werden und wird auch spezieller als *externes Modell* bezeichnet. Wissensbasierte Systeme verfolgen das Ziel, basierend auf diesem Wissen Schlussfolgerungen zu ziehen und somit neues Wissen zu gewinnen.

Modelle können auch die Basis zur Extraktion von Objekten aus Bilddaten liefern. Hierbei werden die geometrischen, radiometrischen oder topologischen Eigenschaften der Objekte beschrieben. Der hierfür auch verwendete speziellere Ausdruck lautet *konzeptionelles Modell*.

In der Semantik der formalen Logik ist der Modellbegriff ebenfalls zentral: eine erfüllende Variablenbelegung einer logischen (z. B. booleschen) Formel oder einer Menge logischer Formeln ist hier als Modell zu verstehen. In der Semantik einiger nichtklassischer Logiken (wie z. B. der intuitionistischen Logik) werden Modelle als die *möglichen Welten* bezeichnet.

Geometrie und Inhalt: In der Luftbildinterpretation spielen die Begriffe Geometrie und Inhalt eine zentrale Rolle. Über die Bildkoordinate erfolgt der mathematische Zugang zur Bildgeometrie, der Zugang zum Bildinhalt wird über das gesamte Bild realisiert. Hier wird deutlich, dass erst die digitale Form der Daten und der damit verbundene mathematische Zugang eine automatische Bearbeitung (z. B. Geocodierung oder Interpretation) ermöglicht, da hier einem durch eine Bildkoordinate beschriebenen Bereich auf automatisierte Weise eine Bedeutung zugewiesen werden kann.

2.2 Anwendungsfelder

Die Interpretation von Fernerkundungsdaten findet Anwendung in der Photogrammetrie, der Kartographie, den Geowissenschaften und den diesen Bereichen angegliederten Disziplinen. Konkret ergeben sich folgende Einsatzgebiete: das Anlegen, Verifizieren und Aktualisieren von Geographischen-Informationssystemen (GIS), die direkte oder über das GIS indirekte Erstellung von Karten unterschiedlicher Thematiken und Maßstäben, die Bestimmung von Veränderungen im Landschaftsbild für den Küstenschutz, die Städteplanung, Katastrophenschutz oder den Umweltschutz. Eine weitere Anwendung ist das Erstellen von 3D-Visualisierungen im Stadt- oder Landschaftsbereich für die Planung von neuen Objekten, wie z. B. Windkraftanlagen oder den Sendestationen von Mobilfunkstationen.

2.3 Datenerfassung

Die automatische Interpretation von Fernerkundungsdaten hat durch die in den 80-er Jahren einsetzende Entwicklung von digital arbeitenden Sensoren immer mehr an Bedeutung gewonnen. Besonders der Einsatz von Satelliten förderte die Entwicklung und den Einsatz von Bildverarbeitungssystemen und -methoden. Als wegweisend sind hier die Satelliten LANDSAT-MSS und dessen Nachfolger LANDSAT-TM zu nennen, die sich damals durch die folgenden Eigenschaften auszeichneten:

1. Die Daten lagen in digitaler und multispektraler Form vor.
2. Es war möglich, sie global in gleicher Qualität verfügbar zu haben, da eine verlustfreie Vervielfältigung möglich war.
3. Das Kopieren der Daten stellt keinen Aufwand dar, wodurch der Preis niedrig gehalten werden konnte.

In diesem Zusammenhang ist auch die gleichzeitige rasante Entwicklung von Rechnerhardware, Speichermedien und von Übertragungsformen zu sehen,

die den Transport, die Speicherung und die Verarbeitung der Daten erst ermöglichen.

Ein Aufnahmesystem besteht hierbei zum einen aus dem Sensor und zum anderen aus einer Plattform. Bei den Sensoren werden passive Sensoren, die bestimmte von der Erdoberfläche reflektierte Wellen absorbieren (z. B. im optischen Bereich), von den aktiven Sensoren unterschieden. Aktive Sensoren empfangen die von ihnen emittierten Signale und werten das von den Oberflächen reflektierte Signal (z. B. bezüglich Intensität, Phasenlage und Laufzeit) aus. Beispiele sind hier der Laserscanner oder der SAR-Sensor (SAR - Synthetic Aperture Radar).

Die Plattformen sind die Träger der Sensoren und sind meistens Flugzeuge, Satelliten oder auch Ballons. Es können aber auch terrestrische Plattformen, wie hohe Gebäude, zum Einsatz kommen. Die satellitengestützten Systeme liefern fast ausschließlich digitale Ausgangsdaten und besitzen oftmals Multispektralabtaster, mit denen mehrere Spektralbereiche zeitgleich erfasst werden.

Bei den neueren Sensoren werden die Ausgangsdaten fast ausschließlich in digitaler Form ausgegeben. Es gibt aber auch analog arbeitende Systeme und eine große Anzahl in Archiven vorliegende analoge Datenbestände. Diese beiden Formen der Daten, digital und analog, sind jedoch ineinander überführbar, sodass analoge Daten beispielsweise durch Scannen digitalisiert werden können.

Bei den Satelliten-Systemen ist zu beachten, dass diese jederzeit Daten aufnehmen können, ihr Observationsbereich jedoch durch die Flugbahn vorgegeben ist und nur gering – indem der Sensor geschwenkt wird – veränderbar ist. Aus den genannten Gründen werden Satelliten eher für kontinuierliche Beobachtungen eingesetzt und Flugzeuge dort, wo seltene Beobachtungen zu bestimmten Zeitpunkten stattfinden sollen. Flugzeuge finden ihren Einsatz also dort, wo kleinräumige Fernerkundungseinsätze mit hohen Auflösungen zu observieren sind oder die zeitliche Flexibilität eine Rolle spielt [12]. Auch ist die Atmosphäre nicht für alle Wellenlängen durchlässig, sodass nicht alle Frequenzen vom Satelliten aus beobachtbar sind.

Ein nicht unwesentlicher Gesichtspunkt bei den Satellitensystemen ist die Übertragung der aufgenommenen Daten zur Erde [45], da die meisten Beob-

achtungssatelliten nicht geostationär sind und ihre Daten nicht immer direkt zur Empfangsstation auf der Erde übertragen werden können. Um dieses Problem zu lösen, gibt es die folgenden Ansätze:

1. Die Daten sind auf dem Satelliten zwischenspeichern und können beim Überfliegen der Empfangsstation (meist an den Polen der Erde) übertragen werden. Nachteilig ist hier die kurze Zeit für die Übertragung im Gegensatz zur möglichen Aufnahmezeit.
2. Durch ein auf der Erdoberfläche verteiltes Netz von Empfangsstationen kann dieser Nachteil behoben werden.
3. Der Einsatz eines geostationären Relaisatelliten bietet dem Beobachtungssatelliten die Möglichkeit, seine Daten über diesen nahezu kontinuierlich abzugeben. Der Nachteil besteht hier in den zusätzlichen Kosten für den oder die Relaisatelliten.
4. Falls eine ununterbrochene Beobachtung der Erdoberfläche gefordert ist, bietet sich ein gleichmäßig verteiltes Netz von Satelliten an. Dieses kann so dimensioniert werden, dass mindestens immer ein Satellit Kontakt zur Empfangsstation hat. Über diesen können dann auch die anderen Satelliten via Intersatellitverbindungen ihre Daten abgeben.

2.4 Geographische Informationssysteme

In den Industrieländern werden raumbezogene Daten seit etwa 30 Jahren in digitaler Form erfasst oder in eine solche überführt. Ziel bei den kartographischen Anwendungen ist die Erstellung von gegebenenfalls vielfältigen thematischen Karten unter der Verwendung einer gemeinsamen Datenbasis, in der alle Daten gemeinsam und konsistent abgespeichert sind. Weiterhin können digital vorliegenden Datenbestände auch automatisch analysiert werden und beispielsweise für Planungen im Bauwesen oder als Entscheidungshilfe bei Verwaltungen oder in der Politik verwendet werden.

GI-Systeme beinhalten ein einheitliches, räumliches Bezugssystem für jedes enthaltene Objekt und raumbezogene Abfragemechanismen. Ein GIS muss hierfür die folgenden Methoden bereitstellen:

- Verfahren zur Erfassung oder zum Importieren von Daten,
- Persistieren von Daten, üblicherweise durch Ablegen in eine Datenbank,
- Manipulation oder Erweiterung des Datenbestandes, um temporal bedingte Veränderungen zu erfassen,
- Verarbeitung und Analyse des Datenbestandes für unterschiedliche Fragestellungen,
- Ausgabe der Daten in tabellarischer Form oder als Karten.

Zur Neuerfassung und zum Erweitern des Datenbestandes wird häufig auf digital vorliegenden Daten zurückgegriffen. Diese sind zuvor zu digitalisieren, falls sie in analoger Form vorliegen und zu geocodieren. Hierbei werden die Daten in der Art resampelt, dass die Erdoberfläche die geometrische Referenz der Ausgangsdaten darstellt. Anschließend werden die Daten manuell oder teilautomatisiert ausgewertet und die Ergebnisse in ein GIS transferiert.

In Deutschland ist bundeseinheitlich für den mittleren und kleinen Maßstabsbereich das "Amtliche Topographisch-Kartographische Informationssystem" (ATKIS) [41][42][44] eingeführt worden und strukturiert die Landschaft in Objekte und Objektteile. Es wird von den Landesvermessungsämtern und dem BKG (Bundesamt für Kartographie und Geodäsie) erstellt und an öffentliche und private Bedarfsträger abgegeben.

2.5 Wissensbasierte Luftbildinterpretation

Die Verfahren der Bildverarbeitung stellen mathematische Methoden dar, die ein Bild (eine Zahlenmatrix) in eine neue Matrix überführen oder dieser ein bzw. mehrere numerische Werte zuordnet. Da derartige Methoden keine inhaltliche Bedeutung mit dem Bild und seinen Inhalten assoziieren, sind sie in einem

gewissen Sinne *bedeutungslos*. Ist jedoch das Bild nur ein Element in einem komplexeren, globaleren Vorgang wie beispielsweise einem Analyseprozess, so kann unter Verwendung von Bildverarbeitungsoperatoren eine Interpretation stattfinden. Eine derartige Bildinterpretation stellt eine semantische Funktion dar, welche einer Festlegung der möglichen Modelle bedarf. Die Interpretation einer Szene wird dann mittels des eingebrachten a priori Wissens aus den gewonnenen Bedeutungen der Bildteile zusammengesetzt.

Im Kapitel 5 wird sowohl die grundsätzliche Modellierung einer Semantikfunktion als auch das konkrete System GeoAIDA zur Bildinterpretation dargestellt. Das Interpretationssystem GeoAIDA vereinigt sowohl symbolische als auch subsymbolischen Verfahren, die sich wechselseitig unterstützen und aufrufen. Die subsymbolischen Verfahren stellen die bereits angesprochenen Verfahren der *bedeutungslosen* Bildverarbeitung dar, welche auch häufig als *Low-Level-Verfahren* bezeichnet werden, während die symbolische Vorgehensweise als *High-Level-Verfahren* bezeichnet wird und an der Modelldefinition orientiert ist. Die Verbindung oder besser die Zusammenarbeit dieser beiden Ansätze wird dadurch gegeben, dass der symbolische Ansatz Erwartungen, also einschränkende Bedingungen, formuliert, die durch die Ergebnisse des subsymbolischen Ansatz verifiziert oder falsifiziert werden. Die allgemeine Strategie realisiert folglich das Alternieren von subsymbolischen und symbolischen Prozessen, wobei auf der symbolischen Ebene unterschiedliche Abstraktionsebenen mit unterschiedlichen Konkretisierungen durchlaufen werden.

Für die bisher dargestellte Aufgabe stehen zweierlei Arten von Wissen zur Verfügung. Dieses ist zum einen das a priori Wissen über die Zusammenhänge des modellierten Problembereiches, häufig auch als *Welt(-ausschnitt)* bezeichnet, und zum anderen das Wissen über die beim Interpretationsprozess eingesetzten subsymbolischen Verfahren.

Im Bereich der Mustererkennung hat sich der Einsatz von Wissen und das Verwenden von wissensbasierten Verfahren als sehr nützlicher Ansatz bei der Modellierung von semantischer Kognition herausgestellt. Die in den hier relevanten Gebieten entwickelten Verfahren beruhen aber oft nur auf der Verwendung einer einzigen Methodik (subsymbolisch oder symbolisch). Durch den Verzicht der Kombination der unterschiedlichen Ansätze wird jedoch der mögliche

– und notwendige – Synergieeffekt vernachlässigt, der beispielweise durch das Zusammenwirken von *Neuronalen Netzen* (ein Verfahren der impliziten Wissensrepräsentation) als auch von *Semantischen Netzen* (explizite Wissensrepräsentation) erzielt werden kann. So führt die sich wechselseitig bedingende Kombination von Verfahren aus den unterschiedlichen Bereichen zu sichereren und konsistenteren Ergebnissen bei der Bildinterpretation.

Der Graphcharakter ist dabei eine vielen Wissensrepräsentationen zugrunde liegende Struktur (Semantische Netze, Bayes-Netze). Vergleicht man die Semantischen Netze mit den Bayes Netzen, so besitzen die Knoten als Repräsentation von Objekten mit ihren Attributen in beiden Darstellungen vergleichbare Eigenschaften. Unterschiede bestehen hier hauptsächlich in der Bedeutung der Kanten. Diese repräsentieren bei Semantischen Netzen die strukturellen Zusammenhänge (*part_of*, *is_a*, *instance_of*) der Knoten, während sie bei den Bayes-Netzen die stochastischen Eigenschaften abbilden [40].

3 Operatoren der digitalen Bildverarbeitung

In diesem Kapitel werden die verschiedenen Klassen von Bildverarbeitungsoperatoren und einige, die im Rahmen dieser Arbeit entstanden und als Module bei der wissensbasierten Interpretation verwendeten Operatoren dargestellt.

Bildbearbeitungsoperatoren, die ein Eingangsbild in ein Ausgangsbild transferieren, werden in der Literatur [60] in die folgenden drei Kategorien eingeteilt:

1. *Punktoperatoren* - diese verwenden zur Bestimmung eines Bildpunktes im Ausgangsbild genau einen Bildpunkt aus dem Eingangsbild. Typische Operatoren dieser Klasse sind Kontrast- oder Helligkeitsanpassungen.

$$\tilde{g}(i, j) = f_n(g(i, j))$$

2. *Lokale Operatoren* - diese benutzen zur Bestimmung eines Bildpunktes im Ausgangsbild eine lokale Umgebung eines Bildpunktes – der Größe L, M im Eingangsbild. Typische Beispiele dieser Klasse sind der Medianfilter oder die Bestimmung des Gradienten, sowie fast alle morphologischen Operatoren.

$$\tilde{g}(i, j) = g(i, j) * * h(i, j) = \sum_{l=-\frac{L-1}{2}}^{\frac{L-1}{2}} \sum_{m=-\frac{M-1}{2}}^{\frac{M-1}{2}} h(l, m) \cdot g(i - l, j - m)$$

3. *Globale Operatoren* - bei dieser Klasse von Operatoren gehen die gewichteten Bildpunkte des gesamten Eingangsbildes in die Bestimmung eines Bildpunktes im Ausgangsbild ein.

$$\tilde{g}(u, v) = \sum_{l=0}^{L-1} \sum_{m=0}^{M-1} g(l, m) W_l^{ui} W_m^{vj}$$

Für die diskrete Fouriertransformation gilt:

$$W_L = e^{\sqrt{-1} \frac{2\pi}{L}}, \quad W_M = e^{\sqrt{-1} \frac{2\pi}{M}}$$

Des Weiteren gibt es *modellbasierte Verfahren*, denen ein mehr oder weniger komplexes Modell inhärent ist. Diese Operatoren durchsuchen systematisch ein Bild nach dem ihnen vorgegebenen Modell und liefern beispielsweise ein Labelbild mit einer Bedeutung oder eine vektorielle Beschreibung der im Bild gefundenen Modellhinweise.

Der erste, im anschließenden Kapitel 3.1 dargestellte Operator zur Extraktion von Straßen aus Fernerkundungsdaten ist ein typisches Beispiel für einen *modellbasierten Operator*. Kapitel 3.2 stellt einen allgemeinen Ansatz vor, wie die Parameter der Straßenextraktion unter Verwendung eines Referenzdatensatz automatisch adaptiert werden kann. Im darauffolgenden Kapitel 3.3 wird ein aus mehreren Schritten bestehender Operator zur Bestimmung von Schattenbereichen und zur radiometrischen Entzerrung beschrieben.

3.1 Extraktion von Straßenverläufen aus Fernerkundungsdaten

Dieses Kapitel behandelt die Extraktion von Straßen aus Fernerkundungsdaten. In der Literatur wurden in der Vergangenheit bereits verschiedene Verfahren zur Extraktion von Straßen vorgestellt. Frühe Ansätze stammen von Aviad und Carnine [8], die eine längere Folge von Pixelpaaren mit entgegengesetzten Gradientenrichtungen als Straßen interpretierten. Durch Barzohare [14] wurde dieser Ansatz um ein stochastisches Modell erweitert, indem Änderungen der Straßenbreite und der Richtung in vorgegebenen Bereichen zugelassen wurden. Die Verwendung der Radiometrie und Geometrie von Straßen wurde in Baumgartner et al. [16] beschrieben. In dem hier vorzustellenden modellbasierten Ansatz gehen die genannten Aspekte mit ein und werden um ein flexibles Suchverfahren

erweitert, mit dem Ziel die gefundenen Straßenbereiche zu einem Straßennetz zu erweitern.

Die verschiedenen Ansätze können in Klassen eingeteilt werden, da die verwendeten Extraktionsprinzipien stark von den zur Verfügung stehenden Auflösungen sind. Bei groben Auflösungen, wie sie bei Satellitenbildern vorliegen, sind Straßen nur als Linien zu erkennen. Hierfür werden in [102] Pixelketten über ein Template-Matching bestimmt und diese regelbasiert gruppiert. Ein adaptiver rekursiver Entscheidungsprozess wird in [36] vorgeschlagen. Ein auf Snakes basierendes Verfahren für Satellitendaten mit einer Auflösung von 20m ist in [104] zu finden. Snakes sind Verfahren, die zwei Punkte (hier im zweidimensionalen Raum) so miteinander verbinden, dass der Wert einer Gütefunktion optimiert wird. Bei den Snakes wird diese Funktion als *Energie* bezeichnet, deren Wert minimiert werden soll. Die Energie wird entsprechend gegebener Maße (Krümmung, Radiometrie) anhand des gesamten Linienverlaufs iterativ bestimmt.

Bei Auflösungen im Bereich zwischen 0,5m und 5m ist eine Straße als flächiges Objekt zu erkennen. Die hier verwendeten Verfahren müssen diesen Sachverhalt berücksichtigen. Die meisten Verfahren bestimmen sicher erkennbare Straßenbereiche, die entweder über eine erweiternde Suche oder über Snakes miteinander verbunden werden. Bei den Snakes wird versucht, benachbarte Straßenstücke miteinander zu verbinden. Bei der Bestimmung der Energie geht hier die Parallelität der Randbereiche der Straßenhypothese mit ein. Beispiele hierfür sind in [78] und [109] zu finden. Bei den erweiternden Suchverfahren [39][46] werden die beiden Straßenränder durch zwei separate Suchprozesse für die Straßenrandbereiche erfasst, deren Ergebnisse über eine verbindende Funktion abgestimmt werden. Die Suche beginnt an den Enden der gefundenen Straßenstücke und versucht, diese zu erweitern und falls möglich mit anderen Straßenstücken zu verbinden.

Hier wird ein modellbasiertes Verfahren vorgestellt, das versucht, die Straßen durch einen erweiternden Suchprozess zu einem Straßennetz zu verbinden. Weiterhin wird ein Qualitätsmaß für die erzielten Ergebnisse unter Verwendung eines Referenzdatensatzes entwickelt.

3.1.1 Das Modell

Das verwendete Modell beschreibt die Eigenschaften der Objekte – hier der Straßen – durch die, ihnen inhärenten Eigenschaften, die sich nach der Aufnahme im Datenmaterial wiederfinden. Zu unterscheiden sind dabei zum einen die sensorspezifischen Eigenschaften und zum anderen die Auflösung der gewonnenen Daten.

Straßen erscheinen bei Auflösungen $\geq 5\text{m}$ nur noch als Linie und prägen sich bei feineren Auflösungen als flächenhafte Objekte aus. Im hier gewählten Ansatz wird von flächenhaften Straßenbereichen ausgegangen, wobei zu erwähnen ist, dass die linienhafte Darstellung für die Erfassung von globaleren Eigenschaften wie etwa der Netzstruktur nützlich sein kann. Diese Darstellung kann bei höherauflösendem Datenmaterial durch einen Resampling-Prozess in einer sogenannten AuflösungsPyramide gewonnen werden. Bei Auflösungen $\leq 0.5\text{m}$ kommen immer mehr Teilstrukturen der Straße zum Vorschein, die als Einzelobjekte modelliert werden müssten. Hierzu gehören die Fahrbahnmarkierungen, Fahrzeuge und Begleitobjekte wie Gehwege oder Gräben. Durch diese Ausführung wird deutlich, dass die Komplexität des Modells von der Auflösung abhängig ist. Der beschriebene Ansatz arbeitet bezüglich der obigen Diskussion in einem Bereich von 0.5m bis 5m optimal.

Eine Straße prägt sich folglich als überwiegend länglicher Bereich aus. Dieser besitzt über eine längere Strecke eine konstante Breite und eine annähernd gleiche Radiometrie in den Daten. Weiterhin bildet die Menge der Straßen in einem Bild ein Straßennetz. Straßen, die nicht mit anderen Straßen verbunden sind, können nur in Randbereichen der Bilddaten auftauchen, da isolierte Verkehrswege nicht den ihnen zugewiesenen Sinn erfüllen. Die Straße hebt sich weiterhin durch Luminanzänderungen an ihren Randbereichen zu ihrer Umgebung ab. Die bisher genannten Eigenschaften sind in allen hier betrachteten Sensoren zu beobachten, jedoch mit unterschiedlichen Ausprägungen der einzelnen Attribute. Abbildung 3.1 zeigt ein Luftbild im visuellen Bereich und einen vergleichbaren Ausschnitt, aufgenommen mit einem SAR Sensor.

Zu sehen ist, dass die Luminanz der Straßen in beiden Sensoren annähernd konstant ist, jedoch im SAR-Bild zu den dunkleren Bildbereichen gehört, während



Abbildung 3.1: Eingangsdaten: Luftbild (visuell) und SAR-Bild

sie im visuellen einen mittleren bis hohen Luminanzwert besitzt. Das parametrisierbare Modell hilft hier diese unterschiedlichen Ausprägungen der gesuchten Objekte für den Operator zu beschreiben.

In den meisten Fällen erscheinen Straßen entweder heller oder dunkler als ihre Umgebung und es genügt, ein Modell für homogen berandete Bereiche zu verwenden. Der Luminanz- und Gradientenverlauf ist für homogen berandete Bereiche im linken Teil der Abbildung 3.2 zu sehen. Zur Detektion von Straßen, deren Luminanz zwischen denen der beiden umgebenden Regionen liegt (s. rechter Teil Abbildung 3.2), kann das Modell des Algorithmus zusätzlich für diesen Helligkeitsverlauf aktiviert werden.

Mit diesem parametrisierten Modell werden anfangs sichere Straßensegmente aus dem Bild extrahiert, die anschließend unter Verwendung eines A* Suchalgorithmus [71], [107] erweitert werden. Das verwendete Modell ist anfänglich sehr strikt und wird bei der Erweiterung immer toleranter, sodass auch gewisse Abweichungen vom Modell zulässig werden. Hierdurch wird es möglich, Straßenbereiche, die durch Teilverdeckungen vom Modell abweichen, zu extrahieren. Gründe für die Abweichung der erkennbaren Straßenbereiche vom Modell können beispielsweise durch Verdeckungen, parkende Autos oder Schattenwurf

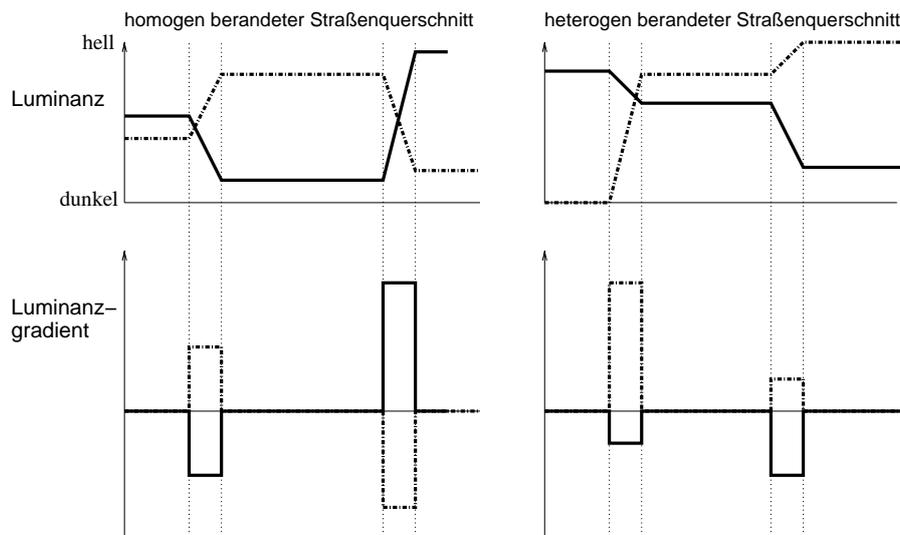


Abbildung 3.2: Schematische Darstellung der Straßenquerschnitte

auftreten.

Zu den bisher beschriebenen Eigenschaften der Modellverifikation können dem Algorithmus noch weitere Vorinformationen übergeben werden, um den Analysevorgang robuster zu gestalten und eine Beschleunigung zu erreichen. Zum einen können vorhandene GIS-Daten über bekannte Straßenverläufe bei der Analyse mit einfließen und zum anderen schränkt ein optionales Binärbild die Bereiche ein, in denen Straßen erwartet werden. Dieses Binärbild kann ebenfalls GIS-Informationen als Eingangsdaten nutzen oder von anderen Bildverarbeitungsoperatoren stammen. Als Beispiel sei hier die Gebäudeerkennung genannt, wodurch der Straßenextraktion die Bereiche, an denen Gebäude erkannt wurden, mitgeteilt werden und diese bei der Suche nach Straßenkandidaten nicht mehr betrachtet werden müssen. Im Folgenden werden die einzelnen Schritte des Algorithmus detaillierter beschrieben.

3.1.2 Extraktion der Basisinformation

In einem ersten Schritt wird unter Verwendung eines Sobel-Operators aus dem Eingangsbild $e(i, j)$ ein Gradientenbild berechnet, dessen Betrag $g(i, j)$ und Richtung $d(i, j)$ bestimmt wird. Mit den in Abbildung 3.3 dargestellten Fal-

tungsmasken wird das Eingangsbild gemäß Gleichung 3.1 und 3.2 gefaltet. Bei stark verrauschten Bilddaten können Masken der Größe $5 * 5$ zur Berechnung ausgewählt werden.

$$\text{Faltung: } e(i, j) ** h(i, j) = \sum_{l=-\frac{L-1}{2}}^{\frac{L-1}{2}} \sum_{m=-\frac{M-1}{2}}^{\frac{M-1}{2}} h(l, m) \cdot e(i-l, j-m) \quad (3.1)$$

$$\text{Komponenten: } g_H(i, j) = e(i, j) ** h_H(i, j) \quad g_V(i, j) = e(i, j) ** h_V(i, j) \quad (3.2)$$

$$h_H = \begin{array}{|c|c|c|} \hline 1 & 0 & -1 \\ \hline 2 & 0 & -2 \\ \hline 1 & 0 & -1 \\ \hline \end{array} \quad h_V = \begin{array}{|c|c|c|} \hline 1 & 2 & 1 \\ \hline 0 & 0 & 0 \\ \hline -1 & -2 & -1 \\ \hline \end{array}$$

Abbildung 3.3: Faltungskerne der Richtungskomponenten einer 3x3 Matrix

Der Betrag des Gradienten, dessen bildliche Darstellung in 3.4 zu sehen ist, ergibt sich nach Gleichung 3.3.

$$g(i, j) = \sqrt{g_H(i, j)^2 + g_V(i, j)^2} \quad (3.3)$$

Die Richtung des Gradienten wird in einem weiteren Bild $r(i, j)$ quantisiert in acht Richtungen gespeichert. Richtungswerte, deren zugehöriger Gradientenbetrag kleiner als eine anzugebende Schwelle ist, erhalten den Wert 0 (keine signifikante Richtung). Die Einteilung in acht Richtungen reicht für die folgenden Betrachtungen aus und vereinfacht die weiteren Analyseschritte. Das Ergebnis der Richtungsbestimmung ist für einen Ausschnitt des Luftbildes aus Abbildung 3.1 in Abbildung 3.5 zu sehen. Die beiden berechneten Bilder und das Eingangsbild dienen als Datengrundlage für alle kommenden Schritte der Straßenextraktion.

3.1.3 Einbringen von Vorinformationen

Vorinformationen können wie bereits angesprochen als Binärbild oder in Form von GIS-Daten in die Analyse einfließen. Um hierfür die anschließenden Analy-



Abbildung 3.4: Gradientenbild $g(i, j)$ des Luftbildes aus 3.1

seschritte nicht speziell anpassen zu müssen, werden diese Informationen bereits an dieser Stelle eingebracht, als Schritt zwischen der Gewinnung der Basisinformationen und der eigentlichen Analyse. Es wird hier der Ansatz verfolgt, dass die einzubringenden Vorinformationen nur das Gradientenbild verändern und so bei der Auswertung sinngemäß Einfluss nehmen. Dies hat den Vorteil, dass zum jetzigen Zeitpunkt keine harten Entscheidungen getroffen werden, sondern nur

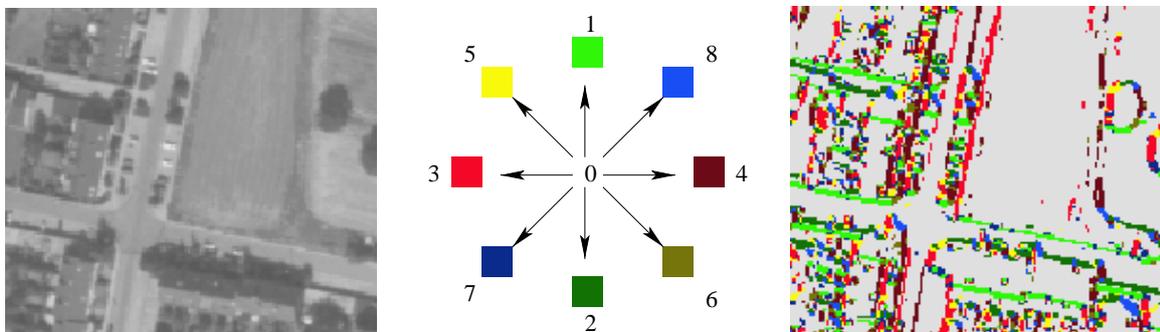


Abbildung 3.5: Quantisierte Richtungen der signifikanten Gradienten

die Hinweise für eine Straße an den gegebenen Stellen verstärkt werden. Auch muss der folgende Auswertungsprozess bei diesem Vorgehen nicht geändert werden.

Die Binärdaten, mit denen Bereiche von der Analyse ausgeschlossen werden, verändern das Gradientenbild so, dass in den Bereichen, in denen nicht nach Straßen gesucht werden soll, alle Werte auf 0 gesetzt werden. Im Gegensatz zu dieser binären Entscheidung verstärken die GIS-Daten nur die Hinweise für das Vorhandensein von Straßenhypothesen. Bei der Einbringung von GIS-Daten ist die zugrunde liegende Idee, die Gradientenwerte an den Stellen, an denen Straßen durch die GIS-Daten prognostiziert werden, anzuheben. Dies geschieht durch die Multiplikation mit einem Gewichtungsfaktor und bietet so der Analyse verstärkte Hinweise, an welchen Stellen sich die Straßenränder befinden sollten.

Der Faktor, mit dem die Gradientenwerte multipliziert werden, ist einem, aus den GIS-Daten erstellten Wahrscheinlichkeitsbild $p(i, j)$ entnommen. In diesem Bild ist für jede Richtung (0° , 45° , 90° , 135°) ein Gewichtungsfaktor eingetragen. Es werden jedoch nur die Gradienten erhöht, deren Richtung mit den von den GIS-Werten vorgegebenen Richtungen korrespondieren. Die Berechnung der neuen Gradienten ergibt sich nach Gleichung 3.4. Hierbei liefert die Funktion $f_r(p)$ die Richtung und die Funktion $f_w(p)$ den Gewichtungsfaktor ≥ 1 für den Bildpunkt (i, j)

$$\bar{g}(i, j) = \begin{cases} g(i, j) & : f_r(p(i, j)) \neq r(i, j) \\ g(i, j) * f_w(p(i, j)) & : f_r(p(i, j)) = r(i, j) \\ 255 & : \text{sonst} \end{cases} \quad (3.4)$$

Abbildung 3.6 veranschaulicht für einen Ausschnitt, wie aus dem Gradientenbild (links) durch die für ihre Richtungen farblich dargestellten Gewichtungen der GIS-Daten ein neues Gradientenbild (rechts) bestimmt wird.

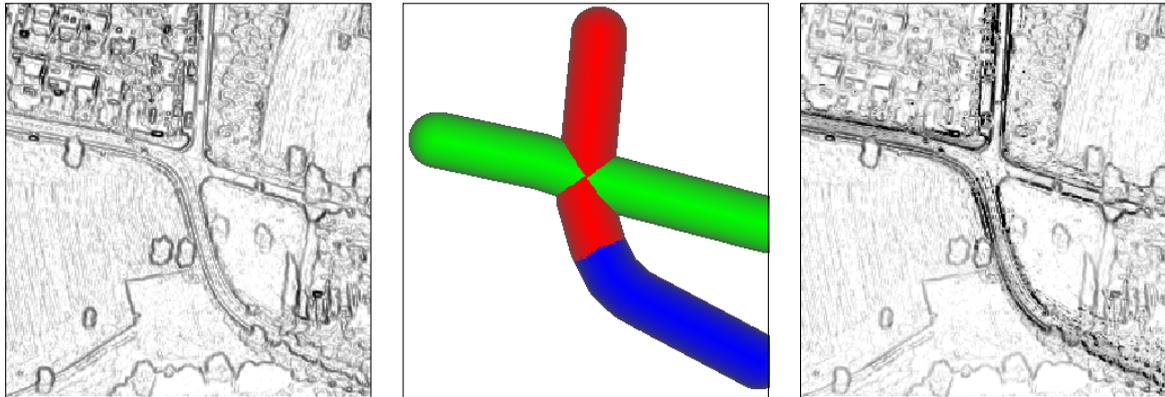


Abbildung 3.6: Auswirkungen der GIS-Daten auf das Gradientenbild

3.1.4 Bestimmung von Startbereichen

Zu bestimmen sind nun Straßenabschnitte, die allen aktuellen Parametervorgaben zum Modell genügen. Diese *Startbereiche* sollen durch eine spätere Suche erweitert werden. Zur Bestimmung von korrespondierenden Pixelpaaren, die einen Straßenrand markieren, sind für jeden Eintrag im Richtungsbild $r(i, j)$ mögliche, zugehörige Pixel zu bestimmen. Hierzu wird in Gradientenrichtung und in der entgegengesetzten Richtung in einem Suchfenster, welches durch die minimale und maximale Straßenbreite gegeben ist, nach möglichen zugehörigen Pixeln gesucht. Die Richtung der möglichen Pixel muss sich um 180° vom betrachteten Pixel unterscheiden. Beim *heterogenen Modus* kommen die identischen Richtungswerte als mögliche Kandidaten noch hinzu.

Für die Menge der korrespondierenden Pixelpaare, die das Merkmal der Richtungszuordnung erfüllen, ist nun im Eingangsbild $e(i, j)$ zu überprüfen, ob die Grauwerte der eingeschlossenen Straßenpixel im zulässigen Bereich liegen. Begrenzt wird dieser Bereich durch die Modellparameter minimaler und maximaler Straßengrauwert, die dem Programm als Parameterwert übergeben werden. Die verbleibenden Paare werden als Menge potentieller Pixelpaare übernommen und für jede Alternative wird der Mittelpunkt und die Richtung der Straße bestimmt. Der ebenfalls in acht Richtungen quantisierte Wert der Straßenrichtung wird an die Stelle des bestimmten Mittelpunktes in der Matrix $q(i, j)$ eingetragen. Das Tripel von zwei Randpunkten und dem richtungsbehafteten Mittel-

punkt wird im Folgenden als *Straßenquerschnitt* bezeichnet.

Bei möglichen, konkurrierenden Eintragungen in $q(i, j)$ für unterschiedliche Straßenquerschnitte wird nur der Mittelpunkt mit den deutlichsten Begrenzungen eingetragen. Bestimmt wird die beste Alternative über die Summe der Gradienten der zugehörigen Randpixel.



Abbildung 3.7: Straßensegmente mit Rand- und Mittelpunkten als Startbereiche

Aus der Gesamtmenge der bestimmten Querschnitte sollen nun möglichst große Teilmengen gebildet werden, die einen Straßenverlauf darstellen. Hierzu werden, beginnend an einer Startposition, benachbarte Straßenquerschnitte schrittweise über eine Bestensuche zu *Straßensegmenten* verbunden. Bei dieser Suche geht die bisherige Richtung der Straße, deren mittlere Breite und der mittlere Grauwert der gefundenen Straßenregion für die Bestimmung eines Kostenwerts mit ein. Dieser Suchvorgang lässt das Überspringen von kleinen Störstellen, in denen es keinen passenden Richtungswert im Bild $q(i, j)$ gibt, zu. Passende Richtungswerte besitzen auf beiden Seiten identische oder einen um maximal 45° abweichenden Richtungswert. Abbruchbedingungen bei dieser Suche sind:

- Es wurden keine Straßenquerschnitte mit passendem Richtungswert gefunden.

- Die Straßenbreite über- oder unterschreitet die bei der Modellinitialisierung vorgegebenen Werte.
- Die den Straßenquerschnitt repräsentierenden Luminanzwerte liegen nicht im vorgegebenen Intervall.

Das bei der Suche gefundene Ergebnis mit den geringsten Kosten wird als akzeptiertes Straßensegment gespeichert, falls die Länge die vorgegebene Mindestlänge übersteigt und die enthaltenden Störstellen weniger als 20% ausmachen. Die gefundenen Straßensegmente erfüllen die durch die Modellvorgaben beschriebenen Bedingungen sehr streng, bilden jedoch noch kein Straßennetz, sondern repräsentieren, wie in Abbildung 3.7 zu sehen ist, gut erkennbare Straßenstücke, von deren Enden aus eine weitere Suche ansetzen kann, um das Straßennetz zu vervollständigen.

3.1.5 Erweitern und Verbinden von Straßensegmenten

Die isolierten Straßensegmente sollen nun erweitert und falls möglich miteinander zu einer Netzstruktur verbunden werden. Beginnend mit den Segmentenden als Startpunkte wird versucht, die Straßensegmente zu erweitern. Da es auch hier, wie im vorangegangenen Kapitel 3.1.4 Alternativen gibt, die beim Erweitern zu betrachten sind, wird ein Suchalgorithmus verwendet, um diese konkurrierenden Ergebnisse zu verwalten. Zum Einsatz kommt ein A* Algorithmus, wobei der insgesamt kostengünstigste Weg die Modellvorstellung am besten erfüllen soll und damit dem Modell am besten entspricht. Die A* Suche ist im Folgenden algorithmisch beschrieben:

Struktogramm des A*-Suchalgorithmus

- 1 Bilde eine Liste aus Teilwegen! Die anfängliche Liste soll die Null-Länge und einen Null-Schritt-Weg von der **Wurzel** aus besitzen.
- 2 Bis die Liste leer oder das **Ziel** erreicht ist, bestimme, ob der erste Weg in der Liste den Zielknoten erreicht.
 - 2a Wenn der erste Weg das Ziel erreicht \Rightarrow Suche beendet.
 - 2b Wenn der erste Weg das Ziel nicht erreicht,
 - 2b1 nehme den ersten Weg aus der Liste,
 - 2b2 bilde durch **Erweiterung** des entfernten Weges um einen Schritt neue Wege,
 - 2b3 füge die neuen Wege in die Liste ein,
 - 2b4 sortiere die Liste nach der Summe der bisher angefallenen **Kosten** inklusive der **Restkosten**, und setze die Wege, die die niedrigsten Kosten verursachen, an die Spitze,
 - 2b5 wenn zwei oder mehr Wege einen gemeinsamen Knoten erreichen, streiche alle diese Wege bis auf den, über den der gemeinsame Knoten mit den geringsten Kosten zu erreichen ist.
- 3 Wenn der Zielknoten gefunden wurde, melde **Erfolg**, sonst **Misserfolg**.

Die Umsetzung des Algorithmus auf das vorliegende Problem geschieht folgendermaßen:

Die **Wurzeln** entsprechen den Straßensegmentenden, also den Startpunkten der kommenden Suche. Die **Erweiterung** des Weges entspricht einer Erweiterung des linken und rechten Straßenrandes um einen Randpixel. Das **Ziel** ist erreicht, wenn durch die schrittweise Erweiterung, beginnend an der Wurzel, das erweiterte Straßensegment mit einem anderen Straßensegment verbunden werden kann. Ein weiteres **Ziel** ist das Erreichen einer vorgegebenen Suchtiefe. An dieser Stelle wird der Suchalgorithmus neu initialisiert, wodurch die aktuelle Position zur neuen Wurzel wird. Hierdurch wird vermieden, dass die Suchbäume des A*-Algorithmus zu groß werden. Da durch diesen Schritt die bis dahin

schlechteren Alternativen verworfen werden und diese nicht weiter betrachtet werden müssen, sinkt auch die Rechenzeit des Verfahrens.

Aus dem Grad der Übereinstimmung einer gefundenen Erweiterung mit dem Modell werden die **Kosten** für diese und ihre Alternativen erstellt. Die beim A*-Algorithmus optimistisch abzuschätzenden **Restkosten** werden hier als Projektion der bisher angefallenen Kosten auf die verbleibende Reststrecke bestimmt. Diese Projektion genügt nicht den theoretischen Erfordernissen des A*-Algorithmus, bei dem die **Restkosten** immer kleiner als die tatsächlichen Restkosten sein sollen, um das Erreichen der besten Lösung zu garantieren. Da es im vorliegenden Fall jedoch nicht möglich ist, die tatsächlichen Restkosten abzuschätzen, bietet der hier gewählte Ansatz einen Kompromiss. Er beschleunigt die Analyse, indem die vielversprechendsten Wege bevorzugt expandiert werden. Der Algorithmus bricht mit **Misserfolg** ab, falls die Kosten des kostengünstigsten Weges über den zulässigen Maximalkosten liegen.

3.1.5.1 Straßensegmenterweiterung

Bei der Erweiterung der Straßensegmente ist weiterhin das zugrunde liegende Modell zu beachten. Hierbei ist besonders der Ort der beiden korrespondierenden Randpunkte, die Richtung der beiden Straßenbegrenzungen und die Richtung der Straße zu beachten. Die zu bestimmenden Richtungen werden auch hier wieder auf acht mögliche Werte quantisiert.

Die Richtung der Straßenränder ergibt sich aus dem letzten und den drittletzten Bildpunkt (siehe Abbildung 3.8 die *blaue* Kennzeichnung) des betrachteten Randes. In dieser Abbildung sind die möglichen nachfolgenden Kandidaten, die sich durch die in *Rot* dargestellte Richtung der Straßenmittellinie ergeben, für die beiden Randbereiche durch ? symbolisiert. Die möglichen Kandidaten weisen von den bereits Gefundenen weg. Durch den Bezug zur Mittellinie wird das *Auseinanderlaufen* der Straße verhindert. Bei der Verfolgung von Kurven ist zu beachten, dass ein zur Innenseite der Kurve zugehöriger Pixel mehreren Pixeln der Außenkurve zugehörig sein kann. Dieses ist in Abbildung 3.9 dargestellt und bedeutet, dass der Algorithmus die unterschiedliche Randpixelanzahl bei Kurven im Straßenverlauf beachten muss.

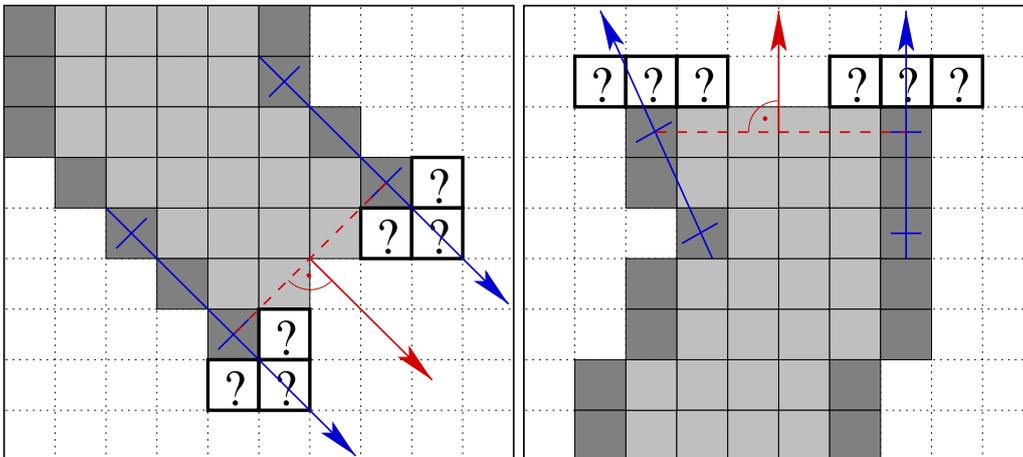


Abbildung 3.8: Potentiell nachfolgende Randpunkte für zwei Situationen

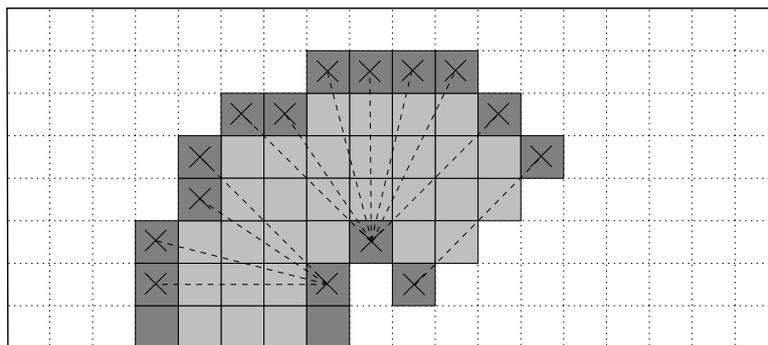


Abbildung 3.9: Zugehörigkeit der Randpixel bei Kurven

3.1.5.2 Bestimmung der Kosten

Die Kosten stellen die zentrale Steuergröße des A*-Algorithmus dar. Sie entscheiden nach jeder Erweiterung eines Teilpfades, welcher Pfad als nächster zu betrachten ist. Am Ende eines Suchprozesses – bei Erreichen der Suchtiefe t – entscheiden die Kosten, welcher Pfad die optimale Erweiterung darstellt. Bei der Bestimmung der Kosten für einen Teilpfad gehen sowohl die Modellparameter als auch die gefundenen Merkmale eines Straßensegmentes mit ein.

Beim verwendeten A*-Algorithmus ergeben sich die Gesamtkosten $\mathcal{C}(\mathcal{N})$ als Summe der bis zu einem Schritt angefallenen Kosten \mathcal{C}_A , zuzüglich der Restko-

sten \mathcal{C}_R . Der Wert w_c im Wertebereich $[0, 1]$ gewichtet hierbei die Restkosten.

$$\mathcal{C}(\mathcal{N}) = \mathcal{C}_A + w_c * \mathcal{C}_R \quad (3.5)$$

Die Kosten \mathcal{C}_A sind die Summe aller Einzelkosten der Wegelemente des bisherigen Teilweges bis zur Position n .

$$\mathcal{C}_A = \sum_{i=1}^n \mathcal{C}_i \quad (3.6)$$

Die Restkosten \mathcal{C}_R müssen optimistisch, das heißt geringer als die tatsächlich anfallenden Kosten bestimmt werden, um das Erreichen des besten Ziels zu garantieren [107]. Da eine solche Abschätzung bei der hier vorliegenden Problemstellung nicht möglich ist, auf die Restkosten aus Effizienzgründen jedoch nicht verzichtet werden soll, werden diese als Projektion der bisher angefallenden Kosten für den bereits gefundenen Bereich auf die Reststrecke $(t - n)$ bestimmt. Durch die Gewichtung von \mathcal{C}_R mit dem Faktor $w_c < 1.0$ in Gleichung 3.5 wird der Forderung nach einer optimistischen Abschätzung nicht immer genüge getan. Diese Wahl der Restkosten gewährleistet folglich nicht, dass immer der optimale Pfad gefunden wird. Die gewählte Art der Restkostenabschätzung erscheint beim vorliegenden Problem jedoch akzeptabel, falls der Gewichtungsfaktor w_c klein genug gewählt wird und führt zu einer Einschränkung des Suchraumes.

$$\mathcal{C}_R = \mathcal{C}_A * \frac{t - n}{t} \quad (3.7)$$

Die Kosten für ein Wegelement \mathcal{C}_i ergeben sich aus dem linken und rechten Rand $(\mathcal{C}_r, \mathcal{C}_l)$. Hierbei werden sowohl die Eigenschaften der einzelnen Randpunkte, als auch die Eigenschaften des gebildeten Straßensegments betrachtet (\mathcal{C}_s) . Die Gewichtung der Randpunkte gegenüber dem Straßensegment wird durch den Parameter w_p eingestellt.

$$\mathcal{C}_i = (1 - w_p) * (\mathcal{C}_l + \mathcal{C}_r) + w_p * \mathcal{C}_s \quad (3.8)$$

Die Kosten \mathcal{C}_l (links) resultieren aus den Übereinstimmungen des linken Randpunktes mit den Modellparametern. Gleiches gilt für die Kosten \mathcal{C}_r des rechten Randes.

$$\mathcal{C}_l = \sum_{E_i \in E_l} E_i * w_e \quad (3.9)$$

$$\mathcal{C}_r = \sum_{E_i \in E_r} E_i * w_e \quad (3.10)$$

Die Kosten \mathcal{C}_s ergeben sich entsprechend aus der Übereinstimmung des Straßensegments mit den zugehörigen Modellparametern.

$$\mathcal{C}_s = \sum_{E_i \in E_s} E_i * w_s \quad (3.11)$$

3.1.5.3 Modellverifikation der Einzelkosten

Die im vorangegangenen Kapitel 3.1.5.2 in Gleichung 3.9-3.10 aufgeführten Einzelkosten E_i sollen im Folgenden näher bestimmt werden. Hierbei gehen die folgenden Modellparameter und die aus den Bilddaten gewonnenen Merkmale zur Bestimmung der Einflüsse ein:

1. Modellparameter

- a) b_{min} - minimale Straßenbreite in Pixel einschließlich Randpunkte
- b) b_{max} - maximale Straßenbreite in Pixel einschließlich Randpunkte
- c) l_{min} - minimaler Luminanzwert eines Straßenoberflächenpixels
- d) l_{max} - maximaler Luminanzwert eines Straßenoberflächenpixels

2. Merkmalbasierte Einflüsse

- a) μ_l - Luminanzmittelwert eines Straßenquerschnitts
- b) σ_l Standardabweichung der Luminanz eines Straßenquerschnitts
- c) δ Straßenbreite
- d) ρ Übereinstimmung der beiden Randrichtungen [0,1]
- e) ϱ Übereinstimmung der Randgradienten und der Straßenrichtung
- f) Δ Randgradienten der Randpunkte
- g) ν Anschluss an eine bereits gefundene Straße
- h) v Potentielle Verbindung zwischen Straßensegmenten

Alle im Folgenden mit E beginnenden Größen symbolisieren Einzelkosten. Diese in den oben genannten Formeln aufgeführten Größen umfassen jeweils den Wertebereich $[0, 1]$, deren Einfluss auf die Kosten durch die zugehörigen Gewichte gesteuert werden kann.

a) Mittelwert der Luminanz μ_l :

Der Mittelwert μ_l ergibt sich aus den zum Straßenquerschnitt S_q gehörenden Pixelwerten des Eingangsbildes $e(i, j)$.

$$\mu_l = \frac{1}{n} \sum_{(i,j) \in S_q} e(i, j) \quad \text{mit} \quad n = \#S_q \quad (3.12)$$

Der Luminanzmittelwert $\bar{\mu}_l$ eines Straßensegments Tw bestimmt sich über alle Straßenquerschnitte dieses Teilweges.

$$\bar{\mu}_l = \frac{1}{n} \sum_{\mu_l \in Tw} \mu_l \quad \text{mit} \quad n = \#Tw \quad (3.13)$$

Der zugehörige Einfluss E_{μ_l} und $E_{\bar{\mu}_l}$ ergibt sich aus dem Vergleich von μ_l mit den Modellparametern l_{min} und l_{max} . Grundlage für die Bestimmung des Einflusses E_{μ_l} bildet die Vorschrift aus Abbildung 3.10.

$$E_{\mu_l} = \begin{cases} 0 & : l_{min} \leq \mu_l \leq l_{max} \\ (\mu_l - l_{max}) \frac{1}{d} & : l_{max} < \mu_l < l_{max} + d \\ (l_{min} - \mu_l) (1 - \frac{1}{d}) & : l_{min} - d < \mu_l < l_{min} \\ 1 & : \text{sonst} \end{cases} \quad (3.14)$$

$$\text{mit} \quad d = \frac{l_{max} - l_{min}}{10}$$

$$E_{\bar{\mu}_l} = \begin{cases} 0 & : l_{min} \leq \bar{\mu}_l \leq l_{max} \\ 1 & : \text{sonst} \end{cases} \quad (3.15)$$

b) Standardabweichung der Luminanz σ_l :

Die Standardabweichung σ_l ergibt sich aus den zum Straßenquerschnitt S_q gehörenden Pixelwert des Eingangsbildes $e(i, j)$ und dem in Gleichung 3.12 bestimmten Mittelwert μ_l .

$$\sigma_l = \sqrt{\frac{1}{n} \sum_{(i,j) \in S_q} (e(i, j) - \mu_l)^2} \quad \text{mit} \quad n = \#S_q \quad (3.16)$$

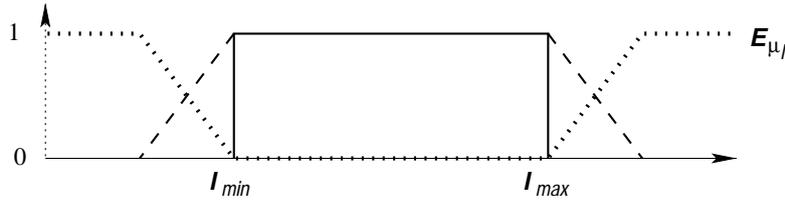


Abbildung 3.10: Trapezfunktion zur Bestimmung der Kosten für E_{μ_l} in Gleichung 3.14

Der zugehörige Einfluss E_{σ_l} berechnet sich direkt aus der Standardabweichung der Luminanz.

$$E_{\sigma_l} = \begin{cases} \frac{\sigma_l}{d} & : \sigma_l < d \\ 1 & : \text{sonst} \end{cases} \quad \text{mit} \quad d = l_{max} - l_{min} \quad (3.17)$$

c) Straßenbreite δ :

Die aktuelle Straßenbreite δ entspricht der Pixelanzahl des Straßenquerschnitts. Der Einfluss E_δ wird ebenfalls über eine Trapezfunktion unter Verwendung der Modellparameter der minimalen Straßenbreite b_{min} und maximalen Straßenbreite b_{max} bestimmt.

$$E_\delta = \begin{cases} 0 & : b_{min} \leq \delta \leq b_{max} \\ (\delta - b_{max}) \frac{1}{d} & : b_{max} < \delta < b_{max} + d \\ (b_{min} - \delta) (1 - \frac{1}{d}) & : b_{min} - d < \delta < b_{min} \\ 1 & : \text{sonst} \end{cases} \quad (3.18)$$

mit $d = b_{max} - b_{min}$

d) Richtungsübereinstimmung der Straßenränder ρ :

Die beiden Richtungen der Straßenränder (s. Kap. 3.1.5.1) sollten in die gleiche Richtung weisen. Unterscheiden sich die in acht Werte quantisierten Richtungen R um den Betrag 1, so werden die Kosten auf 0,5 gesetzt. Der Einfluss E_ρ ist folglich in drei Werte quantisiert.

$$E_\rho = \begin{cases} 0 & : R_{rechts} = R_{links} \\ 0,5 & : |R_{rechts} - R_{links}| = 1 \\ 1 & : \text{sonst} \end{cases} \quad (3.19)$$

- e) Übereinstimmung der Randgradienten und der Straßenrichtung ϱ :
Die Gradienten G der Randpixel sollen orthogonal zur aktuellen Suchrichtung des Straßenverlaufes sein. Der Einfluss E_{ϱ_r} und E_{ϱ_l} bewertet diese Eigenschaft sowohl für den rechten als auch für den linken Rand.

$$E_{\varrho_r} = \begin{cases} 0 & : G_{rechts} \not\perp \text{ Suchrichtung} \\ 1 & : G_{rechts} \perp \text{ Suchrichtung} \end{cases} \quad (3.20)$$

$$E_{\varrho_l} = \begin{cases} 0 & : G_{links} \not\perp \text{ Suchrichtung} \\ 1 & : G_{links} \perp \text{ Suchrichtung} \end{cases} \quad (3.21)$$

- f) Wert des Randgradienten Δ :

Der Wert für den rechten und linken Gradienten (Δ_l, Δ_r) ergibt sich direkt aus dem Gradientenbild $g(i, j)$ zur Bestimmung von E_{Δ_r} und E_{Δ_l} .

$$E_{\Delta_r} = \frac{255 - \Delta_r}{255} \quad (3.22)$$

$$E_{\Delta_l} = \frac{255 - \Delta_l}{255} \quad (3.23)$$

- g) Verbindungen zwischen Straßensegmenten ν :

Die Verbindung des aktuellen Straßensegments mit bereits gefundenen Segmenten ist für die Bildung von Straßennetzen wichtig. Es werden sowohl die Verbindungen einer Straßenseite (E_{ν_r} oder E_{ν_l}) als auch die beider Straßenseiten (E_ν) als Einflussgrößen bestimmt.

$$E_{\nu_r} = \begin{cases} 0 & : \text{Anschluss rechts vorhanden} \\ 1 & : \text{Anschluss rechts nicht vorhanden} \end{cases} \quad (3.24)$$

$$E_{\nu_l} = \begin{cases} 0 & : \text{Anschluss links vorhanden} \\ 1 & : \text{Anschluss links nicht vorhanden} \end{cases} \quad (3.25)$$

$$E_\nu = \begin{cases} 0 & : \text{Anschluss beidseitig vorhanden} \\ 1 & : \text{Anschluss beidseitig nicht vorhanden} \end{cases} \quad (3.26)$$

- h) Potentielle Verbindung zwischen Straßensegmenten ν :

Über eine trichterförmige Suchmaske wird in Richtung der aktuellen Straßenrichtung nach möglichen Verbindungen mit bereits extrahierten Straßensegmenten gesucht. Auch hier werden sowohl die Ränder einzeln (E_{ν_r}

oder E_{v_l}) betrachtet, als auch die Möglichkeit, dass beide Straßenränder verbunden werden.

$$E_{v_r} = \begin{cases} 0 & : \text{Anschluss rechts möglich} \\ 1 & : \text{Anschluss rechts nicht möglich} \end{cases} \quad (3.27)$$

$$E_{v_l} = \begin{cases} 0 & : \text{Anschluss links möglich} \\ 1 & : \text{Anschluss links nicht möglich} \end{cases} \quad (3.28)$$

$$E_v = \begin{cases} 0 & : \text{Anschluss beidseitig möglich} \\ 1 & : \text{Anschluss beidseitig nicht möglich} \end{cases} \quad (3.29)$$

3.1.6 Vektorisieren

Nach der Bestimmung der Straßenverläufe werden die Ergebnisse einer Nachverarbeitung unterzogen. Hierbei werden zu kurze oder isolierte Straßenstücke entfernt, Kreuzungspunkte bestimmt, der Straßenverlauf geglättet und anschließend die Straßenmittellinien vektorisiert.

Zur Bestimmung der Kreuzungspunkte werden die in den Kreuzungsbereich einlaufenden Straßenmittellinien einer Dilatation und anschließenden Erosion unterzogen. Dieses Vorgehen ist in Abbildung 3.11 dargestellt. Ausgehend von den Segmentierungsergebnissen im linken Bild wird durch eine mehrmalige Dilatation der Kreuzungsbereich auf den im mittleren Bild dargestellten Bereich vergrößert. Die Erosion dieses Bereiches auf eine ein Pixel starke Linie liefert den endgültigen Kreuzungspunkt und ist im rechten Bild zu sehen.

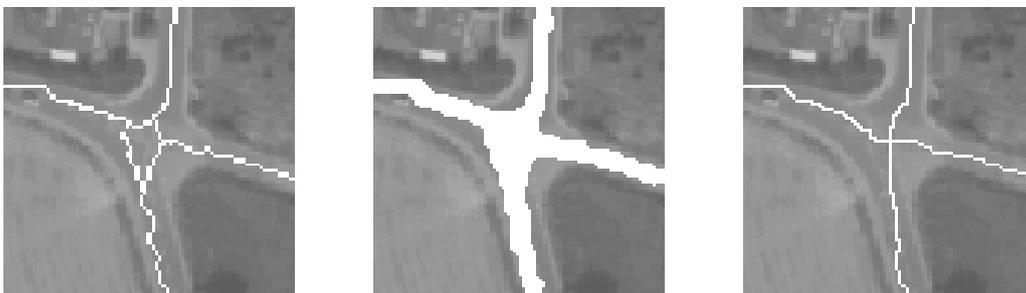


Abbildung 3.11: Bestimmung des Kreuzungsmittelpunktes durch Dilatation und Erosion

Die Vektorisierung der Straßenmittellinien zerteilt das gefundene Straßennetz

in der Form, dass für alle – in eine Kreuzung – einlaufenden Straßen an dieser Stelle ein Startpunkt für eine aus Stützpunkten bestehende Vektorliste erzeugt wird. Weitere Startpunkte sind Straßenenden, wie sie an den Bildgrenzen oder in Sackgassen auftreten. Der Straßenverlauf zwischen zwei Startpunkten wird nun so durch weitere Stützpunkte approximiert, dass die erstellte Vektorbeschreibung die gefundene Straßenmittellinie bis auf einen anzugebenden Restfehler beschreibt.

3.1.7 Ergebnisse der Straßenextraktion

Die Ergebnisse der Straßenextraktion sind stark von den eingestellten Parametern abhängig. Ein Umrechnen (resampeln) der Bilddaten auf eine Auflösung von etwa $2m * 2m$ je Pixel hat sich für das vorgestellte Verfahren als sinnvoll erwiesen. Einige Parameter, wie etwa die minimale und maximale Straßenbreite oder auch der Helligkeitsbereich der Straßen, lassen sich aus dem Datenmaterial ableiten. Andere Parameter müssen vom Benutzer so angepasst werden, dass sie zu den vorliegenden Daten passen.

Eine automatische Adaption dieser Parameterwerte wäre wünschenswert und auch möglich, falls für die vorliegenden Fernerkundungsdaten eine Interpretation beispielsweise in Form eines Geo-Informationssystems vorliegt.

Für die Auswertung von großen Datenmengen kann eine Adaption der Parameter auch anhand eines Teilbereiches geschehen, dessen Straßenverläufe bekannt oder manuell aus den Bilddaten extrahiert wurden. Da bei Bilddaten einer Befliegung die Ausprägung und Struktur der Straßen in den Daten ähnlich ist, kann ein einmal gewonnener Satz von Parametern für die gesamte Auswertung Verwendung finden.

Um eine objektive Adaption der Parameter durchzuführen, muss ein Qualitätsmaß bestimmt werden, welches die Übereinstimmung der beiden Datensätze – der Ergebnisbeschreibung der Straßenextraktion sowie des Referenzdatensatzes aus dem GIS – wiedergibt [48]. In das Qualitätsmaß, das in Kap. 3.1.8 beschrieben wird, gehen die folgenden Maße ein:

- die Länge der übereinstimmenden Straßenbereiche,
- die Länge der in der Straßenextraktion gefundenen Straßen, die in den GIS-Daten keine Entsprechung finden,
- die Länge der im GIS vorhandenen Straßen, die in der Straßenextraktion nicht verifiziert werden konnten.

Ergebnisse der Straßenextraktion für verschieden Sensoren und eine digitalisierte Karte sind in den Abb 3.12 und 3.13 zu sehen. In Schwarz sind die extrahierten Mittellinien und in Rot die vektorisierten Ergebnisse dargestellt.

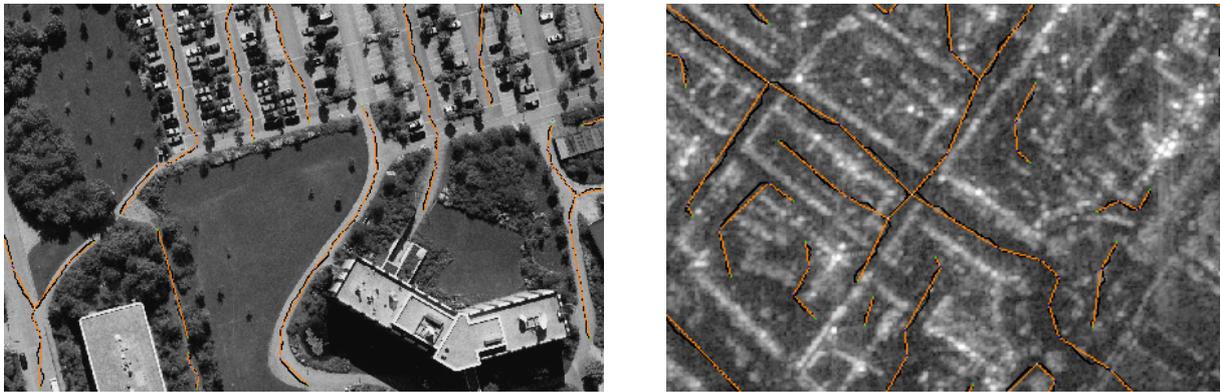


Abbildung 3.12: Ergebnis der Straßenextraktion in einem IR-Bild und einem SAR-Bild

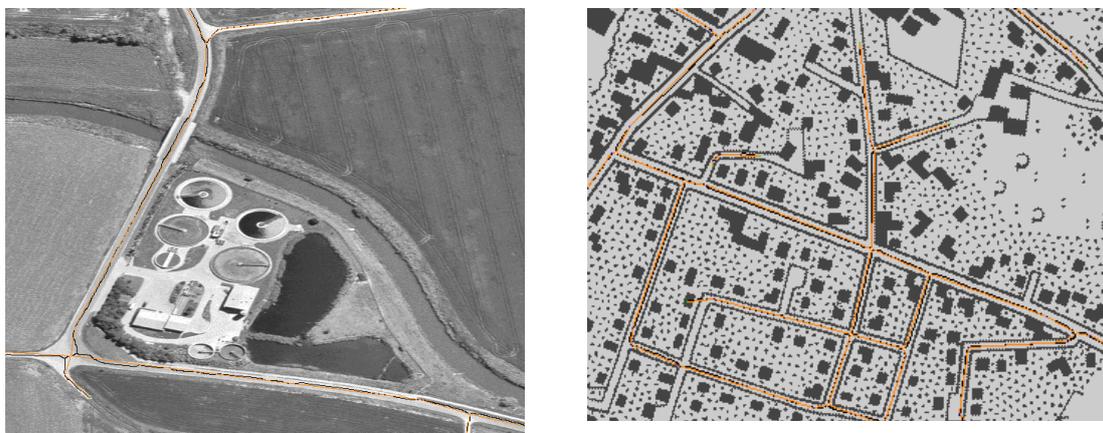


Abbildung 3.13: Ergebnis der Straßenextraktion in einem Luftbild und einer Karte

Die häufigsten verbleibenden Fehler treten bei der Straßenextraktion in Bereichen auf, in denen der Straßenverlauf im Bildmaterial nicht gut zu erkennen ist. Dieses tritt bei Verdeckung der Straße durch andere Objekte wie etwa Alleebäume oder durch Schattenwurf auf. Wie das Problem des Schattenwurfes verringert werden kann, wird im Kapitel 3.3 gezeigt.

3.1.8 Ergebnisbewertung

Die Bewertung der automatisch, extrahierten Straßen geschieht über einen Vergleich mit einem *Solldatensatz* oder *Referenzdatensatz*. Dieser Datensatz ist beispielsweise das Ergebnis einer manuellen Extraktion der Straßenverläufe durch einen Auswerter oder das Ergebnis einer Abfrage eines GI-Systems für den zu untersuchenden Bereich. Um ein Qualitätsmaß zu bestimmen, muss die Übereinstimmung der Straßenmittellinien der beiden Datensätze bestimmt werden [106]. Die zu bewertenden Datensätze liegen hierzu in Form von Vektordaten vor, sodass eine Liste von Stützpunkten ein Straßenstück beschreibt, in der keine Kreuzungspunkte enthalten sind. Kreuzungspunkte oder das Ende von Sackgasen sind somit mögliche Startpunkte der Stützpunktlisten. Es ist jedoch möglich, dass eine zwei Kreuzungen verbindende Straße in mehrere Stützpunktlisten aufgeteilt ist. Mit einem *Straßenteilstück* soll im Folgenden ein Straßenbereich zwischen zwei Stützpunkten einer Stützpunktliste bezeichnet werden.

Die Bestimmung der Übereinstimmung beinhaltet die Festlegung einer geometrischen Genauigkeit, die angibt, ab wann zwei Straßenstücke das gleiche Objekt beschreiben. Anzugeben ist hier der maximale Wert für eine mögliche Translation als auch für die zulässige Rotation. Mögliche Unterschiede, die zwischen den beiden Datensätzen vorliegen können [24], sind hier am Beispiel für die Verwendung der Referenzdaten aus einem GIS aufgeführt:

- Die Straßenverläufe im GIS entsprechen nicht den realen Gegebenheiten, da beim Eintrag der Daten Fehler gemacht wurden (Versatz, Unvollständigkeit, Fehlen von Details).
- Im GIS vorhandene Daten wurden nach der Erfassung zurückgebaut. Diese Daten liegen im GIS vor, lassen sich jedoch nicht mehr mit den aktuelleren

Bilddaten verifizieren.

- Es sind neue Straßen entstanden, die bisher nicht im GIS verzeichnet waren.

Für die Berechnung des Qualitätsmaßes soll die *Vollständigkeit* und die *Korrektheit* verwendet werden [46] [95]. Hierbei gibt die *Vollständigkeit* an, ob im extrahierten Datensatz Straßenstücke fehlen, die in der Referenz vorhanden sind. Hier werden die fehlenden Straßen erfasst. Die *Korrektheit* dagegen erfasst die in den extrahierten Daten zuviel aufgeführten Straßen.

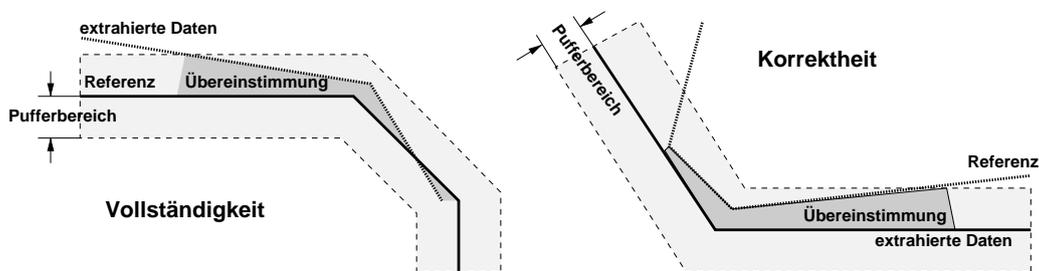


Abbildung 3.14: Bereiche der Zuordnung bei *Vollständigkeit* und *Korrektheit*

Ein graphmatchendes Verfahren ist für den Vergleich nicht geeignet, da hierbei nur ganze Straßenteilstücke miteinander verglichen werden können. Dieses ist jedoch – wie auch in Abbildung 3.14 zu sehen – nicht ausreichend, da auch Übereinstimmung in Teilbereichen von Straßenteilstücken vorliegen können. Aus diesem Grund werden die Vektordaten quantisiert und in einen entsprechenden Bildbereich überführt. Um die Translation berücksichtigen zu können, wird ein Pufferbereich um die zu überprüfenden Daten gelegt. Wie in Abbildung 3.14 dargestellt, lassen sich die übereinstimmenden Bereiche leicht bestimmen. Die Übereinstimmung der Richtung ergibt sich aus den Vektordaten. Hierzu wird jedem – im Bildbereich – entstandenen Pixel die aus den Vektordaten berechnete Richtung zugewiesen. Zwei Pixel stimmen überein, falls die zugehörigen Mittellinien unterschiedlichen Datensätzen angehören, ihr Abstand innerhalb der Translationstoleranz liegt und die Richtungsdivergenz kleiner als ein vorgegebener Maximalwert ist. Die Zuordnung ist als orthogonale Suche um ein Straßenteilstück im Pufferbereich realisiert, wobei bei alternativen Kandidaten

eine Zuweisung zu der mit dem geringsten Abstand umgesetzt wurde. Das Verfahren gewährleistet, dass keine Mittellinieneinheit, die durch die Überführung in den Bildbereich entstanden ist, mehr als einmal zugeordnet wird.

Die geforderten Qualitätsmaße können nun folgendermaßen bestimmt werden:

$$\text{Vollständigkeit} = \frac{\text{Länge der übereinstimmenden Referenz}(Ref_{match})}{\text{Länge der Referenz}(Ref_{all})} \quad (3.30)$$

$$\text{Korrektheit} = \frac{\text{Länge der übereinstimmenden, extrahierten Daten}(Ex_{match})}{\text{Länge der extrahierten Daten}(Ex_{all})} \quad (3.31)$$

$$\text{Qualität} = \frac{Ref_{match} + Ex_{match}}{Ref_{match} + Ex_{match} + (Ref_{all} - Ref_{match}) + (Ex_{all} - Ex_{match})} \quad (3.32)$$

Für die Ergebnisse gilt: *Vollständigkeit* $\in [0, 1]$ (Gleichung 3.30), *Korrektheit* $\in [0, 1]$ (Gleichung 3.31) und *Qualität* $\in [0, 1]$ (Gleichung 3.32). Durch die genormte Darstellung der Ergebnisse ist ein aussagekräftiger Vergleich von Ergebnissen möglich. In Kapitel 3.2 wird der Einsatz des Programms zur automatischen Parameteroptimierung für die Straßenextraktion als Beispielanwendung demonstriert.

3.1.9 Zusammenfassung

Es wurde ein modellbasiertes Verfahren zur Straßenextraktion und die hierbei vorliegenden Teilprobleme besprochen. Dieses Verfahren und die angesprochenen Teilprobleme stehen exemplarisch für Bildverarbeitungsoperatoren, wie sie bei der Bildinterpretation Verwendung finden.

Es zeigte sich, dass die Extraktion in unterschiedlichen Gebietsklassen eine differenzierte Qualität der Ergebnisse aufwies. Grundsätzlich kann zwischen den Klassen *urbaner Bereich*, *ländlicher Bereich*, *Wald* und *Gebirge* unterschieden werden. In den hier gezeigten Beispielen können die im *ländlichen Bereich* gewonnenen Ergebnisse als gut, die im *urbanen Bereich* gewonnenen Gebiete ebenfalls als gut bezeichnet werden, solange die Bebauung nicht zu eng ist. Im innerstädtischen Bereich nimmt die Qualität mit der Dichte der Bebauung und

der zunehmenden Vielfalt der Objekte stark ab. In Waldgebieten ist die Erkennung der Straßenverläufe fast als unmöglich anzusehen, da es bisher keinen Sensor gibt, der durch Bäume verdeckte Straßen abbildet. Die Straßen im bergigen Gelände sind kurvenreicher und es bedarf hier eines angepassten Modells. In dieser Arbeit wurden keine Daten aus dieser Gebietsklasse verwendet.

Störend für die Extraktion sind Objekte, die die Straße direkt verdecken (Bäume, Fahrzeuge und Brücken), aber auch die Schatten der Straßenrandobjekte. Objekte, die in den Fernerkundungsdaten eine ähnliche Ausprägung wie Straßen besitzen (Hausdächer), führen ebenfalls zu einer Verschlechterung der Ergebnisse. Durch die Erkennung dieser Objekte oder deren Einflüsse können diese aus den Bilddaten eliminiert werden, um die Straßenextraktion zu verbessern. Ein andere Alternative bestünde darin, der Straßenextraktion die Information über das Vorhandensein und die Lokalität dieser Objekte mitzuteilen, um so die Erkennung der Straßen kontextabhängig zu steuern.

3.2 Automatisierte Adaption der Parameter von Bildverarbeitungsoperatoren

Die starke Abhängigkeit der Ergebnisse von Bildverarbeitungsoperator von einzustellenden Parametern, erzeugt den Wunsch, eine automatische Adaption der für die Ergebnisse relevanten und nicht aus den Bilddaten bestimmbar Parameter automatisch zu bestimmen. Um dieses Problem zu lösen, wird hier ein allgemeiner Ansatz beschrieben, wie vorhandene Maxima einer parameterabhängigen Gütefunktion bestimmt werden können. Die gerichtete Suche nach den Maxima wird als Adaption bezeichnet. Beschreibt eine konkrete Gütefunktion eine Klasse von Problemen, so können die bei einer Adaption gewonnenen Parameter für die Berechnung aller Klassenelemente verwendet werden.

Eine solche Klasse bilden beispielsweise Bilder eines Sensors mit vergleichbarer Auflösung, die bei ähnlichen Beleuchtungssituationen in einer Messreihe aufgenommen wurden und aus denen vorgegebene Objekte extrahiert werden sollen. Die Gütefunktion liefert einen von den Eingangsdaten abhängigen Qualitätswert und besteht bei der hier gestellten Fragestellung aus einem Bildverarbeitungsoperator, einem Eingangsdatensatz, einem klassifizierten Kontrolldatensatz und einer Funktion zur Bestimmung des Qualitätswertes. Der Eingangsdatensatz ist ein repräsentativer Datensatz einer Klasse von Bilddaten. Der Bildverarbeitungsoperator liefert mit einem gegebenen Eingangsdatensatz und einer Parameterbelegung ein Ergebnis. Dieses Ergebnis wird in der Qualitätsfunktion mit dem klassifizierten Kontrolldatensatz verglichen, die eine Güte als Ergebnis liefert. Diese Güte liefert beispielsweise einen Wert im Bereich $[0, 1]$. Dieser Zusammenhang ist in Abbildung 3.15 als *Gütefunktion* im problemspezifischen Bereich zusammengefasst.

Ist der Parametersatz eines Bildverarbeitungsoperators für einen Eingangsdatensatz bestimmt, so können diese Parameter für alle Fragestellungen beibehalten werden, deren Daten zu der zur Adaption verwendeten Klasse gehören. Für die hier vorliegende Fragestellung können drei Parametertypen unterschieden werden:

- Parameter, deren Werte a priori bekannt sind. Hierzu gehören beispielsweise die Auflösung oder der Sensortyp der Bilddaten.
- Aus den vorliegenden Daten oder Kolateraldaten können Parameter für den Programmaufruf berechnet werden.
- Die zu adaptierenden Parameter eines Programms, deren Werte nicht vor dem Programmaufruf exakt berechnet werden können, jedoch einen mehr oder weniger starken Einfluss auf das Ergebnis haben. Der Wertebereich, in dem ein Parameter liegt, ist häufig eingeschränkt und kann aus oben genannten Gründen gegebenenfalls auf einen Unterbereich eingeschränkt werden.

Das vorliegende Problem lässt sich mathematisch als Suche nach den Maxima in einem n -dimensionalen Raum auffassen. Die Zahl n ist durch die Anzahl der zu adaptierenden Parameter gegeben und das Maximum ist der maximale Wert, den die Gütefunktion liefert. Hier ist diese als Ergebnis des Vergleiches der Referenzdaten mit dem Ergebnis des zu adaptierenden Verfahrens mit einer Parameterbelegung zu sehen. Die im Parameterraum durch die Bewertung beschriebene Fläche kann beliebige Maxima und Submaxima aufweisen und unstetig sein. Einschränkend wird jedoch für die weiteren Betrachtungen vorausgesetzt, dass die Funktion überwiegend stetig ist und der n -dimensionale Parameterraum diskret und endlich ist. Jede Dimension i kann somit als Vektor mit m_i Werten $0, 1, \dots, m_i$ für einen Parameter aufgefasst werden.

Da die Berechnung eines Qualitätsmaßes zeitaufwendig sein kann und eine Optimierung gegebenenfalls viele Berechnungen beinhaltet, ist es wünschenswert, mehrere Berechnungen parallel zu realisieren. Hierbei ist eine Protokollierung der Ergebnisse notwendig, um eine mehrmalige Berechnung von Punkten im Parameterraum auszuschließen.

Abbildung 3.15 zeigt einen Überblick über das Gesamtsystem und den Zusammenhang der Komponenten zueinander. Der als *Interpreter* bezeichnete Teil stellt die Verbindung zwischen dem problemunabhängigen und dem problemabhängigen Bereich her. Die Verbindungen, in denen Daten zwischen den einzelnen Prozessen ausgetauscht werden, sind gestrichelt dargestellt. Die durchgezogenen Linien beschreiben die Verwendung der Teilkomponenten untereinander.

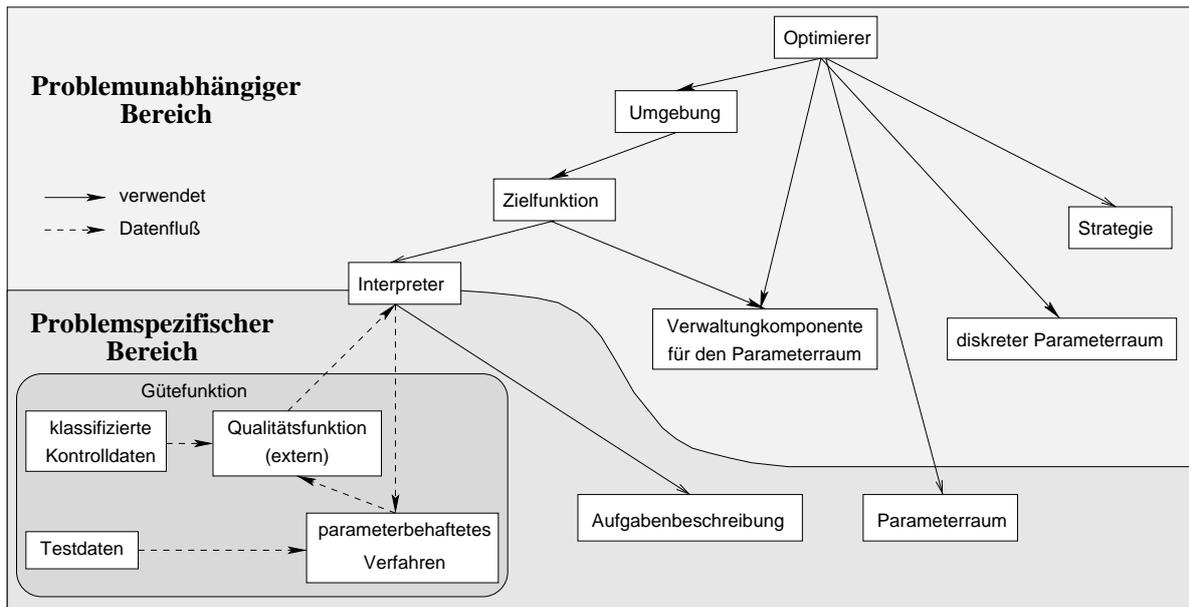


Abbildung 3.15: Überblick über die Komponenten des Adaptionssystems

der. Die Gütefunktion ist an die hier vorliegende Fragestellung angepasst, kann jedoch allgemeiner als Funktion aufgefasst werden, die einen Parametersatz aus dem Parameterraum \mathcal{P} in den Güteraum \mathcal{G} abbildet.

$$f : \mathcal{P} \rightarrow \mathcal{G}$$

3.2.1 Komponenten des Adaptionssystems

Gütefunktion Als *Gütefunktion* soll im Folgenden die Funktion bezeichnet werden, deren Parameter zu adaptieren sind. Dieses kann eine beliebige, als Programm vorliegende nicht zufallsbedingte Funktion sein, die in Abhängigkeit von ihrer Parameterbelegung reproduzierbare Ergebnisse liefert.

Parameter Die *Parameter* der Gütefunktion müssen in einer diskreten Form vorliegen und stellen für den i -ten Parameter einen m_i -dimensionalen Vektor dar, die vom Anwender spezifiziert werden.

Diskreter Parameterraum Der *diskrete Parameterraum* ist, wie bereits angesprochen, ein diskreter, endlicher, n -dimensionaler Raum, dessen Dimension durch die Anzahl der zu adaptierenden Parameter gegeben ist. Er wird aus dem vom Anwender angegebenen *Parameterraum* erzeugt und besitzt eine Inkrement- und Dekrementfunktion, um den Parameterbereich systematisch abzufragen. Für jeden Punkt in diesem diskreten Parameterraum liefert die *Gütefunktion* ein vergleichbares Gütemaß.

Interpreter Der *Interpreter* ist die Schnittstelle zwischen dem problemspezifischen und dem problemunabhängigen Teil der Adaption. Der *Interpreter* ruft die Gütefunktion für den ihm übergebenen Parametersatz auf und gibt das Ergebnis der Gütefunktion weiter. In der *Aufgabenbeschreibung* sind die aufzurufenden Programme angegeben, die vom *Interpreter* zu verwenden sind.

Zielfunktion Die *Zielfunktion* ruft den Interpreter für einen Punkt im Suchraum auf – eine Parameterbelegung – und erhält vom Interpreter das zu diesem Punkt gehörenden Gütemaß. Der Aufruf der *Zielfunktion* ist über eine *Verwaltungskomponente* für den Parameterraum realisiert, um wiederholtes Aufrufen der *Zielfunktion* mit identischen Belegungen zu vermeiden. Bevor die *Zielfunktion* den *Interpreter* für einen Parametersatz aufruft, wird überprüft, ob hierfür bereits ein Gütemaß vorliegt. Ist dieses der Fall, wird das Ergebnis an die *Zielfunktion* geliefert. Anderenfalls wird der zu bearbeitende Punkt im Parameterraum in der *Verwaltungskomponente* als aktiv markiert und der *Interpreter* aufgerufen. Das Ergebnis wird der *Zielfunktion* übergeben und in der *Verwaltungskomponente* für den berechneten Raumpunkt eingetragen.

Verwaltung des Parameterraumes Die berechneten Gütemaße für einen Punkt im Parameterraum sollen in einer zentralen *Verwaltungskomponente* gespeichert werden, um die wiederholte Berechnung für einen Punkt im Parameterraum zu verhindern. Die Anzahl der möglichen Gütemaße für eine Optimierungsaufgabe ergibt sich als Produkt der Mächtigkeit der einzelnen Parameter-

vektoren.

$$N = \prod_{i=0}^{n-1} m_i \quad \text{mit} \quad n \in \mathbf{N}, m_i \in \mathbf{N}$$

Da der benötigte Speicherplatz mit zunehmender Dimension des Parameter-raumes schnell ansteigt und durch die einzusetzenden *Optimierungsstrategien* gegebenenfalls nur eine geringe Anzahl von Punkten innerhalb des Raumes berechnet werden müssen, ist deren Verwaltung über einen binären ausgeglichenen Baum realisiert (red-black-tree). Hierdurch werden nur für die Raumpunkte Speicher angefordert, die auch berechnet wurden. Jeder berechnete Raumpunkt wird in einem Blattknoten des Baumes abgelegt. Die maximale Anzahl von Entscheidungsschritten, ob ein bereits berechneter Punkt im Baum vorliegt, beträgt $\log_2 N$, falls N die Anzahl der Blattknoten im Baum ist.

Umgebung Die *Umgebung* eines Punktes im Parameterraum ist im Allgemeinen ein Hyperwürfel und repräsentiert die Nachbarpunkte des zentralen Punktes. Es wird vorausgesetzt, dass die Gütemaße in der *Umgebung* Aufschluss darüber geben, in welcher Richtung ein Maximum zu erwarten ist (Stetigkeit der *Gütefunktion*). Anhand der Gütemaße der *Umgebung* entscheidet der *Optimierer* entsprechend seiner Strategie, für welchen Punkt aus der aktuellen *Umgebung* die neue *Umgebung*, mit diesem Punkt als Zentrum berechnet werden soll. Um die Gütemaße der *Umgebung* zu bestimmen, ruft die *Umgebung* für die in ihr enthaltenen Punkte die *Zielfunktion* auf.

Im n -dimensionalen Parameterraum besitzt ein Hyperwürfel 2^n Eckpunkte und der zentrale Punkt hat $3^n - 1$ Nachbarpunkte. Im zweidimensionalen Raum entsprechen die Ecken einer Vierer-Nachbarschaft und alle Nachbarpunkte einer Achter-Nachbarschaft. Unter Verwendung der *Zielfunktion* werden für den initialen, vom *Optimierer* gelieferten Parametervektor die Gütemaße aller Punkte innerhalb der *Umgebung* bestimmt, wobei zwischen den beiden Nachbarschaftsformen gewählt werden kann.

Optimierer Der *Optimierer* besitzt eine Optimierungsstrategie, anhand der er entscheidet, für welchen Punkt im Parameterraum die Umgebung als nächstes initialisiert wird. Für diese Entscheidung werden die in der Umgebung vorlie-

genden Gütemaße verwendet. Die verwendete Strategie bei der Suche nach dem Maximum kann durch eine flexible Schnittstelle durch den Anwender vorgegeben werden.

Um den Adaptionprozess zu beschleunigen, können mehrere Rechner für die Berechnung der Gütemaße verwendet werden. Hierzu können mehrere autonome Optimierer, parallel von verschiedenen Startpunkten des Parameterraums aus, eine Optimierung starten. Ihre Ergebnisse werden zentral in der *Verwaltungseinheit* protokolliert, um mehrmaliges Berechnen eines Punktes im Parameterraum zu verhindern.

Strategie Die *Strategie* ist Bestandteil des *Optimierers* und beschreibt das Abbruchkriterium für die von einem Optimierer durchgeführte Suche, den Einsatz und die Art der Sprungfunktion, nachdem ein Maximum gefunden wurde, den Typ der Umgebung und die Auswahl des Folgepunktes beim Suchprozesses.

Die Suche startet in einem vom Optimierer vorgegeben Startpunkt, der an die Umgebung zur Berechnung der benötigten Umgebungswerte übergeben wird. Auf Basis der von der Umgebung berechneten Gütewerte liefert die vom Optimierer verwendete Strategie einen Nachfolgepunkt. Bei der Maximumssuche (engl. hill-climbing) wird der Weg des steilsten Anstiegs solange verfolgt, bis ein Maximum erreicht ist. Dies repräsentiert das Ende einer Suche. Der Optimierer entscheidet nun, ob die Abbruchbedingung für den gesamten Suchprozess erfüllt ist. Diese kann vom Anwender spezifiziert werden und kann beispielsweise das Überschreiten eines Gütemaßes oder die Anzahl der gefundenen Maxima sein. Ist die Abbruchbedingung nicht erfüllt, wird anhand einer Sprungfunktion ein neuer Startwert für eine erneute Suche bestimmt. Hierbei kann beispielsweise auf eine Liste von Startpunkten oder auf eine Zufallsfunktion (engl. simulated-annealing) zurückgegriffen werden. Das Erreichen eines Plateaus – alle Gütemaße der Umgebung sind gleich – oder eines Grats wird als spezielles Maximum behandelt und führt zur Ausführung der Sprungfunktion.

3.2.2 Parameteradaption für die Straßenextraktion

Beispielhaft werden hier die Parameter des in Kapitel 3.1 erläuterten Straßenextraktions-Verfahrens adaptiert. Die Ergebnisbewertung erfolgt mit der im Kapitel 3.1.8 beschriebenen Qualitätsfunktion, wobei hier nur der Ergebniswert *Qualität* aus Gleichung 3.32 verwendet wird.

Die folgenden Parameter des Verfahrens sollen im angegebenen Bereich adaptiert werden:

- w - minimale Straßenbreite in Pixel einschließlich Randpunkte
[3,4,5,6]
- W - maximale Straßenbreite in Pixel einschließlich Randpunkte
[5,6,7,8,9,10,11,12,13]
- b - minimaler Luminanzwert eines Straßenoberflächenpixels
[130,135,140,145,150,165,160,170,180,185,190,195,200,210]
- B - maximaler Luminanzwert eines Straßenoberflächenpixels
[200,210,220,225,230,235,240,250]
- l - maximale Straßenlänge für die sicheren Startbereiche
[6,7,8,9,10,11,12,13,14]
- g - Schwellwert für das Gradientenbild
[20,25,30,35,40]
- d - maximale Suchtiefe des A*-Algorithmus
[6,9,11,13,15,18,23,28,32,37]
- a - maximal erlaubter Fehler bei der Vektorisierung
[3,5,7]

Aus dieser Beschreibung ergeben sich 5443200 mögliche Punkte im achtdimensionalen Parameterraum.

Als Abbruchbedingung wurde das Auffinden von 30 Maxima angegeben. Da der Datensatz viele Straßen im bewaldeten Gebiet enthielt, wurde aus dem vom GIS

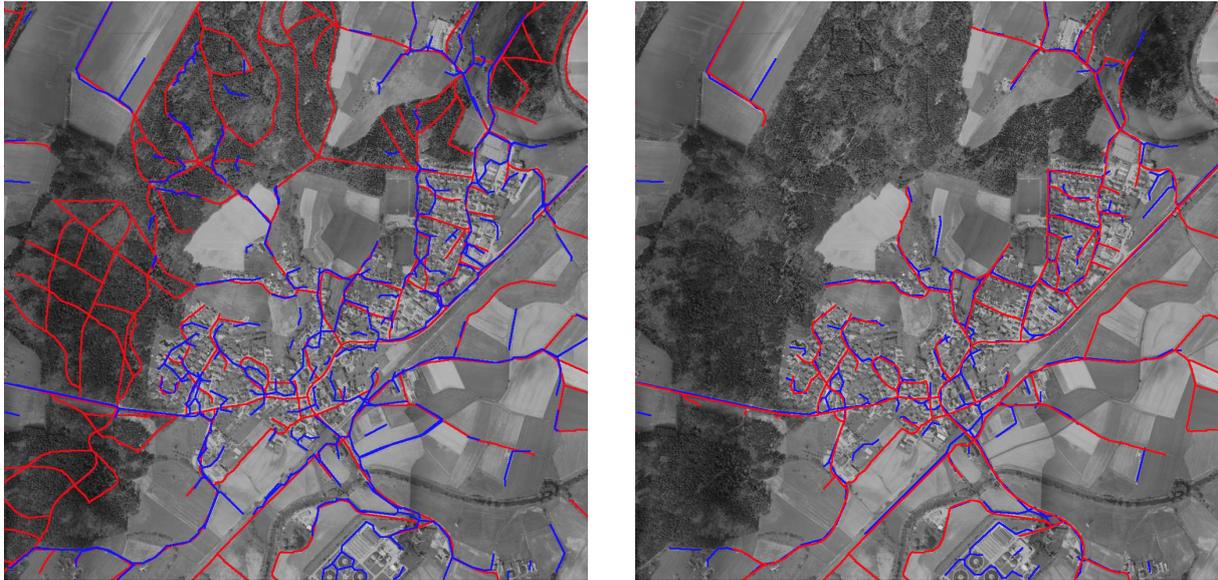


Abbildung 3.16: Ergebnisse der Parameteradaption:

- 1) Referenzdatensatz enthält die im Wald vorhandenen Straßen (links)
- 2) Referenzdatensatz enthält keine im Wald vorhandenen Straßen (rechts)

gelieferten Datensatz ein zweiter Vektordatensatz erzeugt, in dem die Straßen im bewaldeten Gebiet entfernt wurden. Die Ergebnisse der Parameteradaption sind in Abbildung 3.16 zu sehen. Das linke Bild zeigt die Ergebnisse mit dem kompletten GIS-Datensatz und das rechte Bild das Ergebnis mit dem reduzierten Datensatz (ohne die Straßen im Wald). In Rot sind jeweils die GIS-Daten und in Blau die extrahierten Daten dargestellt.

Es wurde auf Grauwertbildern mit einer Auflösung von $2m * 2m$ gerechnet, Zusatzinformation über das Vorhandensein von Häusern oder anderen Objekten wurden nicht verwendet und auch keine radiometrische Entzerrung für die Bilddaten durchgeführt.

Die numerischen Ergebnisse nach den Verfahren aus Kapitel 3.1.8:

Verfahren	Vollständigkeit	Korrektheit	Qualität	Rechenschritte
GIS-Daten (komplett)	0,75	0,69	0,56	1981 (0.026%)
ohne Waldbereiche	0,82	0,88	0,74	2902 (0.053%)

Die Vollständigkeit beschreibt, ob Referenzdaten in den extrahierten Daten fehlen. Eine Vollständigkeit von 1 sagt aus, das alle Referenzdaten auch in den

extrahierten Daten vorliegen. Die zusätzlich extrahierten Daten werden durch den Wert der Korrektheit beschrieben. Hier besagt der Wert 1, dass keine Daten extrahiert wurden, die keine Repräsentation in den Referenzdaten besitzen. Die Qualität nimmt den Wert 1 an, falls beide Datensätze identisch sind. Die Anzahl der benötigten Rechenschritte ist in der letzten Spalte vermerkt, in der sich die Prozentangabe auf den Anteil des expandierten Suchraumes bezieht.

3.3 Schattendetektion unter Verwendung von Laserscandaten

Die Erkennung von Objekten in Fernerkundungsdaten wird häufig durch Beleuchtungseffekte wie Schlagschatten oder Abdunkelung in sonnenabgeneigten Bereichen erschwert oder führt zu falschen Ergebnissen. Um dieses Problem bei der Objekterkennung zu beseitigen, wird in diesem Kapitel ein Ansatz zur Bestimmung von Schattenbereichen in Fernerkundungsdaten vorgestellt. Um die Zuverlässigkeit der Bestimmung der beschatteten Bereiche gegenüber den bisherigen Ansätzen zu erhöhen, werden Höhendaten aus einer Laserscanbefliegung bei der Auswertung mitverwendet. Die unterschiedlichen, mit diesem Ansatz berechneten Ergebnisse werden diskutiert und für eine radiometrische Entzerrung der visuellen Eingangsdaten verwendet.

Die in visuellen Bilddaten durch Schatten auftretenden Störungen erschweren z. B. bei der Detektion von Straßen [17], [22] die Objekterkennung. Daher ist es gerade im Kontext einer wissensbasierten Auswertung [100], [23] wichtig, beschattete Bereiche in den Bilddaten zu kennen, um dann unter Ausnutzung dieses Wissens den Analyseablauf und die damit verbundene Objekterkennung zu steuern. Weiterhin ist es möglich, nach einer Schattendetektion diese Bereiche durch eine radiometrische Entzerrung zu korrigieren, um so die Ergebnisse der anschließenden Objekterkennung zu verbessern.

Hier wird beschrieben, wie eine Fusion von Sensordaten angewendet werden kann, um Schattenbereiche in visuellen Bilddaten zu bestimmen. Voraussetzung für diesen multisensoriellen Ansatz ist die Registrierung aller Bilddaten auf ein gemeinsames Geokoordinatensystem. Zusätzlich zu den visuellen Bilddaten werden Laserscandaten verwendet. Die hier ergänzend verwendeten Laserscandaten werden in naher Zukunft zu Standarddaten in der Fernerkundung gehören und bei der Erstellung eines DEM (digital elevation model) zum Einsatz kommen.

Die verwendeten Bilddaten sind in Abbildung 3.17 zu sehen. Die Schattenbereiche werden aus den Bilddaten auf folgende Weise bestimmt. In einem ersten Schritt werden in den Höhendaten die Gebäude extrahiert, deren Lokalisation verwendet wird, um die zugehörigen Schattenbereiche zu bestimmen. Durch

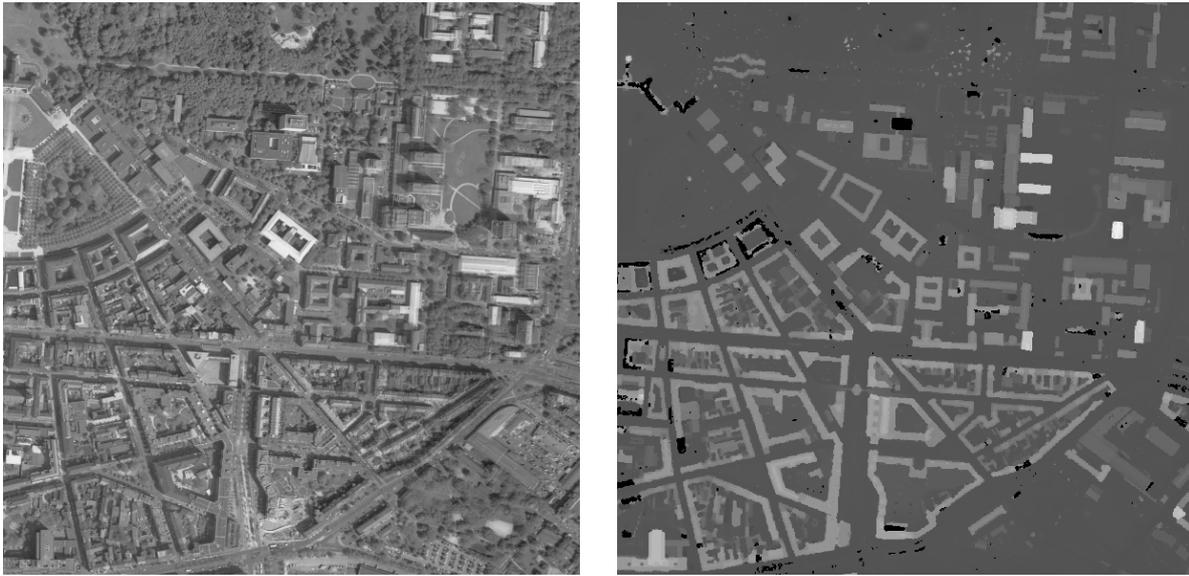


Abbildung 3.17: Eingangsdaten: Luft- und Laserscanbild

Auswertung dieser gefundenen Regionen kann ein Schwellwert bestimmt werden, mit dem das gesamte visuelle Bild binarisiert werden kann. Dieser *erste* Schritt liefert mit dem verwendeten sehr einfachen Ansatz zwar schon die beschatteten Bildbereiche, jedoch auch nicht gewünschte Bereiche, die nur aufgrund ihrer Materialeigenschaften dunkel erscheinen. Um das *erste* Ergebnis zu verbessern, wird in einem zweiten Schritt unter Verwendung des Höhenmodells und der berechneten Schatten der Sonnenstand für den Aufnahmezeitpunkt des visuellen Bildes berechnet. Hiermit kann eine Projektion in das Höhenmodell durchgeführt werden, wodurch man ein *zweites* Ergebnis erhält. Die gewonnenen Ergebnisse können zur radiometrischen Entzerrung oder bei einer wissenschaftlichen Interpretation verwendet werden.

Im folgenden Kapitel werden die verwendeten Daten kurz beschrieben. Anschließend wird ein erstes Schattenbild mit Hilfe eines Schwellwertes aus den visuellen Daten bestimmt. Unter Verwendung dieses Ergebnisses und den Höhendaten wird der Sonnenstand zum Aufnahmezeitpunkt geschätzt. Ein zweites Schattenbild wird durch eine Projektion der Sonne in das Höhenmodell berechnet. Bei der Entzerrung wird eine Kombination beider Ergebnisse verwendet.

3.3.1 Datenmaterial

Als Eingangsdaten liegen visuelle Bilddaten mit einer Auflösung von 0,25m und Laserscandaten mit einer Auflösung von 1m vor. Beide Datensätze wurden vor der Auswertung auf ein gemeinsames Geokoordinatensystem entzerrt, sodass Orthophotos vorliegen und damit eine Korrespondenz der Daten auf Pixelebene gegeben ist. In Abbildung 3.17 sind mögliche Eingangsdaten zu sehen, wobei die Höhen der Laserscandaten im rechten Bild durch Grauwerte codiert sind (helle Regionen repräsentieren hohe Orte, dunkle entsprechend tieferliegende Orte). Die im Laserscanbild zu sehenden schwarzen Bereiche stellen Gebiete dar, an denen kein Messwert bestimmt werden konnte. Die im Herbst aufgenommenen Laserscandaten geben das Gelände ohne die Vegetationshöhen wieder. Die im Sommer durchgeführten visuellen Aufnahmen zeigen belaubte Bäume.

In einem ersten Schritt sollen die Gebäude extrahiert werden (siehe [77]). Da die Laserscandaten die Höhengsprünge der Hauswände gut wiedergeben, ist es möglich, hierüber und über Formparameter Bereiche zu detektieren, die Häuser repräsentieren. Hierzu werden ausgehend von den Laserscan-Höhendaten zunächst Höhengsprünge detektiert. Zusammenhängende Regionen, deren lokale Höhengdifferenzen unter einem vorgegebenen Schwellwert lagen, werden zusammengefasst und mit einem einheitlichen Label versehen. Regionen unterhalb einer Mindestgröße werden der sie einschließenden Region zugeordnet, zu große Regionen dem Hintergrund. Bei den kleinen Regionen werden so Objekte auf Hausdächern (z. B. Schornsteine) dem Dachbereich zugeordnet. Die obere Grenze beschränkt die maximale Größe eines Gebäudes. Durch diese Segmentierung erhält man ein Labelbild, in dem jedes Gebäude durch einen Grauwert dargestellt wird (Abbildung 3.18).

3.3.2 Schattendetektion im visuellen Bild

Die detektierten Gebäude werden nun verwendet, um die Gebäudeschatten zu bestimmen. Ausgehend von den Gebäudegrenzen wird orthogonal in alle vier Richtungen nach dunklen Bereichen gesucht. Da diese Suche an den Gebäudegrenzen beginnt und sich von diesen während der Suche immer wei-

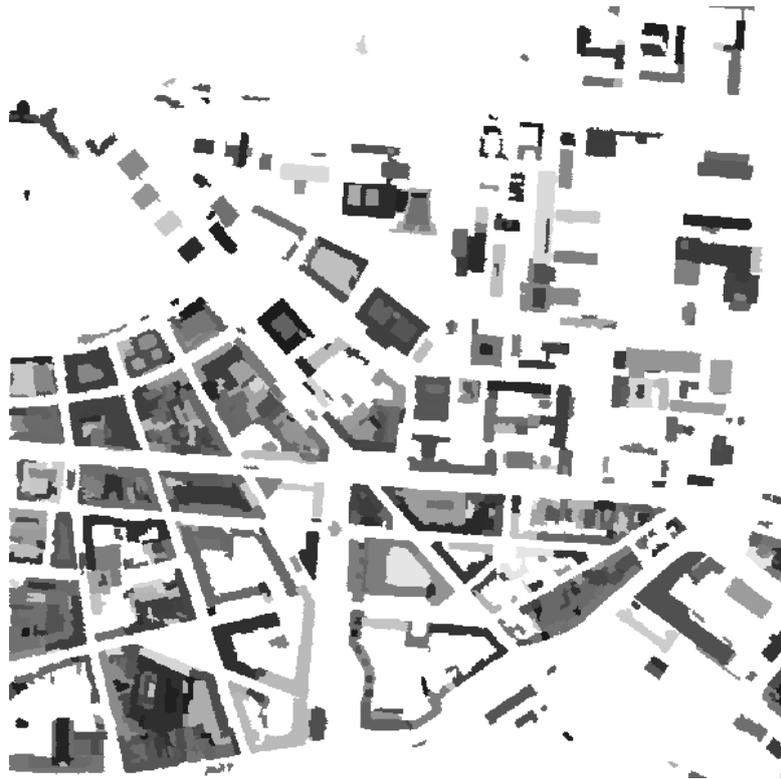


Abbildung 3.18: Labelbild der detektierten Gebäude

ter entfernt, kann das Schattenende durch einen Übergang von dunklen nach hellen Pixeln, also durch einen ausgeprägten Gradienten bestimmt werden. Die Gebäudeseiten, an denen sich keine Schattenbereiche anschließen, können durch ihre Luminanzwerte von den Schattenregionen unterschieden werden.

Die bei dieser Suche betrachteten Pixel können nun benutzt werden, um über eine Histogrammanalyse einen Schwellwert zu bestimmen, mit dem das gesamte visuelle Bild in beschattete und nicht beschattete Bereiche unterteilt werden kann. Neben den tatsächlich beschatteten Bereichen ergeben sich aber auch falsch klassifizierte Gebiete, die aufgrund ihrer Materialeigenschaften oder ihrer Neigung zur Sonne im visuellen Bereich dunkel erscheinen, aber nicht beschattet waren. Dieses Ergebnis dient als Grundlage für die kommenden Schritte, die robust gegenüber den aufgetretenen Fehlern sind. Ein Ausschnitt des Ergebnisbildes des ersten Schrittes ist in der Abbildung 3.20 dargestellt.

3.3.3 Schattenprojektion

Ausgehend von dem bisherigen Ergebnis und den vorliegenden Höhendaten sollen nun die vorhandenen Fehler beseitigt werden. Hierzu wird das aus den Laserscandaten gewonnene Höhenmodell verwendet, welches mit einer künstlichen Sonne zum Aufnahmezeitpunkt der visuellen Daten beschienen wird. Da dieses Höhenmodell die Objekthöhen repräsentiert, wird das entstehende Schattenbild nicht die Fehler enthalten, die bei der oben beschriebenen Schwellwertbildung auftraten.

3.3.3.1 Bestimmung des Sonnenstandes

Falls der Aufnahmezeitpunkt (Ort, Datum, Uhrzeit) bekannt ist, könnte der Sonnenstand direkt hieraus abgeleitet werden. Da diese Angaben nicht immer vorliegen, folgt nun eine Beschreibung, wie der Sonnenstand zum Aufnahmezeitpunkt der visuellen Daten aus dem bereits vorhandenen *ersten* Schattenbild und den Höhendaten zu bestimmen ist.

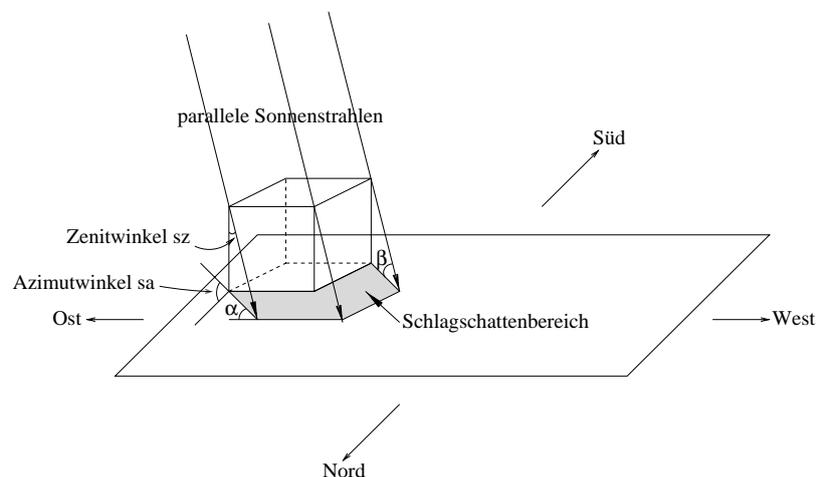


Abbildung 3.19: Darstellung zur Bestimmung des Azimut- und Zenitwinkels

In Abbildung 3.19 sind sowohl die Höhendaten als auch die Schattenbereiche dargestellt. Zu berechnen sind die beiden eingezeichneten Winkel, der Azimut-

und der Zenitwinkel, was nach Gleichung 3.33 und Gleichung 3.34 geschieht.

$$\tan(\alpha) = \frac{\Delta y}{\Delta x} \longrightarrow sa = 90 + \alpha \quad (3.33)$$

$$\tan(\beta) = \frac{h}{l} \longrightarrow sz = 90 - \beta \quad (3.34)$$

Um die beiden Winkel zu bestimmen, müssen sowohl die Ecken der Häuser in den Höhendaten als auch die dazugehörigen Ecken im Schattenbild ermittelt werden. Die Gebäudeecken sind aus dem bereits oben erwähnten Labelbild zu extrahieren. Ausgehend von den gefundenen Gebäudeecken können anschließend die Ecken des dazugehörigen Schattenbereiches ermittelt werden. Zu bestimmen sind jeweils die Punkte der Regionen mit den minimalen bzw. den maximalen Koordinaten. Es ergeben sich für den Suchbereich Einschränkungen, da ein Objektschatten nicht nach Norden ausgerichtet sein kann und der minimale Zenitwinkel durch optional anzugebende Ortsangaben (Breitengrad) bekannt ist. In Abbildung 3.20 sind für einige Gebäude die Gebäudeecken als auch die korrespondierenden Schattenpunkte durch Kreuze markiert.

Der scheinbare Versatz der Gebäudeecken begründet sich in einer nicht völlig korrekten Entzerrung der visuellen Bilddaten, da die Gebäudehöhen nicht berücksichtigt wurden, sondern nur das DGM zu Verfügung stand. Nicht zu jeder Gebäudeecke ließ sich ein korrespondierender Schattenpunkt bestimmen und die gefundenen Paare liefern nicht alle dieselben Winkel. Der resultierende Sonnenstand ergibt sich als Mittelwert aller Einzelmessungen, wobei die Messwerte, die außerhalb der Standardabweichung liegen, beim endgültigen Resultat nicht berücksichtigt werden.

3.3.3.2 Projektion

Mit dem erhaltenen Sonnenstand und den Höhendaten kann nun ein *zweites* Schattenbild bestimmt werden, indem durch ein *ray-tracing* Verfahren für jeden Bildpunkt ermittelt wird, ob dieser bei dem berechneten Sonnenstand von der Sonne beschienen wird oder durch andere Objekte verdeckt wird und daher beschattet ist. Ein Ergebnisausschnitt dieses Verfahrens ist, mit dem *ersten*



Abbildung 3.20: *Schattenbild* mit eingezeichneten Gebäude- und Schattenecken

Schattenbild überlagert, in Abbildung 3.21 dargestellt. Die von beiden Ergebnissen gemeinsam als Schatten bestimmten Bereiche sind in einem dunkleren Grauton dargestellt. Die schwarz dargestellten Bereiche zeigen die von der Projektion zusätzlich als Schatten bestimmten Gebiete und der hellere Grauton kennzeichnet die zusätzlich durch den Luminanzschwellwert detektierten Schattenbereiche.

Die auffälligsten Unterschiede sind in den zusätzlichen Schattenbereichen des Schwellwertverfahrens in den mit Bäumen bestandenen Bereichen zu erkennen. Dies liegt darin begründet, dass die Laserscandaten im Herbst, als die Bäume kein Laub hatten, aufgenommen wurden und daher im Höhenbild nicht zu sehen sind und folglich auch keinen Schatten liefern. Im Gegensatz dazu zeigen die im Sommer aufgenommenen visuellen Daten die, zu den Bäumen gehörenden, Schatten. Anders verhält es sich mit den im Bild vorkommenden immergrünen Nadelgehölz. Wie in Abbildung 3.21 links oben zu erkennen ist, stimmen die Schattenbereiche hier in beiden Ergebnissen überein.

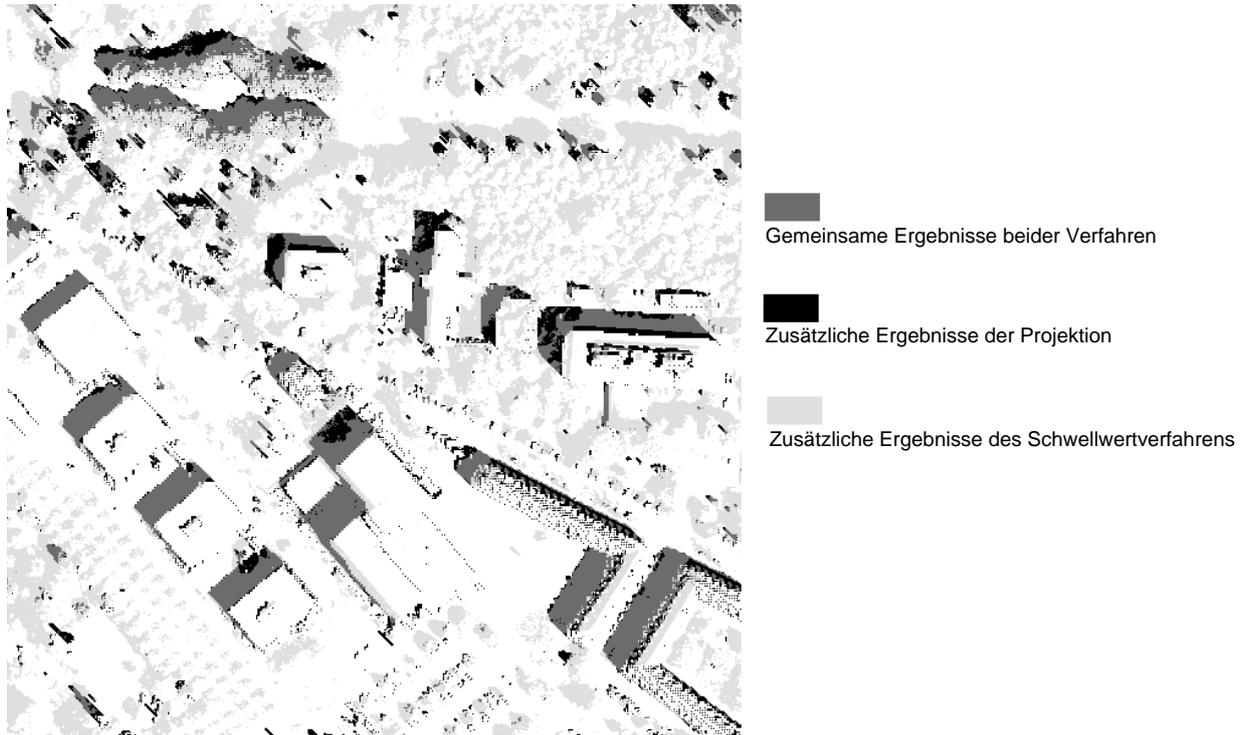


Abbildung 3.21: Gegenüberstellung der Verfahren zur Schattenbestimmung

Weiterhin gibt es bedingt durch die unterschiedlichen Aufnahmezeitpunkte Unterschiede in den beiden Schattenbildern. So ist das in den Laserscandaten (Abbildung 3.22b) als heller Bereich zu erkennende Gebäude, in den zugehörigen visuellen Daten (Abbildung 3.22c), die zu einem späteren Zeitpunkt aufgenommen wurden, bereits abgerissen. Die Unterschiede in den Schattenbildern zu diesem Ausschnitt sind in Abbildung 3.22a zu erkennen. Bewegliche Objekte wie Autos oder Straßenbahnen führen ebenfalls zu Differenzen in den Schatten-

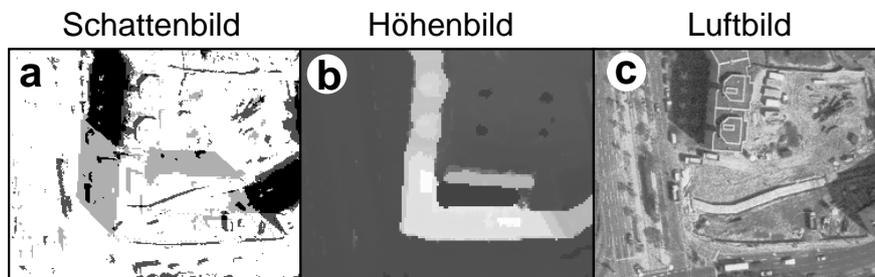


Abbildung 3.22: Beispiel für ein abgerissenes Gebäude

bildern.

3.3.4 Radiometrische Entzerrung

Die radiometrischen Fehler, die durch eine Entzerrung behoben werden können, umfassen zwei Bereiche. Zum einen gibt es Schlagschatteneffekte. Das sind Bereiche, die nicht von der Sonne beschienen werden können, z. B. Bereiche die durch hohe Gebäude oder Geländeerhebungen beschattet werden. Ein anderer Effekt tritt durch die Neigung des Geländes auf. Trotz gleicher Objekteigenschaften erscheinen Gebiete in Abhängigkeit von ihrer Neigung zum Lichteinfall und zum Betrachtungswinkel unterschiedlich. Beide Aspekte sollen bei der radiometrischen Entzerrung Beachtung finden.

3.3.4.1 Neigungsbild

Die Geländeneigung ist basierend auf den Höhendaten in Abbildung 3.23 visualisiert. Die jeweiligen Ausrichtungen des Geländes sind durch unterschiedliche Grauwerte dargestellt, wobei jeder Grauwert eine Richtung kodiert. Zu sehen ist links oben eine von Straßen umgebene Erhebung und die Dächer einiger Häuserzeilen. Im rechten Bild ist ein größerer Ausschnitt eines Innenstadtbereiches mit Straßen und Hausdächern zu sehen.

In der Literatur gibt es verschiedene Ansätze, um das Problem der neigungsbedingten Verzerrung zu lösen. In [53] werden vier Methoden zur Entzerrung vorgeschlagen, die auf dem in Abbildung 3.25 gezeigten Modell basieren. Die Luminanz der geneigten Fläche L_S wird jeweils aus der Luminanz einer horizontalen Fläche L_H berechnet. Ein statistisch-empirischer Ansatz geht von einem Zusammenhang von L_H und L_S nach folgender Beziehung aus:

$$L_H = L_S - \cos(i) \cdot m \quad (3.35)$$

Hierbei ist m ein Maß für die lokalen Reflektionseigenschaften. Je stärker ein Objekt reflektiert, desto höher ist der Wert für m . Bei einem anderen Ansatz,

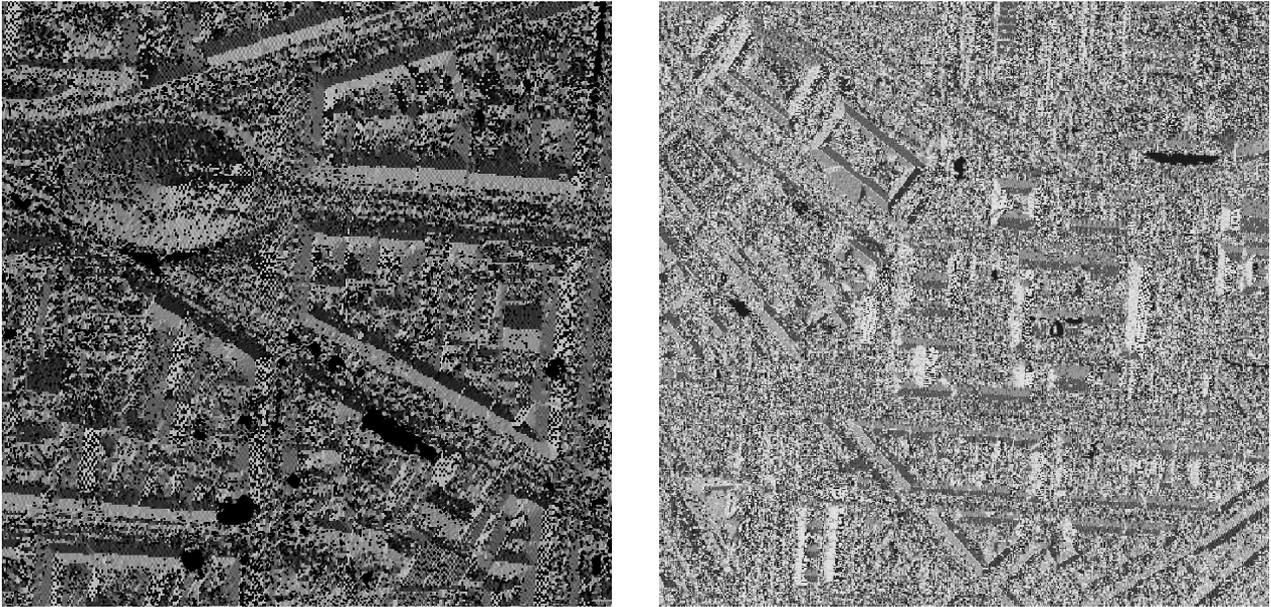


Abbildung 3.23: Aus dem Höhenmodell bestimmte Richtungsbilder

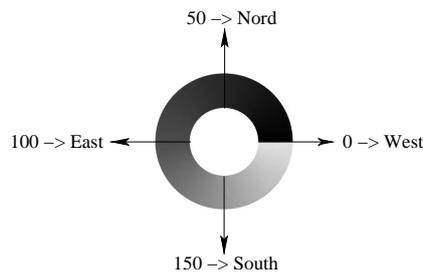


Abbildung 3.24: Korrespondenz der Farben zu den Richtungen

der Cosinus-Korrektur , berechnet sich L_H wie folgt:

$$L_H = L_S \cdot \frac{\cos(sz)}{\cos(i)} \cdot c \quad (3.36)$$

Die einfache Variante der Cosinus-Korrektur setzt die Variable $c = 1$.

Der C-Korrektur-Ansatz wird sowohl in [53] als auch in [99] beschrieben und soll auch hier verwendet werden. Die Luminanz der geneigten Fläche wird dabei folgendermaßen berechnet:

$$L_H = L_S \cdot \frac{\cos(sz) + C}{\cos(i) + C} \quad (3.37)$$

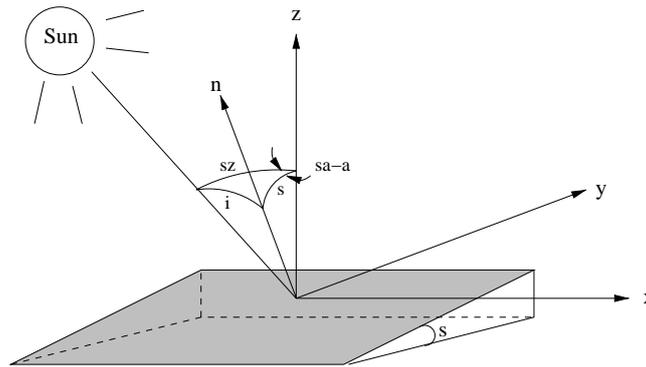


Abbildung 3.25: Herleitung des Zwischenwinkels i aus der Sonnenposition (sz, sa) und der Pixelorientierung (s, a)

Im Modell in Abbildung 3.25 sind drei Richtungen zu unterscheiden: Die Richtung, in der sich die Kamera befindet, die direkte Sonnenrichtung und die Flächennormale der Ebene, auf dem sich der betrachtete Bildpunkt befindet. Der Zenitwinkel kann als bekannt angenommen werden, er wurde in Absatz 3.3.3.1 geschätzt. Im Weiteren kann auch der Flächennormalenvektor aus den Laserscandaten bestimmt werden. Die Illumination $\cos(i)$ berechnet sich aus dem Zenitwinkel sz , dem Azimutwinkel sa , der Neigung der Fläche s und der Ausrichtung der Fläche a zu:

$$\cos(i) = \cos(sz) \cdot \cos(s) + \sin(sz) \cdot \sin(s) \cdot \cos(sa - a) \quad (3.38)$$

Die Korrekturkonstante C ist von den Materialeigenschaften der aufgenommenen Fläche und vom verwendeten Sensortyp abhängig. In [90] wurden für C für die jeweiligen unterschiedlichen Objektklassen Werte von 0,15 bis 11,74 ermittelt. Für Siedlungsflächen, Laub-, Misch- und Nadelwald und Ackerbauflächen wurden Werte für C von 0,25 bis 0,39 verwendet. Man kann prinzipiell für jede Objektklasse eine Korrekturkonstante empirisch ermitteln, die bei der Entzerrung berücksichtigt wird. Die hier verwendete Korrekturkonstante wurde – angelehnt an [90] – für das gesamte Siedlungsgebiet auf $C = 0,28$ gesetzt.

3.3.4.2 Algorithmische Umsetzung

Bereiche, die von Schlagschatten betroffen sind, sind durch die oben beschriebenen Verfahren bestimmt worden. Wir verwenden für die Entzerrung die Kombination beider Ergebnisse, wobei nur die in Abbildung 3.21 dunkelgrau dargestellten Bereiche als Schlagschatten gewertet werden. Durch die Kombinati-

on werden sowohl die Fehler des *ersten* Schattenbildes, die dunklen Bildbereiche, als auch die bei der Projektion auftretenden geringen Unterschiede zu den tatsächlichen Schatten behoben. Die Grauwerte in den Schlagschattenbereichen müssen um einen bestimmten Wert – im Folgenden als *shift* bezeichnet – angehoben werden. Dieser *shift* wird durch den in Kapitel 3.3.2 berechneten Gradienten an den Schattengrenzen der Gebäude zur Umgebung lokal bestimmt.



Abbildung 3.26: Radiometrisch korrigiertes Luftbild aus Abbildung 3.17

Die endgültige Berechnung des entzerrten Bildes soll beide angesprochenen Effekte berücksichtigen. Zur Vereinfachung wird der in Formel 3.37 vorkommende Quotient im Folgenden mit k bezeichnet:

$$k = \frac{\cos(sz) + C}{\cos(i) + C} \quad (3.39)$$

Der Anhebungswert *shift* soll mit dem Faktor k gewichtet werden, sodass sich das entzerrte Bild $z(i, j)$ aus dem Eingangsbild $e(i, j)$ folgendermaßen berech-

net:

$$z(i, j) = \begin{cases} e(i, j) + k \cdot shift & : s(i, j) = 0, \\ e(i, j) + k & : sonst \end{cases} \quad (3.40)$$

Hierbei ist $s(i, j)$ das resultierende Schattenbild, in dem die Werte 0 Schlag Schatten anzeigen. Das so berechnete entzerrte Bild ist in Abbildung 3.26 dargestellt und enthält nur an den Stellen Schatten, die in Absatz 3.3.3.2 diskutiert wurden.

3.3.5 Ergebnisbewertung der radiometrischen Entzerrung

In diesem Kapitel wurde ein Ansatz zur multisensoriellen, radiometrischen Entzerrung von visuellen Fernerkundungsdaten vorgestellt. Hierbei wurde die Neigung des Geländes zur Sonne und zum Betrachter berücksichtigt. Die Schlag Schattenbereiche wurden separat behandelt und bezüglich ihrer lokalen Datenlage radiometrisch entzerrt. Die Sicherheit der Bestimmung der beschatteten Bereiche wurde durch einen multisensoriellen Ansatz erhöht und die Ergebnisse diskutiert. Einsatz findet das Verfahren zur Bestimmung von Schattenbereichen in visuellen Luftbildern. Durch den Schritt der radiometrischen Entzerrung werden die Ergebnisse der folgenden Objektextraktion verbessert. In Abbildung 3.27 ist das Ergebnis der Straßenextraktion aus Kapitel 3.1 mit derselben Parametrisierung für einen Bildbereich vor und nach der radiometrischen Entzerrung dargestellt.

Der Ansatz, die Problembereiche durch die Auswertung der Daten bezüglich spezieller Objekte für die nachfolgende Auswertung zu verbessern, liefert sicherere Ergebnisse als die Anpassung der Parameter. Zum einen ist diese Parameteranpassung oft sehr aufwendig und zum anderen liefert sie meist nur lokal bessere Ergebnisse. So führt zwar eine tolerantere Parameterwahl dazu, dass auch in den Originalbildern der Schattenbereich auf den Straßen als Straße erkannt wird, jedoch ergibt sich an anderen Stellen des Bildes eine Übersegmentierung der Bilddaten. Neben den Schatten können auch andere Objekte, wie etwa Bäume, die mit der Straße eine Allee bilden, zur Erkennung von Straßen verwendet werden.



Abbildung 3.27: Ergebnisse der Straßenextraktion auf dem Originalbild (links) und dem radiometrisch entzerrten Bilddaten (rechts)

4 Wissensrepräsentationsformen und ihre Nutzung

„Artificial Intelligence is the study of ideas which enable computers to do things, that make people seem intelligent. [...] The central goals of Artificial Intelligence are to make computers more useful and to understand the principles which make intelligence possible.”

Winston, 1993

Diese Arbeit befasst sich mit der Interpretation von Fernerkundungsdaten, die mit verschiedenen Sensoren und unterschiedlichen Auflösungen gewonnen wurden. Die Automatisierung der Interpretationsaufgabe ist dabei ein zentraler Punkt der Arbeit. Die vorliegenden Szenen sind oftmals sehr komplex, sodass die Aufgabe der Interpretation nicht unter Verwendung eines Operators – wie in Kapitel 3 beschrieben – gelöst werden kann, sondern nur durch das Zusammenspiel mehrerer Operatoren für die unterschiedlichen Teilaufgaben zu realisieren ist. Dieses Zusammenspiel umfasst dabei das gezielte und problemspezifische Initialisieren und Aufrufen dieser subsymbolischen Operatoren, als auch die Überprüfung, Gruppierung und die Bewertung der gewonnenen Interpretationen. Dem menschlichen Auswerter hilft bei der Interpretation von Fernerkundungsdaten das über die Jahre angesammelte Vorwissen über das Erscheinungsbild und den Zusammenhang von Objekten. Das Vorwissen beinhaltet das deklarative Wissen, das Wissen über die einsetzbaren Verfahren und das darin enthaltene prozedurale Wissen.

Soll nun ein Interpretationssystem eine solche Aufgabe automatisch erledigen, so liegt es nahe, vorhandenes Vorwissen über die Szenenobjekte und ihre Bezie-

hungen zueinander in den Analyseablauf zu integrieren. Ausgehend von dieser Betrachtung werden im folgenden Abschnitt Wissensrepräsentationsformen vorgestellt und diskutiert, die in der “Künstlichen Intelligenz” (KI) Verwendung finden [83] [85] und geeignet sind, Wissen zu repräsentieren, und bei der Analyse von Daten verwendet werden können.

Die Künstliche Intelligenz (Artificial Intelligence) ist eine Fachrichtung der Informatik und befasst sich mit dem Verstehen und Formalisieren von Prozessen, wie sie dem menschlichen Denken eigen sind. Primäres Ziel dieser Forschungen ist es nicht, alle intelligenten Eigenschaften des Menschens zu kopieren, sondern durch Nachahmung oder Adaption von Lösungen und Lösungsfindungsprozessen, mit denen ein Mensch den Gegebenheiten der realen Welt begegnet, die Einsatzmöglichkeiten der automatischen Behandlung von Problemen zu erweitern. Die komplexen und vielseitigen Probleme und Lösungsansätze der realen Welt sind häufig nur recht eingeschränkt zur Formalisierung und Algorithmisierung im herkömmlichen Sinne geeignet. Die vorliegenden, zu verarbeitenden Information sind oft unvollständig oder fehlerhaft, der Lösungsraum zu komplex oder schlicht zu groß, sodass brauchbare Ergebnisse nur geschätzt oder angenähert zu ermitteln sind.

Ziel ist hier die *Beschreibung* einer bildhaften Szene, wobei diese die Konvention einer *Repräsentation* für die Darstellung eines speziellen Gegenstandes nutzt. Die *Repräsentation* ihrerseits besteht aus einer Menge von Vereinbarungen, wie eine Klasse von Objekten zu beschreiben ist.

Die rechnergestützte *Repräsentation* und Verarbeitung kognitiver Fähigkeiten ist der Themenbereich, mit dem sich die Disziplin der KI beschäftigt ([96],[25],[50]). Mit den dort verwendeten Verfahren sollen die für die spezielle Fragestellung relevanten Aspekte der menschlichen Intelligenz für anspruchsvolle zu automatisierende Prozesse nutzbar gemacht werden. Behandelt werden hier üblicherweise einzelne spezielle Themen, sodass die hier behandelte *Welt* nur einen Ausschnitt (*Weltausschnitt*) repräsentiert.

Rein theoretisch sind alle rechnerbasierten Repräsentationen und deren Analysepotentiale gleich, da sie letztendlich auf denselben Informationseinheiten und derselben Funktionsbasis aufsetzen. In der Praxis etablieren sich jedoch nur die Wissensrepräsentationsformen, die die Sachverhalte hervorheben, die zur

Lösung einer Fragestellung wichtig sind.

Eine gute Repräsentation ist folglich in der Lage auch einen schwierigen Zusammenhang verständlich darzustellen [38] und umfasst folgende Aspekte:

- Eine Anzahl von syntaktischen und semantischen Vereinbarungen zur Beschreibung von Fakten ist eine Repräsentation.
- Die verwendeten Symbole und deren mögliche Anordnung beschreibt die *Syntax* einer Repräsentation.
- Die Bedeutung der Symbole sowie die Bedeutung einer Anordnung von Symbolen wird beschrieben durch die *Semantik* einer Repräsentation.

Eine Wissensrepräsentation kann als Abbildung eines Ausschnitts der realen Welt bezeichnet werden, wobei die Darstellung des Wissens implizit oder explizit geschehen kann.

Implizites Wissen

Von implizitem Wissen kann immer dann gesprochen werden, wenn Wissen, z. B. in Form einer Funktion, Verwendung findet, das Bestandteil einer deterministischen Abarbeitungskette ist. Implizites Wissen wird folglich als Bestandteil eines festen Ablaufplans angewendet,

Bei der Anwendung von implizitem Wissen wird eine Funktion folglich nicht aufgerufen, weil ein System oder Programm eine Methode gesucht hat, sondern weil zum Zeitpunkt der Beschreibung eines Lösungsweges oder Programmes feststand, welche Funktion in der vorliegenden Situation zu verwenden ist. Das Wissen, dass eine Funktion an einer bestimmten Stelle aufzurufen ist, ist also direkt im Programm enthalten.

Explizites Wissen

Explizites Wissen zeichnet sich dadurch aus, dass einem System eine Reihe von Problemlösungsmöglichkeiten bekannt ist und es in Abhängigkeit vom vorliegenden Problem die geeignete Funktion zum Lösen auswählt.

Deklarativ:

Deklarative Methoden beschränken sich auf die Beschreibung eines Sachverhaltes und sagen nichts über die Anwendung des Wissens zur Lösung eines Problems aus. Es wird nur der Problemkontext oder Weltausschnitt modelliert, in dem ein Problem vorliegt. Zur Problemlösung werden zusätzlich meist prinzipielle Verfahren definiert, die das Bearbeiten des Weltausschnitts erlauben. Das Verhältnis der modellierten Objekte zueinander für den betrachteten Weltausschnitt wird also getrennt von einem aktiven System, das dieses Wissen in der Analyse verwendet, beschrieben. Der Vorteil dieses Ansatzes liegt in der größeren Übersichtlichkeit und besseren Lesbarkeit des modellierten Weltausschnittes. Ein weiterer Vorteil der deklarativen Wissensrepräsentation liegt darin, dass jedes Wissenselement nur einmal beschrieben werden muss, unabhängig davon, wie häufig und in welcher Situation es zur Lösung des Problems verwendet wird. Neue Objekte, die Wissen repräsentieren, können auf einfachem Wege hinzugefügt werden und bestehendes Wissen kann leicht modifiziert oder entfernt werden. Der Aufwand, um eine konkrete Lösung zu berechnen, ist bei der deklarativen Wissensrepräsentation jedoch oftmals größer als bei der impliziten Darstellung.

Prozedural:

Eine prozedurale Wissensrepräsentation beschreibt den Gebrauch von Wissen, also das *Wie* der Problemlösung und wird oft als die direkteste, natürlichste Art der Wissensdarstellung empfunden, da hier Schritt für Schritt angegeben wird, wie eine Aufgabe zu lösen ist. Dieser Ansatz ist gegenüber dem deklarativen Ansatz effizienter in der Abarbeitung, jedoch nicht so flexibel und versagt daher oft bei komplexen Aufgabenstellungen. Die prozedurale Wissensdarstellung ist daher nur zur Lösung von schon bekannten oder zumindest typisierten Problemen geeignet, deren Lösungsalgorithmus beschreibbar ist.

Die beiden Paradigmen lassen sich folgendermaßen vergleichen: Prozedurale Methoden beschreiben "wie man etwas macht" und deklarative Methoden geben an "wie etwas funktionieren sollte". Genauso, wie eine effiziente Problemlösung in der Realität diese beiden Ansätze vereint, so werden auch in den Methoden der Wissensrepräsentation meist keine rein prozedural oder rein deklarativ

funktionieren Verfahren eingesetzt, sondern eine Kombination der beiden Paradigmen.

4.1 Wissensbasierte Systeme

Systeme, die eine Wissensbasis beinhalten, die Wissen über eine oder mehrere Anwendungsdomänen oder auch über die Welt bzw. einen Weltausschnitt repräsentieren, werden wissensbasiert [51] genannt. Ein Charakteristikum solcher Systeme ist die Trennung von Wissen über die jeweilige Anwendungsdomäne (Faktenwissen) und die Steuerung der Verarbeitung dieses Wissens (Problemlösungswissen). In Abbildung 4.1 ist der Unterschied zwischen konventionellen Systemen, die Wissen implizit nutzen, den wissensbasierten Systemen gegenübergestellt. Die implizite Wissensnutzung erschwert es, Wissen

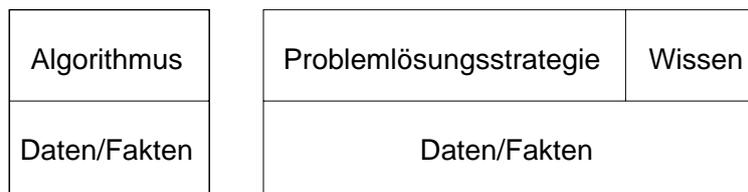


Abbildung 4.1: Gegenüberstellung von impliziter und expliziter Wissensnutzung

nachträglich einzubringen oder zu verändern, da hierzu ein Eingriff in den Algorithmus notwendig ist. Die wissensbasierten Systeme verfügen über eine Schnittstelle zwischen dem anwendungsspezifischen Wissen und der allgemeinen Problemlösungsstrategie, wodurch die Modifikation des Systems leichter und übersichtlicher realisiert werden kann. Die Komponente, die die allgemeine Problemlösungsstrategie umsetzt, wird oft als Inferenzmaschine bezeichnet. Dieser Ansatz ist vor allem von Vorteil, wenn eine Problemlösungsstrategie für mehrere Anwendungsgebiete anwendbar ist. In diesem Fall muss für unterschiedliche Aufgaben nur die entsprechende Wissensbasis ausgetauscht werden.

Zu den Anwendungsfeldern, in denen wissensbasierte Systeme eingesetzt werden, gehören:

- **Diagnoseprobleme:** Hierbei geht es um das Feststellen der Situation, in der sich ein System bzw. ein Teilsystem befindet. Charakteristisch für Diagnoseprobleme ist die vollständig bekannte Lösungsmenge, aus der das zutreffende Element (die zutreffende Diagnose) gefunden werden muss. Einsatz finden diese Ansätze beispielsweise in der Medizin.
- **Planen und Konfigurieren:** Hier sind Lösungswege, Prozessabläufe und Systemzusammenstellungen für ein bestimmtes Problem zu finden. Die zu findende Lösung setzt sich dabei häufig aus Bestandteilen einer Lösungsbibliothek (Baukasten) zusammen oder besteht aus einer Adaption einer bekannten Lösung für eine ähnliche Aufgabenstellung. Im Gegensatz zu Diagnoseproblemen ist hier eine Eigenschaft – wie etwa die Kreativität – von Nutzen. Prädiktives Planen ist die Aufstellung eines Plans im Vorhinein, reaktives Planen ist die Änderung von Plänen bei Eintreten bestimmter Situationen. Zu diesem Aufgabenbereich gehören auch Scheduling-Probleme (Zeitplanung) und Optimierungsaufgaben.
- **Maschinelles Lernen:** In diesem Anwendungsbereich geht es um die Entwicklung von Verfahren, wie ein wissensbasiertes System möglichst automatisch neues Wissen akquirieren, klassifizieren, einordnen und verwenden kann.

4.2 Regelbasierte Systeme

Es gibt eine Reihe von Analysesystemen, die auf einem regelbasierten Problemlösungsparadigma beruhen. So wurde beispielsweise für die Konfiguration von VAX-Computern das regelbasierte System XCON [64] verwendet. Auch in der Medizin wurden und werden eine Reihe von regelbasierten Systemen unterstützend in der Diagnose eingesetzt.

Die atomaren Einheiten eines regelbasierten Systems sind die aus einem Bedingungsteil (Prämisse) und einem Aktionsteil (Konklusion) bestehenden Regeln. Liegt ein Datenbestand (Sachbestand) vor, kann dieser mit den Regeln bearbeitet werden, und es können neue Daten erzeugt werden.

Regel₁: WENN (*Bedingung₁*, *Bedingung₂*, ...) DANN (*Aktion₁*, *Aktion₂*, ...)

Die Aktionen einer Konklusion verändern den Datenbestand, falls alle Bedingungen der Prämisse erfüllt sind. Ein regelbasiertes System kann eingesetzt werden, um zu überprüfen, ob aus einem Anfangsdatenbestand und den Regeln ein bestimmtes Faktum erzeugt werden kann, was der Beantwortung einer Frage zu diesem Faktum entspricht. Kann dieses spezielle Faktum nicht durch wiederholte Regelanwendung erzeugt werden, bleibt die Frage unbeantwortet.

Falls durch den Datenbestand mehrere Regeln erfüllt sind, muss eine Regelauswahl stattfinden, die entscheidet, welche der möglichen Regeln auszuführen ist. Hier gibt es die folgenden alternativen Konfliktlösungsstrategien:

- Spezifität der Regel: Verwende die Regel, deren Bedingungsteil für die aktuelle Datenlage am spezifischsten ist, deren Bedingungen folglich eine Obermenge aller in Frage kommenden Regeln darstellt.
- Priorität der Regeln: Angewendet wird die Regel mit der höchsten Priorität, welche den Regeln bei der Erstellung gegeben wurde oder dynamisch aus der Situation heraus zugeteilt wurde.
- Priorität der Daten: Wie die Regeln könnten auch die Daten in einer Prioritätenliste verwaltet werden. Es ist die Regel anzuwenden, die die höchste Priorität bezüglich ihrer Prämisse aufweist.
- Priorität der Zeit: Verwende die der möglichen Regeln, welche am längsten nicht, oder die, die zuletzt zur Anwendung kam.
- Zufallsprinzip: Wähle rein zufällig eine der möglichen Regeln aus.

Weiterhin sind Kombinationen möglich. Falls die erste Auswahl aufgrund einer Priorität eine Teilmenge liefert, kann die Priorität eines weiteren Kriteriums zur Auswahl herangezogen werden.

Ein Nachteil der expliziten Wissensrepräsentation anhand von Regeln ist, dass bei großen Regelmengen diese schnell unübersichtlich werden und die Gewährleistung der logischen Konsistenz aller Regeln schwer zu überblicken

ist. Gegebenenfalls müssen Methoden zur Konsistenzprüfung eingesetzt werden. Hilfreich ist es hier, die Regeln – falls möglich – in Gruppen zu unterteilen, die dann bei einem spezifischen Datenbestand aktiviert oder deaktiviert werden können.

Nachteilig bei der Verwendung von Regeln zur Wissensrepräsentation ist, dass mit zunehmender Regelanzahl die Übersichtlichkeit über das der Regelbasis innewohnende Verhalten schnell abnimmt. Auch ist die Konsistenz der in der Regelbasis enthaltenen Regeln nicht immer leicht zu überprüfen.

4.3 Logik

Die Logik hat eine lange Tradition in der Mathematik und ist leistungsfähiger als die in 4.2 dargestellten regelbasierten Systeme. Sie basiert auf einer Reihe von Definitionen und einem festen Formalismus, mit denen sich Dinge beweisen lassen. Hierzu muss der zu bearbeitende Themenbereich jedoch in diesen mathematischen Formalismus überführt werden, wodurch das zur Bearbeitung anstehende Problem oft in den Hintergrund rückt und die Logik zum Selbstzweck verfällt. Es gibt eine Reihe von Ansätzen, die Logik zum automatischen Lösen von Problemen einzusetzen, wie es beispielsweise in [57] und [47] dargestellt ist.

Die klassische Logik arbeitet mit booleschen Aussagen. Die Wissensformen vom Typ Faktenwissen und Zusammenhangswissen können als wahre Aussagen in der Logik als Wissensbasis abgelegt werden.

Zum Beweisen werden in der Logik Inferenzregeln wie der *Modus ponens*, *Modus tolens* und *Resolution* [86][87] verwendet, um neue Fakten, auch Ausdrücke genannt, aus der Wissensbasis zu erzeugen. Ausdrücke werden unter der Verwendung von *Prädikaten* repräsentiert. Die Prädikate stellen Formeln dar, die einen Wahrheitswert liefern. Der Ausdruck $Haus(x)$ drückt aus, dass das Objekt x ein Haus ist. Neben den *Junktoren* $\&$, \vee und \neg , um Ausdrücke zu verbinden, existiert die Implikation (\Rightarrow , \Leftrightarrow). So symbolisiert der Ausdruck:

$$Erhebung(x) \& Farbe(x, rot) \& \neg Industriegebäude(x) \Rightarrow Haus(x),$$

dass eine Erhöhung im Höhenbild, die im visuellen Bild eine rote Region dar-

stellt und kein Industriegebäude ist, die Bedeutung Haus erhält.

Um mit Mengen von Objekten zu arbeiten, stellt die Logik den Allquantor \forall und den Existenzquantor \exists zur Verfügung. Solange die Eingangsfakten widerspruchsfrei sind, garantiert der Formalismus der Logik, dass durch die nachfolgende Inferenzschritte die Konsistenz erhalten bleibt. Anfragen an das System werden beantwortet, indem – ausgehend von der Faktenbasis oder der formulierten Anfrage – diese in eine äquivalente Beschreibung überführt werden.

Mit der Logik lassen sich keine zeitlichen Zusammenhänge oder Eventualitäten modellieren und auch die Formulierung von Ausnahmen kann sehr umständlich sein. Ein weiterer Nachteil der Logik ist die kombinatorische Ausprägung des Suchraumes, die bei der Beantwortung einer Anfrage auftreten kann, und gegebenenfalls nicht in endlicher Zeit beantwortet werden kann. Dies wird *Halteproblem* genannt und betrifft die Logiken höherer Ordnung

4.4 Frames

Eine objektorientierte Wissensrepräsentation eines Sachverhaltes aus der realen Welt bieten die von [67] eingeführten *Frames* [72] [54]. Im Gegensatz zum Objekt aus der objektorientierten Programmierung wird bei den Frames auf die Kapselung verzichtet, wodurch die in den Frames definierten Datenelemente einen globalen Charakter erhalten. Die Eigenschaften oder Attribute eines Frames werden in sogenannten Slots gespeichert, wobei ein Slot den Zustand eines Attributs sowie die Prozedur, die dieses Attribut berechnet, enthält. Das Verändern als auch der Zugriff auf das Attribut wird über die dem Slot zugehörige Prozedur realisiert. Ein Frame kann weitere Prozeduren besitzen, die beschreiben, was beim Aktivieren und beim Verlassen geschehen soll.

Die Objekttypen der Frames, auch *Klassen* genannt, repräsentieren Prototypen bzw. *Schablonen* für ein Objekt. Dem Faktum, dass in der realen Welt mehrere unterschiedliche Objekte eines Objekttyps existieren, wird mit Hilfe von generischen Frames und deren Instanzen Rechnung getragen. Ein solcher generischer Frame oder auch Instanz hält für jedes Attribut einen Slot bereit. Ein generischer Frame beinhaltet nicht nur den durch ein Attribut zu belegen-

den Slot, sondern er umfasst zusätzliche Informationen und Programme. Hierzu gehören:

- **Defaultwerte:** Wird eine Instanz eines Objekttypes erzeugt, können Attribute auf Standardwerte gesetzt werden, die im Normalfall zutreffen. Gelten sie für eine spezielle Instanz nicht, können sie auf spezielle Werte abgeändert werden.
- **Bedingungen:** Es können Bedingungen für Attributwerte in Abhängigkeit zu anderen Attributwerten formuliert werden.
- **Funktionen:** Funktionen können lesend oder schreibend Attributeinträge verändern. Es ist möglich, bei jedem Zugriff auf ein Attribut Schlüsse zu ziehen und andere Attribute oder Frames zu modifizieren.

Ein Nachteil dieser Wissensrepräsentationsform ist, dass keine einfache Methode existiert, die bestehende Wissensbasis konsistent zu halten.

Vererbung: Eine weitere wichtige Eigenschaft des Framekonzepts ist die Vererbung, die es ermöglicht, semantische Zusammenhänge zwischen verschiedenen Frames in Form von Hierarchien auszudrücken. Hierbei kann ein generischer Frame Attribute an einen anderen generischen Frame vererben, ähnlich den Vererbungsmechanismen objektorientierter Programmiersprachen. Um diese Beziehung zweier Frames auszudrücken, wird ein spezieller Slot der „A Kind Of“ (AKO-Slot) eingeführt. Andere Arten von Beziehungen als die der Vererbung lassen sich mit Slots nicht explizit modellieren.

4.5 Semantische Netze

Semantische Netze wurden von [84] als einfaches Modell des menschlichen Gedächtnisses eingeführt. Sie können als Erweiterung des Frame-Formalismus aufgefasst werden und präsentieren bestehende Zusammenhänge zwischen Objekten besonders übersichtlich. Semantische Netze stellen gerichtete Graphen

dar, die aus einer Menge von Knoten (engl. Nodes) und Kanten (engl. Links) mit einer Bedeutungen oder auch *Rolle* (engl. Link Label) bestehen. Diese Begriffe sind folgendermaßen definiert

- Knoten: repräsentieren beliebige (primär unabhängige) Sachverhalte, die Ideen, Ereignisse oder Objekte des betrachteten Weltausschnitts.
- Kanten: stellen die Verbindung zwischen zwei Knoten dar, wobei hier im Gegensatz zu den Frames jede Art der Beziehung modelliert werden kann.
- Kantenbedeutung: bezeichnet die Art der Beziehung, die zwei über eine Kante verbundene Knoten haben.

Da die wissenschaftliche Untersuchung dieser Form der Wissensrepräsentation ein Hauptbestandteil der vorliegenden Arbeit ist, wird das Thema im Kapitel 5 detailliert dargestellt. Die Eigenschaften von Semantischen Netzen sind die übersichtliche und verständliche Darstellung des Problembereichs, sowie die Modularität der Wissensdarstellung. Da die Wissensrepräsentation in Form von modularen Knoten geschieht, ist das Wissen flexibel auf neue Aufgaben übertragbar. Die Verarbeitungsstrategie kann durch die Semantik der Kanten spezifiziert und global in einer Analyseeinheit abgelegt werden.

4.6 Agenten

Die Forschung hat in der jüngsten Vergangenheit im Bereich der agentenbasierten Systeme einen starken Aufschwung erlebt. Agentenbasierte Systeme mit ihrem an autonomen Einheiten orientierten Ansatz erscheinen als vielversprechend für Anwendungen im Bereich der Modellbildung und Simulation. Ein Vergleich der Agentensysteme zu den Semantischen Netzen kann insofern erfolgen, als dass die Knoten des Semantischen Netzes als Agenten und die Kanten als die Kommunikationswege zwischen den Agenten angesehen werden können. Diese sind jedoch viel dynamischer als die feste Struktur der Semantischen Netze.

Es gibt viele Definitionen des Agenten-Begriffs, die hier nicht diskutiert werden sollen. Durchgesetzt hat sich die Beschreibung als kommunizierende, autonome

und verteilt arbeitende Einheiten, die über sogenannte *Sensoren* zum Aufnehmen von Informationen und über *Aktoren* zum Agieren innerhalb ihrer *Welt* verfügen.

Ein Nachteil der Agentensysteme ist der oft hohe Kommunikationsaufwand. Dieser führt zu einer zu starken Fokussierung auf diese (interessante) Fragestellung, wodurch das eigentliche zu behandelnde Thema jedoch in den Hintergrund gerät. Bei komplexen Systemen auf der Basis von verteilten Agenten kann es weiterhin zum Verlust der Übersichtlichkeit kommen, da das System von der Idee her kein rein deterministisches Verhalten besitzen soll.

Agenten treten in der Regel nicht isoliert voneinander auf. Eine vorhandene Gruppe von gleichartigen oder verschiedenen Agenten sollen miteinander kommunizieren und interagieren. Solche Systeme werden als Multi-Agenten-Systeme bezeichnet, wobei sich zwei grundlegende Arten der Interaktion unterscheiden lassen. Bei der **Teambildung** und **Kooperation** von Agenten können diese in kleinen oder größeren Gruppen zusammenarbeiten, um ein gemeinsames Ziel, die jeweiligen individuellen Ziele oder eine kurz- bzw. langfristige Planung zu verwirklichen. Bei der **Konkurrenz** oder **Gegnerschaft** konkurrieren Agenten untereinander um Aufgaben. Hierbei können sie sich als Gegner gegenseitig beeinträchtigen und behindern, um ihr individuelles Ziel zu erreichen.

CORBA ist eine für die Realisierung von verteilten Systemen entwickelte Spezifikation. Der im Jahre 1989 gegründete internationale und unabhängige Zusammenschluss von Softwareentwicklern, Netzwerkbetreibern und Hardwareproduzenten zur *Object Management Group* (OMG) spezifizierte mit der *Object Management Architecture* (OMA) eine Softwarearchitektur, die das Zusammenspiel von Anwendungen verschiedener Hersteller ermöglicht. Diese Anwendungen sind dabei – so die Zielsetzung – unabhängig vom Betriebssystem, von der Hardware und auch von der verwendeten Programmiersprache.

Kern dieser Architektur ist die Spezifikation des *Object Request Broker* (ORB), der das universelle Kommunikationsmedium der Object Management Architecture für beliebig geartete Programme oder Programmteile in verteilten und möglicherweise heterogenen Systemen darstellt. Der diese Architektur beschreibende Standard heißt *Common Object Request Broker Architecture* (CORBA) und zeichnet sich durch folgende Eigenschaften aus:

- **Verteilungstransparenz**

Der Zugriff auf Dienste innerhalb von CORBA ist unabhängig von der Lokalisation der dienst anbietenden Objekte. Der tatsächliche Aufenthaltsort eines CORBA Objektes ist für das anfordernde Programm nicht interessant und ist diesem auch meistens nicht bekannt.

- **Unabhängigkeit von Hardware, Betriebssystem und der Programmiersprache**

Die einzelnen CORBA Dienste können auf unterschiedlichen Hardwarearchitekturen, unter verschiedenen Betriebssystemen und in einer Reihe unterstützter Programmiersprachen realisiert werden.

- **Objektorientiert**

Objekte sind die grundlegenden Einheiten der CORBA Architektur. Hierbei ist ein Objekt im Sinne der *Object Management Architecture* eine beliebige, eindeutig identifizierbare Einheit, also nicht notwendigerweise ein Objekt im Sinne einer Programmiersprache.

- **Offenheit**

Da bei der Realisierung von Objekten nur die Struktur der Schnittstellen spezifiziert ist, können über den Kommunikationsmechanismus des *Object Request Broker* die Objekte (Programme) von unterschiedlichen Herstellern bzw. Programmierern auch innerhalb eines Rechnernetzes zusammenarbeiten.

- **Effizienz**

Die *Object Management Group* ist gehalten, eine Architektur zu spezifizieren, die eine effiziente Realisierung unterstützt.

4.7 Unsicheres Wissen

Das verwendete Wissen der in den Kapiteln 4.2 bis 4.4 dargestellten Repräsentationen muss in Form von eindeutigen Fakten, Regeln, Bedingungen oder Abbildungen gespeichert werden. Das Benennen von eindeutigen Aussagen über die reale Welt oder die Darlegung eines festen Zusammenhanges zwischen einer

Prämisse und ihrer Konklusion ist oft nicht möglich. Eine Möglichkeit, solche Unsicherheiten zu modellieren, bietet die Einführung von Wahrscheinlichkeiten oder von Toleranzbereichen für Aussagen bzw. Messergebnisse. Die Unsicherheit von vorliegenden Informationen kann folgende Quellen besitzen:

- **Inhärente Ungenauigkeit:**

Die Messung eines Wertes mit der Angabe einer Toleranz.

- **Inhärente Unsicherheit:**

Ein beobachtetes Muster kann mehreren Klassen zugeordnet werden, wobei es jeder Klasse mit einer bestimmten Wahrscheinlichkeit angehört.

- **Unvollständigkeit:**

Vorliegende Information können lückenhaft sein, d. h., es sind nur Teile der eigentlich vorhandenen oder notwendigen Information verfügbar. Die Lücken innerhalb der Information können dabei offen erkennbar oder verdeckt sein. Das Ersetzen fehlender Informationen durch Defaultwerte führt beispielsweise dazu, dass alle Schlussfolgerungen, die auf dieser Annahme aufbauen, unsicher werden.

- **Unsicherheit von Schlussfolgerungen:**

Es kann kein zwingender Zusammenhang zwischen Bedingung und Schlussfolgerung hergestellt werden. Die aktuelle Datenbasis lässt mehrere konkurrierende Schlussfolgerungen zu.

- **Fehlerhaftigkeit:**

Information kann – auch wenn sie vollständig ist – ganz oder teilweise falsch sein, wenn sie aus falschen oder unvollständigen Annahmen abgeleitet wurde.

4.8 Theorem von Bayes

Die klassische Wahrscheinlichkeitsrechnung definiert die Wahrscheinlichkeit für ein Ereignis als:

$$\lim_{N \rightarrow \infty} p(X) = \frac{\text{Anzahl des Auftretens vom Ereignis } X}{\text{Anzahl aller Ereignisse } N} \quad (4.1)$$

Da hier keine weiteren Bedingungen modelliert werden können, erweist sich diese Definition für die Darstellung von Unsicherheiten als nicht geeignet [83]. Der Ansatz ist hier, den Grad des Vertrauens zu einer Hypothese – als Annahme über einen Sachverhalt – als Wahrscheinlichkeit zu definieren, dessen Werte im Intervall $[0, 1]$ liegen. Folgende Schreibweisen finden im Weiteren Verwendung:

$$\begin{aligned} p(X) &= \text{Wahrscheinlichkeit, dass } X \text{ wahr ist} \\ p(X_1, X_2, \dots, X_n) &= \text{Wahrscheinlichkeit, dass } X_1, X_2, \dots, X_n \\ &\quad \text{wahr sind} \\ p(X_1, X_2, \dots, X_n | Y_1, Y_2, \dots, Y_n) &= \text{Wahrscheinlichkeit, dass } X_1, X_2, \dots, X_n \\ &\quad \text{wahr sind, unter der Voraussetzung,} \\ &\quad \text{dass } Y_1, Y_2, \dots, Y_n \text{ wahr sind} \end{aligned}$$

Das allgemeine Theorem von Bayes definiert eine Menge von Hypothesen $H = h_1, h_2, \dots, h_m$ und eine Menge von Ereignissen $E = e_1, e_2, \dots, e_n$ und baut auf folgenden Voraussetzungen auf:

1. Hypothesen der Menge H schließen sich gegenseitig aus:

$$p(h_i, h_j) = 0 \text{ für } i \neq j \quad (4.2)$$

2. Die Menge H ist erschöpfend:

$$\sum_{i=1}^m p(h_i) = 1 \quad (4.3)$$

3. Jedes Teilergebnis e_i ist bedingt unabhängig von jeder Hypothese:

$$p(e_1, e_2, \dots, e_n | h_i) = \prod_{j=1}^n p(e_j | h_i) \quad (4.4)$$

Das Bayes'sche Theorem besagt nun, dass die a posteriori Wahrscheinlichkeit $p(h_i|e_1, e_2, \dots, e_n)$ einer Hypothese h_i als Funktion der bedingten Wahrscheinlichkeiten $p(e_1, e_2, \dots, e_n|h_i)$ sowie der a priori Wahrscheinlichkeiten $p(h_i)$ folgendermaßen berechnet werden kann:

$$p(h_i|e_1, e_2, \dots, e_n) = \frac{p(e_1, e_2, \dots, e_n|h_i)p(h_i)}{p(e_1, e_2, \dots, e_n|h_k)p(h_k)} \quad (4.5)$$

Für ein spezielles Ereignis E und eine Hypothese H ergibt sich:

$$p(H|E) = \frac{p(E|H)p(H)}{p(E)} \quad (4.6)$$

Der Vorteil des Bayes-Theorems liegt darin, dass die Wahrscheinlichkeiten von verschiedenen, auftretenden Ereignissen zusammengefasst werden können. Die so entstehenden Neubewertungen von Hypothesen können dazu führen, dass eine Hypothese eine höhere Wahrscheinlichkeit erhält als die sie unterstützenden Teilhypothesen. Das bedeutet: "Falls viele vage Hinweise auf einen Sachverhalt vorliegen, steigt seine Wahrscheinlichkeit mit jedem weiteren Hinweis an". Eine Ausbreitung von Unsicherheit ist bestimmbar.

Der Nachteil des Ansatzes ist jedoch der oft hohe Aufwand zur Bestimmung der benötigten Wahrscheinlichkeiten. Für jede Hypothese und jedes Ereignis müssen die entsprechenden bedingten Wahrscheinlichkeiten bestimmt und gespeichert werden. Diese Werte können jedoch meist nicht mit der benötigten mathematischen Exaktheit bestimmt werden, da die hierzu notwendigen, großen Datenmengen nicht vorliegen oder beschafft werden können.

4.8.1 Bayes Netze

Bayes Netze basieren auf dem Bayes-Theorem und bieten einen Ansatz, Abhängigkeiten in einer Menge von Zufallsgrößen einfach und übersichtlich darzustellen. Ein Bayes Netz ist ein azyklischer gerichteter Graph, in dem die Verbundwahrscheinlichkeiten mehrerer Aussagen zu neuen Aussagen verrechnet werden. Die Knoten in einem Bayes Netz repräsentieren die Zufallsgrößen und die diese verbindenden Kanten beschreiben deren Abhängigkeiten. Die Kantenrichtung deutet vom Knoten der Bedingung zum Knoten der Konsequenz.

Die Bayes Methoden bilden die Basis für Methoden zum Lernen von Wahrscheinlichkeiten. Benötigt werden die a priori Wahrscheinlichkeiten $p(H)$ jeder Hypothese H und die bedingte Wahrscheinlichkeit $p(E|H)$ des beobachteten Ereignisses E , um jeder Hypothese eine a posteriori Wahrscheinlichkeit $p(H|E)$ zuordnen zu können.

4.9 Certainty Factors

Die in 4.8 aufgezeigten Probleme, werden bei diesem Ansatz teilweise behoben, indem anstelle der Wahrscheinlichkeiten für jede Regel und jedes Faktum ein Certainty Faktor angegeben wird [13]. Die Regeln erhalten einen festen Certainty Faktor und die Fakten, die eine Regel erfüllen, einen dynamischen Certainty Faktor. Kommt eine Regel zur Anwendung, wird aus den festen und den dynamischen Faktoren für die Folgerung ein neuer Certainty Faktor bestimmt. Es werden hier keine bedingten Wahrscheinlichkeiten, sondern nur ein Certainty Faktor je Regel und Faktum abgelegt. Dieser Ansatz verringert den Aufwand bei der Bestimmung der benötigten Faktoren, der Systemimplementierung und das System ist leichter zu handhaben.

4.10 Fuzzy Logik

Die Fuzzy Logik kann als Kombination der Mengenlehre und der Logik angesehen werden [55] [69] [108]. Ihre Idee ist, durch das Verwalten von unsicherem Wissen, eine Theorie der unscharfen Mengen bereitzustellen. Die binäre Eigenschaft eines Objektes, in der klassischen Mengenlehre ein Element oder kein Element einer Menge zu sein, wird in der Fuzzy Logik aufgeweicht. Eine Menge besitzt somit keine harten Grenzen mehr, und ein Element kann mit einer bestimmten Wahrscheinlichkeit einer Menge zugehören. Der Wert der Zugehörigkeit ist aus dem Intervall $[0,1]$. Als besonders vorteilhaft hat sich die Fuzzy Logik bei der Formulierung von Sachverhalten aus dem natürlichen Sprachgebrauch erwiesen.

4.11 Zusammenfassung

Bei den vorgestellten Ansätzen existieren für den hier vorliegenden Problem-bereich grundsätzlich zwei zentrale Fragestellungen. Zum einen, wie neue Methoden im System integriert werden können. Dieses ist beispielweise bei den regelbasierten Systemen sehr einfach durch das Einfügen einer neuen Regel möglich. Zum anderen stellt sich aber auch die Frage, wie eine neu eingeführte Methode in den Gesamtablauf integriert wird und ob das Zusammenspiel der einzelnen Komponenten noch übersichtlich ist. Hier bieten die Semantischen Netze einen Kompromiss, indem sie die Integration von beliebigen Verfahren ermöglichen, aber gleichzeitig deren strukturierte Einbindung in das Gesamtsystem fordern.

Die vorgestellten Ansätze besitzen die Eigenschaft, aus vorhandenem Wissen neues Wissen zu folgern. Hierbei können zwei Formen, das *monotone* und *nicht-monotone* Schließen, unterschieden werden. Werden aus bestehenden Fakten neue Fakten gefolgert, die zu den vorliegenden nicht im Widerspruch stehen, so nennt man diese Art monoton. Bei einer unscharfen Wissensdarstellung können Schlussfolgerungen entstehen, die Fakten erzeugen, die zu den ursprüngliche Fakten im Widerspruch stehen. Der Wissenszuwachs wird als nicht monoton bezeichnet und es bedarf Methoden, um gezogene Schlussfolgerungen zu revidieren.

5 Wissensrepräsentation mit GeoAIDA

Die **Grundvoraussetzung** jeden **Fortschritts** ist die Überzeugung, dass das Nötige möglich ist.

Norman Cousins

Für die wissensbasierte Auswertung von Fernerkundungsdaten wurde das Interpretationssystem GeoAIDA realisiert. Der Wunsch einer Interpretation setzt die Idee von Objekten voraus, welche in einem gegebenen Datensatz zu finden sind. Eine Menge von benennbaren Objekten kann automatisch nur erkannt werden, falls Eigenschaften dieser beschreibbar und in den vorliegenden Daten verifizierbar sind. Die Objekte bzw. der Weltausschnitt, der für die vorliegende Fragestellung relevant ist, wird auch als *externes Modell* bezeichnet[92].

Welche Eigenschaften erkennbar sind, ist abhängig vom Sensortyp und von der Auflösung. Je höher die Auflösung der Daten ist, desto mehr Details werden erkennbar und sollten bei der Festlegung der Objekteigenschaften berücksichtigt werden. Zu erwähnen ist, dass eine hohe Auflösung nicht zwangsläufig zu einer besseren Interpretation führt. Es ist eher so, dass es für jeden Objekttyp einen Auflösungsbereich oder auch Maßstabsraum gibt, in dem dieses Objekt optimal zu erkennen ist. Durch eine Verfeinerung werden die Bestandteile eines Objektes erkennbar, aufgrund dessen das Objekt dann auf einer strukturellen Ebene erkannt werden kann. Durch die Verwendung unterschiedlicher Maßstabsräume ist es möglich, bei der Erkennung immer die optimale Auflösung für unterschiedliche Objekte und Objektteile auszuwählen.

Werden die Daten unterschiedlicher Sensoren herangezogen, so kann die Güte der Objekterkennung gegenüber der Verwendung von nur einem Sensor nur zunehmen. Dieses begründet sich darin, dass aus diesen Datensätzen Infor-

mationen gewonnen werden können, die zumindest teilweise unabhängige Objekteigenschaften wiedergeben. Ein Beispiel hierfür ist die Repräsentation von Gebäuden in Fernerkundungsdaten, die sich im sichtbaren Spektralbereich in ihren unterschiedlichen Farben und Formen darstellen und dadurch schwer als Gebäude klassifizierbar sind. In einem Laserscandatensatz, aus dem ein Höhenmodell erzeugt werden kann, sind Gebäude leichter und sicherer mit einem automatischen Verfahren zu erkennen, da sie sich durch einen Höhengsprung von ihrer Umgebung abheben. Andere Objekte, wie etwa Straßen, die sich im sichtbaren Bereich von manchen Gebäudedächern kaum unterscheiden, weisen keinen Höhengsprung an den Grenzbereichen auf. Durch die zusätzliche Verwendung von Laserscandaten kann so die Bestimmung von Gebäuden als auch die der Straßen – durch weniger Falschhinweise – merklich verbessert werden, da der Höhengsprung eine sehr markante Eigenschaft von Gebäuden ist.

Wie in den vorangegangenen Kapiteln dargestellt, gibt es zwei grundlegende Ansätze, Mustererkennungsaufgaben zu lösen. Zum einen den holistischen Ansatz, in dem durch Bildverarbeitungsoperatoren spezielle Objekte extrahiert werden. Zum anderen bieten die Wissensrepräsentationsformen mit ihren Inferenzmechanismen eine gute Möglichkeit, strukturelle Gegebenheiten zu formulieren. In GeoAIDA wurde eine Kombination beider Verfahren eingesetzt, um die Eigenschaften der beiden Ansätze und die daraus resultierenden Synergieeffekte nutzen zu können. Durch die sich wechselseitig unterstützenden Verfahren können die Interpretationsergebnisse der vorliegenden komplexen Bilddaten signifikant verbessert werden.

5.1 Systemanforderungen

Aufgrund der Vorbetrachtungen und der Komplexität von Fernerkundungsdaten wurden die folgenden Anforderungen an das zu entwickelnde System gestellt:

- Der zu bearbeitende Weltausschnitt bzw. das Problem soll durch eine geeignete Strukturierung übersichtlich darstellbar sein.

- Der Ablauf einer Analyse soll aus der Strukturierung des Problems ableitbar und verständlich sein.
- Es soll möglich sein, Vorwissen einzubringen und in der Analyse zu nutzen.
- Die Verwendung von multisensoriellen Daten – auch unterschiedlicher Auflösungen – und Vektordaten, wie sie von GI-Systemen bereitgestellt werden, soll vom System unterstützt werden.
- Zu allen Objekten sollen die zugehörigen Geokoordinaten bekannt sein, um einen absoluten Bezug herstellen zu können.
- Die Ergebnisdarstellung soll sowohl in bildlicher als auch in einer strukturellen Form vorliegen.

Aus den Anforderungen lassen sich die beiden Themenbereiche *Darstellung* und *Strategie* der Abarbeitung ableiten. Für die Darstellung und Strukturierung des Problems werden Semantische Netze eingesetzt, wie sie bereits in Kap. 4.5 angesprochen und in Kap. 5.4 näher beschrieben werden. Die Analysestrategie ist im System fest vorgegeben und wird in Kap. 5.6 dargestellt.

5.2 Zielsetzung und Anwendungsgebiete

Ziel des gewählten Ansatzes ist die Interpretation von Fernerkundungsdaten, wodurch aus möglicherweise multisensoriellen Bilddaten und der optionalen Verwendung eines GI-Systems eine strukturelle Beschreibung sowie eine Kartendarstellung in verschiedenen Detaillierungsebenen aus den Eingangsdaten erzeugt werden soll. Der Ansatz soll es ermöglichen, neben dem in die Bildverarbeitungsoperatoren eingeflossenem Wissen, zusätzliches Wissen über die Struktur und die Relationen der zu extrahierenden Objektklassen einzubringen. Einsatzgebiete des Systems:

- Anlegen, Verifikation und Update von GI-Systemen,
- Kartenerstellung,

- Erstellung von Zustandsbeschreibungen (z. B. Waldschäden),
- Änderungsdetektion:
 - Katastrophenschutz, z. B. Hochwasser,
 - Städteplanung, Neubaugebiete,
 - Umweltschutz, Regenwald,
- Verbesserung der Modelle zur 3D-Visualisierung.

Für die Umsetzung der Ziele wurde ein auf Semantischen Netzen und einem problemunabhängigen Analyseansatz basierendes System (GeoAIDA) entwickelt. Im Folgenden wird die, den Semantischen Netzen zugrunde liegende Idee näher dargestellt und deren Umsetzung in das realisierte System beschrieben.

5.3 Analyseablauf

Hier soll ein kurzer motivierender Überblick des Analyseablaufs gegeben werden. Die Analyse startet mit der obersten Instanz des generischen Semantischen Netzes, indem diese modellgetrieben (Top-Down) abgearbeitet wird. Hierbei entstehen hypothetische Instanzen, denen jeweils eine eigene Teilregion der gesamten Szene zugeordnet ist. Das entstehende Instanzennetz ist in dem Sinne hierarchisch, dass eine Subinstanz immer eine Teilregion der zur übergeordneten Instanz gehörenden Region enthält. Während der Top-Down-Analyse können für ein und dieselbe Region konkurrierende Interpretationen entstehen.

Notwendigerweise besitzt jedes Konzept, das einen Blattknoten darstellt, einen holistischen Operator. Besitzen auch andere Konzepte holistische Operatoren, so sind diese in der Lage, hypothetische Instanzen für Teilbereiche der zur übergeordneten Instanz gehörenden Region zu erzeugen bzw. zu entscheiden, dass gewisse Regionen nicht zu dieser Klasse gehören können. Da hierdurch nicht mehr die gesamte Region untersucht werden muss, steigt die Effizienz und Geschwindigkeit der Gesamtanalyse.

Ist die unterste Ebene des Semantischen Netzes erreicht und alle Objekte aus den Daten extrahiert, werden die entstandenen hypothetischen Instanzen datengetrieben gruppiert (Bottom-Up). Instanzen, deren Bestandteile und Attribute den im Semantischen Netz angegebenen Bedingungen genügen, werden bestätigt. Liegen für eine Region konkurrierende Interpretationen vor, so entscheidet ein Bewertungsmaß, welche Instanz als Ergebnis verbleibt, sodass am Ende der Analyse jede Region in den unterschiedlichen Interpretationsebenen eindeutig zugeordnet ist. Das Ergebnis liegt abschließend in einer strukturellen und einer bildhaften Form vor.

5.4 Semantische Netze

Semantische Netze stellen einen Formalismus zur Wissensrepräsentation dar, dessen Grundelemente Knoten und Kanten sind. Knoten repräsentieren hierbei Begriffe, Objekte oder Ideen, während Kanten die Beziehung oder Relation zwischen diesen beschreiben.

In Abbildung 5.1 ist beispielhaft ein Semantisches Netz dargestellt, welches elementares Wissen über einige aus dem Bereich des Kartenwesens bekannte Objekte darstellt. In dieser Wissensbasis unterteilt sich ein Land in verschiedene Bundesländer, die ihrerseits Städte, Verkehrs- und Wasserwege beinhalten. Ein Verkehrsweg kann wiederum eine Autobahn oder auch eine Bundesstraße sein. Das ein Bundesland mehrere Verkehrswege haben kann, ist an dieser Abbildung – stellvertretend für alle anderen *teil_von* Kanten – durch die in eckigen Klammern stehenden Zahlen dargestellt. Hier kann eine minimale und maximale Anzahl von Verkehrswegen spezifiziert werden.

Das abgebildete Netz beinhaltet auf der linken Seite eine Modellbeschreibung, in der allgemeine Begriffe für Objekte aufgeführt sind. Auf der rechten Seite sind real existierende Instanzen angegeben, die über eine spezielle Kante – die *instanz_von*-Kante – mit den Objekten der linken Seite verbunden sind. Jede Kante besitzt eine Rolle, die spezifische Beziehungen zwischen Begriffen oder Entitäten angibt. In einem allgemeinen Semantischen Netz kann jede Art von Beziehung zwischen zwei Knoten durch die Rolle einer verbindenden Kante be-

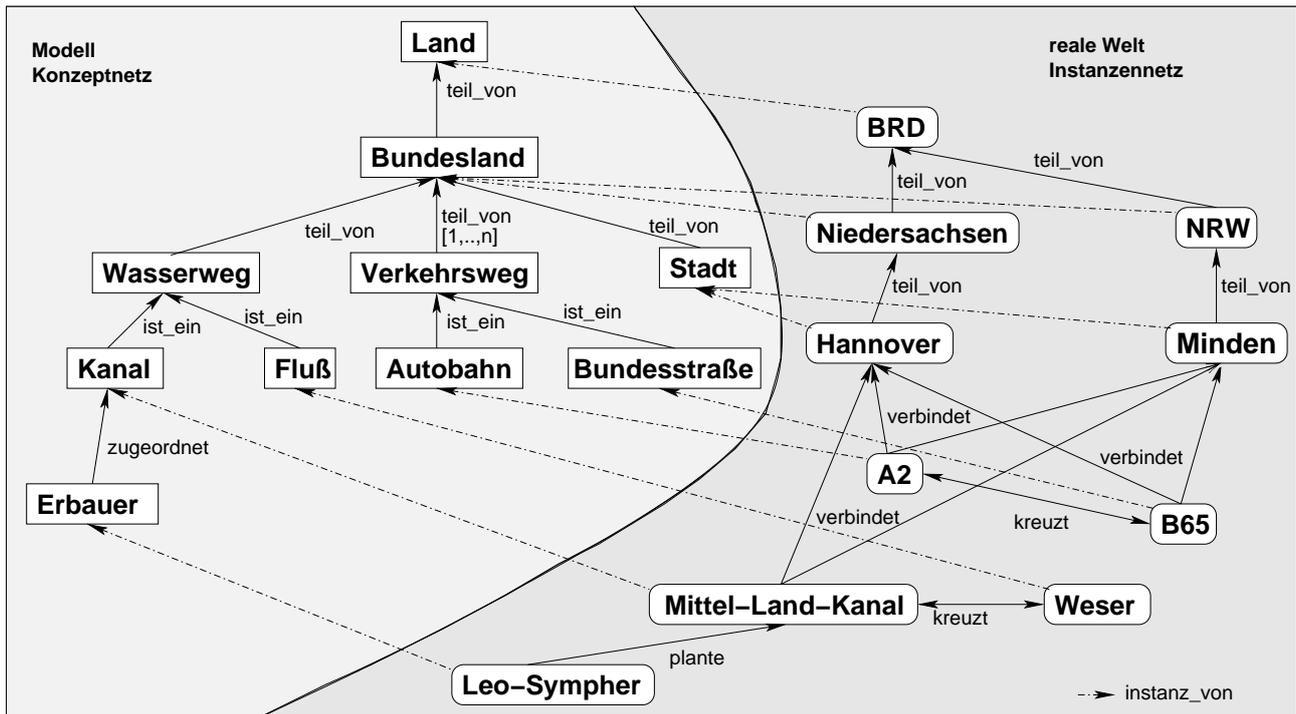


Abbildung 5.1: Beispiel für ein allgemeines Semantisches Netz

schrieben werden.

Die in Abbildung 5.1 angegebenen Rollen haben bei genauer Betrachtung unterschiedliche Qualitäten. Es gibt Kanten, deren Rollen genereller sind und in vielen semantischen Netzwerksprachen Verwendung finden. Hierzu gehören:

- **teil_von**: Diese Kante etabliert eine Bestandteilsbeziehung zwischen dem Ursprungs- und dem Zielknoten der Kante.
- **ist_ein**: Diese Kante beschreibt eine Generalisierung zwischen den durch diese Kante verbundenen Knoten und beinhaltet meist eine Vererbung der allgemeinen Eigenschaften an das speziellere Objekt.
- **instanz_von**: Diese Kante stellt die Verbindung zwischen Modell- und Instanzennetz dar.

Für viele Anwendungen reicht diese Menge von Kanten für die Beschreibung aus und führt zu einem sehr klaren Aufbau der Netze, die durch die festgelegte

Anzahl der Kantentypen und ihrer expliziten Semantik für eine automatische Wissensnutzung geeignet sind.

Hierdurch wird die Schwäche von allgemeinen Semantischen Netzen behoben, die durch die Verwendung von Kanten mit beliebigen Rollen sehr unübersichtlich werden können. Diese allgemeinen Semantischen Netze bergen weiterhin das Problem, dass es nicht möglich ist, einen generellen Analysealgorithmus zu realisieren, durch den das im Semantischen Netz eingebrachte Wissen für eine Anfrage genutzt werden kann. Dieses begründet sich darin, dass das in den Kanten implizit vorhandene Wissen der Analyse nicht bekannt sein kann, falls diese beliebige Rollen besitzen können.

Eine automatische Nutzung des im Semantischen Netz enthaltenen Wissens durch eine Einschränkung auf eine feste Menge von Kantentypen wurde beispielsweise in ERNEST [59] und AIDA [61] umgesetzt. Durch die Einschränkung auf einige Kantentypen ist es somit möglich, in der Analyse Regeln einzubringen, die das implizit in den Kanten enthaltene Wissen nutzen, da es im Analysealgorithmus explizit vorliegt. Alternativ ist es möglich, das Wissen, welches durch die Kanten symbolisiert wird, explizit als zur Kante gehörende Funktion umzusetzen. Dieser Ansatz ist in GeoAIDA realisiert.

Ein weiterer Nachteil der allgemeinen Semantischen Netze besteht in der Gewährleistung der logischen Konsistenz eines Netzes. Diese wird vor allem dadurch hervorgerufen, dass Kanten zwischen beliebigen Konzepten im Netz, unabhängig von ihrer Position existieren können und die möglicherweise auftretenden Schleifen zu unüberschaubaren Netzen und schwer zu erfassendem Verhalten führen. Um dieses Problem zu umgehen, dem Anwender eine klare und überschaubare Struktur zu präsentieren und da es der gewählten Problemstellung – der Interpretation von Fernerkundungsdaten – naheliegt, wurde die Struktur des Semantischen Netzes in GeoAIDA auf eine Baumdarstellung beschränkt.

Die Darstellung der *instanz_von* Kante geschieht klarer in der in Abbildung 5.2 gewählten Form. Hier ist dargestellt, dass das Objekt 'Hannover' eine Instanz des modellierten Begriffs *Stadt* ist mit dem Eintrag 'Hannover' im Attribut *Namen*. Konzepte werden in dieser Arbeit durchgängig in eckigen und Instanzen in abgerundeten Kästchen dargestellt. Bereits in dem in Abbildung 5.1 darge-

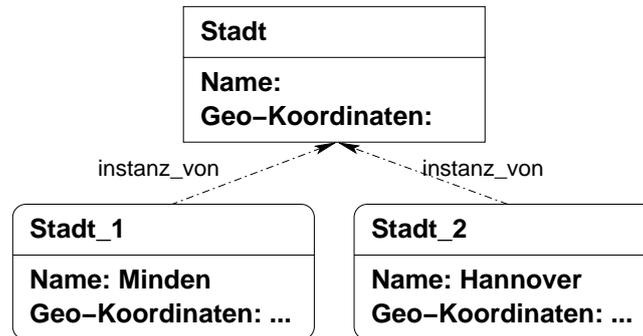


Abbildung 5.2: Mehrere Instanzen eines Konzeptes

stellten Semantischen Netz sind einige Vorzüge dieser Repräsentationsform zu erkennen. Hierzu gehören:

Modularität: Da das gesamte Wissen über einen Begriff zentral im dazugehörigen Konzept abgelegt ist, ist ein modularer Zusammenbau eines Semantischen Netzes möglich. Diese Eigenschaft ist beim Erweitern und Verändern eines vorhandenen Netzes von Vorteil, da die einzelnen Objekte in Form von Bausteinen aus einem Netz entfernt oder in dieses eingefügt werden können.

Strukturiertheit: Die gewählte Struktur erleichtert dem Anwender die Erzeugung von umfangreichen Wissensbasen, da der verwendete Formalismus dem menschlichen Strukturierungsschema ähnelt.

Übersichtlichkeit: Durch den klaren Aufbau des Netzes durch Knoten und Kanten und der genannten Strukturiertheit ergibt sich ein übersichtlicher Aufbau der Wissensbasis.

Kompaktheit: Die Darstellung des Wissens ist kompakt, da sie sich auf die wesentlichen Dinge konzentriert und unwichtige Details, wie etwa die Interna einzelner Bildverarbeitungsoperatoren, verdeckt, jedoch die Möglichkeit des Zuganges erhält.

5.5 Strukturelle Komponenten von GeoAIDA

Die grundlegenden Komponenten zur Wissensrepräsentation in GeoAIDA bilden, wie in Semantischen Netzen üblich, die Knoten und Kanten. Diese und die weiteren für eine Interpretation von komplexeren Eingangsdaten nötigen Strukturen werden im Folgenden vorgestellt. In Kapitel 6 wird die Funktion und deren Aufgabenbereich motiviert.

5.5.1 Konzeptdefinitionen

Die Modellierung eines Begriffs oder einer Idee geschieht in GeoAIDA in einem Konzept. Diese Konzeptdefinition beinhaltet Methoden, Attribute und Bestandteile, um das zu diesem Begriff im verwendeten Kontext benötigte Wissen zu modellieren. Diese Struktur ist konzeptionell für den Begriff *Siedlung* in Abbildung 5.3 dargestellt.

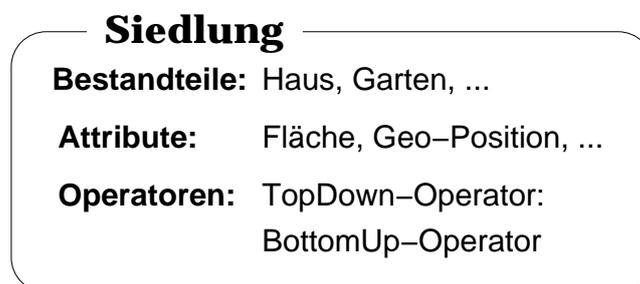


Abbildung 5.3: Darstellung des Konzeptes *Siedlung*

Die Gesamtheit der Konzeptdefinitionen mit den verbindenden Kanten, bildet das Konzeptnetz und die Struktur der Wissensbasis. Während der Analyse entstehen aus dieser generischen Wissensbasis – entsprechend den vorliegenden Sensorsignalen – Instanzen. Dieser Prozess ist anschaulich mit dem Prozess vergleichbar, durch den in einer objektorientierten Programmiersprache Objekte auf Grundlage einer Klassendefinition entstehen.

Konzepte umfassen hierbei nur die allgemeine Beschreibung des repräsentierten Begriffs und seiner Eigenschaften, unabhängig von einem konkreten Signal. In-

stanzen hingegen können nur in Verbindung mit konkreten Signalen erzeugt werden.

5.5.2 Instanzen

Instanzen sind im Gegensatz zu den Konzepten die Objekte, die nicht mehr nur ein Modell darstellen, sondern mit konkreten Sensorsignalen verbunden sind.

Der Schritt der Instantiierung wurde bereits in Abbildung 5.2 durch die *instanz_von*-Kanten motivierend dargestellt. Hierbei handelt es sich um eine Instanzenbildung auf der Grundlage der Konzeptdefinitionen, wobei diese mit konkreten Werten eines Signalbereiches belegt wird. Die Werte, die Attribute der Instanz, werden durch die mit dem Konzept verbundenen Bildverarbeitungsoperatoren bestimmt oder ergeben sich aus der Struktur des vorliegenden Instanzennetzes. Während also im Konzept *Siedlung* vermerkt ist, dass einer *Siedlung* eine geographische Position zugeordnet ist, ist in einer Instanz *Siedlung* ein konkreter – von den jeweiligen vorliegenden Sensordaten abhängiger – Wert eingetragen. Die Instanzen können im – aus der Struktur des Konzeptnetzes entstehenden – Instanzennetz fünf Zustände annehmen.

HI steht für hypothetische Instanz, hierbei wird ein Sensorsignal oder ein Teil des Sensorsignals einer Instanz zugeordnet, ohne dass diese Zuordnung bereits eindeutig festgelegt wird. Hieraus folgt, dass für den verwendeten Signalbereich noch weitere Hypothesen existieren können, die einer alternativen Interpretation entsprechen.

PI repräsentiert eine partielle Instanz. Diesen Status erhält eine hypothetische Instanz, nachdem weitere Informationen der Instanz hinzugefügt wurden, aber ein weiterer, meist iterativer Analyseablauf die Instanz bestätigen oder verwerfen muss. Hier wird also zum Ausdruck gebracht, dass es innerhalb eines Schritts nicht möglich war, alle notwendigen Analyseergebnisse zu bestimmen und vorerst ein erreichtes Zwischenergebnis vorliegt.

CI bezeichnet eine komplette Instanz, die aus einem der beiden vorangegangenen Zustände entstehen kann. Für eine komplette Instanz liegen alle

bestimmbaren Informationen vor. Spätestens zu diesem Zeitpunkt sollte für die Instanz eine Bewertung vorliegen, falls es zum entsprechenden Bildausschnitt konkurrierende Interpretationen gab.

MI ist der Zustand von Instanzen, die aufgrund der Anwendung eines holistischen Operators verworfen wurden und steht für fehlende (engl. missing) Instanz.

TRASH kennzeichnet eine Instanz, deren Unterknotenstruktur dazu führte, dass diese Instanz verworfen (wertlos engl. trash) wurde. Hier hat folglich die strukturelle und nicht die holistische Analyse zum Verwerfen einer Instanz geführt.

Das Setzen der Attributwerte einer Instanz geschieht über Funktionen, die bei der Konzeptdefinition über eine definierte Schnittstelle spezifiziert werden. Hierdurch ist es möglich, beliebige vorliegende Programme zur Attributberechnung im Gesamtsystem zu integrieren. Diese Funktionen werden im Folgenden unter dem Begriff Bottom-Up-Operatoren zusammengefasst.

Falls es vom Anwender gewünscht ist, nicht das gesamte Semantische Netz zu expandieren, können Äste des Baumes deaktiviert werden, indem Konzepte als virtuelle Blattknoten definiert werden. Dieses geschieht durch einen Schalter, über den jedes Konzept verfügt und der alle Subknoten im Analyseablauf abschaltet.

Weiterhin verfügt eine Instanz über einen Analysestatus, der protokolliert, in welchem Analysezustand sich die Instanz befindet.

TD steht für Top-Down-Analyse. In diesem Status befindet sich eine Instanz zum Zeitpunkt ihrer Erzeugung und behält diesen bis die Top-Down-Analyse abgeschlossen ist.

BU kennzeichnet eine Instanz während der Bottom-Up-Analyse.

Bei den Attributen eines Konzeptes können die für die Analyse obligatorischen von den optionalen Attributen eines Konzeptes unterschieden werden. Zu den obligatorischen Attributen einer Instanz gehören:

- der Klassenname, den die Instanz vom zugehörigen Konzept erhält,
- ein eindeutiger Name, der die Instanz identifiziert und sich aus dem Klassennamen und einer Nummer als Suffix ergibt,
- die Geokoordinaten, die die zur Instanz gehörende Region als umschließendes Rechteck beschreibt,
- die Angabe eines Labelbildes, in dem die Region symbolhaft dargestellt wird,
- das Label, mit dem die Region im Labelbild kodiert ist,
- die das Labelbild beschreibenden Geokoordinaten,
- die Bildkoordinaten des umschließenden Rechtecks für die durch den Label gegebene Region.

Die Menge der obligatorischen Attribute ist beliebig und von den jeweiligen Aufgabenstellungen abhängig.

5.5.3 Kanten

Kanten stellen die Relationen zwischen Konzepten explizit dar. In GeoAIDA werden keine speziellen Kantentypen vorgegeben. Es ist dem Benutzer freigestellt, die Semantik der – die Konzepte verbindenden – Kanten selbst funktional zu beschreiben. Selbstverständlich sind die bereits oben angesprochenen generellen Kantentypen realisierbar und werden im Folgenden detaillierter beschrieben.

Bestandteilkanten: Eine wichtige Relation etablieren die Bestandteilkanten, oft mit *teil_von*, *bst_von* oder *part_of* bezeichnet, die eine Dekomposition eines Begriffs in Subbegriffe realisiert. In der hier vorliegenden Arbeit ist diese Zerlegung hauptsächlich als eine Unterteilung einer Region in Teilregionen zu verstehen. Dieser Zusammenhang kommt beispielsweise bei den Konzepten

Wasserweg, *Verkehrsweg* und *Stadt* zum Konzept *Bundesland* in Abbildung 5.1 zum Ausdruck. Um den Sachverhalt ausdrücken zu können, dass es innerhalb eines Bundeslandes mehrere Objekte vom Typ *Stadt* geben kann, ist eine solche Kante möglicherweise mehrdimensional und etabliert eine Abbildung der Bestandteile zu dem zusammenfassenden Konzept. Hierbei kann in manchen Fällen gefordert sein, dass eine Mindestanzahl von obligatorischen Bestandteilen eines Typs vorliegen muss, damit das übergeordnete Konzept bestätigt werden kann. Beispielsweise müsste ein *Quadrat* genau vier *Eckpunkte* als Bestandteile besitzen. Andererseits kann ein Objekt neben den obligatorischen Bestandteilen weitere optionale Bestandteile besitzen. Ein Beispiel hierfür ist die Anzahl der *Bundesländer* in einem *Land*. Es sollte mindestens zwei Bundesländer in einem Land geben, da eine begriffliche Verfeinerung anderenfalls überflüssig wäre, es können aber mehr als zwei – also weitere obligatorische Bundesländer – existieren.

Die dargestellte Semantik der Bestandteilkante wird auch oft als kontextabhängige Bestandteilkante gekennzeichnet, da der Ursprungsknoten der Kante nur im Kontext des übergeordneten Zielknotens zu verstehen ist. Eine Stadt würde in der hier dargestellten Wissensbasis ohne die Zuordnung zu einem Bundesland keinen Sinn ergeben. Über kontextabhängige Bestandteilkanten ist folglich ein Datenaustausch wünschenswert. Dieser Datenaustausch kann sowohl modellgetrieben, *von oben nach unten*, stattfinden beispielsweise durch die Einschränkung des Zuständigkeitsbereichs der untergeordneten Knoten, als auch datengetrieben, *von unten nach oben*, um etwa die gefundenen Attribute an das übergeordnete Zielkonzept weiterzureichen.

Spezialisierungskanten: Die Spezialisierungskanten verbinden einen spezielleren Begriff mit seiner generelleren Repräsentation. Der Ausdruck *Kanal* ist durch den Begriff *Wasserweg* genereller beschrieben. Dieser Kantentyp wird häufig als *ist_ein*, *is_a* oder bei inverser Betrachtung der Beziehung als *kon_of* bezeichnet. Auch in diesem Fall ist ein Austausch von Attributen zwischen den durch diese Kante verbundenen Knoten in beide Richtungen möglich. Dieser Kantentyp wird hauptsächlich für eine begriffliche Strukturierung der Wissensbasis verwendet, wodurch deren Übersichtlichkeit erhöht und ein Übergang zwi-

schen verschiedenen Begriffswelten realisiert wird. Verschiedene Abstraktionsebenen trennen dabei unterschiedliche Begriffsebenen und ermöglichen verschiedene Sichten auf die vorliegende Wissensbasis. Weiterhin wird ein Konzept A konkreter als ein Konzept B bezeichnet, falls dieses direkt oder indirekt über eine *ist_ein* Kante mit B verbunden ist. Folglich ist A konkreter als B , wenn B abstrakter als A ist. Konkretisierungskanten sind im Gegensatz zu den Bestandteilkanten eine eindeutige Abbildung zwischen den verbundenen Instanzen.

Bei der Definition der Kanten wurde ähnlich vorgegangen wie bei den Attributberechnungen der Konzeptdefinitionen. Die nötigen Funktionen, die die Semantik einer Kante beschreiben, können außerhalb des Analyseprogramms realisiert und anschließend in das System integriert werden. Diese Funktionen – im Folgenden als Top-Down-Operatoren bezeichnet – übernehmen vorrangig die Gruppierung von Instanzen.

Instantiierungskanten: Sie bilden die Modellwelt in die reale vorgefundene Welt ab und stellen eine Abbildung von dem Konzeptnetz in das Instanzennetz dar und werden als *inst_of* oder *inst_von* bezeichnet. Diesen Kanten ist keine benutzerdefinierte Funktion zuweisbar und werden vom Analyseprozess automatisch erzeugt. Über diese Kanten wird die Erzeugung der Instanzen inklusive der Generierung der benötigten Attribute realisiert, als auch die Anbindung der notwendigen Funktionen an die Instanz vorgenommen.

5.5.4 Attribute

Die individuellen Eigenschaften von Instanzen werden in GeoAIDA in Attributen abgelegt und liefern einen wesentlichen Beitrag zur Erreichung des Analyseziels. Die Attribute können dabei anfänglich leer oder mit Default-Werten versehen sein. Weitere, während der Analyse benötigte Attribute müssen nicht zwingend in der Konzeptdefinition enthalten sein, sondern können dynamisch während der Analyse ermittelt und in einer Attributliste verwaltet werden.

Durch die Attributberechnungsfunktionen – im Top-Down- und Bottom-Up-

Operator – erfolgt die Ermittlung der konkreten Attributwerte für die Instanzen, auf deren Grundlage diese bewertet und damit verifiziert oder falsifiziert werden können. Hier wird deutlich, dass ein Großteil der Leistungsfähigkeit des gesamten Analysesystems auf den Bildverarbeitungsoperatoren und den Berechnungsfunktionen für die Attribute aufbaut und diese die Basis für die Qualität der Analyse darstellen.

5.5.5 Relationen

Ein weiteres wichtiges Kriterium der strukturellen Mustererkennung ist die Betrachtung von Relationen zwischen den Instanzen. Diese Relationen vergleichen Attribute von Instanzen und liefern zusätzliche Informationen über die Struktur von Objekten und bilden eine weitere Grundlage, die Plausibilitäten dieser zu überprüfen. Die Ergebnisse des Vergleiches können ebenfalls für die Bestätigung oder Verwerfung einer Hypothese, aber auch für die Bewertung einer Instanz Verwendung finden. Ein durch eine Relation modellierter Zusammenhang wäre, eine zu einem Haus benachbarte Wiese als Garten zu interpretieren und eine Wiese, die in einem landwirtschaftlich genutzten Bereich liegt, als Weideland zu bezeichnen.

5.5.6 Modalitäten

Die Formulierung unter Verwendung von obligatorischen und optionalen Kanten ist nicht für alle Objekte ausreichend, da ein Begriff gegebenenfalls verschiedene Ausprägungen besitzen kann. Ein Industriegebiet kann aus einem Hafenbereich und einer Ansammlung von runden Tanks zur Gaslagerung oder aus einer Gruppierung von großen Gebäuden bestehen. Diese beiden Ausprägungen ein und desselben Begriffes geschehen über die Modalitäten. Diese stellen verschiedene Gruppierungen von Bestandteilen dar, die einen übergeordneten Begriff bestätigen können.

5.5.7 Bewertungen

Die Interpretation von komplexen Daten, wie die in digitaler Form vorliegenden Fernerkundungsdaten, ist nicht fehlerfrei möglich. Dieses ist leicht verständlich, bedenkt man die unterschiedlichen Ausprägungen von Objekten in den Sensordaten bei unterschiedlichen Beleuchtungsbedingungen, zu unterschiedlichen Jahreszeiten oder die große Varianz von Repräsentationen ein und desselben Objekttypes, wie beispielsweise bei Häusern. Folglich sind die der Interpretation zugrunde liegenden Daten stets mit einer gewissen Unsicherheit behaftet, der während der Analyse Rechnung getragen werden muss. Dieses betrifft sowohl die Entscheidung, ob ein Objekt an einer gewissen Stelle in den Daten vorliegt, als auch die Behandlung von Konflikten, in denen sich zwei oder mehrere Objekte gegenseitig ausschließen. Die Bewertung und deren Funktionen sind in GeoAIDA benutzerdefiniert zu realisieren. Beispiele für die Verwendung von Bewertungen im Kontext der Auswertung von Fernerkundungsdaten werden im Kapitel 6.3 angegeben. Im Folgenden werden Bewertungsansätze dargestellt, die in GeoAIDA vom Benutzer verwendet werden können.

Diskrete Entscheidung: Hier wird eine einfache binäre Entscheidung $\{0, 1\}$ getroffen, ob ein Objekt erkannt wurde oder nicht.

Kontinuierliche Bewertung: Anhand von Attributwerten wird ein Bewertungsmaß bestimmt, welches in den Wertebereich $[0, 1]$ abgebildet wird.

lexikographische Ordnung: Verschiedene Bewertungen innerhalb eines Objektes können in einem hierarchischen Vektor abgelegt werden. Ein Vergleich mit einem konkurrierendem Objekt geschieht über die verschiedenen Ebenen des Bewertungsvektors, bis eine Entscheidung getroffen werden kann.

Einbeziehung von Unsicherheit: Bei der Bewertung einer Aussage können auch vage Vertrauensmaße Verwendung finden [101]. So kann etwa der zur Verfügung stehende Bereich $[0, 1]$ in die disjunkten Bereiche *Glaubwürdigkeit*, *Unwissenheit* und *Unglaubwürdigkeit* eingeteilt werden. Die Glaubwürdigkeit ist der Wert, der für eine Aussage spricht, während die Unglaubwürdigkeit angibt, was gegen diese Aussage spricht. Die Unwissenheit ist der verbleibenden Bereich zwischen der Glaubwürdigkeit und der Unglaubwürdigkeit.

Es gilt: $Glaubwürdigkeit + Unglaubwürdigkeit + Unwissenheit = 1$.

Als *Plausibilität* wird die Summe von Glaubwürdigkeit und Unwissenheit bezeichnet.

Erfassung von Ungenauigkeit: Neben der Unsicherheit einer Messung kann diese auch über eine Ungenauigkeit verfügen. So beschreibt eine ungenaue Messung einen Toleranzbereich für einen Attributwert. Unter Verwendung von Fuzzy-Mengen kann sowohl der mit einer Toleranz behaftete Messwert als auch die im Modell festgelegten Erwartungswerte für ein Attribut als Trapez dargestellt werden. Die Fuzzy-Theorie liefert Methoden, um mit diesen Trapezfunktionen zu rechnen, aus denen man die *Möglichkeit* und die *Notwendigkeit* für eine Aussage bestimmen kann.

Zusammenfassung von Bewertungen:

Die Zusammenfassung von Bewertungen zu einer gemeinsamen Aussage ist erforderlich, falls die Bewertung eines Objektes anhand der Bewertungen seiner Bestandteile geschieht, oder ein Objekt mehrere bewertete Attribute besitzt, die zu einer Aussage vereinigt werden. Zu unterscheiden ist die Kombination komplementärer Bewertungen und redundanter Bewertungen.

Zusammenfassung komplementärer Bewertungen: Hier ist die Maximum-, Minimum- oder Mittelwertbildung ein gängiges Verfahren, um komplementäre Bewertungsmaße zu vereinigen. Es können aber auch gewichtete Produkte der Einzelbewertungen gebildet werden. Dieser Ansatz ist bei der Bewertung eines Objektes anhand der Bewertungen seiner Bestandteile anzuwenden.

Zusammenfassung redundanter Bewertungen: Liegen unabhängige Aussagen mit unterschiedlichen Sicherheiten für ein Objekt vor, so wird bei der Verwendung der Möglichkeit und Notwendigkeit in [30] für die Möglichkeit Po der minimale Wert $Po(s) = \min_i Po_i(s)$ und für die Notwendigkeit Ne der maximale Wert gewählt $Ne(s) = \max_i Ne_i(s)$. Durch die Anwendung dieser Regel kann es bei sich widersprechenden Aussagen dazu kommen, dass die Notwendigkeit größer als die Möglichkeit wird und die Voraussetzung $Po(s) \geq Ne(s)$ nicht erfüllt ist. Dieses Problem wird durch die Kombination der Aussagen nach der Dempster-Shafer Evidenztheorie [28] [93] [37], die ein numerisches Verfahren zur Verarbeitung von unsicherem Wissen bereitstellt, behoben. Die fehlende

Möglichkeit der klassischen Wahrscheinlichkeitstheorie, das Fehlen von Informationen in adäquater Weise auszudrücken, führte zur Evidenztheorie, in der auf die Additivität der Wahrscheinlichkeitsmaße verzichtet wurde. Dieses Maß wurde durch ein Vertrauensintervall ersetzt, dessen Grenzen für die pessimistische und die optimistische Einschätzung eines Ereignisses stehen. Mit der Evidenztheorie ist es nun möglich, die mit Plausibilität und Glaubwürdigkeit bezeichneten Werte konsistent zu verarbeiten.

5.6 Analysestrategie

Nachdem nun die Bildverarbeitungsoperatoren und die Definition des Semantischen Netzes – als problemspezifische Komponenten von GeoAIDA zur Wissensrepräsentation – vorgestellt wurden, soll nun beschrieben werden, wie das modellierte Wissen zur Interpretation zu verwenden ist, das heißt, in welcher Weise die Zuordnung der Sensordaten zu den Konzeptdefinitionen geschieht. Einen Überblick über das Gesamtsystem zeigt Abbildung 5.4. Auf der linken Seite sind die Eingangsdaten zu sehen. Sie umfassen die zu analysierenden Daten – multisensorielle Sensordaten auch unterschiedlicher Auflösung, Vektordaten – und das hier symbolisch dargestellte Semantische Netz als problemspezifische Wissensbasis. Auf der rechten Seite sind die Ergebnisse des Systems angedeutet, die aus thematischen Karten in unterschiedlichen Abstraktionsebenen und einer strukturellen Beschreibung in Form eines Instanzennetzes besteht. Jedem erkannten Objekt ist in der strukturellen Beschreibung eine Instanz zugeordnet, die wiederum einem Label in der thematischen Karte entspricht. Die Ebenen des Instanzennetzes korrespondieren mit den Abstraktionsebenen der Kartendarstellung.

Die in der Mitte dargestellte, problemunabhängige Systemsteuerung besteht aus zwei zentralen Analyseschritten. Die Top-Down-Analyse zerlegt iterativ Regionen in Teilregionen, während die Bottom-Up-Analyse gefundene Teile zu Begriffen der darüberliegenden Instanzen gruppiert. Die Operatoren, die für die problemunabhängige Analysestrategie die problemspezifischen Daten bestimmen, werden als externe Top-Down- und Bottom-Up-Operatoren aufgerufen. Diese Operatoren sind in den Konzepten des Semantischen Netzes eingetragen.

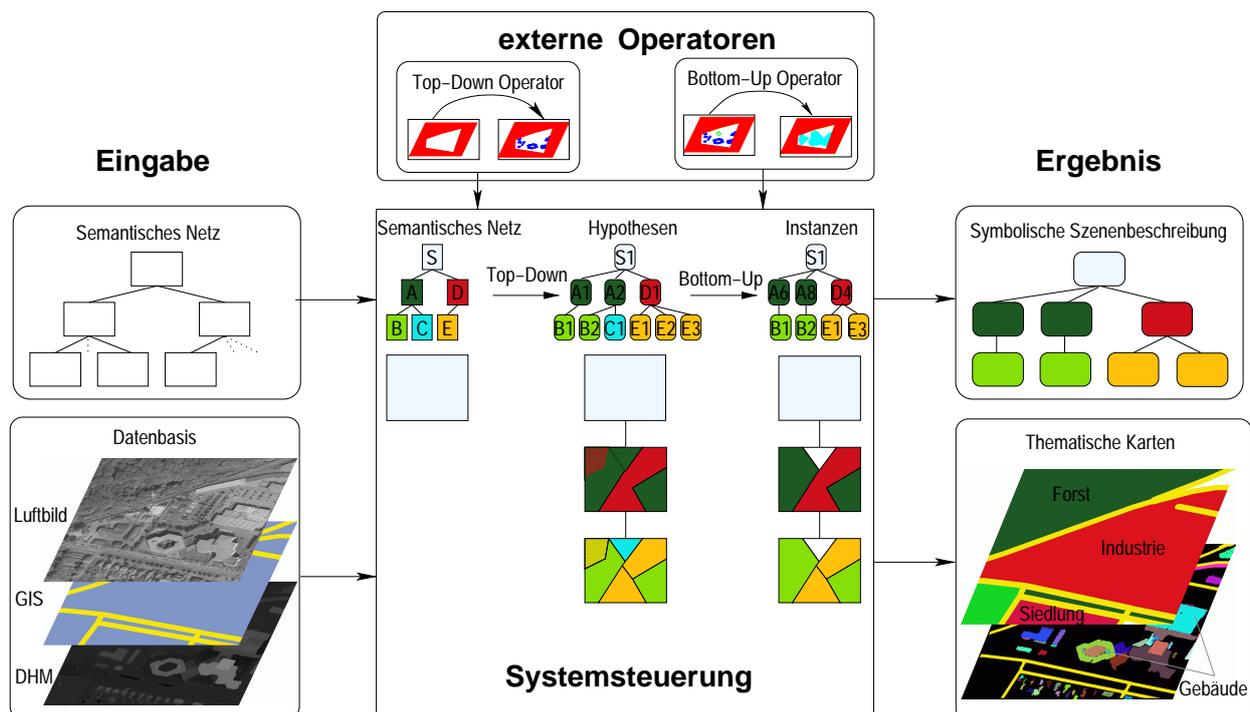


Abbildung 5.4: Systemüberblick GeoAIDA

Durch die Definition einer klaren Schnittstelle für diese Operatoren ist es leicht möglich, vorhandene oder aus fremden Bildverarbeitungssystemen stammende Operatoren im System zu integrieren. Diese beiden Analyseschritte werden im Folgenden genauer beschrieben.

5.6.1 Top-Down-Analyse

Die Aufgabe der Top-Down-Analyse ist die Dekomposition einer Region in Teilregionen. Hierzu startet im obersten Konzept des Semantischen Netzes die Top-Down-Analyse, die einer modellgetriebenen Vorgehensweise entspricht und generiert aus diesem Konzept S eine hypothetische Instanz (HI). Dieser obersten Instanz sind die gesamten Eingangsdaten zugeordnet. In Abbildung 5.4 ist dies durch das helle Rechteck mit der Bezeichnung S_1 als hypothetische Instanz und dem hellen Rechteck unterhalb des mittleren Netzes – welches den gesamten Bildbereich symbolisiert – angedeutet. Ausgehend von dieser obersten, hypothetischen Instanz wird für alle darunterliegenden Konzepte im Konzeptnetz,

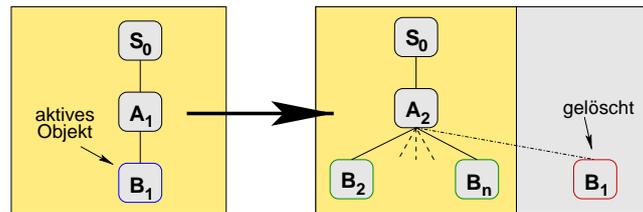


Abbildung 5.5: Funktionsweise des Top-Down-Operators

die durch Kanten mit dem Konzept der Instanz S_1 verbunden sind, der Top-Down-Operator – wie in Abbildung 5.5 dargestellt – aufgerufen.

Wird von der Instanz A_1 vom Typ A in Abbildung 5.5 der Top-Down-Prozess für ein Konzept vom Typ B aufgerufen, so erzeugt dieser Prozess eine temporäre hypothetische Instanz B_1 vom Typ B . Für die Instanz B_1 wird nun der Top-Down-Operator aufgerufen, der als zu bearbeitende Region die gesamte, der hypothetischen Instanz A_1 zugehörige Region erhält. Der Top-Down-Operator liefert als Ergebnis eine Liste von Instanzen vom Typ B , die Teilregionen der durch A_1 repräsentierten Region bilden und in Abbildung 5.5 durch die hypothetischen Instanzen B_2 bis B_n dargestellt sind. Die Instanzen B_2 bis B_n sind disjunkte Teilregionen der Eingangsregion. Das temporäre Objekt B_1 wird nach dem Verbinden der Instanzen B_2 bis B_n mit der darüberstehenden Instanz A_1 gelöscht.

Falls ein Konzept A im Semantischen Netz mehrere Unterknoten besitzt, wird der beschriebene Prozess für alle Unterknoten parallel gestartet. Die Disjunktheit bezüglich der den Instanzen zugeordneten Flächen gilt zwischen den Instanzen desselben Typs, jedoch nicht zwingend für Instanzen unterschiedlichen Typs. Liegt eine Überschneidung für zwei Instanzen unterschiedlichen Typs vor, so repräsentieren diese Instanzen konkurrierende Interpretationen bezüglich ihrer Schnittmenge. Zu beachten ist hierbei, dass den Ergebnissen der Top-Down-Operatoren bei den unterschiedlichen Instanztypen auch unterschiedliche Sensordaten zugrunde liegen können, hierbei also nicht zwingend eine Überdeckung in den Sensordaten vorliegt. Eindeutig festzustellen ist eine Überschneidung daher nur aufgrund des gemeinsamen Koordinatensystems für alle Sensordaten. Das Auflösen dieser konkurrierenden Interpretationen ist an dieser Stelle der Analyse nicht gewünscht und ist Aufgabe des Bottom-Up-Operators.

Als Top-Down-Operator können beliebige Bildverarbeitungsoperatoren eingesetzt werden, die eine Region in Teilregionen unterteilen. Diese werden im Folgenden als *holistische Operatoren* bezeichnet, da sie Hinweise auf komplexere Objekte liefern, ohne diese explizit strukturell zu beschreiben und somit im Gegensatz zu den strukturellen Verfahren der Bottom-Up-Analyse stehen. Der Einsatz von holistischen Operatoren führt zu einer Effizienzsteigerung der Analyse. Einerseits werden Regionen als nicht zur aktuellen Klassen zugehörig markiert und müssen daher auch nicht weiter untersucht zu werden. Andererseits können unterschiedliche, entstandene Teilregionen parallel abgearbeitet werden. Die Funktionsweise der holistischen Operatoren wird in Kapitel 5.6.4 näher beschrieben. Liegt für ein Konzept kein holistischer Operator vor, so wird für die gesamte Region eine Hypothese vom Typ des Konzeptes gebildet.

Der aufgezeigte Prozess eines Top-Down-Schrittes wird nun iterativ für alle entstandenen Unterknoten ausgeführt. Die Top-Down-Analyse endet an den Blättern des Semantischen Netzes oder auch in höhergelegenen Konzepten, falls der Anwender die Tiefe der gewünschten Interpretation eingeschränkt hat.

5.6.2 Bottom-Up-Analyse

Der Bottom-Up-Analyseschritt entspricht einer datengetriebenen Auswertung der Szene. Wie bereits angesprochen, ist die Aufgabe der Bottom-Up-Operatoren das Gruppieren von Instanzen und das Bewerten dieser Gruppierung. Gestartet wird der Bottom-Up-Operator für eine bestehende Instanz, falls für alle Subinstanzen der Bottom-Up-Operator aufgerufen wurde. Für die Blattknoten wird der Bottom-Up-Operator direkt nach dem Top-Down-Operator aktiviert.

Der Vorgang der Bottom-Up-Analyse ist in Abbildung 5.6 dargestellt. Hierbei zeigt der obere Knoten eine vorliegende Interpretation, die noch mögliche konkurrierende Teilinterpretationen enthält. Der Aufruf des Top-Down-Operators für die Instanz A_1 könnte dabei beispielsweise zu einem der unten dargestellten Ergebnisse führen. Der Ergebnisknoten **I** enthält alle Subknoten, da alle Objekte den in A_1 spezifizierten Bedingungen genügen und diese als komplette Instanz bestätigt wurde. Im Ergebnisknoten **II** wurden die Instanzen C_2 und C_3 für die

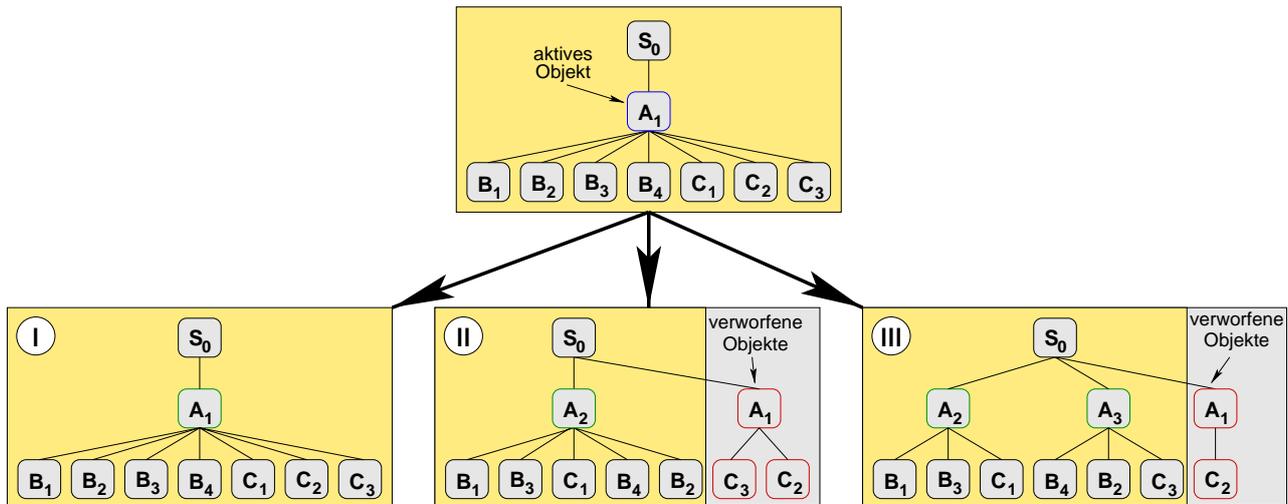


Abbildung 5.6: Funktionsweise des Bottom-Up-Operators mit drei alternativen Ergebnissen

Bildung der Instanz A_2 nicht benötigt bzw. ihre Eigenschaften genügten nicht denen, die im Bottom-Up-Operator spezifiziert wurden. Auch eine Aufspaltung der durch A_1 repräsentierten Region in Teilregionen ist, wie im Ergebnisknoten **III** dargestellt, möglich. Die den Instanzen A_2 und A_3 zugeordneten Subknoten genügen den im Konzept A vorliegenden Beschreibungen. Die Instanzen A_1 und C_2 konnten nicht bestätigt werden. Praktische Beispiele für die hier dargestellten Situationen werden im Kapitel 6.2 aufgezeigt.

Das Auflösen von Konflikten, also die Zuweisung einer Flächeneinheit zu einer eindeutigen Instanz, ist neben der Gruppierung der Unterknoten Aufgabe des Bottom-Up-Operators. Der Konfliktbereich kann dabei eine ganze Region oder auch Teilregionen, die eine Schnittmenge von zwei alternativen Regioneninterpretationen darstellt, umfassen. Zum Auflösen der Konflikte wird eine Bewertung der Alternativen benötigt, die durch eine vom Benutzer zu definierende Funktion (siehe Kap. 5.5.7) zu berechnen ist. Bei der Bestimmung der Bewertung sollten die Attribute der Unterknoten und ihre Beziehung zueinander Verwendung finden.

5.6.2.1 Generischer Bottom-Up-Operator

Da der Bottom-Up-Operator für die meisten Anwendungen einen ähnlichen Aufbau besitzt, wurde ein generischer Bottom-Up-Operator realisiert, der über eine funktionale Beschreibung an die zu realisierenden Aufgaben angepasst werden kann. Hierbei handelt es sich um eine Beschreibungssprache, die dazu dient, Operationen auf Mengen auszuführen. Die Mengenelemente repräsentieren hypothetische Instanzen, die den übergeordneten Begriff konkurrierend oder komplementär beschreiben. Zu den Mengenoperationen gehören das Vereinigen von Mengen und das Bilden von Teilmengen durch Selektieren anhand von Attributwerten der vorliegenden Instanzen oder deren Relationen zu anderen Instanzen.

Die Beschreibung der gewünschten Funktionalität geschieht in einer Postfix-Notation und wird einem speziellen Operator übergeben. Mit Hilfe des generischen Bottom-Up-Operator lassen sich eine Vielzahl von Attributen für die vorliegenden Instanzen berechnen, aus denen auch ein Bewertungsmaß zur Auflösung von möglichen alternativen Gruppierungen bestimmt werden kann. Dieser Operator gestattet es, ohne Programmieraufwand einen problemangepassten Bottom-Up-Operator zu generieren. Eine Anwendung des Operators ist in Kapitel 6.2 zu finden.

5.6.3 Analysestruktur

Die Analyse ist im Folgenden in einer strukturellen Form dargelegt und bezieht die im Programmpaket realisierten Funktionen mit ein. Die Analyse kann nach dem Laden einer Projektbeschreibung mit der Funktion *StarteAnalyse*, die als Struktogramm in Abbildung 5.7 dargestellt ist, aktiviert werden. Die Projektbeschreibung beinhaltet ein Semantisches Netz und eine Referenz auf den zu bearbeitenden Datensatz. Die Analyse startet im obersten Konzept des Semantischen Netzes – dem Wurzelknoten – und erzeugt aus diesem eine Instanz, für die anschließend die Funktion *childTopDown* aufgerufen wird. Nach der gesamten Analyse werden die Ergebnisse für die Ergebnisdarstellung aufbereitet und dem Benutzer präsentiert.

Funktion <i>StarteAnalyse</i>	
Bestimme den Wurzelknoten des Konzeptnetzes	
IF	\exists Wurzelknoten
THEN	<ol style="list-style-type: none"> 1. Erzeuge eine Instanz von dem Typ des Wurzelknoten 2. Rufe für diese Instanz die Funktion <i>childTopDown</i> 3. Nach Ablauf der Analyse erzeuge Ergebnisdarstellung
ELSE	Beende die Analyse

Abbildung 5.7: Struktogramm der Analyse

Die iterativ durch das gesamte Netz laufende Funktion *childTopDown* ist in Abbildung 5.8 dargestellt. Hierbei wird für alle Unterknoten die Funktion *execTopDown* parallel aufgerufen, sodass auch die als externe Prozesse laufenden Bottom-Up-Operatoren parallel ablaufen. Dies führt zu einem effizienten und zügigen Fortschritt der Analyse, die durch den Einsatz von Mehrprozessor-Rechnern beschleunigt werden kann.

Es gibt jedoch Situationen, in denen eine parallele Bearbeitung aller Knoten nicht sinnvoll ist. Dieser Fall liegt z. B. dann vor, wenn bekannt ist, dass spezielle obligatorische Bestandteile zu dieser Instanz vorliegen müssen. In diesem Fall ist es wünschenswert, zuerst zu überprüfen, ob die obligatorischen Bestandteile etablierbar sind, um im Falle des Zutreffens auch die weiteren Bestandteile zu expandieren. Um eine solche Steuerung zu erreichen, ist es möglich, jedem Konzept eine Priorität zuzuordnen, die die Reihenfolge regelt, in der die einzelnen Subknoten aufgerufen werden. Erfüllt das Ergebnis die Bedingung, dass die Anzahl der Objekte für dieses Konzept zwischen der minimalen und maximalen Anzahl liegt, so wird die Analyse mit den restlichen Konzepten der nächsten Priorität fortgeführt. Ist die angegebene Bedingung nicht erfüllt, so wird die Analyse dieses Teilbaumes nicht weiterverfolgt und der Status der aufrufenden Instanz auf den Wert *TRASH* gesetzt.

Die Erzeugung von neuen hypothetischen Instanzen geschieht in der Funktion *execTopDown* (siehe Abbildung 5.9). Hier wird der durch den Benutzer angegebene externe Top-Down-Operator für den aktuellen Bildausschnitt aufgerufen. Als Ergebnis liefert dieser Operator eine Liste von hypothetischen Instanzen, die

Funktion: <i>childTopDown</i>	
Bestimme alle Unterknoten des aktuellen Knoten <i>akt</i> im Konzeptnetz	
IF	\exists mindestens ein Unterknoten \wedge der Knoten <i>akt</i> ist nicht als letzter zu instantiierender Knoten markiert
THEN	Sortiere diese Liste <i>L</i> von Konzepten entsprechend ihrer Priorität
	FOR \forall Prioritäten
	FOR \forall Konzepte einer Priorität
	<ol style="list-style-type: none"> 1. Erzeuge Instanzen für diese Konzepte 2. Setze Instanzattribute: Status = HI, Analysestatus = TD 3. Wende die Funktion <i>execTopDown</i> an (<i>execTopDown</i> ruft den Top-Down-Operator auf, der im zur Instanz gehörigen Konzept durch den Benutzer eingetragen wurde. Der Top-Down-Operator liefert eine Liste von Instanzen vom aktuellen Typ zurück)
ELSE	<ol style="list-style-type: none"> 1. Setze Instanzattribute: status = Status = CI, Analysestatus = BU 2. Führe die Funktion <i>execBottomUp</i> aus

Abbildung 5.8: Struktogramm der Funktion *childTopDown*

mit der aufrufenden Instanz verbunden werden. Anschließend wird für die neu entstandenen Instanzen die Funktion *childTopDown* aufgerufen, wodurch eine iterative Abarbeitung des Semantischen Netzes von oben nach unten erfolgt.

Für eine Instanz, die keine Unterknoten besitzt (Blattknoten) oder an denen der Benutzer keine tiefere strukturelle Analyse wünscht, wird nach der Abarbeitung der Top-Down-Analyse die Funktion *execBottomUp* – wie in Abbildung 5.10 beschrieben – aufgerufen. Nachdem für alle Unterinstanzen einer Instanz die Funktion *childTopDown* abgearbeitet wurde, wird für diese Instanz die Bottom-Up-Funktion aktiviert. Dieser Vorgang durchläuft das ent-

Funktion: <i>execTopDown</i>	
<i>execTopDown</i> ruft den Top-Down-Operator auf, der im zur Instanz gehörigen Konzept durch den Benutzer eingetragen wurde. Der Top-Down-Operator liefert eine Liste von Instanzen vom aktuellen Typ zurück.	
IF	Liegt die Anzahl der vom Top-Down-Operator erzeugten Instanzen innerhalb des Intervalls der obligatorisch geforderten und der obligatorisch plus optional möglichen Instanzenanzahl
THEN	Setze die Analyse dieses Teilbaumes fort <ol style="list-style-type: none"> 1. Verbinde die neu entstandenen Instanzen mit dem bestehenden Instanzenbaum 2. Rufe für alle entstandenen Instanzen die Funktion <i>childTopDown</i> auf
ELSE	Die Analyse des Teilbaumes wird für die aktuelle Hypothese verworfen. (Status = TRASH)

Abbildung 5.9: Struktogramm der Funktion *execTopDown*

standene Instanzennetz datengetrieben von unten nach oben. Es können hierbei zusätzliche Instanzen entstehen, falls die Unterinstanzen wie in Abbildung 5.6 dargestellt mehrere disjunkte Gruppen bilden, die den Gruppenbeschreibungen der aktuellen Instanz genügen. Auch diese Bearbeitung geschieht für alle Instanzen einer Ebene parallel.

5.6.4 Holistische Operatoren

Wie bereits angesprochen, können Bildverarbeitungsoperatoren in GeoAIDA auch als Top-Down-Operatoren an Konzepten angebunden werden, die keine Blattknoten darstellen. Diese Operatoren werden als holistische Operatoren bezeichnet, da sie durch einen ganzheitlichen, nicht expliziten Ansatz die Existenz eines Objektes vom Typ eines bestimmten Konzept anzeigen. Ein holistische Operator ist somit im Gegensatz zu einer strukturellen Erkennung in der Lage, ein Objekt bzw. eine Instanz als einheitliches Ganzes zu erkennen, ohne dass

Funktion: <i>execBottomUp</i>	
IF	der Status einer Instanz ist <i>BU</i> \wedge für alle unter dieser Instanz befindlichen Instanzen wurde der Bottom-Up-Operator bereits ausgeführt
THEN	rufe den Bottom-Up-Operator auf, der im zur Instanz gehörigen Konzept durch den Benutzer eingetragen wurde.

Abbildung 5.10: Struktogramm der Funktion *execBottomUp*

die Bestandteile explizit als Instanzen erzeugt und hierfür die Substrukturen im Semantischen Netz angelegt werden müssen. Jeder Blattknoten des Konzeptnetzes muss einen Bildverarbeitungsoperator besitzen, der hier ebenfalls als holistischer Operator bezeichnet wird, da er Objekte dieser Instanz liefert, ohne deren Substrukturen als Instanzen zu erzeugen.

Eine weitere strukturelle Analyse anhand der Bestandteile ist für Objekte mit holistischem Operator folglich nicht zwingend notwendig. Daher kann die Analyse von Substrukturen bei Bedarf durch das Setzen eines Eintrages an dieser Stelle unterbunden werden. Hierzu existiert im Konzeptnetz die Möglichkeit, in den entsprechenden Konzepten die weitere Analyse der Teilstrukturen zu unterbinden. Dieses ist jedoch nur sinnvoll, falls der verwendete holistische Operator sichere Ergebnisse liefert. Um die Analyse robuster und fehlerfreier zu realisieren, können holistisch erkannte Objekte zusätzlich strukturell untersucht werden, wodurch die Sicherheit einer Aussage zunimmt, falls beide Verfahren das Objekt bestätigen.

Der Einsatz von holistische Verfahren kann weiterhin zu einer Effizienzsteigerung der Analyse eingesetzt werden. Diese wird dadurch erreicht, dass der holistische Operator nicht mehr alle Hypothesen erzeugt, die dem generischen Semantischen Netz innewohnen, sondern nur die, die einen Hinweis darauf liefern, dass das gesuchte Objekt in der zu untersuchenden Region vorliegt.

Ein Beispiel für den Einsatz eines holistischen Operators soll anhand des in Abbildung 5.11 dargestellten Semantischen Netzes erfolgen. Der linke Teil der Abbildung zeigt ein Semantisches Netz, in dem die mit einem holistischen Operator versehenen Instanzen mit einem Kreis dargestellt sind. Auf der rechten Seite der Abbildung sind zwei Bilder dargestellt, die zwei Interpretationen zu unter-

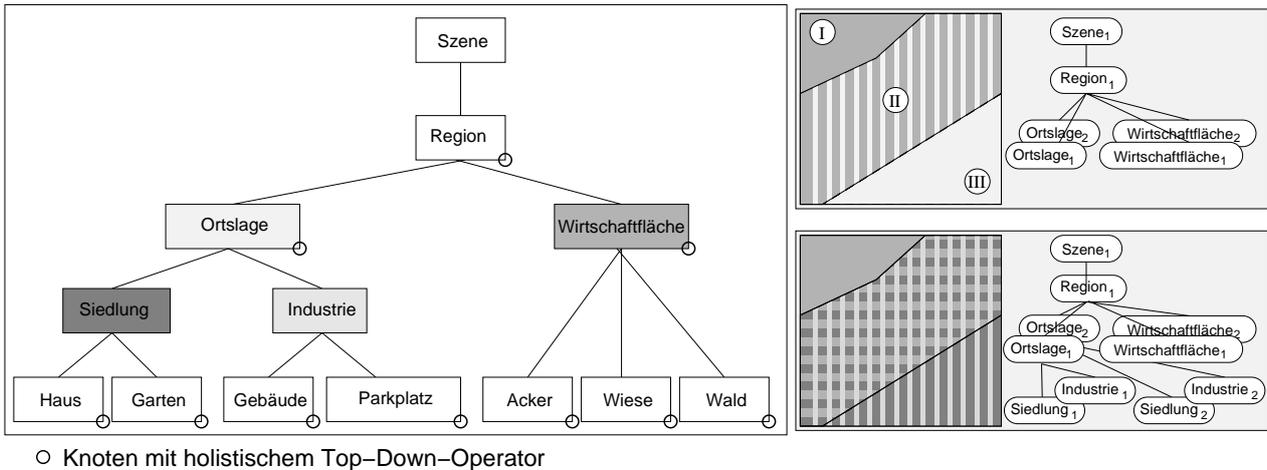


Abbildung 5.11: Beispiel für ein Konzeptnetz mit holistischen Top-Down-Operatoren

schiedlichen Zeitpunkten innerhalb des Analyseablaufes mit den dazugehörigen Instanznetzwerken darstellen.

Nachdem die Analyse gestartet wurde und die *Szene* mit den Eingangsdaten initialisiert wurde, teilt der holistische Operator vom Konzept *Region* den gesamten Bereich – beispielsweise durch die Verwendung der Straßenverläufe aus einem GIS – in drei Teilregionen *I*, *II* und *III* auf. Diese drei Teilregionen werden an die holistischen Operatoren von *Ortslage* und *Wirtschaftsfläche* übergeben. Die Fläche *I* wird nur der Klasse *Wirtschaftsfläche*, die Fläche *III* nur der Klasse *Ortslage* und die Fläche *III* wird als Hypothese beiden Klassen zugeordnet. Es entstehen somit zwei hypothetische Instanzen für die Klassen *Ortslage* und *Wirtschaftsfläche*. Die Instanzen der Klasse *Ortslage* werden im nächsten Schritt den konkreteren Begriffen *Siedlung* *Industrie* zugeordnet. Da in diesem Schritt keine holistischen Operatoren vorliegen, werden für jede hypothetische Instanz *Ortslage* eine hypothetische Instanz der Klasse *Siedlung* und *Industrie* erzeugt. Die nun vorliegende Situation ist im rechten unteren Teil der Abbildung 5.11 dargestellt. Die alternativen Interpretationen einer Region sind durch die schraffierte Überlagerung der Farben der Konzepte aus dem Semantischen Netz dargestellt.

Der Gewinn durch den Einsatz der holistischen Operatoren ist dort zu erkennen, wo die möglichen Hypothesen für eine Region eingeschränkt wurden. Dieses ist

für die Regionen *I* und *III* der Fall, da Region *I* nicht auf das Vorliegen einer *Ortslage* untersucht werden muss. Für Region *III* gilt dieses entsprechend für die Klasse *Wirtschaftsfläche*.

5.7 Zusammenfassung der Eigenschaften von GeoAIDA

GeoAIDA ist ein System, das speziell für die Interpretation von Fernerkundungsdaten entwickelt wurde und die Integration von Zusatzwissen in Form eines Semantischen Netzes ermöglicht.

Die Grundlage bilden spezielle Bildverarbeitungsoperatoren, die es ermöglichen, einzelne Objekttypen aus den vorliegenden Fernerkundungsdaten zu extrahieren. Die Ergebnisse dieser Bildverarbeitungsoperatoren können unter Berücksichtigung des Kontextes, in dem ein Objekt erkannt wurde, und den Nachbarschaftsbeziehungen zu anderen Objekten verbessert werden. Zur Einbringung dieses strukturellen Wissens wurde ein System entwickelt, das es ermöglicht, Wissen in die Ablaufsteuerung, die Initialisierung der Bildverarbeitungsoperatoren und die Behandlung der Ergebnisse unter Berücksichtigung von Objektrelationen einzubringen.

Als besondere Eigenschaften einer solchen semantischen Netzwerksprache ist die Modularität, Wohlstrukturiertheit und Kompaktheit zu erwähnen. Hierbei werden Begriffe in GeoAIDA als Konzeptdefinitionen in einem generischen Semantischen Netz modelliert. Während der Analyse der Daten entstehen aus diesen Konzepten Instanzen, die mit Bildbereichen der Eingangsdaten korrespondieren. Hierbei arbeitet das System streng hierarchisch, sodass die Region einer Instanz nur als Teilmenge in einer Subinstanz auftreten kann.

Instanziinterne Merkmale werden in GeoAIDA als Attribute realisiert, die durch Relationen in Beziehung zueinander gesetzt werden können. Die Bewertung der Sicherheit einer Interpretation kann durch eigene Bewertungsfunktionen erfolgen. Durch die Definition einer festen Grundstruktur der Konzepte und einer definierten Schnittstelle der Kantensemantik ist GeoAIDA in der Lage, eine

problemunabhängige Analysestrategie bereit zu stellen, wodurch eine wesentliche Schwäche von allgemeinen semantischen Netzwerken behoben wird. Die Verarbeitungsstrategie basiert auf einer modellbasierten Top-Down- und einer datengetriebenen Bottom-Up-Analyse. Die modulare und explizite Wissensrepräsentation einerseits und die flexible Anbindung von Bildverarbeitungsoperatoren andererseits machen GeoAIDA zu einem Werkzeug zur automatischen Interpretation von Fernerkundungsdaten mit den folgenden Eigenschaften:

- Übersichtliche Strukturierung des zu bearbeitenden Weltausschnittes bzw. Problems durch die Verwendung eines semantischen Netzwerkformalismus.
- Klare Ablaufsteuerung, die auch bei komplexen Problemen eine klare Bearbeitungsstrategie bereit stellt.
- Einbringen und Nutzen von Vorwissen in Form von:
 - Relationen der Objekte zueinander,
 - Integration von extern vorhandenem Wissen (z. B. GIS, Planungsdaten oder ehemalige Interpretationen).
- Der Bezug der Objekte zu ihren Geokoordinaten ist durchgängig gegeben.
- Umfangreiche Ergebnisdarstellung:
 - Bildhafte Darstellung der Ergebnisse in einer hierarchischen Kartenform,
 - strukturelle Darstellung der Ergebnisse inklusive der Objektattribute,
 - Zuordnung der strukturellen und bildhaften Ergebnisse zueinander.

5.7.1 Systemvergleich

Hier werden noch einige Designaspekte angesprochen und ein Vergleich des Systems GeoAIDA mit den Ansätzen in ERNEST und AIDA aufgezeigt. Diese

beiden Systeme wurden gewählt, da sie dem Realisierten vom Ansatz her am ähnlichsten sind.

Die Modellierung der Semantischen Netze in GeoAIDA geschieht nicht in Form eines allgemeinen Netzes, sondern als Baumstruktur. Diese Darstellung wurde gewählt, da sie eine übersichtliche Wissensrepräsentation ermöglicht und den hierarchischen Analyseansatz besser widerspiegelt. Der Unterschied zu anderen Realisierungen (ERNEST, AIDA), dass gleiche Konzepte in unterschiedlichen Pfaden nun nicht nur einmal sondern mehrmals vorliegen, ist für den Anwender nicht nachteilig zu bewerten, da die Modellierung des Konzeptes an einer anderen Stelle im Konzeptnetz durch eine Kopie erzeugt werden kann. Für den Analyseablauf ist der Unterschied in der Darstellung irrelevant, da zum Zeitpunkt der Instantiierung in beiden Varianten ein neues Objekt angelegt wird und sich für die Analyse kein Unterschied ergeben würde, ob Instanzen desselben Types einen gemeinsamen Ursprung besitzen oder nicht.

Im Gegensatz zu vielen anderen Realisierungen von Semantischen Netzen [59], [61] ist in GeoAIDA die Semantik der im Netz verwendeten Kanten nicht explizit vorgegeben. Es ist jedoch, bedingt durch die behandelte Fragestellung und dem daraus resultierenden hierarchischen Vorgehen, eine Mischung aus Konkretisierungs- und Bestandteils-Kante. Ihre konkrete, semantische Funktion kann jedoch vom Benutzer bestimmt werden.

Die Unterschiede zwischen GeoAIDA und dem Vorgängermodell AIDA lassen sich folgendermaßen zusammenfassen:

- In GeoAIDA existiert keine Regelbasis mit benutzerdefinierten Inferenzregeln, die während der Analyse in der durch den Benutzer angegebenen Situation angewendet werden. Die Analyse in GeoAIDA besitzt eine feste Struktur. Die objektspezifischen Funktionalitäten sind über den Top-Down- und Bottom-Up-Operator angebunden.
- In AIDA existieren kein holistischen Operatoren.
- Die holistischen Operatoren in GeoAIDA ermöglichen eine Datenbindung nicht erst (wie in AIDA) in den untersten Konzepten (den Blättern), sondern in jeder Instanz. Die Ebene, in der sich das zur Instanz gehörende

Konzept befindet, ist hierfür nicht relevant.

Der wichtigste Unterschied zwischen GeoAIDA und ERNEST ist, dass in ERNEST die Analyse als Suchproblem realisiert ist, während sie in GeoAIDA als lokales Entscheidungsproblem umgesetzt wurde. Der Suchbaum in ERNEST hält in seinen Suchbaumknoten in sich konsistente Interpretationen vor, wobei die Blätter des Baumes konkurrierende Beschreibungen darstellen, die über eine Bewertung miteinander verglichen werden.

6 Interpretationsschritte

Mache die Dinge so einfach wie möglich – aber nicht einfacher. *Albert Einstein*

Die Funktionalität des Systems soll hier anhand repräsentativer Analyseschritte aus dem Bereich der Landnutzung demonstriert werden. Grundlage des in der Wissensbasis realisierten Semantischen Netzes waren die im ATKIS vorliegenden Objekthierarchien, von denen hier wichtige Teilhierarchien umgesetzt wurden. Die einzelnen Teilaspekte werden der Übersicht wegen separat demonstriert. Das den meisten Interpretationen zugrunde liegende Semantische Netz ist in Abbildung 6.1 zu sehen, wobei für die Darstellung einiger Analyseschritte nur Ausschnitte oder leicht modifizierte Varianten des Netzes benötigt werden.

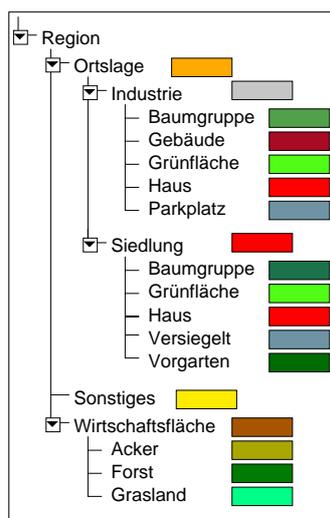


Abbildung 6.1: Semantisches Netz zur Bestimmung der Landnutzung

Für die initiale Aufteilung des gesamten Bildes werden hier die Straßenverläufe verwendet, da diese in vielen Bereichen Grenzen zwischen verschiedenen Gebietsklassen bilden. Die Straßenverläufe können einem GIS entstammen oder durch eine modellbasierte Straßenextraktion (siehe Kap. 3.1) aus den Fernerkundungsdaten gewonnen werden. Es ist auch möglich, weitere Operatoren, beispielsweise eine Varianzanalyse, zusätzlich zur Aufteilung der gesamten Region in Teilregionen zu verwenden.



Abbildung 6.2: Szene Koblenz

6.1 Analyseablauf

Das Analyseprinzip soll anhand einer einfachen Teilregion motiviert werden. In Abbildung 6.3 sind für eine Region einige während der Interpretation vorliegende Repräsentationen abgebildet. Die Analyse beginnt modellbasiert mit der

Top-Down-Analyse.

Im Bild 6.3.a ist die Entstehung der Teilregion durch die begrenzenden Straßen (schwarzen Linien) dargestellt. Die Aufteilung einer Szene in Regionen unter der Verwendung von Straßen liefert eine gute initiale Unterteilung, da Straßen häufig identisch mit den Grenzen zwischen unterschiedlichen Landnutzungs-klassen sind. Für die hier betrachtete zentrale Region wird von der Analyse die Hypothese generiert, dass diese Teilregion eine *Ortslage* repräsentiert. Die Hypothese für eine Instanz vom Typ *Wirtschaftsfläche* ist vom holistischen Top-Down-Operator aufgrund der Ergebnisse einer Varianzuntersuchung der Bild-daten dieser Teilregion auszuschließen.

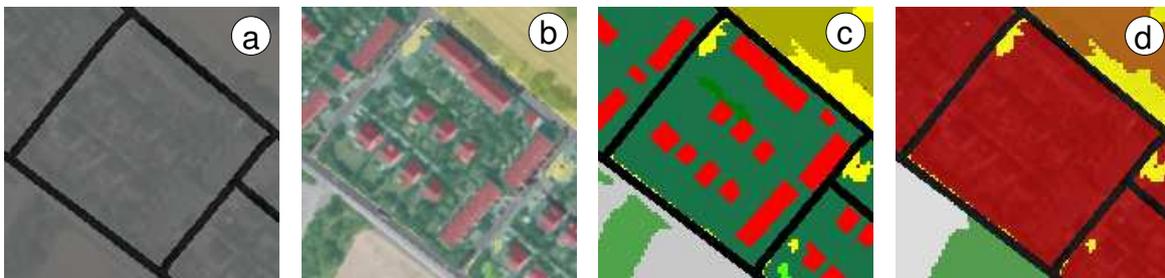


Abbildung 6.3: Interpretationsschritte einer Teilregion

Im darauffolgenden Schritt entstehen die Hypothesen *Siedlung* und *Industrie* für diese Region gemäß dem Semantischen Netz in Abbildung 6.1. Im Bild 6.3.b ist die visuelle Luftbildszene, in der die erkannten Objekte für die Hypothese *Siedlung* eingeblendet sind, zu erkennen. Diese Objekte sind im darauffolgenden Labelbild 6.3.c noch einmal dargestellt, wobei jedem Objekttyp ein eigenes Label (Farbe) zugewiesen ist.

Mit dem Erreichen der untersten Ebene im Semantischen Netz beginnt die datengetriebene Analyse (Bottom-Up) der entstandenen Hypothesen. Die Hypothese *Siedlung* kann als komplette Instanz für diese Teilregion durch die vorhandenen Bestandteile bestätigt werden. Es ergibt sich als Resultat der in Bild 6.3.d in rot dargestellte und von den Straßen eingeschlossene Bereich. Die Gruppierung und Bewertung geschieht mit dem in Kapitel 6.2 beschriebenen Operator. Hierzu wird für die Häuser bezüglich ihrer Größe und Rechteckigkeit ein Bewertungsmaß für jedes Haus bestimmt. Der Mittelwert dieser Bewertungen wird

dann mit einem Faktor gewichtet, der die bebaute Fläche in Relation zur Gesamtfläche der Region setzt. Die Bewertung der Region für die Annahme, dass es sich um eine *Siedlungsfläche* handelt, ist hoch, da relativ viele Häuser eng beieinander stehen. Die Alternative, diese Region als *Industrie* zu interpretieren, erhält die Bewertung 0, da keine Objekte vom Typ *Gebäude* gefunden werden, die obligatorischer Bestandteil einer *Industrie* sind. Der Unterschied zwischen den Konzepten *Haus* und *Gebäude* liegt zum einen in der Grundfläche, die bei Gebäuden größer sein kann und zum anderen in der Formgebung. Ein Wohnhaus ist im Normalfall so gebaut, dass kein Punkt innerhalb des Hauses eine vorgegebene Entfernung zur Außenwand nicht überschreitet, um alle Stellen des Hauses mit Sonnenlicht beleuchten zu können. Industriegebäude hingegen sind oft künstlich oder von oben beleuchtet und besitzen daher diese Einschränkung nicht.

Durch die oben gezeigte Aufteilung der gesamten Szene in Teilregionen sinkt auch die Anzahl der darin auftretenden Objekte, wodurch der Interpretationsprozess robuster wird, da bei der anschließenden Gruppierung nur die innerhalb der Region vorliegenden Hypothesen zu beachten sind. Die Reduktion der kombinatorischen Komplexität hat weiterhin den Effekt, dass die Rechenzeit für die Interpretation einer Teilregion sinkt. Durch die Möglichkeit, die vorliegenden Teilregionen parallel abzuarbeiten, wird die Analyse zusätzlich beschleunigt.

6.2 Generischer Bottom-Up-Operator

Die Aufgabe der Bottom-Up-Operatoren ist die Zusammenfassung und Gruppierung der als Knoten in einer Liste vorliegenden hypothetischen Instanzen und die Bewertung der neu entstandenen Kompositionen. Da dieser Vorgang im Allgemeinen ein jeweils für die aktuelle Aufgabe spezieller Vorgang ist, sind hierfür auch grundsätzlich spezielle Operatoren notwendig. Die vorliegende Aufgabe ist jedoch hauptsächlich ein mengentheoretisches Problem und kann durch Listenoperationen und einigen spezielle Funktionen allgemein beschrieben werden.

Um dem Anwender eine einfache Möglichkeit zu bieten, diese Gruppierung oh-

ne die Kenntnis spezieller Programmierkenntnisse zu ermöglichen, wurde eine funktionale Beschreibungssprache realisiert, mit deren Hilfe Operationen auf Listen von hypothetischen Instanzen und Funktionen auf den Hypothesen ausgeführt werden können. Die Funktionen berechnen Attribute der Hypothesen, die aus den Bilddaten bestimmt werden können. Hierzu gehören eine Reihe von Formparametern, Flächenparameter, Positionsangaben der zugeordneten Region im Bild, Geokoordinaten der Region und deren Flächenmomente. Zusätzlich benötigte Funktionen können den Bestehenden bei Bedarf hinzugefügt werden. Weiterhin existieren eine Reihe von arithmetischen Funktionen, um mit den vorhandenen Attributwerten neue Attribute zu bestimmen. Vergleichende Funktionen auf Attributwerten innerhalb oder zwischen Hypothesen liefern boolesche Ergebnisse, die mit logischen Funktionen weiter verarbeitet werden können. Neu berechnete Attribute können den Hypothesen oder den Hypothesenlisten als Wert zugewiesen werden.

Das Selektieren von einzelnen Objekten, Objektklassen oder Teillisten anhand von Attributwerten zu neuen Listen sowie das Verbinden von Listen ist möglich. Auf diesen neu entstandenen Listen können die oben genannten Funktionen ebenfalls angewandt werden. Die berechneten Ergebnisse liefern Instanzen und lassen sich in dem in GeoAIDA verwendeten XML-Format abspeichern. Das Ergebnis umfasst neben den Beschreibungen der Instanzen auch die bildlichen Repräsentationen in Form von Labelbildern.

Die Formulierung einer Gruppierungsanweisung geschieht in einer UP-Notation (Umgekehrt Polnische Notation oder auch Postfix Notation - Waclaw Sierpinski) unter Verwendung eines Stackmechanismus.

Der Ausdruck:

Knotenliste "Haus" selectClass

liefert beispielsweise eine Teilliste von *Knotenliste*, die nur noch Objekte vom Typ *Haus* enthält. Die einzelnen Anweisungen werden auf den Stack gelegt. Ist dieses eine Funktion, so erwartet diese ihre Argumente auf dem Stack, entfernt diese und schreibt ihr Ergebnis wieder auf den Stack.

Jeder Knoten repräsentiert eine Instanz und besitzt einen eigenen *lokalen Stack*, auf dem Werte gespeichert werden können. Dieser lokale Stack findet Anwendung, wenn für die Knoten spezifische Werte gespeichert oder weitere Attribute

berechnet werden sollen. Der Wert eines neu berechneten Attributes liegt nach der Funktionsanwendung auf diesem lokalen Stack.

Der lokale Stack wird bei dem folgenden Aufruf verwendet:

```
Knotenliste "Gebäude" selectClass
"area" calc 675 > select
"orthogonality" calc "orthogonality" set
```

In der ersten Zeile wird wie in dem ersten Beispiel eine Teilliste gebildet. Hier werden alle Objekte vom Typ *Gebäude* extrahiert. Die Zeichenkette "area" ist ein Parameter für die Funktion *calc*. Diese berechnet für alle *Gebäude* die Grundfläche in m^2 , die auf dem lokalen Stack der Knoten vom Typ *Gebäude* gespeichert wird. Die Zahl 675 wird von der Funktion > für den Vergleich mit dem obersten lokalen Stackelement verwendet, das hierbei vom Stack entfernt wird. Das boolesche Resultat des Vergleichs liegt als Ergebnis wiederum auf dem lokalen Stack. Die Funktion *select* bildet die Teilmenge der *Gebäude*, deren Stackeintrag eine 1 enthält, also eine Grundfläche von mehr als $675m^2$ hatten. In der letzten Zeile wird die *Rechtwinkligkeit* (engl. orthogonality) berechnet und den Instanzen durch die Funktion *set* als zusätzliches Attribut übergeben. Dieses Attribut kann für weitere Berechnungen verwendet werden und ist Bestandteil der strukturierten Ergebnisbeschreibung der Gesamtanalyse.

6.3 Behandlung von Alternativen

In Abbildung 6.4 sind für eine Region die während der Analyse erkannten Objekte zweier alternativer Interpretationen für den bebauten Teil der Abbildung 6.5 dargestellt. Die Legende der farblichen Ergebnisdarstellung ist dem Semantischen Netz in der Abbildung 6.1 zu entnehmen.

Zu erkennen sind auf der linken Seite die Objekte, die zur Interpretation *Industrie* und auf der rechten Seite, die zur Interpretation *Siedlung* gehören. Die Bestimmung der Bewertung soll für die Hypothese *Industrie* motiviert werden. Die als *Gebäude* erkannten Bereiche besitzen Attribute, die eine gute Bewertung für diese Objekte zulassen. So besitzen die Flächen der einzelnen *Gebäude* einen Wert zwischen $120m^2$ und $460m^2$, die Orthogonalität, die die Rechtwink-

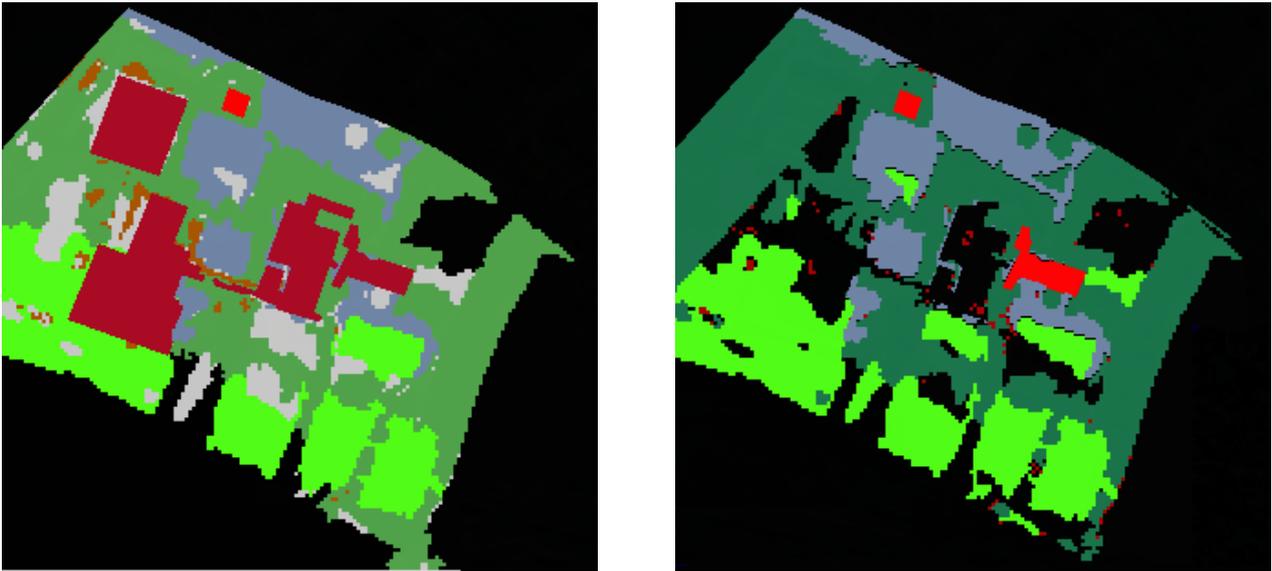


Abbildung 6.4: Alternative Interpretationen einer Teilregion vom Typ *Ortslage* aus Abbildung 6.5. Legende siehe Abbildung 6.1

ligkeit der Objektkanten beschreibt, liegt zwischen 0,82 und 0,97. Die weiteren Objekte, wie *Parkplätze* und *Baumgruppen*, besitzen Werte innerhalb der erwarteten Größen. Bei der Bestimmung der Bewertung für die betrachtete Region gehen die Einzelbewertungen der Objekte und das Verhältnis der Gesamtfläche zu der mit Gebäuden bebauten Fläche ein. Für die Interpretation der Region als *Siedlung* wird ein ähnliches Verfahren verwendet, wobei jedoch die verwendeten Attribute und Objekttypen an die Klasse *Siedlung* angepasst sind. Es ergibt sich eine Bewertung für die Interpretation *Industrie* von 0,92, für die Interpretation *Siedlung* nur von 0,42. Die Analyse entscheidet sich aufgrund der Bewertung für die Interpretation als *Industrie*. Die Vereinigungsmenge aller Flächen der farblich dargestellten Teilobjekte bilden gemeinsam den Bereich der Klasse *Industrie*.

6.4 Datengetriebene Aufspaltung von Regionen

Das zweite Beispiel setzt die im vorangegangenen Beispiel beschriebene Entscheidung voraus und soll verdeutlichen, wie eine Region durch die weitere strukturelle Analyse in mehrere alternative, disjunkte Klassen aufgeteilt wer-

den kann. Die Abbildung 6.5 zeigt eine Region aus der Abbildung 6.2, die durch die Verwendung der Straßenverläufe zur Aufteilung der Szene entstanden ist.



Abbildung 6.5: Durch Straßen (gelb) begrenzte Teilregion aus Abbildung 6.2

Die möglichen alternativen Interpretationen für den bebauten Bereich – der einen Teilbereich der gesamten Region darstellt – wurden bereits in Kapitel 6.3 unter Verwendung einer Bewertung dargestellt (siehe Abbildung 6.4) und sollen hier nicht mehr behandelt werden. Zu beachten ist jedoch, dass bei einer Interpretation der gesamten Region die möglichen Alternativen berücksichtigt werden und die Entscheidung als auch die nun folgende Aufspaltung in einem Analyseablauf stattfinden. Die beiden Entscheidungen sind Bestandteil der Bottom-Up-Analyse, finden jedoch an unterschiedlichen Stellen des in Abbildung 6.1 dargestellten Netzes statt. Die Bewertung und die Auswahl zwischen den Alternativen *Industrie* und *Siedlung* geschieht in der darüberliegenden Instanz *Ortslage*, und die Aufspaltung der Region in zwei disjunkte Teilregionen in der Instanz *Region*, die ursprünglich durch den Verlauf der Straßen entstanden war.

Zum Industriegebiet gehören neben den Gebäuden die Objekte, die diesen benachbart sind und den im Semantischen Netz beschriebenen Klassen und deren Kriterien genügen. Durch die im Semantischen Netz enthaltenen Restriktionen entfallen die als *Forst_1* und *Acker_1* (Abbildung 6.6) gekennzeichneten Flächen

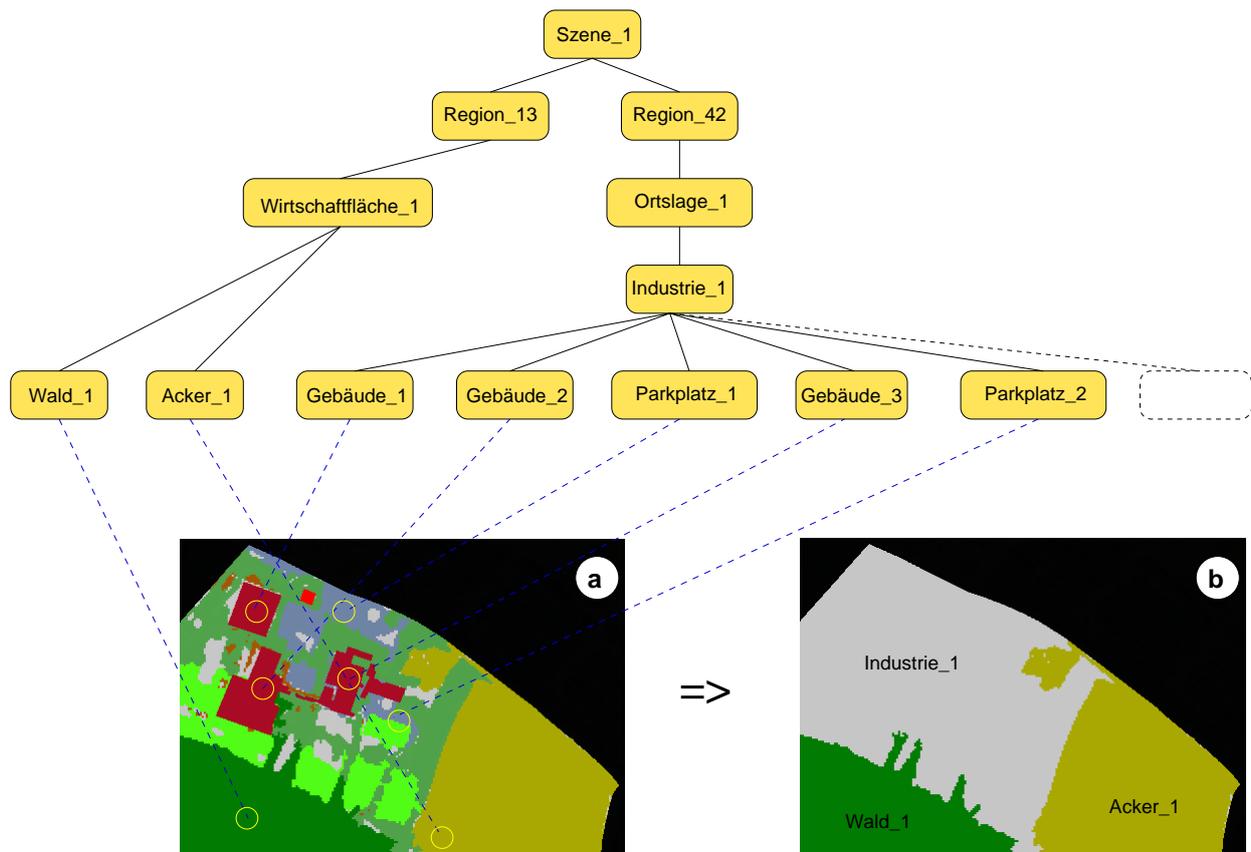


Abbildung 6.6: Instanzennetz der Interpretation für die in Abbildung 6.5 dargestellte Region

als Bestandteil der Instanz *Industrie_1*. Diese, im Kapitel 6.3 nicht behandelten und in Abbildung 6.6 als *Wald_1* und *Acker_1* bezeichneten Bereiche, werden von der Instanz *Industrie_1* nicht als ihr zugehörige Objekte verifiziert. Die Bewertung dieser Teilflächen ist in der Alternative, diese Bereiche als Bestandteil der *Wirtschaftsfläche_1* zu interpretieren, am größten. Die *Wirtschaftsfläche_1* unterteilt sich in *Acker_1* und *Wald_1*. Dieser Zusammenhang ist im Instanzennetz der Abbildung 6.6 durch die Verbindungslinien zwischen den Instanzen und den farblich markierten Teilregionen zum Bild 6.6.a veranschaulicht. Das Bild 6.6.b zeigt das Ergebnis der Interpretation zusammenfassend für die Klassen *Industrie*, *Acker* und *Wald*.

6.5 Topologische Beziehungen

Hier soll die Nutzung topologischer Beziehungen von Objekten zu deren Gruppierung aufgezeigt werden. Begriffe wie beispielsweise *Forst* oder auch *Siedlung* deuten implizit an, dass die diese Begriffe bildenden Objekte gewissen räumlichen Relationen genügen. Bei den genannten Beispielen besitzen die Häuser oder Bäume Abstände, die unterhalb einer angebbaren Distanz liegen. In Abbildung 6.7 ist beispielhaft das Ergebnis einer Extraktion von Bäumen aus einem Laserscanbild abgebildet. Die Region *Forst_1* ist in hellgrün ange deutet und umfasst alle Einzelbäume (*Einzelbaum_1* bis *Einzelbaum_n*). Diese Interpretation ist im *Instanzenetz 1* in Abbildung 6.7 dargestellt.

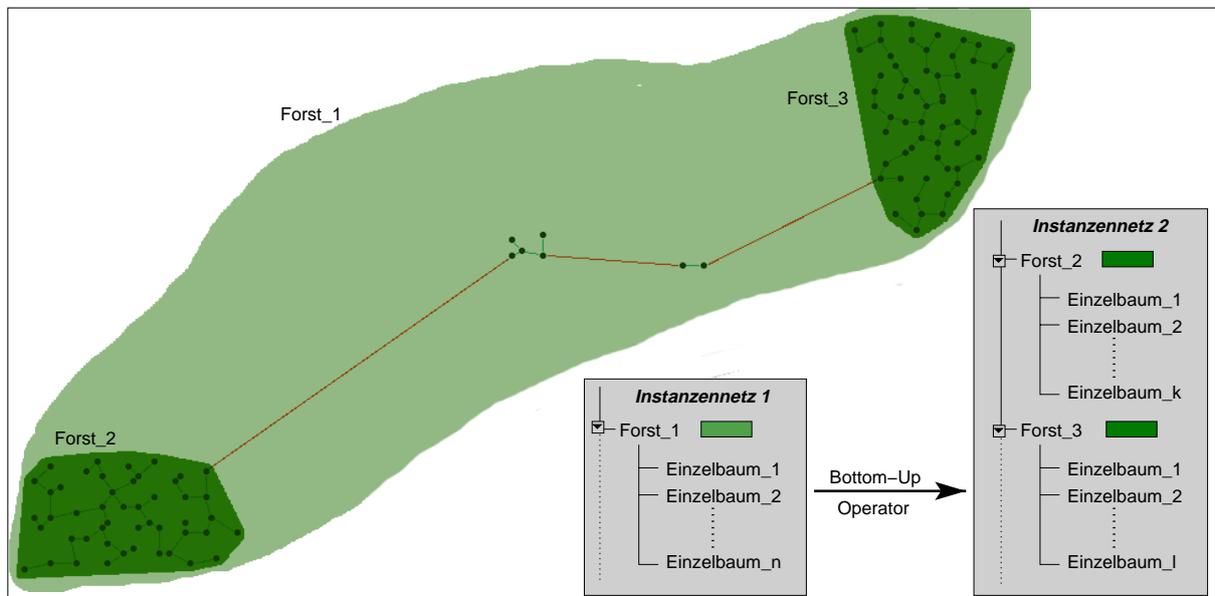


Abbildung 6.7: Minimum-Spanning-Tree über eine Menge von detektierten Bäumen

Zusätzlich zu den Zentren der erkannten Bäume ist der Graph eingezeichnet, der alle Einzelbäume so miteinander verbindet, dass die Summe der Distanzen aller Verbindungen minimal wird (Minimum-Spanning-Tree). Durch Auftrennen des Graphen an den rot eingezeichneten Verbindungslinien – sie überschreiten einen vorgegebenen Maximalabstand – entstehen vier Einzelgruppen. Da die Flächen der beiden mittleren Baumgruppen zu gering sind, um diese als *Forst* zu interpretieren, entfallen diese Bäume für die weiteren Betrachtungen.

Die beschriebene und ähnliche Funktionen sind in GeoAIDA im Bottom-Up-Operator (siehe Kapitel 6.2) realisiert und können so auf einfachem Wege eine Menge von Objekten (hier *Bäume*) unterteilen, der auch Instanzen unterschiedlichen Typs angehören können. Die Erzeugung von Teilmengen ist gleichbedeutend mit der Unterteilung einer Region in Teilregionen, die getrennte Instanzen des übergeordneten Konzeptes etablieren. Das Ergebnis des Bottom-Up-Operator ist im *Instanzenetz 2* in Abbildung 6.7 dargestellt. Die Instanzen *Forst_2* und *Forst_3* entsprechen den dunkelgrünen Bereichen der Abbildung 6.7. Der Bereich der Instanz *Forst_1* mit den n Einzelbäumen wurde durch den unten stehenden Aufruf in die Teilregionen *Forst_2* und *Forst_3* aufgeteilt, die jeweils disjunkte Teilmengen der n Einzelbäume bilden.

Für das hier gezeigte Beispiel lautet der Aufruf des generischen Bottom-Up-Operators folgendermaßen:

```
Knotenliste "Einzelbäume" selectClass 50 maxDist "area" calc 500  
> select
```

Hierbei gibt die 50 das Kriterium für das Aufspalten des Minimum-Spanning-Tree an. Der Minimum-Spanning-Tree wird durch die Funktion *maxDist* aufgespannt, dessen Objekte entsprechend gruppiert und das Ergebnis als Stack mit den entstandenen Teillisten zurückgegeben wird. Dieses Vorgehen kann ebenso verwendet werden, um getrennt stehende Siedlungen u.ä. in einer Region zu gruppieren.

6.6 Spezialisierung von Objekten

Die Spezialisierung von Objekten kann in GeoAIDA auf zwei Wegen realisiert werden. Zum einen kann im Semantischen Netz ein neues Konzept eingeführt werden, das spezieller als ein bereits vorhandenes ist, zum anderen kann ein Attribut in einer Instanz anzeigen, das diese Instanz spezieller ist als ihre allgemeine Objektbezeichnung. Die Ansätze sollen an einem Beispiel, dessen Bildbereich in Abbildung 6.8 dargestellt ist, demonstriert werden.

Die Eingangsdaten für dieses Beispiel umfassen ein visuelles Bild und ein IR-Bild mit einer Auflösung von 0.25cm Kantenlänge der Pixel. Weiterhin liegt ein



Abbildung 6.8: Szene Ravensburg, multisensorielle Daten (VIS, IR, Laserscan)

Laserscandatenatz mit $1m$ Auflösung vor, in dem das erste und letzte Signal des reflektierten Laserimpulses (first-pulse, last-pulse) unterschieden werden kann.

Aus den Eingangsdaten lassen sich die in Abbildung 6.9 dargestellten Objekte extrahieren. Es sind hier nur die Objekte abgebildet, die im Ergebnis enthalten sind. Die möglichen alternativen Objekte werden hier nicht betrachtet, da sie für die hier vorliegende Darstellung nicht relevant sind. Die bei der Objektextraktion verwendeten Eigenschaften und das Ergebnis sollen kurz erläutert werden.

Die in schwarz eingezeichneten Straßen wurden mit dem in Kapitel 3.1 beschriebenen Verfahren gewonnen und dienen auch hier zur initialen Aufteilung der Szene in Teilregionen. Die Gebäude wurden anhand der Laserscandaten bestimmt [77]. Es geht die Steilheit der im Höhenbild enthaltenen Wände, die Größe der Grundfläche, die Höhe und die Form der Objekte in die Klassifizierung ein. Die als Gebäude erkannten Objekte sind durch die rot markierten Bereiche symbolisch angezeigt. In diesem Fall soll keine Unterscheidung zwi-

schen Häusern und Gebäuden gemacht werden, da es hier nicht von Belang ist. Die Wasserbereiche bilden sich im IR-Bild als dunkle Bereiche ab und sind in den Laserscandaten als Fehlstellen markiert. Da das vom Laser ausgesandte Signal von den Wasseroberflächen absorbiert wird, liegt an diesen Stellen kein Messwert vor, was eine Evidenz für Wasserflächen liefert. *Klärbecken* wurden als vereinzelte Wasserflächen modelliert, die eine runde Form aufweisen und deren Radius zwischen 10m und 30m liegt.

Die Klasse *Versiegelt* sind Gebiete, die im IR-Bild keinen Hinweis auf Vegetation liefern und repräsentieren Asphaltwege oder Parkplätze.



Abbildung 6.9: Ergebnisse der Objekterkennung für die in Abbildung 6.8 dargestellten Daten

Die Vegetationsbereiche lassen sich aus dem IR-Bild und dem visuellen Bild bestimmen. Die Klasse *Baumgruppe* lässt sich von den *Grünflächen* durch die Verwendung der Laserscandaten unterscheiden. Da der Laserstrahl im Vegetationsbereich nur teilweise von den höherliegenden Blättern reflektiert wird, gibt es in diesen Bereichen eine hohe Varianz der Einzelwerte bei der Entfernungsmessung, da auch Teile des Laserstrahls an tieferliegenden Vegetationsteilen oder erst auf Bodenniveau reflektiert werden. Dieses "Rauschen" in der Entfernungsmessung kann nun dazu verwendet werden, die kürzeste (die Baumkrone) und die längste Entfernung (der Waldboden) über die Laufzeitdifferenzen zu ermitteln. Vegetationsbereiche mit hohen Objekten (z. B. Bäume) sind an großen lokalen Differenzen in den Laserscandaten zu erkennen. Über diese Differenz ist

die Unterscheidung der Klassen *Grünfläche* und *Baumgruppe* realisiert. Baumgruppen unterscheiden sich von der Klasse *Forst* durch ihre Form und Fläche.

Die in Gelb dargestellten Bereiche konnten keiner Klasse sicher zugeordnet werden. Besonders auffällig ist hier der zum Klärwerk gehörenden Bereich, in dem nur eine Wasserfläche sicher erkannt wurde. Für die in dieser Region vorliegenden Objekte gab es in der Modellierung keine adäquaten Entsprechungen.

6.6.1 Spezialisierung durch neue Konzeptdefinition

Die Erweiterung des Konzeptnetzes aus Abbildung 6.1 um den Eintrag *Klärwerk*, wie in Abbildung 6.10 dargestellt, ist eine explizite Möglichkeit, Spezialisierungen im Semantischen Netz einzuführen. Ziel ist es, das mit der bisherigen Wissensbasis als Industriegebiet erkannte Gebiet spezieller zu beschreiben und als Klärwerk zu interpretieren. Die Erweiterung besteht aus dem Konzept *Klärwerk* und den dazugehörigen Bestandteilen. Die Existenz von Klärbecken ist für die Instantiierung einer Region als *Klärwerk* notwendig. Daher sind die *Klärbecken*, von denen mindestens zwei und die *Gebäude*, von denen mindestens eins vorhanden sein muss, als obligatorische Bestandteile des Konzeptes *Klärwerk* modelliert. Die weiteren Bestandteile *Grünfläche*, *Versiegelt*, *Parkplatz* und *Baumgruppe* sind optional.

Jedes Gebiet, für das während der Analyse eine Hypothese des Konzeptes *Ortsslage* (siehe Abbildung 6.1) generiert wurde, wird nun zusätzlich überprüft, ob dieses Gebiet ein Klärwerk repräsentiert. Hierzu wird während der Top-Down-Analyse aus jeder *Ortsslage* eine Hypothese für das Vorliegen einer Instanz vom Typ *Klärwerk* generiert. Die Entscheidung, ob eine Region zu der Klasse *Klärwerk* gehört, fällt anhand der Bewertung, die während der Bottom-Up-Analyse berechnet wird und in 6.3 dargestellt wurde. Dieser Ansatz setzt voraus, dass die durch den Anwender realisierte Bewertung die speziellere Region besser bewertet als die allgemeinere Alternative, falls genügend Hinweise für das speziellere Objekt vorliegen.

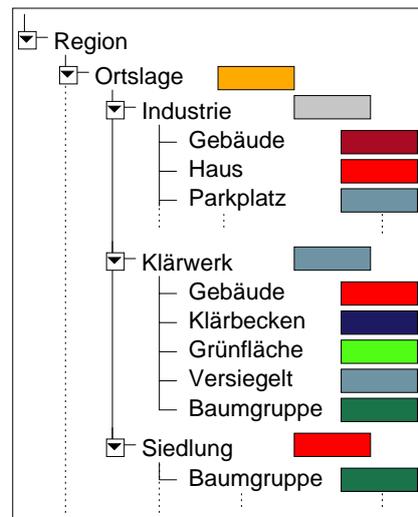


Abbildung 6.10: Erweiterung des Semantischen Netzes aus Abbildung 6.1

6.6.2 Spezialisierung über Attribuierung

Die Spezialisierung einer Instanz kann alternativ über die Attribuierung des Objektes geschehen. Hierbei wird nur der Bottom-Up-Operator der entsprechenden Instanz in der Form erweitert, dass alle als Industriegebiet erkannten Gebiete auf die speziellen Eigenschaften, die auf ein Klärwerk hinweisen, untersucht werden. Durch dieses Vorgehen ist keine doppelte Interpretation wie in 6.6.1 notwendig. Die speziellere Instanz ist nach der Analyse an dem Wert des zusätzlichen Attributs *Spezialisierung* zu erkennen.

Das Ergebnis beider Ansätze ist gleichbedeutend, wobei der Ansatz aus 6.6.1 expliziter ist, jedoch einen höheren Analyseaufwand verursacht. Der zusätzliche Aufwand kann jedoch als gering angesehen werden, da hier keine aufwendigen Bildverarbeitungsoperatoren zum Einsatz kommen, sondern nur eine Gruppierung und Bewertung von schon erkannten Objekten erfolgt. Das für beide Verfahren gleiche Ergebnis für die Klasse *Klärwerk* ist in Abbildung 6.11 zu sehen. Der nicht zum *Klärwerk* gehörenden Bereich ist leicht verschattet dargestellt.



Abbildung 6.11: Der in Abbildung 6.9 als Klärwerk erkannte Bereich

7 Zusammenfassung

Es hört doch jeder nur, was er versteht.

Johann Wolfgang von Goethe

Die Zielsetzung dieser Arbeit war die Entwicklung einer wissensbasierten Methodik, die es erlaubt, komplexe Fragestellungen der raumbezogenen Bilddeutung basierend auf Fernerkundungsdaten zu lösen.

Dieser Ansatz sollte die mit dem aktuellen Stand der Technik verbundenen Fragestellungen im Bereich der Bildinterpretation, wie Flexibilität, Parameterabhängigkeit sowie der Fehlerbehandlung aus dem Gebiet der Luftbildinterpretation in urbanen Gebieten so weit wie möglich erfassen. Mit den bisherigen Methoden, die meist implizite Verfahren der Bildinterpretation verwenden, konnten gute Ergebnisse nur mit einem hohen Grad an manueller Zusatzarbeit erreicht werden. Dieser Nachteil konnten hier durch den zusätzlichen Einsatz expliziter Verfahren und das Einbringen von Wissen über die zu extrahierenden Objekte und die zum Einsatz kommenden Verfahren weitestgehend behoben werden und führte zu einer signifikanten Steigerung des Automatisierungsgrades.

Um dieses Ziel zu erreichen, wurde zunächst exemplarisch für die zum Einsatz kommenden subsymbolischen Verfahren ein modellbasierter Ansatz zur Extraktion von Straßen aus Luftbildern und ein Bewertungsverfahren für die hiermit erzielten Ergebnisse vorgestellt. Durch das Einbringen von weiteren Informationen, hier die Detektion von bei der Straßenextraktion störenden Schattenbereichen, konnten die zuvor erzielten Ergebnisse noch weiter verbessert werden. Die Abhängigkeit der Qualität der Ergebnisse von Extraktionsverfahren von manuell einzustellenden Parametern wurde durch ein allgemeines Verfahren zur Adaption von Parametereinstellungen behoben.

Im zweiten Teil wurden wissensbasierte Ansätze beschrieben, die bei der Bildinterpretation eingesetzt werden können. Das realisierte Bildinterpretationssystem (GeoAIDA) verwendet zur Wissensrepräsentation die übersichtliche Struktur von Semantischen Netzen und ermöglicht die Beschreibung von komplexen Fragestellungen. Subsymbolische Verfahren, wie sie im ersten Teil beschrieben wurden, können im Bildinterpretationssystem integriert werden und liefern die notwendige Verbindung zwischen den Daten und der symbolischen Verarbeitung.

GeoAIDA realisiert die Integration von subsymbolischen und symbolischen Verfahren, indem wissensbasierte Methoden eingesetzt werden, um Bildverarbeitungsoperatoren gezielt auszuwählen, zu initialisieren, anzuwenden, deren Ergebnisse auszuwerten und die Einzelobjekte zu komplexeren Einheiten zu gruppieren. Hiermit wurde ein hybrider Ansatz vorgestellt, der die Probleme vieler herkömmlicher Bilddeutungsverfahren durch die Synergie dieser beiden Ansätze löst und den Automatisierungsgrad bei der Interpretation entscheidend erhöht. Die wichtigsten Eigenschaften des wissensbasierten Systems sind:

- Die einfache Integration und Initialisierung von beliebigen Bildverarbeitungsoperatoren zur Segmentierung in GeoAIDA.
- Die Modellierung der Reihenfolge und des Kontextes, in dem ein Bildverarbeitungsoperator auszuführen ist.
- Die Verwaltung der von den Bildverarbeitungsoperatoren gelieferten Ergebnisse.
- Die Möglichkeit einer übersichtlichen graphischen Darstellung von Wissen über die einzelnen modellierten Objekte und deren Relationen untereinander.
- Die Auswahl der besten Interpretation, falls alternative Beschreibungsmöglichkeiten vorliegen. Hierzu ist es nötig, geeignete Bewertungen für die erkannten Objekte umzusetzen.
- Eine übersichtliche und detaillierte Ergebnisdarstellung.

Das erstellte System GeoAIDA bietet sowohl die Möglichkeit der Erstellung und Modifikation von Semantischen Netzen, die als Wissensbasis des Systems zu sehen sind, als auch die Ausführung der wissensbasierten Bildinterpretation von Fernerkundungsdaten in einer übersichtlichen Ablaufsteuerung. Weiterhin liefert es die Präsentation der Ergebnisse in einer strukturierten Darstellung in Form eines Instanzenbaumes und einer XML-Beschreibung sowie einer Kartendarstellung mit unterschiedlichen Abstraktionsebenen.

Während der Implementierung und Erprobung des Systems zeigte sich, dass ein nicht unerheblicher Anteil des Aufwands zur Umsetzung einer kompletten Lösungsstrategie für die Realisierung der zur Segmentierung der Fernerkundungsdaten benötigten Bildverarbeitungsoperatoren und der Auswahl einer geeigneten Bewertung benötigt wird. Die von GeoAIDA bereitgestellten Möglichkeiten erleichtern jedoch durch die klare Strukturierung der Problemlösungsstrategie die Modellierung komplexer Probleme erheblich. Die Notwendigkeit, in vielen Fällen spezielle Verfahren zu entwickeln, um bestimmte Teilprobleme zu lösen, wie für die Straßenextraktion gezeigt, ist jedoch nicht beseitigt. Ähnliches gilt auch für die bei der Analyse benötigten Bewertung, die oft sehr spezielle Aspekte einer vorliegenden Situation zu erfassen hat.

Zusammenfassend lässt sich sagen, dass die Flexibilität und die durch die klare Strukturiertheit entstandene Übersichtlichkeit des Systems GeoAIDA es gestattet, komplexe Probleme modular in Teilprobleme zu zerlegen und zu beschreiben. Weiterhin bietet die wissensbasierte Repräsentation – wie sie in GeoAIDA stattfindet – die Grundlage, dass Verfeinerungen oder Erweiterungen einer Aufgabe einfach und übersichtlich möglich sind und ihre Auswirkungen auf das Gesamtsystem überschaubar bleiben. Dem Anwender ist somit ein Werkzeug gegeben, komplexe Aufgaben übersichtlich zu beschreiben und eingebrachtes Wissen für eine automatische Interpretation von raumbezogenen Inhalten aus Fernerkundungsdaten zu nutzen.

Als Ausblick lassen sich die multitemporale Interpretation und die gemeinsame Interpretation bezüglich unterschiedlicher Fragestellungen für eine Szene nennen. Die multitemporale Interpretation kann als Ziel ein Monitoring für eine Region beinhalten oder die Erkennung von speziellen Objektklassen, die – wie beispielsweise in [40] für ein Messegelände gezeigt – nur über eine zeitliche Se-

quenz von Bildern erkennbar sind. Beim Monitoring ist die Veränderung einer Szene über die Zeit zu erfassen. Multitemporale Daten können auch zu sichereren Ergebnissen bei der Interpretation führen, da bereits Hypothesen für Objekte vorliegen und eine Verifikation der Objekte als Bestätigung betrachtet werden kann. Weiterhin sind gewisse Objekte zu manchen Zeitpunkten besser zu extrahieren, deren Existenz zu einem späteren Zeitpunkt vorausgesetzt werden kann. Ein Beispiel hierfür liefert die Straßenextraktion in Neubaugebieten. In den Bauphasen können die Straßenverläufe sicherer erkannt werden, da sie zu diesem Zeitpunkt noch nicht von Bäumen verdeckt und von Straßenrandobjekte beschattet werden. Mit dem Wissen über den Verlauf der Straßen kann zu einem späteren Zeitpunkt bei unsicherer Datenlage das Vorhandensein einer Straße an diesen Stellen angenommen werden.

Der strikte hierarchische Ansatz lässt für ein Flächenelement in den Eingangsdaten bisher nur eine eindeutige Interpretation zu. Durch verschiedene Fragestellungen kann es jedoch notwendig sein, diese Beschränkung aufzuheben. Als Beispiel sei hier eine Wiese genannt, die aus der Sicht einer Verwaltungsbehörde eine Wirtschaftsfläche verkörpert, für einen Biologen aber ein Feuchtwiesenbiotop repräsentiert. Die Grenzen dieser beiden genannten Objekte sind nur in den seltensten Fällen identisch, sodass das Problem nicht über eine zusätzliche Attribuierung eines Objektes gelöst werden kann. Ein anderes Beispiel ist die Zugehörigkeit von Objekten zu zwei Gruppierungen unterschiedlicher Bedeutung. Diese müssen bisher schnittmengenfrei formuliert werden. Die genannten Probleme wären im bisherigen System nur durch mehrere separate Analyse-durchläufe mit unterschiedlichen Wissensbasen möglich. Eine gemeinsame Analyse wäre aber gegebenenfalls wünschenswert, da sich die verschiedenen Interpretationen gegenseitig unterstützen können oder Konflikte auf Fehler in den Interpretation hinweisen können.

Literaturverzeichnis

- [1] Abelson, H., Sussman, G. J., Sussmann, J.: *Structure and Interpretation of Computer Programs* Cambridge, 1996
- [2] Abmayr, W.: *Einführung in die digitale Bildverarbeitung* Stuttgart, 1994
- [3] Albayrak, S., Bussmann, S.: *Kommunikation und Verhandlungen in Mehragenten-Systemen* Verteilte Künstliche Intelligenz: Methoden und Anwendungen; S. 55 - 91; BI-Wissenschaftsverlag; Mannheim, 1993
- [4] Altenkrüger, D., Büttner, W.: *Wissensbasierte Systeme*, Vieweg Verlag, 1992
- [5] Anderberg, M. R.: *Cluster Analysis for Applications*, Academic Press 1973
- [6] Andresen, T., Mott, C., Schüpferling, R., Zimmermann, S., Schmeider, T.: *Objektorientierte Analyse von Fernerkundungsdaten zur Erfassung aquatisch/terrestrischer Parameter*, Blaschke, T. (Hrsg.), GIS und Fernerkundung, Neue Sensoren - Innovative Methoden, Wichmann Verlag, S. 222-232, Heidelberg, 2002
- [7] Argialis, D. P., Harlow, C. A.: *Computational Image Interpretation Models: An Overview and a Perspective*, Photogrammetric Engineering and Remote Sensing, Volume 56, No. 6, S. 871 - 886, 1990 1.3

-
- [8] Aviad, Z., Carnine, P.: *Road Finding for Road-Network Extraction*, Proceedings, Computer Vision and Pattern Recognition, IEEE, S. 814 - 819, 1988 3.1
- [9] Baatz, M., Schäpe, A.: *Object-oriented and multi scale image analysis in semantic networks*, Proc. 2nd International Symposium on Operationalisation of Remote Sensing Conference and Exhibition, Enschede, ITC, 1999 1.3
- [10] Baatz, M., Schäpe, A.: *Multiresolution segmentation – an optimization approach for high quality multi-scale image segmentation*, Strobl, J., Blaschke, T., Griesebner, G., (Hrsg.) *Angewandte Geographische Informationsverarbeitung – Beiträge zum AGIT-Symposium Salzburg*, S. 12 - 23, Karlsruhe, 2000 1.3
- [11] Barr, S., Barnsley, M.: *A region-based, graph-theoretic data model for the inference of second-order thematic information from remotely-sensed images*, International Journal of Geographical Information Science, Vol. 11, No. 6, S. 555 - 576, 1997 1.3
- [12] Bähr, H.-P., Vögtle, T.: *Digitale Bildverarbeitung*, Wichmann Verlag, Karlsruhe, 1991 2.3
- [13] Barclay Adams, J.: *Probabilistic reasoning and certainty factors*, Buchanan, B., Shortliffe, E. (Hrsg.). *Rule based expert systems*, S. 263 - 271, Addison-Wesley, Massachusetts, 1984 4.9
- [14] Barzohar, M., Cooper, D.: *Automatic Finding of Main Road in Areal Images by Using Geometric - Stochastic Models and Estimation*, Proceedings, Computer Vision and Pattern Recognition, IEEE, S. 459 - 464, 1993 3.1
- [15] Bauer, T., Steinnocher, K.: *Per-parcel land use classification in urban areas applying a rule-based technique*, GIS 6/2001, Hüthig GmbH & Co. KG, Heidelberg, 2001 1.3
- [16] Baumgartner, A., Steger, C., Weidemann, C., Mayer, H., Eckstein, W., Ebner, H.: *Update of Roads in GIS from Aerial Imagery: Ve-*

- rification and Multi-Resolution Extraction*, International Archives of Photogrammetry and Remote Sensing, Vol.: 31, I.: 3, S. 53 - 58, 1996
3.1
- [17] Baumgartner, A., Eckstein, W., Heipke, C., Hinz, S., Mayer, H., Radig, B., Steger, C., Wiedemann, C.: *TUM-Research on Road Extraction*, publisher: Christian Heipke and Helmut Mayer, S. 43 - 64, 1999 3.3
- [18] Blaschke, T.: *Objectextraktion und regelbasierte Klassifikation von Fernerkundungsdaten: Neue Möglichkeiten für GIS-Anwender und Planer*, 5. Symposium Computergestützte Raumplanung CORP 2000, S. 153 - 162, 1999
- [19] Blaschke, T.: *Objektextraktion und regelbasierte Klassifikation von Fernerkundungsdaten: Neue Möglichkeiten für GIS-Anwender und Planer*, Schrenk, M. (Hrsg.) Beiträge zum 5. Symposium zur Rolle der Informationstechnologie in der und für die Raumplanung, S. 153 - 162, Wien, 2000 1.3
- [20] Blaschke, T., Lang, S., Lorup, E., Strobl, J., Zeil, P.: *Object-oriented image processing in an integrated GIS/remote sensing environment for environmental applications*, Cremers, A., Greve, K. (Hrsg.), Umweltinformatik '00, Bonn, 2000 1.3
- [21] Brooks, R. A.: *Model-Based Three-Dimensional Interpretations of Two-Dimensional Images*, IEEE Trans. on Pattern Analysis and Machine Intelligence (PAMI), Vol. 5, S. 140 - 150, 1983 1.3
- [22] Bückner J.: *Model Based Road Extraction for the Registration and Interpretation of Remote Sensing Data* ISPRS Commission IV Symposium on 'GIS - Between Visions and Applications', Stuttgart, Germany, 7.9.-10.9.98, 1998 3.3
- [23] Bückner J., Pahl M., Stahlhut O.: *GEOAIDA - A Knowledge Based Automatic Image Data Analyser for Remote Sensing Data* CIMA 2001, Second International ICSC Symposium on ADVANCES IN INTELLIGENT DATA ANALYSIS, June 19-22 2001, Bangor, Wales, U.K., 2000 3.3

-
- [24] Busch, A., Willrich, F.: *Quality Management of ATKIS Data*, Proceeding CD of the OEEPE/ISPRS Joint Workshop on Spatial Data Quality Management, S. 12, Istanbul, Turkey, 2002 3.1.8
- [25] Charniak, E., McDermott, D.: *Introduction to Artificial Intelligence* Reading, Massachusetts, Addison-Wesley, 1989 4
- [26] Clement, V., Giraudon, G., Houzelle, S., Sadakly, F.: *Interpretation of Remotely Sensed Images in a Context of Multisensor Fusion Using a Multispecialist Architecture*, IEEE Trans. on Geoscience and Remote Sensing, Vol. 31, No 4, S. 779 - 791, 1993 1.3
- [27] Davis, R., Smith, G.: *Negotiation as a Metaphor for Distributed Problems Solving* Readings in Distributed Artificial Intelligence; S. 333 - 356; San Mateo, 1988
- [28] Dempster, A.: *Upper and lower probabilities, inferences based on a sample from a finite univariate population* *Biometrika*, Vol. 55, S. 515 - 528, 1967. 5.5.7
- [29] Dreyfus, H., Dreyfus, S.: *Künstliche Intelligenz*, Rowohlt Verlag, Hamburg, 1987.
- [30] Dubois, D., Prade, H.: *Possibility Theory: An Approach to Computerized Processing of Uncertainty* Plenum Press, New York and London, 1988 5.5.7
- [31] Englisch A., Heipke C.: *Erfassung und Aktualisierung topographischer Geo-Daten mit Hilfe analoger und digitaler Luftbilder*, PFG 3, S. 133 - 149, 1998
- [32] *ERDAS Field Guide* 5th Edition, Revised and Expanded, Version 8.4, ERDAS Inc. - Atlanta, 1999 1.3
- [33] Ernst, H.: *Einführung in die digitale Bildverarbeitung*, Franzis-Verlag GmbH, München, 1991
- [34] Frey, H.: *Digitale Bildverarbeitung in Farbräumen*, Dissertation, Technische Universität München, 1988

-
- [35] Foley, J. D., Dam, A. van, Feiner, St. K., Hughes, J. F.: *Computer Graphics*, 2nd Edition, Addison-Wesley Publishing Company, 1990
- [36] Geman, D., Jedynak, B.: *An active testing model for tracking roads in satellite images*, PAMI 18(1): 14., 1996 3.1
- [37] Gordon, J., Shortliffe, E.: *The Dempster-Shafer theory of evidence*, Buchanan, B., Shortliffe E. (Hrsg.), Rule based expert systems, S. 272 - 292, Addison-Wesley, Massachusetts, 1984 5.5.7
- [38] Gottlob, G., Frühwirth, T., Horn, W.: *GraphicsExpertensysteme*, Springer Verlag Wien, 1990 4
- [39] Groch, W.-D.: *Extraction of line shaped objects from aerial images using a special operator to analyze the profiles of functions*, Computer Graphics and Image Processing 18: S. 347 - 358., 1982 3.1
- [40] Growe, S.: *Wissensbasierte Interpretation multitemporaler Luftbilder* Dissertation, VDI Verlag GmbH, Düsseldorf, 2001 1.3, 2.5, 7
- [41] Grünreich, D.: *Das Amtliche Topographisch-Kartographische Informationssystem (ATKIS) - Konzeption und Realisierungsstand*, Seminar Geographische Informationssysteme in der öffentl. Verwaltung, IPF Universität Karlsruhe, Band C, Karlsruhe, 1988 2.4
- [42] Grünreich, D.: *ATKIS - A Topographic Information System as a Basis for a GIS and Digital Cartography in West Germany*, Geol. Jb. Vol. A122, S. 207 - 215, Hannover, Germany, 1992 2.4
- [43] Hanson, A. R., Riseman, E. M.: *VISION: A Computer System for Interpreting Scenes*, Hanson, A. R., Riseman, E. M. (Hrsg.), Computer Vision Systems, Academic Press, S. 303 - 334, New York, 1978 1.3
- [44] Harbeck, R.: *Das AdV-Vorhaben "Amtliches Topographisch-Kartographisches Informationssystem (ATKIS)" Inhaltliche Konzeption*, Nachrichten aus dem Karten- und Vermessungswesen der Universität Hannover, Heft 166, S. 130, Hannover, 1990 2.4

-
- [45] Hartl, P.: *Fernwirktechnik der Raumfahrt* Reihe Nachrichtentechnik, Band 2, Springer Verlag, 1980 2.3
- [46] Harvey, W. A.: *Performance evaluation for road extraction*, in International Workshop on 3D Geospatial Data Production: Meeting Application Requirements (International Workshop on 3D Geospatial Data Production: Meeting Application Requirements 1999), S. 175 - 184., 1999 3.1, 3.1.8
- [47] Hayes, P. J.: *In Defense of Logic* Fifth International Joint Conference on Artificial Intelligence, Cambridge, MA, 1977 4.3
- [48] Heipke C., Mayer H., Wiedemann C., Jamet O.: External evaluation of automatically extracted road axes, PFG 2, S. 81 - 94., 1998 3.1.7
- [49] Heipke C.: Stand und Perspektiven der digitalen Photogrammetrie zur Erfassung und Aktualisierung von Geodaten, VGI (86) 1, S. 17 - 22., 1998
- [50] Heyer, G., Krems, J.: *Wissensarten und ihre Darstellung*, Springer Verlag, G. Görz (Hrsg.), Bd. 169 von Informatik-Fachberichte, Berlin, Heidelberg, New York, Tokyo, 1988 4
- [51] Herrmann, J.: *Maschinelles Lernen und Wissensbasierte Systeme*, Springer Verlag Berlin Heidelberg, 1997 4.1
- [52] Hill, B., Roger, Th., Vorhagen, F. W.: *Comparative Analysis of the Quantization of Color Spaces on the Basis of the CIELAB Color-Difference Formula*, ACM Transactions on Graphics, Vol. 16, No. 2, S. 109-154, 1997
- [53] Itten Klaus I., Meyer Peter: *Geometric and Radiometric Correction of TM Data of Mountainous Forests Areas* IEEE Transactions on Geoscience and Remote Sensing, Vol. 31, No. 4, S. 764-770, 1993 3.3.4.1, 3.3.4.1
- [54] Jackson, P.: *Introduction to expert systems*, second edition, Addison-Wesley, Massachusetts, 1990 4.4

-
- [55] Kaufman, A.: *Introduction to the theory of fuzzy subsets*, Vol 1, S. 1 - 15, Academic Press, New York, 1975 4.10
- [56] Kettig, R. L., Landgrebe, D. A.: *Computer Classification of Remotely Sensed Multispectral Image Data by Extraction and Classification of Homogeneous Objects*, IEEE Transactions on Geoscience Electronics, Volume GE-14, No. 1, S. 19-26, 1976 1.3
- [57] Kowalski R.: *Logic for Problem Solving*, North Holland, Amsterdam, 1979 4.3
- [58] Kummert F.: *Flexible Steuerung eines sprachverstehenden Systems mit homogener Wissensbasis*, Infix Verlag, Sankt Augustin, 1992 1.3
- [59] Kummert, F., et al.: *Control and explanation in a signal understanding environment*, Signal Processing 32, S. 111-145, 1993 1.3, 5.4, 5.7.1
- [60] Liedtke, C.-E., Ender, M.: *Wissensbasierte Bildverarbeitung*, Springer Verlag, Berlin, Heidelberg, New York, 1989 3
- [61] Liedtke, C.-E., Bückner, J., Grau, O., Growe, S., Tönjes, R.: *AIDA: A system for the knowledge based interpretation of remote sensing data*, 3rd International Airborne Remote Sensing Conference and Exhibition, Copenhagen, Denmark, 1997 1.3, 5.4, 5.7.1
- [62] Lütjen, K.: *BPI: Ein Blackboard-basiertes Produktionssystem für die automatische Bildauswertung*, Hartmann, G. (Hrsg.), Mustererkennung 1986, 8. DAGM-Symposium, S. 164 - 168, Springer-Verlag, Berlin, 1986 1.3
- [63] Mayer H., Heipke C., Ebner H.: *Wissensbasierte Bildanalyse zur Extraktion von topographischen Objekten - Grundlagen und Interpretation von Katasterkarten*, ZfV (120) 6, S. 285 - 298, 1995
- [64] McDermott, J.: *R1: A Rule-Based Configurer of Computer Systems*, Artificial Intelligence, Bd. 19, Nr. 1, 1982 4.2

-
- [65] McKeown, D., Harvey, W. A., McDermott, J.: *Rule-Based Interpretation of Aerial Imagery*, IEEE Trans. on Pattern Analysis and Machine Intelligence (PAMI), Vol. 7, S. 570 - 585, 1985 1.3
- [66] Meinel, G., Neubert, M., Reder, J.: *Pixelorientierte versus segmentorientierte Klassifikation von IKONOS-Satellitenbilddaten – ein Methodenvergleich*, Zeitschrift für Photogrammetrie, Fernerkundung und Geoinformationen, Nr. 3, 2001 1.3
- [67] Minsky Marvin: *A framework for representing knowledge*, In P. Winston, *The Psychology of Computer Vision*, S. 211 - 277, McGraw, New York, NY, 1975 4.4
- [68] Müller, J.: *Verteilte Künstliche Intelligenz. Methoden und Anwendungen*, BI Wissenschaftsverlag. 1993
- [69] Negoita, V.: *Expert systems and fuzzy systems*, S. 47 - 60, Benjamin/Cummings Publishing Company, California, 1985 4.10
- [70] Neubert, M., Meinel, G.: *Segmentbasierte Auswertung von IKONOS-Daten – Anwendung der Bildanalyse-Software eCognition auf unterschiedliche Testgebiete*, Blaschke, T. (Hrsg.), *Fernerkundung und GIS: Neue Sensoren – innovative Methoden*, H. Wichmann Verlag, S. 108 - 117, Heidelberg, 2002 1.3
- [71] Nilsson, N. J.: *Principles of Artificial Intelligence*, Springer Verlag, Berlin, 1982 3.1.1
- [72] Nilsson, J.: *Principles of Artificial Intelligence*, Springer Verlag Berlin, S. 361 - 415, Chap. 9: Structured object representations, 1980 4.4
- [73] Niemann, H., Bunke, H.: *Künstliche Intelligenz in Bild- und Sprachanalyse*, B. G. Teubner Verlag, Stuttgart, 1987 1.3
- [74] Niemann, H., Sagerer, G., Schröder, S., Kummert, F.: *ERNEST: A Semantic Network System for Pattern Understanding*, IEEE Trans. On Pattern Analysis and Machine Intelligence, Volume 12. No. 9., 1990 1.3

-
- [75] Matsuyama, T.: *Knowledge-Based Aerial Image Understanding Systems and Expert Systems for Image Processing*, IEEE Transactions on Geoscience and Remote Sensing, Volume GE-25, No. 1, S. 305 - 316, 1987 1.3
- [76] Matsuyama, T., Hwang, V. S. - S.: *SIGMA: A Knowledge-Based Aerial Image Understanding System*, Plenum Press, S. 277 ff, New York, London, 1990 1.3
- [77] Pahl, M. *Architektur eines wissensbasierten Systems zur Interpretation multisensorieller Fernerkundungsdaten*, Dissertation, 2003 3.3.1, 6.6
- [78] Mayer, H., Laptev, I., Baumgartner, A.: *Multi-scale and snakes for automatic road extraction*, Fifth European Conference on Computer Vision, Nummer 1407 in Lecture Notes in Computer Science, Springer-Verlag, S. 720 - 733., 1998 3.1
- [79] Pakzad, K., Growe, S., Heipke, C., Liedtke C.-E.: *Multitemporale Luftbildinterpretation: Strategie und Auswertung*, Künstliche Intelligenz, Vol.: 15, S. 10 - 16, 2001
- [80] Pakzad, K. *Wissensbasierte Interpretation von Vegetationsflächen aus multitemporalen Fernerkundungsdaten*, Dissertation, Bayerischen Akademie der Wissenschaften, München, 2001 1.3
- [81] Paulus, W. R.: *Objektorientierte und wissensbasierte Bildverarbeitung*, Vieweg Verlag, 1992
- [82] Pearl, J.: *Heuristic Intelligent Search Strategies for Computer Problem Solving*, Addison-Wesley, 1984
- [83] Puppe, F.: *Einführung in Expertensysteme*, In Studienreihe Informatik, W. Brauer, G. Goos, (Hrsg.), Springer-Verlag, Berlin Heidelberg New York, 1991 4, 4.8
- [84] Quilian, M. R., Minsky, M.: *Semantic Memory*, in MIT Press, Semantic Information Processing, S. 227 - 270, Cambridge, 1968 4.5

-
- [85] Rich, E.: *Artificial Intelligence*, McGraw-Hill Verlag New York, 1983. 4
- [86] Robinson, J. A.: *Machine-Oriented Logic Based on the Resolution Principle*, Machine Intelligence 3, Donald Michie, New York, 1968 4.3
- [87] Robinson, J. A.: *The Generalized Resolution Principle*, Journal of the Association for Computing Machinery, Bd. 12, Nr. 1, 1965 4.3
- [88] Sagerer, G.: *Darstellung und Nutzen von Expertenwissen für ein Bild-analysesystem*, Dissertation, Springer Verlag, Berlin, Heidelberg, New York, 1985 1.3
- [89] Sagerer, G.: *Automatisches Verstehen gesprochener Sprache*, BI Wissenschaftsverlag Mannheim, Wien, Zürich, 1990 (Reihe Informatik; Bd. 74) 1.3
- [90] Sandmeier Stefan: *Radiometrische Korrektur des Topographieeffekts in optischen Satellitenbildern Vergleich eines semi-empirischen Verfahrens, mit einem physikalisch-basiertem Modell* Photogrammetrie, Fernerkundung, Geoinformation, No. 1, S. 23-32, 1997 3.3.4.1
- [91] Schettini, R.: *A segmentation algorithm for color images*, Pattern Recognition Letters 14, S. 499 - 506, 1993
- [92] Sester, M.: *Lernen struktureller Modelle für die Bildanalyse*, Doktorarbeit, Universität Stuttgart, Deutsche Geodätische Kommission, Reihe C, Nr. 441, München, 1995 5
- [93] Shafer. G.: *A Mathematical Theory of Evidence*, Princeton University Press, Princeton, 1976 5.5.7
- [94] Shapiro, St. C.: *Encyclopedia of artificial intelligence*, Stuart C. Shapiro, editor-in-chief, vol. 1, vol. 2, John Wiley & Sons, Inc., New York, Chichester, Singapore, 1992
- [95] Shufelt, J.: *Performance evaluation and analysis of monocular building extraction from aerial imagery*, IEEE Transactions on Pattern Analysis and Machine Intelligence 21(4), S. 311 - 326, 1999 3.1.8

-
- [96] Sowa, J. F.: *Conceptual Structures: Information Processing in Mind and Machine*, Addison-Wesley, Massachusetts, The System Programming Series, 1984 4
- [97] Steinnocher, K., Bauer, T.: *Objektbasierte Auswertung von hochauflösenden Fernerkundungsdaten in urbanen Räumen*. Angewandte geographische Informationsverarbeitung XII: Beiträge zum AGIT-Symposium Salzburg 2000, S. 31 - 38, 2000 1.3
- [98] Steinle E., Bähr H.-P.: *Laserscanning for change detection in urban environment* Altan & Gründig (eds.): Third Turkish-German Joint Geodetic Days 'Towards A Digital Age', Volume I, S. 147 - 156, Istanbul, Turkey, ISBN 975-561-159-2 (Vol. I) 1999
- [99] Teillet P. M., Guindon B., Goodenough D.G.: *On the slope-aspect correction of multispectral scanner data* Canadian Journal of Remote Sensing, Vol. 8, No. 2, S. 84-106 1982 3.3.4.1
- [100] Tönjes, R., Growe, S., Bückner, J, Liedtke, C.-E.: *Knowledge-Based Interpretation of Remote Sensing Images Using Semantic Nets*, Photogrammetric Engineering and Remote Sensing, Vol. 65, No. 7, S. 811 - 821, 1999 3.3
- [101] Tönjes, R. *Wissensbasierte Interpretation und 3D-Rekonstruktion von Landschaftsszenen aus Luftbildern* Dissertation, VDI Verlag GmbH, Düsseldorf, 1999 1.3, 5.5.7
- [102] Ton, J., Jain, A., Enslin, W., Hudson, W.: *Automatic road identification and labeling in Landsat 4 TM Images*, Photogrammetria 43: S. 257 - 276., 1989 3.1
- [103] Ton, J., Sticklen, J., Jain, A. K.: *Knowledge-Based Segmentation of Landsat images*, IEEE Transactions on Geoscience and Remote Sensing, Volume GE-29, No. 2, S. 223 - 231, 1991 1.3
- [104] Wang, J., Treitz, P., Howarth, P.: *Road network detection form SPOT imagery for updating geographical information systems in the rural-*

- urban fringe*, International Journal of Geographical Information Systems 6(2): S. 141 - 157., 1992 3.1
- [105] Wharton, S.: *A Spectral-Knowledge-Based Approach for Urban and Land-cover Discrimination*, IEEE Transactions on Geoscience and Remote Sensing, Volume GE-25, No. 3, S. 272 - 282, 1987 1.3
- [106] Wiedemann, C., Heipke, C., Mayer, H. und Jamet, O.: *Empirical Evaluation of Automatically Extracted Road Axes*, Bowyer, K. J., Phillips, P. J. (Hrsg.), Empirical Evaluation Methods in Computer Vision, IEEE Computer Society Press, Los Alamitos, California, S. 172 - 187, 1998 3.1.8
- [107] Winston, P. H.: *Artificial Intelligence*. Massachusetts, Addison-Wesley, 1989 3.1.1, 3.1.5.2
- [108] Zadeh, L.: *An Introduction to Fuzzy Logic Applications in Intelligent Systems*, Kluwer Academic Press, 1992 4.10
- [109] Zafiropoulos, P., Schenk, T.: *Color-based contour strategies for road extraction*, International Archives of Photogrammetry and Remote Sensing, Band XXXII, part 3-2W5, S. 203 - 208., 1999 3.1

Index

Symbole	
&	82
\Leftrightarrow	82
\Rightarrow	82
\exists	83
\forall	83
\neg	82
\vee	82
A	
A*-Algorithmus	36
ACRONYM	4
Adaption	52
Agenten	85
AIDA	5
Aktionsteil	80
Allquantor	83
Artificial Intelligence	76
Attribut	106
Aufnahmesystem	19
Ausdruck	82
azyklischer gerichteter Graph	90
B	
Baum	56
Bayes Netz	90
Beschreibung	76
Bestandteilkante	104
Beweis	82
Bewertung	108
Bild	12
Bildbearbeitung	14
Bildbeschreibungssprache	13
Bilddeutung	14
Bildelement	12
Bildklassen	15
Bildobjekt	13
Bildverarbeitung	15
Bottom-Up-Analyse	113
BPI	4
bst_von	104
C	
C-Korrektur-Ansatz	70
Certainty Faktor	91
CORBA	86
Cosinus-Korrektur	70
D	
Deklaratives Wissen	78
digitales Bild	12
Diskrete Entscheidung	108
E	
ECHO	4
eCognition	5
Entitäten	97
Entzerrung	69
ERNEST	4

- Existenzquantor 83
 ExpertClassifier 5
 explizite 16
 Explizites Wissen 77
 Extraktion 13
- F**
- Fakten 82
 first-pulse 136
 Fractal Net Evolution-Technik 5
 Frames 83
 fuzzy logic 5
 Fuzzy Logik 91
- G**
- Generischer Bottom-Up-Operator
 128
 geoAIDA 93
 Geometrische Auflösung 16
 Glaubwürdigkeit 108
 Globale Operatoren 25
- H**
- Halteproblem 83
 Hierarchien 84
 hill-climbing 57
 holistischer Operator 113, 118
 Hybride Systeme 16
 Hyperwürfel 56
- I**
- Implikation 82
 implizit 16
 Implizites Wissen 77
 Inferenzmaschine 79
 Inferenzregeln 82
 inst_of 106
 inst_von 106
 Instantiierungskante 106
 Instanz 102
- instanz_von 98
 Interbildklassenabhängigkeit 15
 Interpretation 14
 Interpretationsprozess 14
 Interpretierer 55
 Intrabildklassenabhängigkeit 15
 is_a 105
 ist_ein 98, 105
- J**
- Junktoren 82
- K**
- Künstliche Intelligenz 76
 Kante 97, 104
 Kante
 instanz_von 98
 Kante
 ist_ein 98
 Kante
 teil_von 98
 KI 76
 Klassifikation 13
 Knoten 97
 Konfliktlösungsstrategien 81
 Konklusion 80
 Kontinuierliche Bewertung 108
 Konzeptdefinitionen 101
 Kurve 17
- L**
- last-pulse 136
 lexikographische Ordnung 108
 Link 85
 Link Label 85
 Logik 82
 Lokale Operatoren 25
- M**
- Möglichkeit 109

-
- Maßstabsraum 93
 Maximumssuche 57
 Mengen 115
 MESSIE 4
 Modalität 107
 Modell 17, 26
 Modell
 externes 17, 93
 Modell
 konzeptionelles 17
 modellbasierte Verfahren 26
 Modus ponens 82
 Modus tolens 82
 monotones Schließen 92
- N**
- nicht-monotones Schließen 92
 Node 85
 Notation 129
 Notwendigkeit 109
- O**
- Objekt 13
 obligatorisch 105
 OMA 86
 OMG 86
 Optimierer 56
 optional 105
 ORB 86
- P**
- Parameter 15, 54
 Parameterabhängigkeit 15
 Parameteradaption 52
 Parameterempfindlichkeit 15
 Parameterraum 54, 55
 Parametertypen 52
 part_of 104
 Pixel 12
- Plattform 19
 Plausibilität 109
 Postfix Notation 129
 Prädikat 82
 Prämisse 80
 Prozedurales Wissen 78
 Punktoperatoren 25
- Q**
- Qualitätsmaß 53
 Quantisierung 12
 Quantoren 83
- R**
- Radiometrische Auflösung 16
 Radiometrische Entzerrung 69
 Rasterung 12
 ray-tracing 66
 red-black-tree 56
 Reflektion 69
 Regel 80
 Regelauswahl 81
 Regelbasierte Systeme 80
 Relation 97, 107
 Repräsentation 76
 Resolution 82
 Rolle 97
- S**
- Satellit 18
 Schattendetektion 61
 Schattenprojektion 65
 Schlagschatten 72
 Schließen
 monoton 92
 Schließen
 nicht-monoton 92
 Semantik einer Repräsentation ... 77
 Semantikfunktion 13

Semantische Netze	84
Sensor	19
SIGMA	4
simulated-annealing	57
Slot	83
SPAM	4
Spezialisierungskante	105
Straßenteilstück	48
Strategie	57
subsymbolisch	15
symbolisch	16
Syntax einer Repräsentation	77
Systemvergleich	122

T

teil_von	98, 104
Theorem von Bayes	89
Top-Down-Analyse	111

U

Umgebung	56
Umgekehrt Polnische Notation ..	129
Ungenauigkeit	109
Unglaubwürdigkeit	108
Unsicherheit	87, 108
Unwissenheit	108
UPN	129

V

Vektor	16
Vektordaten	16
Verfahren	
subsymbolisch	15
Verfahren	
symbolisch	16
VISION	4

W

Weltausschnitt	76
Wissensbasis	82

Wissensrepräsentationsformen	75
------------------------------------	----

Z

Zielfunktion	55
--------------------	----

Lebenslauf

Persönliche Daten:

Name: Jürgen Bückner
Geburtsdatum: 29. Oktober 1964
Geburtsort: Bad Oeynhausen

Schul- und Berufsausbildung:

09/1980 - 06/1982 Ausbildung zum Elektroanlageninstallateur
07/1982 - 06/1983 Ausbildung zum Energieanlagenelektroniker
08/1983 - 01/1985 Fachoberschulreife, Kreisberufsschule Minden
02/1985 - 12/1987 Allgemeine Hochschulreife, Weser Kolleg, Minden
04/1990 - 12/1995 Studium der Mathematik mit Wechsel im zweiten Semester
in den Studiengang Naturwissenschaftliche Informatik an der
Universität Bielefeld, Thema der Diplomarbeit:
*Objektorientierte Realisierung der Kontrollstrategie
einer Wissensrepräsentationssprache*
10/1996 - 05/1998 Fernuniversität Hagen, wirtschaftswissenschaftliches
Zusatzstudium zum Diplom-Wirtschaftsingenieur, 3 Semester

Berufstätigkeit:

07/1989 - 02/1990 Berufspraxis als Energieanlagenelektroniker
11/1992 - 12/1995 studentische Hilfskraft in der Arbeitsgruppe

01/1996 - 07/1996 Angewandte Informatik, Prof. Dr. Sagerer
Wissenschaftliche Hilfskraft an der Universität Hannover,
Institut für Theoretische Nachrichtentechnik

07/1996 - 06/2002 Wissenschaftlicher Mitarbeiter an der Universität Hannover,
Institut für Theoretische Nachrichtentechnik,

02/2001 - Wissenschaftliche Beratertätigkeit für die
Deutsche Messe AG bei der Hannover Messe,

Wehrersatzdienst:

03/1988 - 06/1989 Landbauforschungsanstalt, Hardebek bei Hamburg