Deutsche Geodätische Kommission
bei der Bayerischen Akademie der Wissenschaften

Reihe C | Dissertationen | Heft Nr. 633

Janet Heuwold

Automatische auflösungsabhängige Anpassung von Bildanalyse-Objektmodellen am Beispiel Straßen

München 2009

Verlag der Bayerischen Akademie der Wissenschaften
in Kommission beim Verlag C. H. Beck

ISSN 0065-5325 ISBN 978-3-7696-5045-7

Diese Arbeit ist gleichzeitig veröffentlicht in:
Wissenschaftliche Arbeiten der Fachrichtung Geodäsie und Geoinformatik der Leibniz Universität Hannover
ISSN 0174-1454, Nr. 279, Hannover 2009

Deutsche Geodätische Kommission
bei der Bayerischen Akademie der Wissenschaften

Reihe C Dissertationen Heft Nr. 633

Automatische auflösungsabhängige Anpassung von Bildanalyse-Objektmodellen am Beispiel Straßen

Von der Fakultät für Bauingenieurwesen und Geodäsie
der Gottfried Wilhelm Leibniz Universität Hannover
zur Erlangung des Grades
Doktor-Ingenieur (Dr.-Ing.)
genehmigte Dissertation

von

Dipl.-Ing. Janet Heuwold
geboren am 07.04.1976 in Berlin

München 2009

Verlag der Bayerischen Akademie der Wissenschaften
in Kommission bei der C. H. Beck'schen Verlagsbuchhandlung München

ISSN 0065-5325 ISBN 978-3-7696-5045-7

Diese Arbeit ist gleichzeitig veröffentlicht in:
Wissenschaftliche Arbeiten der Fachrichtung Geodäsie und Geoinformatik der Leibniz Universität Hannover
ISSN 0174-1454, Nr. 279, Hannover 2009

Adresse der Deutschen Geodätischen Kommission:

Deutsche Geodätische Kommission

Alfons-Goppel-Straße 11 ● D – 80 539 München

Telefon +49 – 89 – 23 031 1113 ● Telefax +49 – 89 – 23 031 - 1283 / - 1100

e-mail hornik@dgfi.badw.de ● http://www.dgk.badw.de

Prüfungskommission

Vorsitzender:	Univ.-Prof. Dr.-Ing. Steffen Schön
Referent:	Univ.-Prof. Dr.-Ing. habil. Christian Heipke
Korreferenten:	Univ.-Prof. Dr.-Ing. habil. Olaf Hellwich
	Univ.-Prof. Dr.-Ing. habil. Monika Sester

Tag der Einreichung der Arbeit: 13. März 2009

Tag der mündlichen Prüfung: 06. Mai 2009

© 2009 Deutsche Geodätische Kommission, München

Alle Rechte vorbehalten. Ohne Genehmigung der Herausgeber ist es auch nicht gestattet,
die Veröffentlichung oder Teile daraus auf photomechanischem Wege (Photokopie, Mikrokopie) zu vervielfältigen

ISSN 0065-5325 ISBN 978-3-7696-5045-7

Zusammenfassung

Landschaftsobjekte bilden sich in Luft- und Satellitenbildern unterschiedlicher Auflösung unterschiedlich ab. Während in Bildern hoher Auflösung viele Objektdetails erkennbar sind, können diese in der niedrigen Auflösung teilweise nicht erkannt oder voneinander getrennt werden. Objektteile, die in der hohen Auflösung als Flächen erscheinen, sind in der niedrigen Bildauflösung eventuell nur noch als Punkte oder Linien zu erkennen. In einer noch niedrigeren Auflösung sind dieselben Objektteile möglicherweise überhaupt nicht mehr erkennbar. Methoden der wissensbasierten Bildanalyse zur automatischen Extraktion von Landschaftsobjekten aus Luft- und Satellitenbildern nutzen Modelle, die die Objekte in den Bilddaten beschreiben. Wenn sich die Darstellung der Objekte in Bilddaten unterschiedlicher Auflösung unterscheidet, müssen sich die Modelle in Abhängigkeit von der Auflösung der Bilder ebenfalls unterscheiden.

In dieser Arbeit wird ein neues Verfahren zur automatischen Anpassung von Bildanalyse-Objektmodellen, die für eine gegebene Auflösung erstellt wurden, an eine niedrigere Auflösung vorgestellt. Für die Analyse des Verhaltens des Landschaftsobjektes in unterschiedlichen Bildauflösungen wird die Theorie des Linearen Skalenraumes angewendet. Die verwendeten Objektmodelle, die über Semantische Netze repräsentiert sind, beschreiben das Aussehen der Landschaftsobjekte in der realen Welt sowie im Bild über geometrische und radiometrische Attribute, wie z. B. Objekttyp, Größe und Intensität. Die Modelle beinhalten jedoch auch Operatoren zur Extraktion der jeweiligen Objektteile und stellen damit eine direkte Schnittstelle zu Algorithmen der Merkmalsextraktion her. Der Fokus in dieser Arbeit liegt auf Modellen zur Straßenextraktion aufgrund der enormen gesellschaftlichen Relevanz von Straßen und ihrer damit einhergehenden Bedeutung für die Objektextraktion im Rahmen der Aktualisierung von Geodatenbeständen.

Das entwickelte Verfahren zur automatischen Anpassung folgt einem Prozess in drei Stufen: *Zerlegung, Analyse des Skalenverhaltens* und *Fusion*. Zur Vereinfachung der automatischen Anpassung wird das Objektmodell in der ersten Stufe in einzelne Objektteile und Gruppen von Objektteilen *zerlegt*, deren Skalenverhalten zusammen betrachtet werden muss, da sie sich während der Auflösungsänderung gegenseitig beeinflussen. Die *Analyse des Skalenverhaltens* prädiziert nicht nur das Aussehen aller Objektteile über ihre Attributwerte in der niedrigeren Bildauflösung unter Berücksichtigung von Skalenraumereignissen, sondern auch die Operatoren zur Extraktion der Objektteile werden an die niedrigere Bildauflösung angepasst. Zum Schluss werden die angepassten Objektteil(-gruppen) in der letzten Stufe, der *Fusion*, wieder zu einem einheitlichen Objektmodell zusammengeführt, das für die Objektextraktion in der niedrigeren Bildauflösung geeignet ist. Die Methoden dieser drei Stufen unterscheiden sich je nach Typ des gegebenen Objektmodells. Die Typen, die im Anpassungsprozess unterschieden werden, sind Modelle mit parallelen linienartigen Objektteilen, Modelle mit parallelen linienartigen Objektteilen variabler Anzahl, Modelle mit flächenhaften Objektteilen und Modelle mit lokalen Kontextobjekten.

Zur Demonstration des neuen Anpassungsverfahrens werden drei Beispiele zur Anpassung verschiedener Typen von Objektmodellen an mehrere niedrigere Bildauflösungen präsentiert. Anschließend wird die Leistungsfähigkeit des entwickelten Verfahrens über den Vergleich der Extraktionsergebnisse der automatisch angepassten Objektmodelle aus realen Bilddaten mit den erzielten Ergebnissen der gegebenen Objektmodelle in der hohen Bildauflösung bewertet. Basierend auf den Ergebnissen der einzelnen Beispiele werden Probleme der einzelnen Methoden diskutiert und Lösungsvorschläge aufgezeigt. Die Arbeit schließt mit einer Schlussbetrachtung des entwickelten Verfahrens und gibt einen Ausblick auf mögliche Weiterentwicklungen.

Schlagworte: Objektextraktion, Linearer Skalenraum, Bildauflösung

Abstract

The appearance of landscape objects in aerial and satellite images depends on the resolution of these images. Whereas in images of high resolution many object details can be recognized, they might not be distinguishable from each other in images having a lower resolution. Object parts that appear as areas in high resolution images might merely be recognizable as points or lines in low resolution images. If the resolution is reduced even further, some object parts might not be identifiable at all. Methods of knowledge-based image analysis for the automatic extraction of landscape objects from aerial and satellite images use models to describe the appearance of the objects in the images. Models for image analysis have to take into account the fact that the appearance of the objects depends on the respective image resolution.

In this PhD thesis a new approach for the automatic adaptation of object models for image analysis to a lower image resolution is presented. Thereby, it is assumed that the model to be adapted was originally created for a certain high image resolution. The model for the object extraction at any lower resolution is to be predicted. For the analysis of the object's appearance depending on the image resolution the theory of linear scale-space is employed. The object models are represented as Semantic Nets and describe the appearance of the landscape objects in the real world and in the image. The description uses geometrical and radiometrical attributes such as object type, size, and intensity. The models also contain operators for the extraction of the respective object parts and thereby establish a direct connection to algorithms for feature extraction. The focus in this work is on models for road extraction due to the high social relevance of roads and to their importance in the context of geographical information systems.

The new approach for the automatic model adaptation follows a process of three stages: *decomposition*, *scale behaviour analysis*, and *fusion*. In order to simplify the automatic adaptation, the object model is at first *decomposed* into individual object parts and groups of object parts. Adjacent object parts can influence each other during scale change. The scale behaviour of such interacting object parts has thus to be analyzed together. Subsequently, in the *scale behaviour analysis* stage a prediction of the appearance of all object parts including their attribute values in the lower image resolution is carried out, taking into account possible scale-space events. Moreover, the feature extraction operators are adapted to the lower resolution. Finally, the predicted object parts and groups are combined to a complete object model in the *fusion* stage. This adapted model together with the adapted feature extraction operators can be used for the extraction of the landscape object in the lower resolution. The methods of the three stages differ depending on the type of the given object model. The types distinguished for the adaptation process are models with parallel linear object parts, models with parallel linear object parts of variable number, models with area-type object parts, and models incorporating local context objects.

Several examples for the automatic adaptation of different types of object models are presented for a demonstration of the new approach. In order to evaluate the performance of the developed adaptation technique the extraction results of the automatically adapted object models from real images are subsequently compared to the extraction results gained with given object models for high image resolution. Based on the results of the individual examples in different lower image resolutions, problems of the developed methods are discussed and solutions are presented. The thesis finishes with concluding remarks and gives prospects for further developments.

Keywords: object extraction, linear scale-space, resolution

Inhaltsverzeichnis

1 Einleitung 7

1.1 Motivation 7

1.2 Zielsetzung 8

1.3 Gliederung der Arbeit 10

2 Grundlagen und Stand der Forschung 11

2.1 Wissensrepräsentation mit Semantischen Netzen 11
2.1.1 Aufbau und Eigenschaften Semantischer Netze 12
2.1.2 Anwendungen Semantischer Netze in der Bildanalyse 14
2.1.2.1 Semantische Netze im wissensbasierten System geoAIDA 15
2.1.2.2 Semantische Netze für die Straßenextraktion 16

2.2 Repräsentation und Analyse von Daten in Skalenräumen 19
2.2.1 Vektordaten im Skalenraum 20
2.2.2 Bilddaten im Skalenraum 22
2.2.2.1 Prinzip des Skalenraumes 22
2.2.2.2 Theorie des Linearen Skalenraumes 26
2.2.2.3 Anwendung des Linearen Skalenraumes in der Bildanalyse 33

2.3 Diskussion 35

3 Ein neues Verfahren zur automatischen auflösungs-abhängigen Anpassung von Bildanalyse-Objektmodellen 37

3.1 Voraussetzungen 37
3.1.1 Allgemeine Randbedingungen an den Aufbau der Semantischen Netze 37
3.1.2 Knoten 38
3.1.3 Relationen 39
3.1.4 Extraktionsstrategie 40

3.2 Strategie der automatischen Anpassung 41

3.3 Methodik 43
3.3.1 Objektmodelle mit parallelen linienartigen Objektteilen 44
3.3.1.1 Zerlegung 44
3.3.1.2 Analyse des Skalenverhaltens 45
3.3.1.3 Fusion 51
3.3.2 Objektmodelle mit parallelen linienartigen Objektteilen variabler Anzahl 52
3.3.2.1 Zerlegung 53
3.3.2.2 Analyse des Skalenverhaltens 53
3.3.2.3 Fusion 56
3.3.3 Objektmodelle mit flächenhaften beliebig orientierten Objektteilen 57
3.3.3.1 Zerlegung 57
3.3.3.2 Analyse des Skalenverhaltens 58
3.3.3.3 Fusion 63
3.3.4 Objektmodelle mit lokalen Kontextobjekten 64
3.3.4.1 Konzept der Anpassung 64
3.3.4.2 Kontextmodellierung 65
3.3.4.3 Extraktionsstrategie 66
3.3.4.4 Methode der Anpassung 67

3.4 Diskussion **68**

4 Bewertung des neuen Verfahrens **70**

4.1 Beispiele zur automatischen Anpassung von Objektmodellen **70**
4.1.1 Beispiel 1 – Einfache Straße 71
4.1.1.1 Objektmodell für die hohe Bildauflösung 71
4.1.1.2 Automatische Anpassung 72
4.1.2 Beispiel 2 – Kreuzungsarm 79
4.1.2.1 Objektmodell für die hohe Bildauflösung 79
4.1.2.2 Automatische Anpassung 82
4.1.3 Beispiel 3 – Straße mit Fahrzeug als lokales Kontextobjekt 90
4.1.3.1 Objektmodell für die hohe Bildauflösung 90
4.1.3.2 Automatische Anpassung 92

4.2 Implementierung der Beispielsysteme **98**

4.3 Verifikation **99**
4.3.1 Methode der Verifikation 99
4.3.2 Bilddaten 101
4.3.3 Ergebnisse 103
4.3.3.1 Beispiel 1 – Einfache Straße 103
4.3.3.2 Beispiel 2 – Kreuzungsarm 108
4.3.3.3 Beispiel 3 – Lokale Kontextobjekte 110

4.4 Fazit **113**

5 Schlussbetrachtungen und Ausblick **115**

Literaturverzeichnis **119**

Dank **127**

Lebenslauf **128**

1 Einleitung

1.1 Motivation

Der Maßstab einer Beobachtung spielt bei der menschlichen Wahrnehmung eine zentrale Rolle. Beim Blick aus dem Flugzeug während des Startvorgangs verändert sich das Aussehen der Landschaft merklich. Kurz nach dem Start sind noch viele Details bis hin zu den Formen der einzelnen Gebäude und Straßenflächen mit ihren Markierungen und einzelnen Fahrzeugen erkennbar. Später ist bei steigender Flughöhe eine differenzierte Abgrenzung der einzelnen Objekte kaum noch möglich bis nur noch eine grobe Einteilung in Waldflächen, offene Landschaft und Siedlungsgebiete erkennbar ist. Allerdings erhält man dann einen guten Überblick über die gesamte Landschaft. Die steigende Flughöhe erzeugt einen Maßstabsraum, in dem ein Landschaftsobjekt verschiedene Erscheinungsformen durchläuft. Die Beschreibung eines Landschaftsobjektes würde bei einer detaillierten strukturellen Beschreibung beginnen und mit einer rein texturellen oder spektralen Beschreibung enden.

Landschaftsobjekte in Bilddaten verschiedener Auflösungen, die von Fernerkundungssensoren aus aufgenommen wurden, von einem Satelliten oder einem Flugzeug aus, weisen naturgemäß dieselben Maßstabsunterschiede auf. Je nachdem aus welcher Entfernung bzw. Flughöhe sie erfasst wurden, werden bei gleicher Brennweite unterschiedliche Bildauflösungen erzielt, die der Größe eines Pixels am Boden entsprechen. Abbildung 1.1 zeigt einen Luftbildausschnitt in drei verschiedenen Bildauflösungen. In der höchsten Auflösung sind von einer vierspurigen Straße die einzelnen Fahrbahnmarkierungen als klare Linien sowie die Fahrbahn selbst und Fahrzeuge als flächenhafte Strukturen sehr gut erkennbar. In der mittleren Bildauflösung dagegen ist die Unterscheidung der linienartigen Fahrbahnmarkierungen von der Fahrbahn schwieriger. Gleichzeitig wird die Textur der an die Straße angrenzenden Flächen wichtiger für die Beschreibung der Landschaftsszene. In der niedrigsten Bildauflösung schließlich verlieren sich schmale linienhafte Strukturen, wie die Fahrbahnmarkierungen, komplett. Die Fahrbahn erscheint nicht mehr als Fläche, sondern als kurvige breite Linie und der Netzwerkcharakter der Straße tritt deutlich hervor. Dieses Merkmal kann zusätzlich für eine hochsprachliche Beschreibung der Straße in dieser Auflösung herangezogen werden.

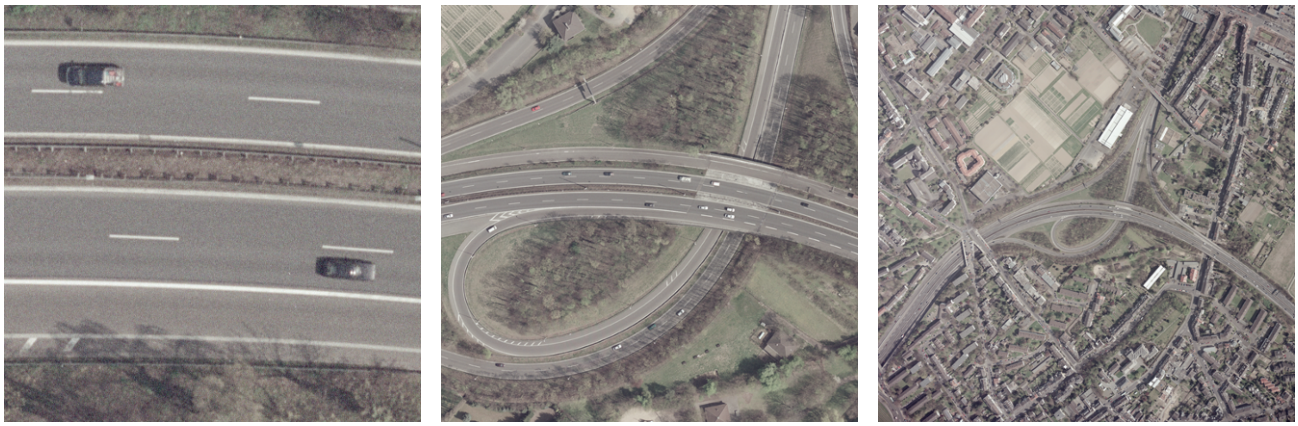


Abbildung 1.1: Luftbildausschnitt in verschiedenen Bildauflösungen: von links nach rechts 0.08 m, 0.5 m und 2.0 m

Die Art und Weise wie wir die Welt visuell wahrnehmen, hängt direkt von dem Maßstab ab, der durch die physikalischen Verhältnisse zum Zeitpunkt der Beobachtung bestimmt wird. Dieses fundamentale Maßstabskonzept wird auch in anderen Wissenschaftsdisziplinen aufgegriffen: In der Physik beispielsweise reicht der Umfang der unterschiedlichen Theorien von Teilchenphysik und Quantenmechanik in sehr feinen Maßstäben bis hin zu Astrophysik und Relativitätstheorie in sehr groben Maßstäben. Es müssen unterschiedliche Konzepte eingesetzt werden, da die Form der Beschreibung durch den Maßstab bestimmt wird. Selbst das menschliche visuelle System (einschließlich der Anatomie des menschlichen Auges), mit dem wir die Umgebung wahrnehmen, spiegelt das Maßstabskonzept auf seine Weise wider. Eine Reihe psychophysiologischer Studien zeigte, dass die menschliche Retina das Bildsignal in einem großen Bereich verschiedener Maßstäbe empfängt und verarbeitet [MARR 1982, KOSSLYN 1994, TER HAAR ROMENY 2002].

Wenn sich jedoch das Aussehen der Landschaftsobjekte in Bilddaten verschiedener Auflösungen ändert, ist auch eine Änderung der Objektbeschreibung nötig, die in der wissensbasierten Bildanalyse in den Modellen der Objektextraktion für die jeweiligen Bildauflösungen formuliert ist. Die Bedeutung dieser Änderung der Objektbeschreibung in verschiedenen Maßstäben wird beispielsweise von LINDEBERG [1998A] explizit hervorgehoben:

> *„The fact that objects in the world appear in different ways depending on the scale of observation has important implications if one aims at describing them."*

Der Umfang der verfügbaren Bildauflösungen von Fernerkundungsbilddaten ist in den letzten Jahrzehnten aufgrund der Entwicklung und des Betriebs einer steigenden Anzahl optischer Fernerkundungssensoren, größerer Sensorchips (und damit Anzahl der Pixel im Bild), sowie unterschiedlicher Sensoren mit komplexeren Bildaufnahmetechniken enorm gestiegen. Die verfügbare Spannweite an Bildauflösungen von unterschiedlichsten Sensoren reicht von hochaufgelösten Luftbildern mit wenigen Zentimetern Bodenpixelgröße bis hin zu Kilometern Bodenpixelgröße in niedrig aufgelösten Satellitenbildern. Eine weitere Steigerung der Vielfalt verfügbarer Sensoren ist mit den derzeit geplanten neuen Fernerkundungssatelliten und den Unmanned Aircraft Vehicle (UAVs) zu erwarten. Diese Vielzahl der optischen Sensoren bewirkt eine Datenflut, die wirtschaftlich nur noch automatisch mit effizienten Mitteln ausgewertet werden kann. Weiterhin benötigen die neuen Internet-Dienste wie Google Earth und Microsoft Windows Live Search Maps aktuelle Daten. Der Bedarf an einer automatischen effizienten Aktualisierung von Geoinformationssystemen (GIS) mit Landschaftsobjekten aus optischen Bilddaten ist groß. Dieser Bedarf muss über neue automatische Verfahren gedeckt werden, die in der Lage sind, eine große Spannweite von Bildauflösungen auszuwerten.

Die wissensbasierte Bildanalyse beschäftigt sich u. a. mit der Entwicklung von geeigneten Modellen, die der automatischen Extraktion von Landschaftsobjekten aus Bilddaten dienen. Modelle der wissensbasierten Bildanalyse sind aufgrund der beschriebenen Veränderung der Objekte im Maßstabsraum üblicherweise nur auf einen engen Auflösungsbereich erfolgreich anwendbar. Aktuelle Bilddaten sind jedoch nicht immer in der gewünschten Auflösung verfügbar oder wirtschaftlich zu beschaffen. Eine automatische Anpassung von bereits existierenden Modellen an das verfügbare Bildmaterial würde daher die automatische Objektextraktion aus Luft- und Satellitenbilddaten vereinfachen, da einmal (manuell) erstellte Modelle für eine hohe Bildauflösung automatisch für die Extraktion der modellierten Landschaftsobjekte aus niedrigeren Bildauflösungen verwendet werden könnten.

1.2 Zielsetzung

Das *Ziel dieser Arbeit* ist die Entwicklung eines automatischen Verfahrens zur Anpassung von wissensbasierten Bildanalyse-Objektmodellen zur Extraktion von Landschaftsobjekten an niedrigere Bildauflösungen, so dass diese über einen größeren Auflösungsbereich einsetzbar sind. Die Anpassung erfolgt ausgehend von einer hohen Bildauflösung, die viele Details enthält, zu niedrigeren Bildauflösungen, die durch einen geringeren Detailgrad gekennzeichnet sind. Eine automatische Anpassung bestehender Objektmodelle an eine niedrigere Auflösung kann die Erzeugung geeigneter Objektmodelle für die wissensbasierte Extraktion von Landschaftsobjekten aus Luft- und Satellitenbildern verschiedener Auflösungen vereinfachen. Die Arbeit ist daher ein Beitrag zur Automatisierung der Objektextraktion, deren Weiterentwicklung nach MAYER [2008] eine primäre Fragestellung in der Bildanalyse einnimmt und die Erstellung von geeigneten Objektmodellen mit einschließt.

Die Modelle der Bildanalyse zur Extraktion von Landschaftsobjekten aus optischen Bilddaten, die Gegenstand dieser Arbeit sind, werden in Form von Semantischen Netzen repräsentiert. Allgemein beschreiben Modelle in der wissensbasierten Bildanalyse spezifische Phänome bzw. Objekte in der realen Welt und/oder in Bildern, jedoch beschreiben sie keine allgemeingültigen Konzepte wie in der Mathematik oder Naturgesetze wie z. B. in der Physik. Die hier verwendeten Modelle umfassen die Modellierung des Aussehens der Objekte in der realen Welt und im Bild über geometrische und radiometrische Attribute, z. B. Objekttyp, Größe, Intensität. Sie beinhalten jedoch auch Operatoren zur Extraktion der jeweiligen Objektteile und stellen damit eine direkte Schnittstelle zu Algorithmen der Merkmalsextraktion her. Der Fokus in dieser Arbeit liegt auf Modellen zur Straßenextraktion aufgrund der enormen gesellschaftlichen Relevanz von Straßen und der damit einhergehenden Bedeutung für die Objektextraktion im Rahmen der Aktualisierung von Geoin-

formationssystemen. Das Konzept der entwickelten Methode zur automatischen Anpassung ist jedoch auf andere Landschaftsobjekte übertragbar.

Die Generalisierung in der Kartographie verfolgt ein auf den ersten Blick ähnliches Ziel: Die Bereitstellung von geeigneten Modellen zur Beschreibung von Landschaftsobjekten in verschiedenen Maßstäben. Im Gegensatz zu den kartographischen Modellen, wie beispielsweise Digitale Landschaftsmodelle (DLMs), beschreiben die hier genutzten Modelle für die Bildanalyse die Objekte wie sie in den Bilddaten erscheinen, einschließlich der Objekte, die für die kartographische Repräsentation in einem GIS irrelevant sind, z. B. Fahrzeuge oder auch unterschiedliche Texturen von Dächern. Ein GIS kann dagegen auch zusätzliche Attribute beinhalten, die aus Bildern nicht ableitbar sind, wie beispielsweise Straßennamen.

Die *Teilziele dieser Arbeit* umfassen nicht nur eine Anpassung der Beschreibung der Objektdarstellung in der niedrigeren Bildauflösung, sondern auch eine Adaption der Algorithmen, die zur Extraktion eines Objektes aus Bildern verwendet werden. Die angepassten Objektmodelle sollen in einem Bild der niedrigeren Auflösung wirklich anwendbar sein und nicht nur zur bloßen Beschreibung der Objekte dienen. Dafür ist auch eine Anpassung der Bildanalyse-Operatoren nötig oder gegebenenfalls eine Auswahl für die Zielauflösung geeigneter Operatoren zu treffen.

Aus der Zielsetzung ergeben sich die folgenden *Kernproblematiken*:

1. automatische Prädiktion des Objektaussehens in der niedrigeren Bildauflösung
2. Bestimmung der Extrahierbarkeit des Objektes bzw. Objektteils in der niedrigeren Auflösung
3. automatische Anpassung der Operatoren zur Extraktion der Objektteile oder Auswahl geeigneter Operatoren
4. Repräsentation der Objektbeschreibung in den angepassten Modellen, so dass diese für eine automatische Objektextraktion einsetzbar sind

Die vorliegende Arbeit vereint die Modellierung von Landschaftsobjekten über Semantische Netze mit der Untersuchung des Objektverhaltens im Maßstabsraum. Das Konzept der Semantischen Netze enthält bereits eine enge Verbindung zum Maßstabsraum, da die Bestandteilshierarchien des Semantischen Netzes im weitesten Sinne verschiedenen Skalen entsprechen. Der Aufbau der Semantischen Netze gibt damit Hinweise auf die Modellierung in unterschiedlichen Maßstäben, jedoch ist durch eine einfache Reduktion auf die obere Hierarchieebene keine ausreichend gute Modellierung in niedrigerer Auflösung gegeben. Für eine geeignete Modellierung des Objektes in einer niedrigeren Bildauflösung sind vielmehr die Ermittlung des genauen Bezugs der einzelnen Hierarchieebenen untereinander und die Beschreibung der Merkmale des Objektes in der niedrigeren Auflösung notwendig.

Die Analyse des Verhaltens der Objekte im Maßstabsraum nutzt die Theorie des Linearen Skalenraumes, die größtenteils im Bereich der Computer Vision entwickelt wurde. Die *Innovation in dieser Dissertation* ist die Verknüpfung dieser Theorie mit der Modellierung zur Extraktion von Landschaftsobjekten aus optischen Bilddaten. Die Konzepte und Methoden der Modellierung stammen aus dem Bereich der Künstlichen Intelligenz und der wissensbasierten Bildanalyse. Durch eine Einbettung der gut erforschten Methoden des Linearen Skalenraumes in die wissensbasierte Objektextraktion erscheint dies möglich. Eine Verbindung dieser beiden Bereiche bietet Potenzial für eine Vereinfachung der automatischen Erzeugung von geeigneten Modellen für die Extraktion von Landschaftsobjekten aus Bilddaten. Das Ausschöpfen dieses Potenzials ist das Ziel dieser Dissertation.

Semantische Netze dienen dem entwickelten Verfahren als Repräsentationsform für die Objektbeschreibung. Der Kern des Verfahrens ist jedoch nicht auf die spezielle Form der Semantischen Netze beschränkt, sondern betrifft vielmehr die Prädiktion des Verhaltens der Objekte bei einer Auflösungsänderung bzw. die Anpassung der Objektbeschreibung an eine niedrigere Bildauflösung. Zur Beschreibung der Objekte dienen Semantische Netze.

1.3 Gliederung der Arbeit

Zur verständlichen Vermittlung, Analyse und Bewertung des entwickelten Konzeptes und der Methode zur automatischen Anpassung ist die Arbeit folgendermaßen gegliedert:

In Kapitel 2 werden die Grundlagen und Vorarbeiten der wissensbasierten Objektextraktion mit Semantischen Netzen, speziell zur Straßenextraktion, sowie die Theorie der Skalenräume, insbesondere des Linearen Skalenraumes, erläutert. Das Kapitel zieht am Ende ein Fazit über den Stand der Forschung auf diesen beiden Gebieten.

In Kapitel 3 werden die in dieser Arbeit entwickelten neuartigen Konzepte und Methoden zur automatischen Anpassung von Semantischen Netzen an eine niedrigere Bildauflösung einschließlich theoretischer Überlegungen vorgestellt. Zuerst werden die Methoden für einfache Objektmodelle mit festen relativen Positionen und parallelen linienartigen Objektteilen dargelegt, dann werden die Methoden für die automatische Anpassung komplexer Objektmodelle mit flächenhaften und linienartigen beliebig orientierten Objektteilen sowie für Objektmodelle, die lokale Kontextobjekte mit einbeziehen, detailliert beschrieben.

Kapitel 4 beschäftigt sich mit der Bewertung der entwickelten Methoden. Dazu werden verschiedene repräsentative Beispiele für die Anpassung von Objektmodellen mit dem entwickelten Verfahren gezeigt. Die daran anschließende Verifikation dieser beispielhaft angepassten Modelle anhand realer Bilddaten analysiert die Machbarkeit der automatischen Anpassung von Objektmodellen mit der Theorie des Linearen Skalenraumes. Das Kapitel endet mit Schlussfolgerungen über den Erfolg der entwickelten Methoden.

Den Abschluss der Arbeit bildet Kapitel 5 mit einer ausführlichen Diskussion des entwickelten Verfahrens und die Übertragbarkeit der Methoden auf andere Landschaftsobjekte und Modelltypen. Ein Ausblick zeigt mögliche sinnvolle Erweiterungen auf und gibt Vorschläge für Verbesserungen des entwickelten Verfahrens.

2 Grundlagen und Stand der Forschung

Der Fokus dieses Kapitels liegt auf zwei Schwerpunkten der Bildanalyse, die in dieser Dissertation miteinander verbunden werden: einerseits Modelle, die der Bildinterpretation dienen, und andererseits die Analyse von räumlichen Daten in verschiedenen Skalen.

Zuerst werden in diesem Kapitel die Grundlagen der Modellierung von Wissen in der Bildanalyse, speziell die in der Arbeit verwendeten Semantischen Netze, erläutert. Anschließend wird der Stand der Forschung bei der Modellierung zur Straßenextraktion, dem Anwendungsbeispiel für das entwickelte Verfahren, diskutiert. Der Aspekt der Datenanalyse in verschiedenen Skalen wird als zweiter Schwerpunkt thematisiert. Dabei werden die existierenden Verfahren zur Repräsentation und Analyse von Vektor- und Bilddaten in Skalenräumen vorgestellt. Das Kapitel schließt mit einer Diskussion zum Stand der Forschung der Verknüpfung zwischen wissensbasierter Bildanalyse und der Repräsentation von Bilddaten in verschiedenen Skalen.

2.1 Wissensrepräsentation mit Semantischen Netzen

Die *Wissensrepräsentation* entspricht der Form der Datenstruktur, die für die Organisation des Wissens in einem bestimmten Problembereich genutzt wird [SUNDERMEYER 1991]. Diese Datenstruktur hat die Aufgabe, das Objekt, seine Eigenschaften und die Beziehungen zu anderen Objekten zu beschreiben. Die Wissensrepräsentation ist eines der zentralen Probleme in der Künstlichen Intelligenz (KI) [STRUß 1991]. Semantische Netze sind eine Art der Wissensrepräsentation. Frames, Prädikatenlogik und Produktionsregeln [REIMER 1991, REICHGELT 1991] sind Beispiele weiterer Repräsentationsformen. Das in einer bestimmten Form repräsentierte Wissen wird in einer *Wissensbasis* gespeichert, die in *wissensbasierten Systemen* zur maschinellen Lösung spezifischer Probleme genutzt wird. In der Wissensbasis ist das vom System benutzte Wissen explizit codiert [REIMER 1991]. In wissensbasierten Systemen wird grundsätzlich die Wissensrepräsentation von der Wissensverarbeitung getrennt [BEIERLE & KERN-ISBERNER 2003]. Während die Wissensbasis anwendungsspezifisches Wissen enthält, realisiert die Wissensverarbeitung die Lösung eines bestimmten Problems mit Hilfe der in der Wissensbasis bereitgestellten Information. Das Spektrum des praktischen Einsatzes wissensbasierter Systeme ist extrem breit und reicht von der Konfigurierung von Maschinen über die Organisation des industriellen Einkaufs bis zur Planung für den Umweltschutz [KREMS 1989, PESCHEL 1992].

Wissensbasierte Systeme werden auch im Bereich der Bildanalyse angewendet. Das Ziel der *Bildanalyse* ist die Extraktion von Objekten aus gegebenen Bilddaten. In der wissensbasierten Bildanalyse wird bekanntes Wissen für die Interpretation von Bildern zur Steigerung der Extraktionsgüte genutzt. Das Wissen über die zu detektierenden Objekte wird in einem Modell formuliert. Zu einem Extraktionsalgorithmus in der wissensbasierten Bildanalyse gehört neben dem Modell jedoch immer auch eine Extraktionsstrategie [MAYER 1998A]. Während das Modell das Objekt mit seinen Charakteristika beschreibt, steuert die Strategie den Ablauf des Extraktionsprozesses durch die Festlegung der Abfolge der einzelnen Bildverarbeitungsoperationen. Die Strategie regelt daher die Nutzung des Modellwissens für die Objektextraktion. Grundsätzlich werden die *datengetriebene (bottom-up)* und *modellgetriebene (top-down)* Strategie unterschieden: In der datengetriebenen Strategie werden einzelne in den Bildern extrahierte Merkmale bewertet und zu einem Objekt zusammengefügt. Dagegen generiert die modellgetriebene Strategie Hypothesen für die einzelnen Objektteile aus einem Modell, deren Merkmale in den Bildern verifiziert werden müssen. In der Praxis kommt meist eine Mischung mit unterschiedlicher Gewichtung beider Ansätze zur Anwendung.

Das Modellwissen, das während der Bildinterpretation zum Einsatz kommt, ist sehr vielschichtig und umfassend. Das Wissen im Modell wird nach FÖRSTNER [1993] in die folgenden Kategorien unterteilt:

- *Objektmodell*: Im Objektmodell wird das Objekt in der realen Welt mit der Hierarchie seiner Teilobjekte durch Konzepte beschrieben.
- *Sensormodell*: Das Sensormodell enthält Informationen über die physikalischen und geometrischen Eigenschaften des Bildaufnahmeprozesses.
- *Bildmodell*: Das Bildmodell beschreibt die Darstellung des aufgenommenen Objektes im Bildraum.

- *Analysemodell*: Dieses Modell legt die Techniken zur Analyse der Bildinformation fest.
- *Interpretationsmodell*: Das Interpretationsmodell schließlich enthält die Instanzen der Objekte einer spezifischen Szene – das Ergebnis des Vorganges der Bildanalyse.

Semantische Netze können das Objektmodell, das Bildmodell sowie das Interpretationsmodell beschreiben. Die Semantischen Netze, die in dieser Arbeit zur Objektbeschreibung genutzt werden, integrieren sowohl das Objekt- als auch das Bildmodell.

2.1.1 Aufbau und Eigenschaften Semantischer Netze

Beim Vorgang der Bildinterpretation, die der Mensch vollzieht, kann zwischen *Relationen* und *Objekten* unterschieden werden [KOSSLYN 1994]. Analog dazu rücken in *relationsbezogenen* Repräsentationsformen die Beziehungen zwischen Objekten in den Vordergrund, während bei den *objektbezogenen* Darstellungen die Objekte selbst im Mittelpunkt stehen [LUSTI 1990]. Semantische Netze folgen der objektbezogenen Repräsentationsform und wurden in die KI in den 1960er Jahren von QUILLIAN [1968] eingeführt. Die Idee war, die Struktur semantischer Information in menschlicher Langzeiterinnerung auf Maschinen zu simulieren.

Ein Semantisches Netz repräsentiert Wissen explizit als einen Graphen bestehend aus *Knoten* und *Kanten*. Während die Knoten die einzelnen mehr oder weniger komplexen Objekte oder Klassen von Objekten beschreiben, stellen die Kanten die Beziehungen zwischen diesen Objekten dar und verknüpfen sie miteinander. Zwischen den Knoten werden meist Vererbungsbeziehungen durch hierarchische *Relationen* aufgebaut, aber aufgrund der Vielfalt der möglichen Beziehungen können auch Informationen über andere Relationen zwischen den Objekten dargestellt werden. Der Graph ist gerichtet und azyklisch. Von dem Ausschluss von Zyklen abgesehen, können Semantische Netze beliebig vernetzt werden. Diese hohe Vernetzung zur Verbindung der einzelnen Objekte untereinander ist u. a. die große Stärke von Semantischen Netzen im Vergleich zu anderen Repräsentationsformen, die auf Produktionsregeln oder Logiken basieren [REICHGELT 1991]. Semantische Netze sind als Grafik sehr anschaulich und können durch ihre hohe Flexibilität in der Darstellung auch sehr komplexe Beziehungsstrukturen repräsentieren; nachteilig sind jedoch ihre schnelle Unübersichtlichkeit bei größeren Netzen und das Fehlen einer Standardterminologie [ALTENKRÜGER & BÜTTNER 1992]. Abbildung 2.1 zeigt ein einfaches Beispiel für ein Semantisches Netz. Die Definition der verwendeten Relationen ist auf den folgenden Seiten gegeben. Semantische Netze können in verschiedenen Arten ausgeprägt sein. Beispielsweise sind *Ausführbaren Netzen* im Bereich der Bildanalyse ausführbare Prozeduren angefügt, die einen Algorithmus der Merkmalsextraktion o. ä. durchführen. Nach SOWA [1992] zählen zu Semantischen Netzen auch *Lernende Netze*, z. B. Neuronale Netze, deren Struktur und Gewichte der Kanten aus Trainingsbeispielen gelernt werden.

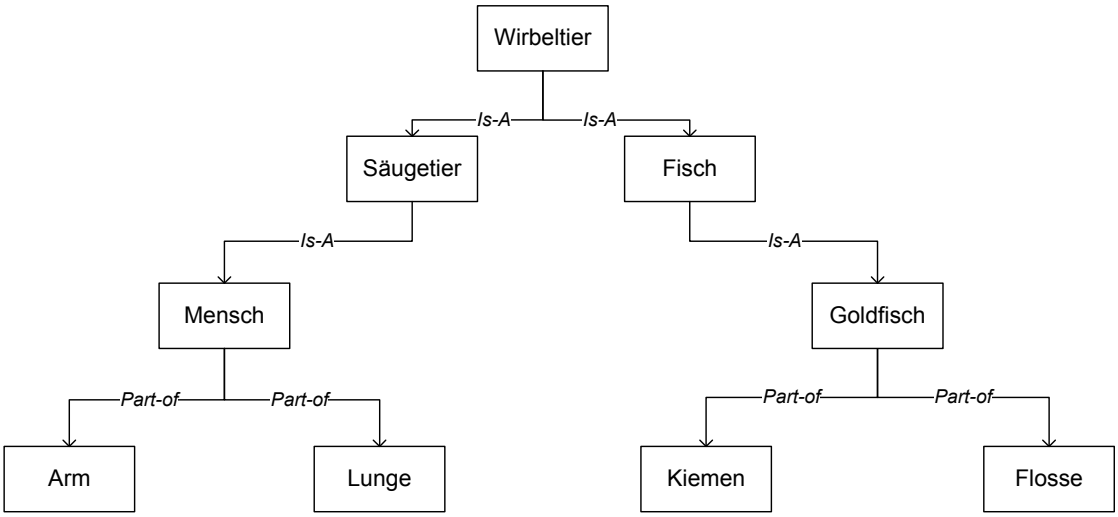


Abbildung 2.1: Beispiel eines Semantischen Netzes

Konzepte und Instanzen

Das Wissen über ein Objekt wird in *Konzepten* mit bestimmten Attributen abgelegt, die die Eigenschaften des Objektes wiedergeben. Konzepte repräsentieren das *a priori* Wissen über ein Objekt. Die *Instanzen* dagegen werden aus spezifischen Daten mit bestimmten Prozeduren erzeugt und repräsentieren das *a posteriori* Wissen. Demzufolge wird zwischen *Konzeptnetz* und *Instanzennetz* unterschieden. Das Instanzennetz setzt sich aus individuellen Knoten und Kanten zusammen, die von konkreten Objekten mit ihren Attributen aus den gegebenen Daten abgeleitet werden. Der Prozess der Erzeugung der Instanzen wird als Instanzierung bezeichnet. Instanzen sind damit Realisierungen der allgemeinen Konzepte aus spezifischen Daten. Da die Instanzen aus Daten abgeleitet werden, existieren unterschiedliche Instanzennetze – für jeden Datensatz ein eigenes Instanzennetz. Dagegen steht in der Regel nur ein einziges Konzeptnetz, das für diesen Objekttyp gültig ist. Das Prinzip von Konzept und Instanz veranschaulicht Abbildung 2.2.

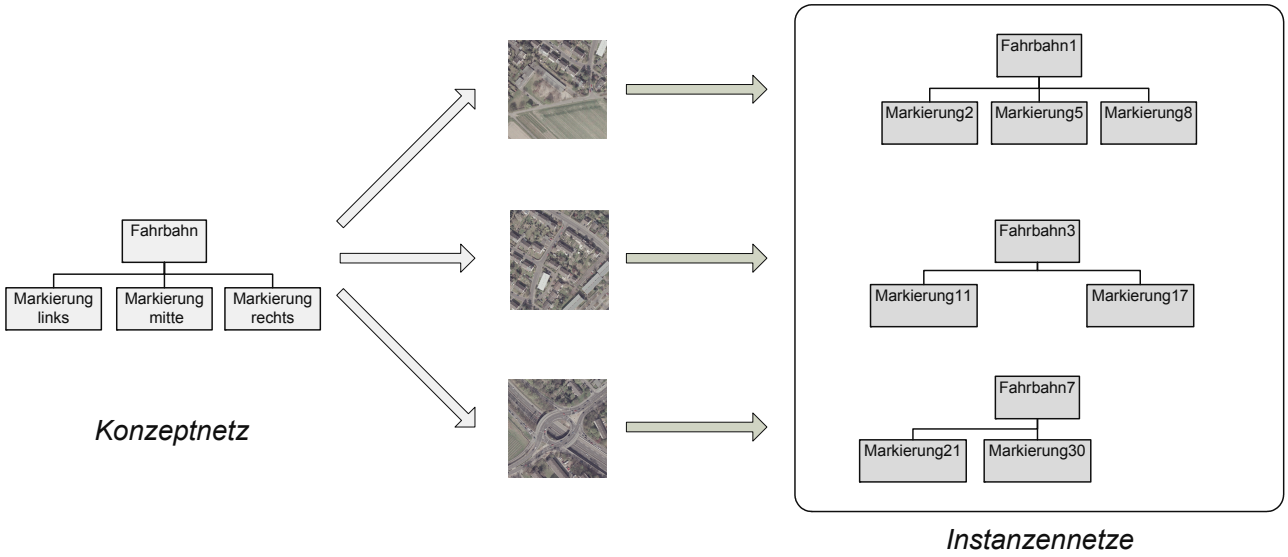


Abbildung 2.2: Konzeptnetz und Instanzennetze: Die Knoten der Markierungen sind optionale Teilobjekte und können daher zu einer unterschiedlichen Anzahl Objektteile in den Instanzennetzen führen.

Attribute der Knoten und Relationen

Attribute in den Knoten beschreiben die Eigenschaften der einzelnen Objektteile, während *Relationen* ihre Beziehungen untereinander beschreiben. Im Bildmodell spezifizieren die Attribute eines Knotens eine bestimmte Eigenschaft des Erscheinungsbildes des betreffenden Objektteils in den gegebenen Bilddaten. Sie können als Einzelwerte oder als Wertebereich angegeben sein. Mögliche Attribute der Objektteile sind z. B.:

- Geometrie
  - Größe
  - Form
  - Orientierung
- Radiometrie
  - Farbwert
  - Textur

Relationen verbinden die Knoten eines Semantischen Netzes miteinander. Es wird zwischen *hierarchischen* und *räumlichen Relationen* unterschieden. Hierarchische Relationen verbinden Knoten unterschiedlicher Ebenen miteinander, während räumliche Relationen die nachbarschaftlichen Beziehungen von Knoten derselben Hierarchieebene darstellen. Zur besseren Strukturierung des Wissens werden häufig mehrere Ebenen innerhalb eines Semantischen Netzes genutzt. Beispielsweise wird ein Objekt mit seinen Teilobjekten in der realen Welt in natürlicher Sprache in allgemeiner Form auf einer Ebene beschrieben, während die Materialebene und die Bildebene die Objektteile mit ihren Eigenschaften in den Bilddaten detaillierter spezifizieren, siehe z. B. [NIEMANN et al. 1990, TÖNJES et al. 1999].

Als Relationen zwischen Objektteilen stehen u. a. zur Verfügung:

- Hierarchisch
  - *Part-of:* Diese Verbindung bezeichnet eine Bestandteilsrelation zwischen übergeordnetem Objektteil und seinen Teilobjekten *(Aggregation)*. Angaben zu minimaler und maximaler Anzahl der Teilobjekte sind zusätzlich möglich. Damit kann im Konzeptnetz eine unbestimmte Anzahl von Teilobjekten auftreten. Das schließt auch die Angabe eines optionalen Teilobjektes im Konzeptnetz ein (häufig als *Optional part-of* benannt), das im Instanzennetz jedoch nicht zur Bestätigung des übergeordneten Objektes vorliegen muss. Die Existenz eines solchen optionalen Objektteils gibt oft einen wichtigen Hinweis für die Extraktion des übergeordneten Objektes, z. B. befindet sich ein Fahrzeug meist auf einer Straße.
  - *Is-a* (auch *Specialization-of*)*:* Diese Verbindung dient zur Spezialisierung der Objekte innerhalb einer allgemeinen Objektklasse, die sich auf zwei verschiedenen Ebenen befinden, z. B. ist ein Pkw die Spezialisierung eines Fahrzeugs. Diese Relation entspricht dem Konzept der *Generalization*.
  - *Instance-of:* Diese Relation dient der Verbindung eines Objektteils zwischen seinem allgemeinen Konzept und einer konkreten Instanz und stellt damit eine Klassifikation dar *(Classification)*.
  - *Concrete-of:* Zur Konkretisierung eines Objektteils zwischen verschiedenen Abstraktionsebenen wird eine abstrakte Repräsentation mit einer konkreteren Repräsentation verbunden, z. B. entspricht eine Dachfläche aus Ziegeln in der Materialebene einer dunklen kompakten Fläche in der Bildebene. Auf diesem Wege können verschiedene Darstellungen eines Objektes in unterschiedlichen Sensoren modelliert werden. Die Konzepte der Bildebene beschreiben Merkmale der Objektteile, die mit Methoden der Bildverarbeitung extrahiert werden können.
- Räumlich
  - *adjacent:* Die mit der Relation verbundenen Objektteile grenzen direkt aneinander.
  - *parallel*: Die Objektteile befinden sich parallel zueinander.
  - *perpendicular*: Die Objektteile befinden sich senkrecht zueinander.

Relationen können auch mit Attributen belegt werden. Auf diese Weise werden noch weitere Informationen zur Verbindung zwischen zwei Objektteilen hinzugefügt. Ein Beispiel für eine attributierte Relation sind *Distanzen* als Attribut zu einer räumlichen Relation, mit der die Entfernung zwischen Objektteilen angegeben werden kann und damit ein metrischer Bezug neben der topologischen Beziehung hergestellt wird.

Die menschliche Wahrnehmung strukturiert Bilder in einzelne Objekte. Ein Objekt wiederum ist aus Teilen zusammengesetzt. Allerdings kann auch ein Objektteil ein eigenes Objekt darstellen. Aus diesem Zusammenhang lassen sich Hierarchien von Objekten aufstellen. Semantische Netze sind aufgrund ihrer Struktur hervorragend zur Repräsentation von Objekten bzw. ihren Teilen und Relationen in einer Hierarchie geeignet und dienen als Modell zur Objektextraktion in der Bildanalyse. Objekte werden in der Bildanalyse häufig über Bestandteilshierarchien modelliert, die die Zusammensetzung des Objektes widerspiegeln. Insbesondere künstliche Objekte und Strukturen wie Gebäude und Straßen können durch eine Kombination einzelner Objektteile ausgezeichnet dargestellt werden. Aufgrund der hierarchischen, objektbezogenen Strukturierung des Wissens ermöglichen Semantische Netze eine computergerechte Formulierung zur Implementierung in objektorientierte Klassenhierarchien.

2.1.2 Anwendungen Semantischer Netze in der Bildanalyse

Die Anwendung der ersten Semantischen Netze von QUILLIAN [1968] fokussierte auf das Verstehen von Texten. Semantische Netze werden seit langer Zeit erfolgreich auf den Gebieten der Psychologie und Linguistik genutzt und untersucht [SOWA 1992]. Eine Verbindung zwischen Sprache und Raum mit Hilfe Semantischer Netze erzeugen BÄHR & MÜLLER [2004] und LUCAS et al. [2008] für die Abbildung verbaler Aussagen aus dem brasilianischen Kataster und Katastrophenmanagement in einer Karte.

Auf dem Gebiet der Bildanalyse wurden Semantische Netze zur Extraktion von verschiedenen Landschaftsobjekten entwickelt, beispielsweise für Gebäude [MÜLLER et al. 2003]. Zur Fokussierung auf das Anwendungsbeispiel „Straße“ in der vorliegenden Arbeit werden im Abschnitt 2.1.2.2 jedoch nur Ansätze zur Straßenextraktion vorgestellt. Nicht alle der vorgestellten Ansätze verwenden Semantische Netze zur getrennten Wissensrepräsentation in einem wissensbasierten System zur Straßenextraktion. Für einige Ansätze dienen Semantische Netze nur als Mittel zur Darstellung des im Extraktionsalgorithmus implizit codierten Wissens.

Wissensbasierte Systeme wurden bereits seit 1985 erfolgreich für die automatische Bildinterpretation eingesetzt [MCKEOWN et al. 1985]. Diese ersten Systeme basierten auf Regeln und Frames [MATSUYAMA 1987]. Semantische Netze wurden erstmals im wissensbasierten System ERNEST zur Mustererkennung in Bild und Sprache genutzt [NIEMANN et al. 1990]. TRINDER & WANG [1998] entwickelten ein wissensbasiertes System zur Straßenextraktion aus verschiedenen Bildauflösungen mit einem Semantischen Netz. Das System verwendet sowohl die Top-Down- als auch die Bottom-Up-Strategie (hybride Strategie). Ebenso wird in AIDA[1], das einen Ansatz zur strukturellen Bildinterpretation verfolgt und für die Bildanalyse optimiert wurde, das Objektwissen in einem Semantischen Netz repräsentiert [LIEDTKE et al. 1997, TÖNJES et al. 1999]. Alle Primitive der untersten Ebene des Semantischen Netzes werden in den gegebenen Bilddaten über Bildverarbeitungsoperatoren gesucht. Diese Objektteile werden dann in einem Bewertungsprozess zu einem einheitlichen Objekt kombiniert. Die Bewertung basiert auf den Informationen über die Eigenschaften des gesuchten Objektes, die in den Konzeptknoten abgelegt sind.

2.1.2.1 Semantische Netze im wissensbasierten System geoAIDA

Das wissensbasierte Bildinterpretationssystem *geoAIDA* [LIEDTKE et al. 2001, BÜCKNER et al. 2002, PAHL 2003] ist eine Weiterentwicklung von AIDA und wird in dieser Arbeit zur Verifikation der entwickelten Methoden an realen Luftbildern verwendet. geoAIDA stellt ein Werkzeug für die Bildanalyse unter Einbeziehung von a priori Wissen in Form von Semantischen Netzen dar. Abbildung 2.3 skizziert das Interpretationskonzept von geoAIDA. Neben Semantischen Netzen und Bilddaten verschiedener Sensoren kann das System auch mit GIS-Daten als zusätzlicher Information arbeiten.

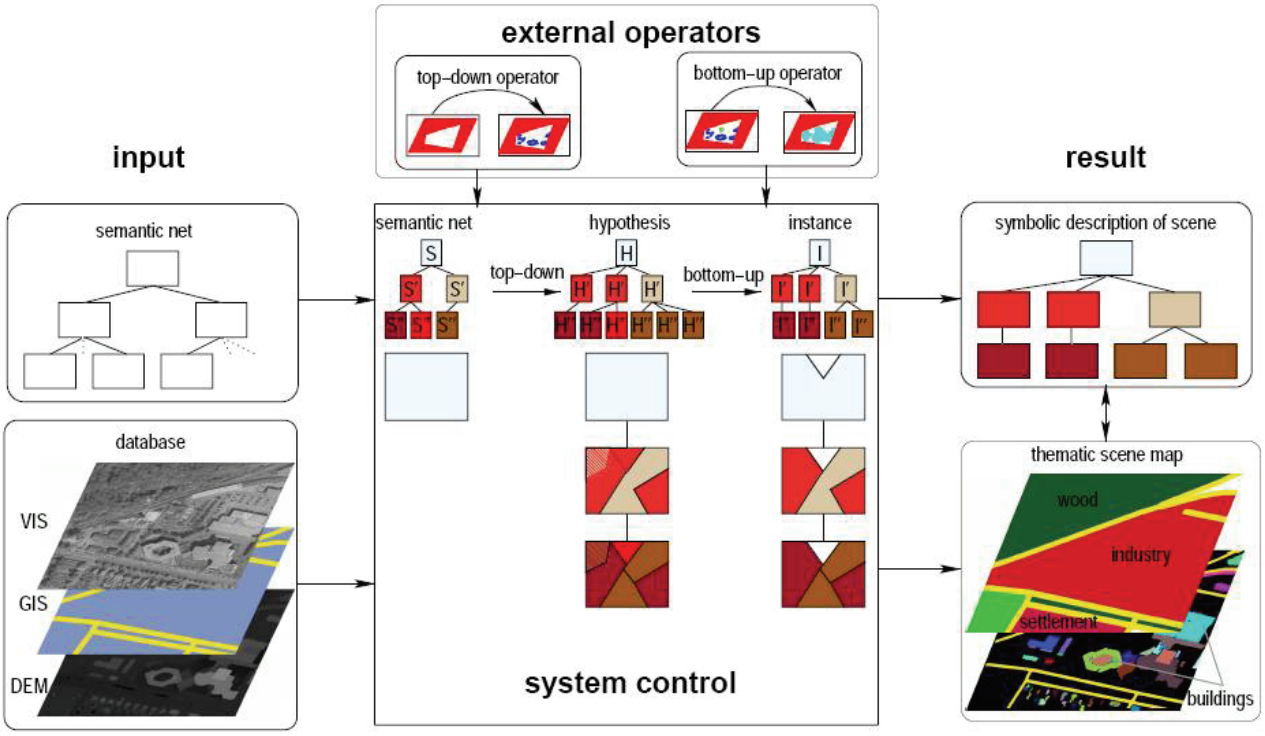


Abbildung 2.3: Interpretationskonzept von geoAIDA, aus [BÜCKNER et al. 2003]

[1] **A**utomatic **I**mage **D**ata **A**nalyser

Auch in geoAIDA werden grundsätzlich zwei Arten von Knoten unterschieden: Konzeptknoten und Instanzenknoten (siehe Abschnitt 2.1.1). Das Konzeptnetz repräsentiert das *a priori* Wissen über die Objekte, die in der Szene erwartet werden. Instanzenknoten entstehen dagegen als Resultat der Objektextraktion in der beobachteten Szene und sind von der jeweiligen Szene abhängig. Das Instanzennetz stellt somit eine symbolische Szenenbeschreibung dar. Mit Hilfe der Wissensbasis werden Hypothesen für die einzelnen Objektteile aus dem Konzeptnetz erzeugt, die dann in den gegebenen Bilddaten bewertet werden. geoAIDA verwendet sowohl *Top-Down-* als auch *Bottom-Up-Operatoren*. Aus der Wissensbasis wird für jedes Objektteil der jeweilige Top-Down-Operator aufgerufen, der aus den gegebenen Bilddaten die Kandidaten für das Objektteil durch Bildverarbeitungsalgorithmen extrahiert. Im Gegensatz zu AIDA kann geoAIDA sog. holistische Bildverarbeitungs-Operatoren als Top-Down-Operatoren zur Extraktion der einzelnen Objektteile auf allen Ebenen des Semantischen Netzes einbinden. Holistische Operatoren können mehrere über Bildverarbeitungsalgorithmen extrahierte Merkmale bzw. Primitive zu einem Objektteil, z. B. einem Haus, kombinieren. Die Ergebnisse der Top-Down-Operatoren entsprechen den Hypothesen, die dann aufgrund der gegebenen Attribute und Relationen in den Konzeptknoten durch die Bottom-Up-Operatoren bewertet und zu übergeordneten Objekten, beispielsweise einer Häuserreihe, gruppiert werden. Resultat der Bildinterpretation mit geoAIDA ist eine symbolische Szenenbeschreibung und ein segmentiertes Bild, das die extrahierten Objekte mit ihren Relationen in einer spezifischen Bildszene beschreibt.

2.1.2.2 Semantische Netze für die Straßenextraktion

Straßen sind komplexe Objekte in einem sehr variablen Umfeld. So variiert beispielsweise die Breite von 3 m für schmale Anliegerstraßen bis hin zu 45 m für achtspurige Autobahnen. Der Belag kann aus unterschiedlichen Materialien, wie Betonplatten, Kopfsteinpflaster oder Asphalt, bestehen. Auch die Form und Farbe der Fahrbahnmarkierungen ist von Staat zu Staat verschieden. In Städten besitzen Straßen oftmals einen angrenzenden Rad- und Fußweg, der in ländlichen Regionen bei Fernstraßen fast nie anzutreffen ist. Auch die Aufnahmeverhältnisse können sehr unterschiedlich sein. Vor allem die Auswirkungen der Wetterbedingungen, des Sonnenstandes auf die erzeugten Schatten und der Jahreszeiten mit der einhergehenden unterschiedlichen Ausprägung der Vegetation können stark variieren. Hinzu kommen die individuellen Eigenschaften des Sensors und die Auflösung des Bildmaterials. All diese Faktoren beeinflussen das radiometrische und geometrische Erscheinungsbild von Straßen in den Bilddaten. Daher können Straßen in optischen Fernerkundungsdaten sehr unterschiedliche Darstellungen aufweisen, die sowohl im Bildanalyse-Objektmodell als auch in der Extraktionsstrategie berücksichtigt werden müssen. Die Modellierung von Straßen mit wissensbasierten Methoden wurde bereits in zahlreichen Forschungsarbeiten untersucht und eine Vielzahl von Veröffentlichungen ist verfügbar.

Eines der ersten Systeme zur Straßenextraktion aus Luftbildern niedriger Auflösung modelliert Straßen auf einfache Weise als dünne helle Linien mit homogener Oberfläche und annähernd konstanter Breite [FISCHLER et al. 1981]. Mit Linienoperatoren unterschiedlicher Zuverlässigkeit extrahierte Straßensegmente werden durch den von Fischler angewendeten F*-Algorithmus über den kostengünstigsten Pfad zu einem Straßennetz verbunden. In RUSKONÉ et al. [1994] werden als Saatpunkte die Mittelpunkte lang gestreckter Regionen durch Wasserscheidentransformation im Gradientenbild ermittelt. Diese werden anschließend unter Ausnutzung der Homogenität der Straßenoberfläche zu Straßensegmenten ausgedehnt. Die Konstruktion des Straßennetzwerkes erfolgt dann über Verbindungshypothesen, die bestimmte geometrische Bedingungen erfüllen müssen. In einem anderen Ansatz wurden die Straßenmodelle als Semantische Netze für die Extraktion von Hauptverkehrsstraßen und Autobahnen aus den amtlichen Vorgaben zum Straßenentwurf erstellt [DE GUNST 1996]. Die Modellierung ist mit einzelnen Fahrspuren und Fahrbahnmarkierungen sehr detailliert. Durch die Fokussierung auf die geometrischen Eigenschaften aus dem Straßenentwurf ist dieser Ansatz jedoch auf den Straßentyp beschränkt, der diesen Vorgaben entspricht.

Die Forschungsgruppe Bildverstehen der TU München entwickelte bereits relativ früh erfolgreich Verfahren zur Extraktion von Straßen aus Luft- und Satellitenbildern niedriger (2 m/Pixel) als auch hoher Auflösung (0.05–0.15 m/Pixel) [HEIPKE et al. 1995]. Es wurden Semantische Netze zur Modellierung verschiedener Straßentypen in niedriger Auflösung mit einer variablen Anzahl der Fahrspuren unter Berücksichtigung der Netzwerkeigenschaft von Straßen eingesetzt [WIEDEMANN 2002]. Da Straßen in einer sehr groben Auflösung nur als Linien erscheinen, führt insbesondere die Ausnutzung der Netzwerkeigenschaft zu einer Verbesserung der Extraktionsergebnisse für Straßen in Bilddaten eines sehr niedrigen Auflösungsbereiches. Kreuzungen im Netzwerk werden hier als punktförmige Objekte modelliert. Für

die Extraktion der Straßensegmente wird der Operator zur Linienextraktion nach STEGER [1998A, 1998B] genutzt. Für hohe Bildauflösungen wurden verschiedene Straßentypen einschließlich Kreuzungen mit einer unterschiedlichen Anzahl der Fahrspuren und Fahrbahnmarkierungen mit Beschreibungen ihrer physikalischen, geometrischen und radiometrischen Eigenschaften modelliert [BAUMGARTNER 2003, HINZ 2003B]. Das umfangreiche Wissen über hoch aufgelöste Straßen in urbanen Gebieten wird in Objektmodell, Sensormodell und Bildmodell unterteilt. Außerdem wird das resultierende Semantische Netz durch die drei Ebenen Reale Welt, Geometrie/Material und Daten strukturiert.

Aufgrund der immer größeren Verfügbarkeit von hochauflösenden Bilddaten sind in den letzten Jahren auch Arbeiten zur detaillierten Modellierung von Kreuzungen entstanden. RAVANBAKHSH [2008] modelliert Kreuzungen in Luftbildern von 0.4 m/Pixel Auflösung als flächenhafte Objekte und nutzt GIS-Information zur Lokalisierung des Kreuzungsbereiches durch Erzeugung von Hypothesen für Kreuzungen. Die Straßenarme sind als homogene längliche Bänder zwischen (anti-)parallelen Kanten mit bestimmten geometrischen und topologischen Eigenschaften modelliert. Das verwendete Modellwissen wird über Semantische Netze dokumentiert. Verschiedene Arten von Snakes (Ziplock Snake, Balloon Snake und GVF[2]) dienen der exakten Rekonstruktion der Kreuzungsfläche. Kreuzungsinseln werden über Level Sets [MALLADI et al. 1995] extrahiert [RAVANBAKHSH et al. 2008A, 2008B].

Straßenextraktion wird heutzutage oftmals zur Aktualisierung von GIS-Daten angewendet, die inzwischen in den meisten Fällen flächendeckend vorliegen. Informationen aus einem GIS können die Extraktion von Straßen sehr erleichtern, da der Suchraum für die Extraktion eingeschränkt werden kann und die Parameter für das Modell aus dem GIS-Vorwissen sehr gut an die vorliegende Umgebung und Straßenform angepasst werden können. Beispielsweise wird in den für die Praxis entwickelten Systemen WIPKA[3] [GERKE et al. 2004] für die Aktualisierung des deutschen ATKIS[4] und ATOMI[5] [ZHANG 2004] für die Schweizer topographische Datenbank GIS-Information zur Unterstützung der Straßenextraktion verwendet. Im Folgenden wird jedoch nur auf Ansätze der Straßenmodellierung ohne Verwendung von GIS-Vorwissen eingegangen, da die Modelle für eine Extraktion ohne GIS-Information das Objekt umfassender beschreiben – eine Voraussetzung für den in dieser Arbeit entwickelten Algorithmus zur Anpassung von Objektmodellen an eine niedrigere Bildauflösung.

Der Kontext, in dem die Bildinterpretation stattfindet, spielt eine große Rolle bei der Darstellung von Straßen in Bilddaten. Allgemein beschreibt Kontext eine Einteilung des Bildes in Mengen von Objekten und ihre charakteristischen Beziehungen untereinander. Es wird *globaler* und *lokaler Kontext* unterschieden. Globaler Kontext bezieht sich auf eine Einteilung in größere landschaftliche Gebiete, während der lokale Kontext andere Objekte in der direkten Nachbarschaft der Straße betrifft. BAUMGARTNER et al. [1997, 1999] unterteilen für den globalen Kontext eine Szene in Siedlung, Wald und offene Landschaft. In der offenen Landschaft, die z. B. durch Felder geprägt ist, sind Straßen am wenigsten durch Verdeckungen anderer Objekte gestört. Die Extraktion von Straßen erreicht dementsprechend in diesem Gebiet die höchste Qualität. In Waldgebieten sind oftmals große Teile der Straße durch Bäume nicht erkennbar. Dasselbe gilt für Alleen, die von Bäumen (teilweise) verdeckt werden. In Siedlungsgebieten stehen Gebäude oftmals nah und parallel zur Straße, Fahrzeuge können die Fahrbahn verdecken oder Schatten werfen. Außerdem ist die Homogenität der Straßenoberfläche häufig gestört. Auch die Straßenform ist extrem variabel und ändert sich in der Regel an Kreuzungen, Einmündungen oder Parkbuchten. Diese Variabilität erschwert die Modellierung von Straßen in urbanen Gebieten und trägt maßgeblich zu einer verminderten Extraktionsqualität bei. Daher lässt sich die Straßenextraktion in Stadtgebieten durch die Einbindung von lokalen Kontextobjekten in das Modell wirkungsvoll unterstützen [RUSKONÉ et al. 1994, RUSKONÉ 1996].

Lokale Kontextrelationen modellieren die möglichen Beziehungen zwischen lokalen Kontextobjekten und der Straße bzw. ihrer Teilobjekte. Sie unterscheiden sich in Abhängigkeit vom jeweiligen globalen Kontext und dem damit einhergehenden Erscheinungsbild der Straße in den Bilddaten (siehe oben). Abbildung 2.4 zeigt lokale Kontextrelationen für Teilobjekte von Straßen in Stadtgebieten auf.

---

[2] **G**radient **V**ector **F**low
[3] **Wi**ssensbasierter **P**hotogrammetrisch-**K**artographischer **A**rbeitsplatz
[4] **A**mtliches **T**opographisch-**K**artographisches **I**nformations**s**ystem
[5] **A**utomated reconstruction of **T**opographic **O**bjects from aerial images using vectorised **M**ap **I**nformation

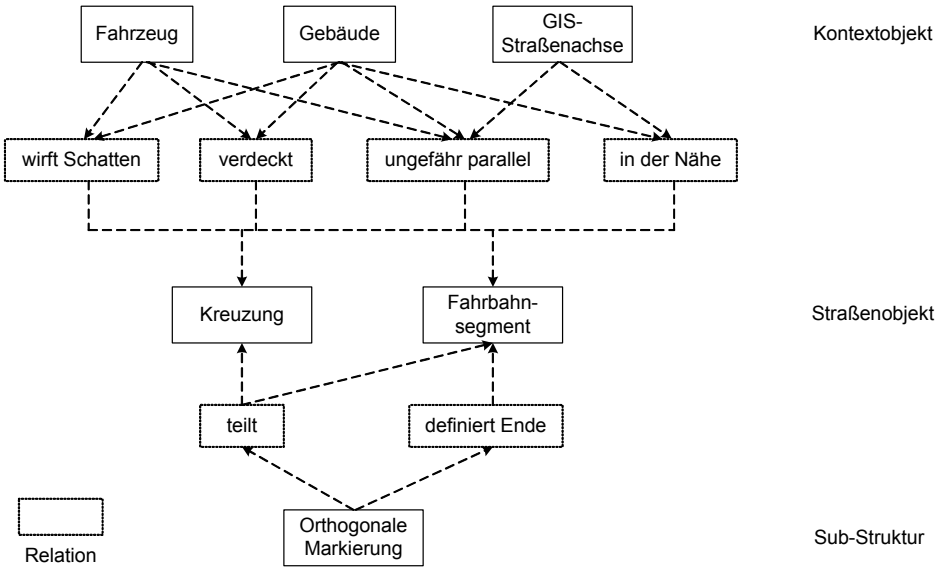


Abbildung 2.4: Beziehungen lokaler Kontextobjekte für Straßen in urbanen Gebieten, nach [Hinz & Baumgartner 2000]

Zur Modellierung relevanter lokaler Kontextobjekte für die Straßenextraktion in Siedlungsgebieten mit Semantischen Netzen existieren umfassende Arbeiten von HINZ & BAUMGARTNER [2000] sowie von HINZ [2003B]. Baumreihen und hohe Gebäude bzw. ihre Schatten wurden über ein Digitales Oberflächenmodell (DOM) in den Extraktionsprozess für Straßen eingebunden [BAUMGARTNER 2003]. Diese zusätzliche Datenquelle spiegelt sich mit einer weiteren Datenebene im Semantischen Netz wider. Fahrzeuge wurden sehr detailliert über ein Semantisches Netz zur Unterstützung der Straßenextraktion aus hochauflösenden Luftbildern modelliert [HINZ 2003A]. Im Allgemeinen werden die Modelle als 2D- oder 3D-Objekte beschrieben und entweder für das Modell top-down in dem Bild Übereinstimmungen gesucht, oder extrahierte Objektmerkmale wie Kanten oder Linien werden bottom-up mit Hilfe verschiedener Algorithmen zu fahrzeugähnlichen Strukturen zusammengefügt und mit dem Modell verglichen. In [HINZ 2004] wird ein Verfahren zur automatischen Fahrzeugerkennung mit einem expliziten 3D-Modell mit geometrischen und radiometrischen Eigenschaften vorgestellt. Eine Unterteilung des Verkehrs über die *Is-a*-Relation in verschiedene Verkehrssituationen (Verkehrsstau, zäh fließender und frei fließender Verkehr) in Abhängigkeit der Fahrzeugdichte nehmen HINZ et al. [2008] vor. Es existiert eine Vielzahl weiterer Ansätze zur Fahrzeugextraktion mit impliziter Modellierung, z. B. [RUSKONÉ et al. 1996]. Diese stützen sich jedoch auf zuvor ermittelte Merkmale in Trainingsdaten und sind daher nur eingeschränkt auf andere Umgebungsbedingungen übertragbar. Diese Einschränkung kann jedoch für die große Vielzahl unterschiedlicher Fahrzeuge in realen Bilddaten bedeutend sein.

Multiskalige Ansätze

Wie bereits in der Einleitung geschildert, variiert das Erscheinungsbild von Straßen in Luft- und Satellitenbildern unterschiedlicher Auflösungen stark. Straßen in niedriger Bildauflösung werden meist durch einfache Linien konstanter Breite modelliert. In einer hohen Bildauflösung müssen die Modelle für Straßen sehr viel detaillierter die Charakteristik der Straße, als flächenhaftes Objekt mit Teilobjekten wie Fahrspurmarkierungen und angrenzenden Kontextobjekten, wiedergeben. Diese sehr verschiedenen Darstellungen von Straßen in Bilddaten unterschiedlicher Auflösung können jedoch auch für eine Verbesserung der Extraktionsqualität genutzt werden. Die Repräsentationen einer Straße in mehreren Auflösungen wurden in verschiedenen Systemen in ein einziges Modell eingebunden, z. B. von BAUMGARTNER et al. [1997] und MAYER [1998A] (vgl. Abbildung 2.5). Ein multiskaliger Ansatz verbessert die Straßenextraktion durch eine Interpretation in niedriger und hoher Auflösung. In hoher Auflösung sind sehr viele Details erkennbar. Die Objektbeschreibung in niedriger Auflösung liefert dagegen zusätzliche Evidenz und der Suchbereich kann in der hohen Auflösung eingeschränkt werden. Vor allem wird hier der Vorteil ausgenutzt, dass kleine lokale Kontextobjekte, wie Pkws, in niedriger Bildauflösung verschwinden und die Extraktion aufgrund der Abstraktion vereinfacht wird [MAYER & STEGER 1998]. BAUMGARTNER [2003] rekonstruiert das Straßennetz über Kreuzungen und Straßenabschnitte. Kreuzungen werden in der hohen Auflösung als Region und in der reduzierten Auflösung als punktförmiger Objekttyp modelliert. Die Straßenabschnitte werden in der hohen Auflösung aus Fahrbahn und Markierungen modelliert, die Fahrbahn erscheint

als Fläche und die Markierungen als Linien. In der reduzierten Auflösung erscheinen die Straßensegmente als Linien. Hier wird im Modell die Objekttypänderung, die mit der Abstraktion zur niedrigeren Bildauflösung einhergeht, berücksichtigt. Aufbauend auf dem Modell von Baumgartner präsentiert HINZ [2003] ein Semantisches Netz zur Straßenextraktion in Stadtgebieten mit drei verschiedenen Bildauflösungen (siehe vorherigen Abschnitt zu lokalem Kontext).

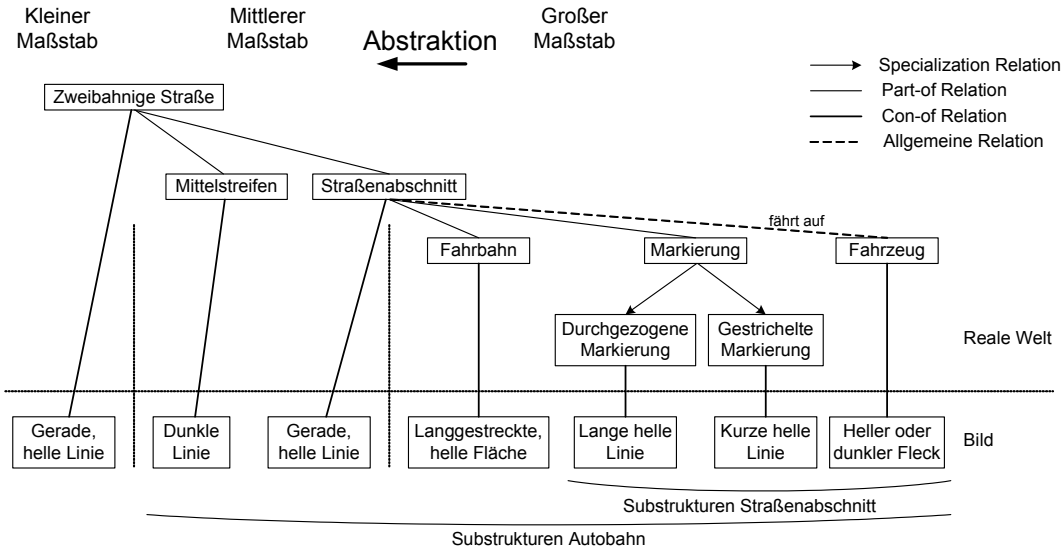


Abbildung 2.5: Multiskaliges Straßenmodell, nach [MAYER 1998A]

Ein multiskaliges Objektmodell als Semantisches Netz für Fahrzeuge in verschiedenen Verkehrssituationen zeigen HINZ et al. [2008]. Die Fahrzeuge sind als farbige Blobs in einer hohen Bildauflösung (0.3–0.8 m) modelliert. In der niedrigen Bildauflösung (1–2 m) dagegen sind keine einzelnen Fahrzeuge modelliert, sondern Fahrzeugkolonnen als glatte kurvige Bänder. Allerdings sind auch hier wie im obigen Modell für Straßen die Repräsentationen der Fahrzeuge in verschiedenen Bildauflösungen zur Steigerung der Evidenz bei der Extraktion in einem Modell integriert.

Weitere Beispiele für multiskalige Ansätze zur Straßenextraktion präsentieren DE GUNST & VOSSELMANN [1997] und TRINDER & WANG [1998]. DE GUNST & VOSSELMANN [1997] zeigen ein Semantisches Netz, das ein Straßennetzwerk in drei festgelegten Bildauflösungen umfasst und in diesen in verschiedenen Detailgraden modelliert wird. TRINDER & WANG [1998] modellieren Straßen sehr ähnlich zu Baumgartner in zwei verschiedenen Bildauflösungen, in denen flächenhafte bzw. linienhafte oder punktförmige Objekttypen je nach Auflösung zur Beschreibung der Objektteile verwendet werden. Die Objektteile zwischen den Bildauflösungen sind hier über *Part-of*- und *Is-a*-Relationen miteinander verbunden.

2.2 Repräsentation und Analyse von Daten in Skalenräumen

Der Begriff *Skala*, *Maßstab* oder englisch *scale* wird in den verschiedenen Gebieten der Geowissenschaften oft unterschiedlich genutzt. Skalen bezeichnen eine Variabilität von Ereignissen, Phänomenen, Mustern und sogar Konzepten in Raum und Zeit [FÖRSTNER 2003]. Ein Geologe versteht unter Maßstab die Korngröße einer Bodenprobe, während der Geophysiker an den Wirkungsbereich seismischer Bewegungen in der Erdkruste denkt, da der geographische Maßstab die Größe eines Untersuchungsgebietes charakterisiert. Der kartographische Maßstab dagegen entspricht dem Verhältnis einer bestimmten Länge in einer Karte zum selben Abstand in der Wirklichkeit. Analog dazu kennzeichnet die *Auflösung* eines Bildes die Rasterweite oder Bodenpixelgröße in x und y in der Realität, angegeben in einer Längeneinheit. In dieser Arbeit wird der Begriff der Skala bzw. des Maßstabes in Karten und Bildern im kartographischen Sinne verwendet. Dementsprechend bezeichnet eine *feine Skala* einen großen Kartenmaßstab und eine hohe Auflösung mit einer kleinen Rasterweite, während eine *grobe Skala* einem kleinen Kartenmaßstab und einer geringen Auflösung mit einer großen Rasterweite entspricht. Raumbezogene Daten in unterschiedlichen Skalen können in einem *Skalenraum* repräsentiert werden, in dem die Strukturen in den Daten abhängig von der jeweiligen Maßstabsebene dargestellt sind.

Der Vorteil einer Analyse von Daten in Skalenräumen besteht in der Möglichkeit zur Abstraktion. Man erhält durch Abstraktion Zugang zum Wesentlichen, das in einem bestimmten Kontext der Kern der Sache ist. Das Sprichwort „den Wald vor lauter Bäumen nicht sehen“ benennt die Notwendigkeit zur Abstraktion. Manchmal ist es nicht möglich, einen übergeordneten Zusammenhang zu erkennen, wenn zu viele Details vorliegen. In diesem Fall ergibt sich der Wald erst aus dem Gefüge der Bäume, die aus einer Menge von Blättern und Ästen bestehen. Bei Betrachtung eines einzelnen Blattes ist der Wald nicht zu erkennen, da Informationen über die Nachbarschaften zu anderen Blättern fehlen. Erst durch die Betrachtung der Umgebung des Blattes erschließt sich das Objekt „Wald“. Genau diesen Überblick ermöglichen gröbere Maßstabsebenen, die in einem Skalenraum bereitstehen.

Für die Darstellung räumlicher Information in Skalenräumen existieren verschiedene Datentypen: Bilddaten als Rasterdaten, die die Thematik der vorliegenden Arbeit betreffen, Vektordaten für GIS sowie Oberflächenmodelle, die beispielsweise über Dreiecksmaschen repräsentiert werden [SESTER 2001]. Im Folgenden soll nur auf Rasterdaten und Vektordaten eingegangen werden. Vektordaten im Skalenraum werden wegen der Ähnlichkeit der Modellgeneralisierung zur Zielsetzung dieser Arbeit behandelt. Im Anschluss werden die Eigenschaften der multiskaligen Ansätze für beide Datentypen verglichen und diskutiert, ob und inwieweit eine Verwendung der Methoden zur Modellgeneralisierung für die Anpassung von Bildanalyse-Modellen an eine niedrigere Bildauflösung möglich ist.

2.2.1 Vektordaten im Skalenraum

Eine Karte entspricht einem spezifischen Modell der Realität. Früher dienten gedruckte Karten als zweidimensionales Abbild der Welt. Heute werden in der digitalen Kartographie die Landschaftsobjekte (z. B. Straßen, Ortschaften, Wälder, Flüsse, etc.) als *Digitale Landschaftsmodelle* (*DLM*) durch Vektordaten repräsentiert. Zu jedem Objekt im DLM werden die räumliche Lage in Form von Koordinaten, der geometrische Typ (Punkt, Linie, Fläche) und beschreibende Attribute (z. B. Straßenname, Fahrbahnbreite) gespeichert. Es besteht die Möglichkeit zweidimensionale (druckbare) Karten als Rasterdaten aus den DLMs zu generieren.

Mit zunehmend kleinerem Maßstab wird die geometrisch exakte Wiedergabe eines Objektes in einer Karte erschwert, da das Objekt schrumpft. Der Kartenmaßstab muss daher das kartographische Erscheinungsbild zwangsläufig beeinflussen. Wenn das *Prinzip der Lesbarkeit* auch unterhalb der graphischen Mindestgröße erhalten bleiben soll, muss für kleinere Maßstäbe das *Prinzip der geometrischen Richtigkeit* eingeschränkt werden. Wenn auf die Wiedergabe des Objektes in der Karte ganz verzichtet wird, wird außerdem das *Prinzip der Vollständigkeit* verletzt. Die Abwägung dieser notwendigen Einschränkungen bei der Kartenerstellung charakterisiert die Generalisierung in der Kartographie [HAKE et al. 2002].

Grundsätzlich kann zwischen *Objektgeneralisierung* und *kartographischer Generalisierung* unterschieden werden. In der digitalen Kartographie umfasst die Objektgeneralisierung, die die Generalisierung vom Objekt zum DLM vornimmt, sowohl die *Erfassungsgeneralisierung* als auch die *Modellgeneralisierung*. Die kartographische Generalisierung dagegen führt zu digitalen kartographischen Modellen oder Folgekarten. Sie tritt meist bei der unmittelbaren Herstellung von Folgekarten, z. B. aus der Grundkarte, von Karte zu Karte (analog oder digital) auf, manchmal jedoch auch bei der Erstellung kleinmaßstäbiger Karten aus Luftbildern. Hier sind vor allem semantische und graphikbedingte Einschränkungen für den Gestaltungsprozess der Karte ausschlaggebend. Abbildung 2.6 skizziert die Zusammenhänge der verschiedenen Typen der Generalisierung.

Während die Erfassungsgeneralisierung das Objekt selbst betrifft, bei der bereits bei der Erstellung eines digitalen kartographischen Objektmodells die Daten gezielt ausgewählt und verändert werden (oft auf der Grundlage des Objektartenkatalogs), wird bei der Modellgeneralisierung ein vorhandenes Modell für einen kleineren Maßstab mit geringerer semantischer und geometrischer Auflösung aufbereitet. Die Modellgeneralisierung zielt damit auf eine Vermeidung einer erneuten Erfassung von Geodaten für einen kleineren Kartenmaßstab [HAKE et al. 2002]. Sie wird auch als Datenbankgeneralisierung bezeichnet, da ein reduzierter Datensatz für einen gröberen Maßstab erzeugt wird, der späteren Analyseaufgaben als Grundlage dient [SESTER 2001].

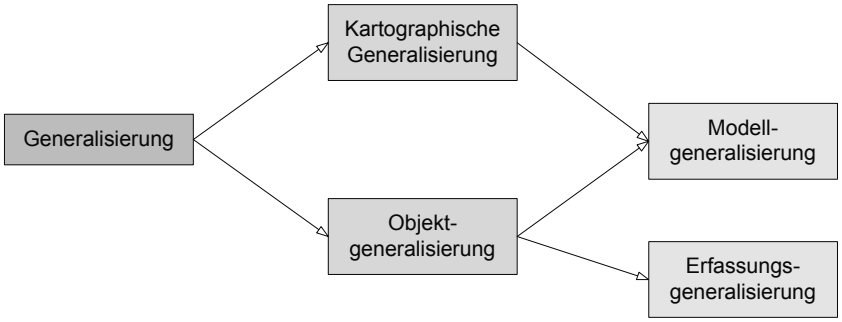


Abbildung 2.6: Typen der Generalisierung in der Kartographie

In der klassischen Kartographie wurde die Generalisierung durch einen kartographischen Experten vorgenommen. Im Zuge des vermehrten Einsatzes automatisierter Generalisierungsverfahren werden verstärkt verbindliche Regeln mit festen Vorgaben für die Generalisierung eingesetzt, die sich beispielsweise aus den Objektartenkatalogen der jeweiligen Maßstäbe ergeben. In der bisherigen Praxis wird jedoch meist eine Mischung aus beiden Ansätzen angewendet [HAKE et al. 2002].

Ein besonderes Konzept zur Speicherung digitaler Karten in unterschiedlichen Maßstäben verfolgt die *Multiple Resolution/Representation Data Base* (*MRDB*), die es erlaubt, Daten unterschiedlicher Maßstäbe miteinander zu verknüpfen sowie unterschiedliche Repräsentationen eines Objektes (z. B. DLM und DTK) vorzuhalten [SESTER et al. 2003]. Die individuellen Objekte in den unterschiedlichen Maßstabsebenen sind über explizite Links miteinander verbunden, die zusätzlich Informationen bzgl. des Übergangs der Objekte zwischen den Maßstäben enthalten und damit zur direkten Erzeugung neuer generalisierter Datensätze genutzt werden können [DEVOGELE et al. 1996, SESTER et al. 1998]. Nach MENG [2000] ist der angestrebte Ablauf der Generalisierung innerhalb einer MRDB durch eine klare Trennung zwischen Modellgeneralisierung und kartographischer Generalisierung charakterisiert. Die Modellgeneralisierung fungiert eher als Vorstufe für die anschließende kartographische Aufbereitung zu den jeweiligen Karten nach visuellen Kriterien. Abbildung 2.7 verdeutlicht das Konzept zur Ableitung von DLMs in kleineren Maßstäben aus einem Basis-DLM mittels Methoden der Modellgeneralisierung sowie deren anschließende getrennte kartographische Visualisierung.

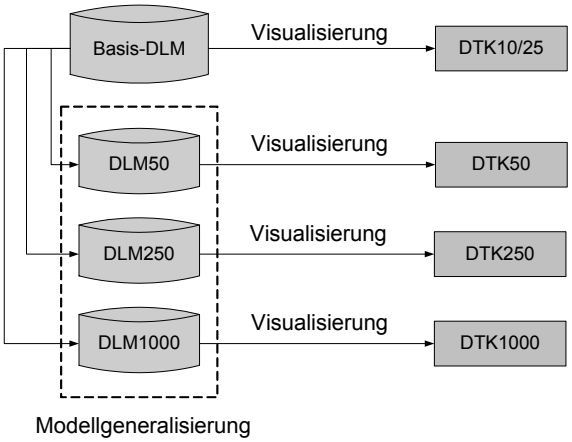


Abbildung 2.7: Konzept zur Modellgeneralisierung für eine MRDB, nach [MENG 2000]

Modellgeneralisierung stellt Methoden für den Übergang von DLMs zwischen verschiedenen Maßstäben bereit. Ziel der Generalisierung in der Kartographie ist vorrangig eine lesbare Karte in einem gröberen Maßstab abzuleiten, nicht die möglichst detailgetreue Wiedergabe und Beschreibung des repräsentierten Objektes im Bild, die in Bildanalyse-Objektmodellen angestrebt wird. Sowohl die kartographische Generalisierung als auch die Modellgeneralisierung kann vor allem das Prinzip der geometrischen Richtigkeit verletzen, um die Lesbarkeit der Karte sicherzustellen.

Ein weiterer wichtiger Unterschied zu den Anforderungen in der Wissensrepräsentation ist die Umgebung, also die Szene, in der das Objekt von Interesse im Bild wiedergegeben wird. Dabei können sehr viele andere Objekte in der

Nachbarschaft des eigentlich interessierenden Objektes auftreten, die dessen Extraktion erschweren. Bildanalyse-Verfahren müssen auch unter diesen Umständen zuverlässige Ergebnisse liefern, so dass oftmals eine umfangreichere Modellierung einer Szene nötig ist. GIS-Objekte aus einem DLM dagegen können als eigenständige vorher definierte Objekte betrachtet werden, die ausschließlich mit anderen, jedoch bekannten, GIS-Objekten in Beziehung stehen. Die exakten Lagebeziehungen der einzelnen GIS-Objekte untereinander sind gut bekannt, während in fernerkundlichen Bilddaten realer Szenen die Lagebeziehungen zu anderen Nachbarschaftsobjekten stark variieren können. Dem muss das Bildanalyse-Objektmodell Rechnung tragen. Übertragen auf das Konzept der Semantischen Netze kann die Modellgeneralisierung als maßstabsabhängige Anpassung eines durch die GIS-Objekte gebildeten Instanzennetzes betrachtet werden, während das Ziel dieser Arbeit der auflösungsabhängigen Anpassung eines Konzeptnetzes entspricht, in dem die konkreten Ausprägungen der modellierten Objekte variieren können.

2.2.2 Bilddaten im Skalenraum

2.2.2.1 Prinzip des Skalenraumes

Jedes Bild entsteht aus einer physikalischen Beobachtung mit einem durch die Aufnahmekonstellation gegebenen Maßstab. Der Maßstab als eine Eigenschaft des Bildes entspricht der Größe der kleinsten Struktur, die im Bild eindeutig aufgelöst werden kann. Dieses Maß wird auch als *innere Skala* bezeichnet, während die Gesamtgröße des Bildes der *äußeren Skala* entspricht [KOENDERINK 1984, FLORACK et al. 1992]. Skalenräume repräsentieren ein gegebenes Bild in verschiedenen inneren Skalen in einer auf einem eindimensionalen kontinuierlichen Skalenparameter $t \in \mathbf{R}_+$ basierenden Familie von abgeleiteten Signalen bzw. Bildern. Die Skalenraumebene mit dem Parameter $t=0$ entspricht dem Originalbild. Mit steigendem Skalenparameter t werden die Bilder immer stärker geglättet. Alle Bilder eines Skalenraumes sind aus derselben Anzahl Pixel gebildet, d. h. alle Bilder sind gleich groß.

Skalenräume und Bildpyramiden

Bildpyramiden dagegen bestehen aus einer Serie von Bildern unterschiedlicher Pixelanzahl in jeder einzelnen Pyramidenstufe, die aus einem Originalbild abgeleitet werden. Bei der klassischen Bildpyramide, wie sie erstmals von Adelson und Burt 1984 vorgeschlagen wurde [ADELSON et al. 1984], entspricht die Bildgröße der jeweils nächsten Stufe nur noch 1/4 der vorhergehenden feineren Stufe. Der gesamte Speicherbedarf einer Bildpyramide ist daher nur um 1/3 im Vergleich zum Originalbild erhöht. Die nächste Stufe der Bildpyramide wird aus der vorhergehenden Bildstufe durch Faltung mit einem Glättungsfilter und einer anschließenden REDUCE-Operation gebildet, in der nur noch die Pixel jeder zweiten Zeile und Spalte für das neue Bild verwendet werden. Die Tiefpassfilterung soll Aliasing bei der nachfolgenden Abtastung vermeiden. ADELSON et al. [1984] verwenden ein Binomialfilter B^4 der Größe 5x5. Wie Jähne jedoch zeigen konnte, ist für die Einhaltung des Abtasttheorems (siehe Abschnitt „Skalenebene und Bildauflösung“) mindestens ein Binomialfilter B^{16} mit einer Filtergröße von 17x17 anzustreben [JÄHNE 1991].

Die Bilder einer Bildpyramide besitzen aufgrund der geringeren Anzahl der Bildpixel unterschiedliche Auflösungen. Der Begriff *Bildauflösung* trifft nur auf diskrete Signale zu, wohingegen *Maßstab* oder *Skala* sowohl auf diskrete als auch auf kontinuierliche Signale zutreffen kann. Damit können Skalenräume für diskrete und kontinuierliche Signale erstellt werden. Bilder in Skalenräumen können durch geeignete Abtastung in eine niedrigere, der Glättung entsprechenden, Auflösung überführt werden. Die Stärke der Tiefpassfilterung ist durch den Skalenparameter t gegeben und bestimmt die im Bild enthaltenen Frequenzen, die durch die Filterung eliminiert werden. Die äquivalente Bildauflösung ist daher von der Größe des Skalenparameters t direkt abhängig. Den Unterschied zwischen Skalenraum und Bildpyramiden skizziert Abbildung 2.8.

Als Bildauflösung wird weiterhin auch eine Angabe zur Anzahl der Pixel für eine Rastergrafik bezeichnet. Den Grafiken werden die jeweiligen Werte zur Anzeige zugewiesen mit Einheiten wie dpi (dots per inch) oder ppi (pixel per inch). Diese Maße betreffen jedoch die relative Dichte der Bildpunkte auf Ein- und Ausgabegeräten, wie z. B. Scanner, Drucker, Bildschirme, Belichter, und haben keinen metrischen Bezug zu aufgenommenen Objekten in den Bildern.

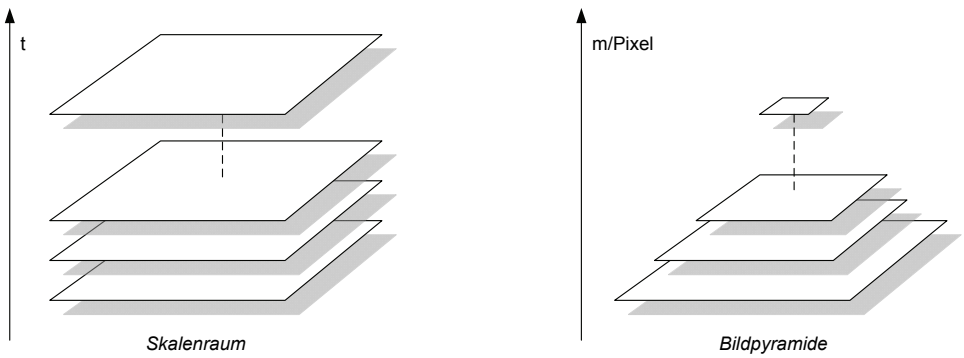


Abbildung 2.8: Prinzip des Skalenraumes im Vergleich zur Bildpyramide

Skalenräume und Fraktale

Einen anderen Ansatz zum Umgang mit Daten in verschiedenen Skalen verfolgt das Konzept der *Fraktale*, das von Mandelbrot intensiv erforscht wurde, z. B. [MANDELBROT 1977, 1987]. Ein Fraktal ist eine Figur, die aus Teilen gebildet wird, die in ihrer Form ähnlich zur Gesamtform selbst sind – diese Figuren besitzen eine Selbstähnlichkeit. Für Fraktale gibt es keine charakteristische Größe, sondern diese Figuren gehen bei allen Verkleinerungen (bzw. Vergrößerungen) in sich selbst über. Viele fraktale Formen sind skaleninvariant. Diese Fraktale besitzen in allen Größenbereichen denselben Grad an Irregularität und Zersplitterung [MANDELBROT 1987]. Viele physikalische Prozesse und Strukturen sind selbstähnlich und lassen sich durch eine bestimmte Klasse von nichtdifferenzierbaren Funktionen sehr gut modellieren. In der Natur finden sich daher viele selbstähnliche Strukturen: beispielsweise Dünen, Farnblätter oder Fjorde an der Küste, deren Unterarme sehr ähnlich zum Hauptarm sein können. Bildstrukturen in Fernerkundungsdaten lassen sich häufig durch fraktale Maße beschreiben. Fraktale Maße beruhen jedoch oft auf unterschiedlichen Definitionen, die einen Vergleich der einzelnen Größen unmöglich machen.

Überblick über Skalenräume

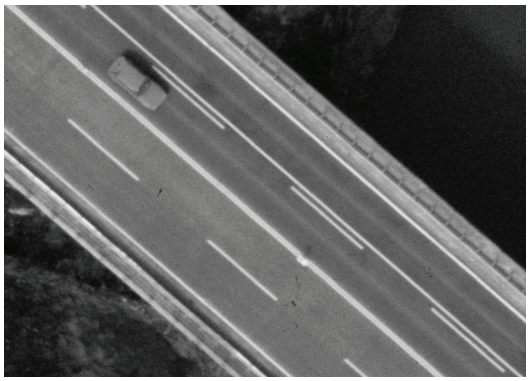
Es existiert eine Vielzahl verschiedener Typen von Skalenräumen, die unterschiedlich gebildet werden und sich in ihren Eigenschaften unterscheiden. Wichtige Eigenschaften der Skalenräume betreffen die Homogenität, d. h. die Ortsunabhängigkeit, sowie die Kausalität, die fordert, dass sämtliche Strukturen in höheren Skalenebenen ihren Ursprung im Originalbild haben müssen. So existieren neben dem Linearen Skalenraum [WITKIN 1983, KOENDERINK 1984] noch der Morphologische Skalenraum [CHEN & YAN 1989, VAN DEN BOOMGAARD & SMEULDERS 1994] und der Anisotrope Skalenraum [PERONA & MALIK 1990]. Weitere Skalenräume sind z. B. der Reaktions-Diffusionsraum [KIMIA et al. 1995, MAYER 1998B], der affin-invariante Skalenraum [SAPIRO & TANNENBAUM 1993] und der Skalenraum, der aus dem Beltrami Flow [SOCHEN et al. 1998, KIMMEL et al. 2000] gebildet wird. Die einzelnen Skalenräume unterscheiden sich insbesondere durch die Art der Glättung bzw. die Bedeutung des Skalenparameters. Abbildung 2.9 illustriert einen Vergleich der Wirkung der unterschiedlichen Glättung einiger Skalenräume. Die drei erstgenannten Skalenräume werden hier kurz vorgestellt:

1. *Linearer Skalenraum*: Im Linearen Skalenraum ergibt sich die Glättung aus der Dämpfung hoher Bildfrequenzen bei der Tiefpassfilterung mit dem Gauß-Kernel. Die Standardabweichung σ bestimmt dabei die Bandbreite der unterdrückten Frequenzen. Der Skalenparameter t entspricht der Varianz der Gauß-Funktion ($t = \sigma^2$). Im Frequenzraum behält die Gauß-Funktion ihre Form, die Standardabweichung verhält sich jedoch reziprok. Je größer σ bzw. t, desto kleiner sind die Frequenzen, die durch die Filterung nicht unterdrückt werden und desto stärker ist die resultierende Bildglättung. Der Lineare Skalenraum bewirkt eine gleichmäßige kontinuierliche Glättung der Bildfunktion, die von der Ausdehnung und dem Grauwert der Objekte gebildet wird. Die Veränderung eines Objektes im Linearen Skalenraum erfolgt daher abhängig von seiner Größe und Helligkeit [MAYER 1998B].

2. *Anisotroper Skalenraum*: Der Lineare Skalenraum verwischt bei zunehmender Glättung Kanten, die meist Objektgrenzen darstellen. Der Anisotrope Skalenraum wird analog zur Idee der anisotropen Diffusion [PERONA & MALIK 1990] erzeugt. Die Glättung des Ausgangsbildes erfolgt hier kantenerhaltend. Kantenbereiche sind durch Grauwertunterschiede mit einem hohen Grauwertgradienten gekennzeichnet, die möglichst wenig Glättung erfahren sol-

len, während innerhalb homogener Flächen im Bild die Glättung verstärkt werden soll. Dazu wird die Diffusionsgleichung um eine Gewichtungsfunktion *c(x,y,t)* zur anisotropen Diffusionsgleichung erweitert. Die Funktion *c* wird abhängig von der Stärke des Gradienten im Bild definiert und bewirkt eine geringe Glättung an Orten mit hohem Gradienten (Kanten) und eine stärkere Glättung bei kleinem Gradienten (homogene Flächen). Bei dieser Definition wird auch die Eigenschaft der Kausalität beibehalten, jedoch nicht die der Homogenität.

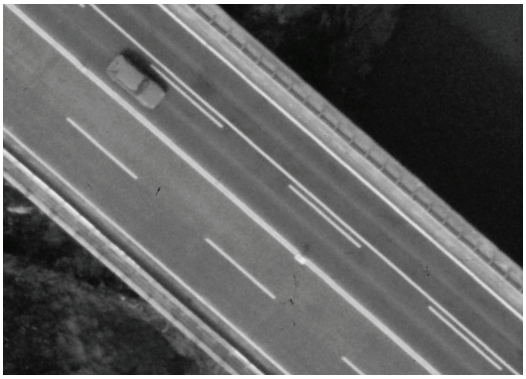
3. *Morphologischer Skalenraum*: Dieser Skalenraum basiert auf mathematischer Morphologie. Er wird gebildet aus einer Folge von *Opening*- und *Closing*-Schritten angewendet auf das Ausgangsbild, die sich aus den morphologischen Grundoperationen *Erosion* und *Dilatation* zusammensetzen. Die Stärke der Glättung des Ausgangsbildes ist von der Größe des verwendeten Strukturelementes abhängig. Für die mathematische Definition wird das Strukturelement durch die strukturierende Funktion *g(x,y)* ersetzt. Eine Besonderheit des Morphologischen Skalenraumes ist, dass der Skalenparameter auch negativ sein kann [MAYER 1998B]. Ein positiver Wert bewirkt ein Closing, ein negativer Wert ein Opening. Objekte im Bild werden durch die Glättung im Morphologischen Skalenraum je nach Verhältnis ihrer Größe zur Größe des Strukturelementes nicht verändert oder fallen vollständig weg. Dabei ist ausschließlich die geometrische Ausdehnung der Objekte für die Glättung von Bedeutung. Wenn das Strukturelement größer ist, wird ein Objekt unabhängig von seinem Grauwertunterschied zum Hintergrund eliminiert.



Originalbild (0.03 m/Pixel)

1.

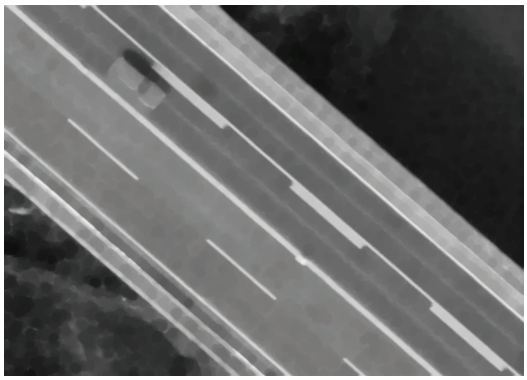
Linearer Skalenraum (σ=7)

2. 

Anisotroper Skalenraum

3a.

Morphologischer Skalenraum (Opening *s*=-7)

3b. 

Morphologischer Skalenraum (Closing *s*=7)

Abbildung 2.9: Luftbild in verschiedenen Skalenräumen

Der Lineare Skalenraum garantiert eine gleichmäßige kontinuierliche Glättung. In den anderen Skalenräumen findet eine ungleichmäßige Glättung statt (insbesondere im Anisotropen Skalenraum). Im Morphologischen Skalenraum ist ausschließlich die Größe der Objekte für die Glättung ausschlaggebend. Die Eigenschaften der unterschiedlichen Glättung können für viele Anwendungen in der Bildanalyse sehr vorteilhaft sein, z. B. bei der Extraktion von lang gestreckten Flächen oder Objektgrenzen über Kanten im Bild. Selbst bei sehr starker Glättung im Anisotropen Skalenraum bleiben die Kanten erhalten (bei geeigneter Wahl der Gewichtungsfunktion *c*) [BLACK et al. 1999], der daher zur selben Segmentierung in allen Skalenebenen führt. Diese Verfahren sind bedeutend für die semantische Beschreibung von Bildern. Die Bedingungen beim Aufnahmeprozess optischer Bilddaten berücksichtigen jedoch keine Semantik, sondern ausschließlich die Bandbegrenzung der Optik. In der vorliegenden Arbeit soll das Verhalten von Landschaftsobjekten in Bilddaten unterschiedlicher Auflösungen prädiziert werden. Daher muss der Aufnahmeprozess für Bilder in niedrigerer Auflösung so gut wie möglich approximiert werden.

Für die Untersuchung des Verhaltens der Objekte beim Übergang von der hohen zur niedrigeren Auflösung in dieser Arbeit wird der Lineare Skalenraum aufgrund der Analogie seiner Glättungseigenschaften zu den physikalischen Gegebenheiten bei der Bildaufnahme gewählt. Der Aufnahmeprozess eines Objektes durch ein bildgebendes System entspricht im Wesentlichen einer Tiefpassfilterung in der Optik und anschließend einer Digitalisierung an 2D regelmäßigen Abtaststellen des geglätteten kontinuierlichen Signals der Reflektion des Objektes. Das resultierende Bild ist mit verschiedenen Abbildungsfehlern behaftet. Die Einflüsse der Abbildungsfehler können mit Hilfe der Modulationsübertragungsfunktion (MTF) mathematisch beschrieben werden. Die MTF für Linsensysteme optischer Sensoren mit kreisförmiger Apertur wird im Allgemeinen über Bessel-Funktionen approximiert [CASTLEMAN 1996]. Die Gauß-Funktion ähnelt sehr stark der Bessel-Funktion 1. Ordnung, vor allem für $\sigma > 1$ [TER HAAR ROMENY 2002]. Auch als optimaler Kernel für eine diskrete Version der Gauß-Funktion zur Generierung des diskreten Linearen Skalenraumes wird eine modifizierte Bessel-Funktion 1. Ordnung empfohlen [LINDEBERG 1990]. Der Lineare Skalenraum ist deshalb für die Zielsetzung dieser Dissertation am besten geeignet. Der Grund ist die durch die Gauß-Funktion garantierte gleichmäßige Glättung der Bildfunktion, die die physikalischen Eigenschaften beim Aufnahmeprozess von optischen Bilddaten geeignet approximiert. Im Abschnitt 2.2.2.2 werden die Theorie des Linearen Skalenraumes und ihre Anwendung im Bereich der Bildanalyse detailliert vorgestellt.

Skalenebene und Bildauflösung

Das mit dem Gauß-Filter geglättete Bild enthält sämtliche Informationen, die auch in einem Bild der korrespondierenden Bildauflösung vorhanden wären. Das geglättete Signal kann jedoch noch zur Simulation eines in einer niedrigeren Auflösung aufgenommenen Bildes abgetastet und so auf die endgültige Größe (Pixelanzahl) verkleinert werden. Die kleinstmögliche Abtastung eines bandbegrenzten Signals gibt das *Abtasttheorem* vor. Danach kann eine Struktur im Signal nur dann richtig rekonstruiert werden, wenn sie pro Wellenlänge mehr als zweimal abgetastet wird [SHANNON 1949]. Wenn diese Abtastrate unterschritten wird, kann im erzeugten Signal Aliasing auftreten und das ursprüngliche Signal falsch wiedergegeben werden. Die kürzeste Wellenlänge bzw. die höchste Frequenz im Signal bestimmt daher die kleinstmögliche (mindeste) Abtastrate und damit die Bildauflösung. Die *Grenzwellenzahl* k_{Grenz} ist die maximale Wellenzahl, die fehlerfrei abgetastet werden kann [JÄHNE 1992]. Die korrespondierende Abtastweite entspricht der Bildauflösung. Die maximale Abtastweite Δx ist durch die Grenzwellenzahl gegeben:

$$\Delta x < \frac{2\pi}{k_{Grenz}} \tag{2.1}$$

Die Standardabweichung der Gauß-Funktion im Ortsraum σ_{Gauss} verhält sich zu ihrer Standardabweichung im Frequenzraum k_{Gauss} reziprok ($k_{Gauss} = 1 / \sigma_{Gauss}$). Zur Vermeidung von Aliasing bei der Abtastung dürfen nur noch sehr niedrige Frequenzen im geglätteten Bild vorhanden sein. Daher wird eine Grenzwellenzahl k_{Grenz} entsprechend einem Vielfachen $2d$ von σ_{Gauss} im Frequenzraum gewählt, an dem die Gaußfunktion genügend klein geworden ist:

$$k_{Grenz} = 2d \cdot k_{Gauss} \tag{2.2}$$

Damit ergibt sich für eine Auflösungsänderung F_A, die der maximalen Abtastweite Δx entspricht:

$$F_A = \frac{2\pi}{k_{Grenz}} = \frac{2\pi}{2d} \cdot \sigma_{Gauss} \tag{2.3}$$

In dieser Arbeit wurde im entwickelten Verfahren zur automatischen Anpassung von Bildanalyse-Objektmodellen an eine niedrigere Bildauflösung $d=\pi$ zur Berechnung der zur Überführung in die Zielauflösung nötigen Glättung gewählt. Für diese Zahl sollte die Gauß-Funktion genügend klein sein, um Abtastfehler zu vermeiden.

2.2.2.2 Theorie des Linearen Skalenraumes

Die Idee des Linearen Skalenraumes wurde in der englischsprachigen Literatur umfassend erstmals von Witkin und Koenderink für eindimensionale [WITKIN 1983] und zweidimensionale Signale [KOENDERINK 1984] eingeführt. Allerdings gab es laut WEICKERT et al. [1999] bereits seit 1959 in der japanischen Literatur erste Veröffentlichungen zu weit reichenden Untersuchungen des Linearen Skalenraumes.

Der Lineare Skalenraum ist der am häufigsten verwendete Maßstabsraum. Er wird durch Faltung eines *n*-dimensionalen Originalsignals $f : \mathbf{R}^n \to \mathbf{R}$ mit dem Gauß-Kernel $g(\vec{x};t) : \mathbf{R}^n \times \mathbf{R}_+ \to \mathbf{R}$ der zugehörigen Dimension *n* gebildet:

$$g(\vec{x};t) = \frac{1}{(2\pi t)^{n/2}} e^{\frac{-(x_1^2+\ldots+x_n^2)}{2t}} \tag{2.4}$$

Der eindimensionale und der zweidimensionale Gauß-Filter sind nach (2.4) gegeben durch:

$$g(x;t) = \frac{1}{\sqrt{2\pi t}} e^{\frac{-x^2}{2t}} \tag{2.5}$$

$$g(x,y;t) = \frac{1}{2\pi t} e^{\frac{-(x^2+y^2)}{2t}} \tag{2.6}$$

Die Varianz dieser Kernel ist der so genannte *Skalenparameter t*, wobei $t > 0$. Die Standardabweichung σ der zugehörigen Gauß-Funktion entspricht daher $\sigma = \sqrt{t}$.

Das Ergebnis der Faltung des Originalsignals *f* mit dem Gauß-Kernel *g* ist eine Familie von geglätteten Signalen $L : \mathbf{R}^n \times \mathbf{R}_+ \to \mathbf{R}$, die aus dem Ausgangssignal abgeleitet wurden:

$$L(\vec{x};t) = f(\vec{x}) * g(\vec{x};t) \tag{2.7}$$

Der so entstandene Lineare Skalenraum $L(\mathbf{x};t)$ ist eine Lösung der Diffusionsgleichung

$$\frac{\partial L}{\partial t} = \Delta \mathrm{L} \tag{2.8}$$

mit dem Laplace-Operator $\Delta = \nabla^2 = \frac{\partial^2}{\partial x_1^2} + \ldots + \frac{\partial^2}{\partial x_n^2}$ unter der Anfangsbedingung $L_0(\mathbf{x};t=0) = f$.

Die Anfangsbedingung $L_0 = f$ garantiert, dass das Ausgangssignal im Skalenraum für den Skalenparameter $t=0$ erhalten bleibt. Die Diffusionsgleichung ist eine partielle Differentialgleichung und stellt einen Zusammenhang her zwischen der Zeit, in der ein Prozess wirkt, und dem Ort, an dem dieser Prozess wirkt. Sie wird in der Physik auch genutzt, um den Wärmeaustausch in homogenen Medien mit gleichmäßiger Wärmeleitfähigkeit zu beschreiben (Wärmeleitungsgleichung): Ein Stück Blech, das über einen begrenzten Zeitraum an verschiedenen Stellen erhitzt wird, erwärmt sich ausgehend von den Wärmequellen gleichmäßig. Auf diese Weise werden nach und nach die Temperaturunterschiede innerhalb des Blechs ausgeglichen. Mit der Wärmeleitungsgleichung lässt sich der Zeitpunkt, an dem das gesamte Blech dieselbe Temperatur erreicht hat, bestimmen. Die verschiedenen Repräsentationen *L* eines Signals *f* in den einzelnen Maßstabsebenen können daher auch als verschiedene Zeitpunkte *t* im Prozess des Wärmeaustauschs einer ursprünglich gegebenen Temperaturverteilung in dem Medium Bild (für zweidimensionale Signale) angesehen werden.

Wie beispielsweise BABAUD et al. [1986] und FLORACK et al. [1992] zeigen konnten, ist die Gauß-Funktion die einzige Lösung für die Diffusionsgleichung. Im Bereich der Signalverarbeitung wird die glättende Wirkung bei der Verwendung als Filter vielfach genutzt. Die Gauß-Funktion behält bei der Fourier-Transformation ihre Form bei, d. h. die Fourier-Transformierte einer Gauß-Funktion ist wiederum eine Gauß-Funktion. Allerdings ist ihre Amplitude im Frequenzraum reziprok zur Standardabweichung σ: Die Amplitude steigt mit zunehmendem σ im Frequenzraum an. Aufgrund der inversen Proportionalität nimmt die Amplitude des Gauß-Kernels im Bildraum dagegen mit steigendem σ ab. In ihrer Charakteristik als Exponentialfunktion, ihrer Symmetrie und stetigen Abnahme von der Mitte aus (Unimodalität) garantiert die Gauß-Funktion ein „gutartiges Verhalten" („well-behavedness") bei der Filterung mit dem Originalsignal [WITKIN 1986]. Außerdem sind Differenzierung und Integration eines diskreten Bildes im Linearen Skalenraum wohl definiert, da der Gauß-Kernel aufgrund der Linearität der Faltungsoperation statt des diskreten Bildes selbst differenziert werden kann [FLORACK et al. 1992]. Eine Differenzierung des Eingangsbildes ist damit durch die Faltung mit der zugehörigen partiellen Ableitung des Gauß-Filters möglich. Weiterhin ist eine Separierung der einzelnen Dimensionen möglich ($g_{xy}(x,y;t)=g_x(x;t)*g_y(y;t)$), so dass eine Faltung mit einer zweidimensionalen partiellen Ableitung des Gauß-Filters durch eine sequenzielle Faltung der beiden eindimensionalen abgeleiteten Gauß-Kernel vereinfacht und daher effizient implementiert werden kann.

Die ersten beiden Ableitungen des zweidimensionalen Gauß-Filters sind gegeben durch:

$$\frac{\partial g(x,y;t)}{\partial x} = -\frac{x}{2\pi t^2} e^{\frac{-(x^2+y^2)}{2t}}, \quad \frac{\partial g(x,y;t)}{\partial y} = -\frac{y}{2\pi t^2} e^{\frac{-(x^2+y^2)}{2t}}. \tag{2.9}$$

Der Laplace-Operator enthält die zweiten Ableitungen. Die Anwendung des Laplace-Operators auf die Gauß-Funktion wird aufgrund der entstandenen charakteristischen Form auch *Mexican Hat* genannt (vgl. Abbildung 2.10):

$$\Delta g(x,y;t) = \frac{\partial^2 g(x,y;t)}{\partial x \partial x} + \frac{\partial^2 g(x,y;t)}{\partial y \partial y} = \frac{1}{2\pi t^2} e^{\frac{-(x^2+y^2)}{2t}} \left(\frac{x^2+y^2}{t} - 2 \right) \tag{2.10}$$

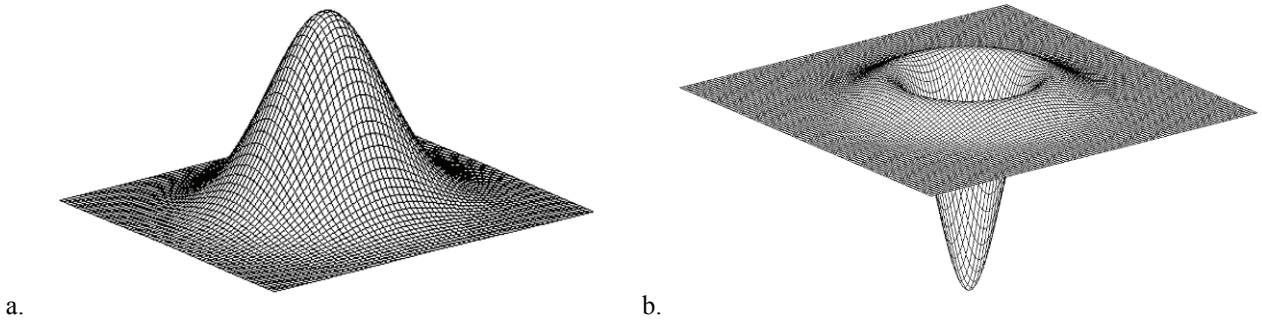


Abbildung 2.10: Gauß-Funktion in 2D (a) und der Laplacian of Gaussian (LoG) Δg (b)

Die Gauß-Funktion ist nicht nur für die Approximation des Aufnahmeprozesses von optischen Bilddaten geeignet (wie in 2.2.2.1 ausgeführt), sondern auch für die Approximation des menschlichen visuellen Systems. Der Mensch nutzt zum Sehen nicht direkt die Signale der einzelnen Sensoren (Stäbchen und Zapfen) in der Retina zur Weiterverarbeitung im Gehirn, sondern eine gewichtete Summe der Signale einer lokalen Nachbarschaft innerhalb der Sensoren, die mit der Gauß-Funktion verschiedener σ modelliert werden [TER HAAR ROMENY 2002]. Erstaunlicherweise ähneln auch die ersten bis vierten Ableitungen des Gauß-Kernels stark der räumlichen Anordnung der Sensoren in der menschlichen Retina [FLORACK et al. 1992], [GRIFFIN & LILLHOLM 2006]. Auch MARR [1982] beschreibt die Ähnlichkeit der Δg-Filterantworten für Kanten und Linien mit elektrophysiologischen Antworten von X-Zellen[6] in der Retina. Diese Phä-

[6] Die größeren physiologischen Zellen der Sensoren der menschlichen Retina werden in zwei Klassen nach ihren zeitlichen Antworten auf einen Reiz eingeteilt: X und Y. Während X-Zellen eine länger anhaltende Antwort geben, zeigen Y-Zellen eine relativ kurze Antwort [MARR 1982].

nomene lassen darauf schließen, dass der Mensch beim Sehen dem Linearen Skalenraum ähnliche Prinzipien zur Wahrnehmung nutzt (wie auch das Beispiel des Steigfluges in der Einleitung zeigt). KOSSLYN [1994] schildert ebenfalls Hinweise, dass das menschliche visuelle System verschiedene Skalen zur Bildinterpretation nutzt.

Eigenschaften des Linearen Skalenraumes

Wichtige Eigenschaften des Linearen Skalenraumes sind [KOENDERINK 1984, LINDEBERG 1994]:

- *Homogenität und Isotropie:*

  Homogenität und Isotropie sagen aus, dass weder Ort noch Richtung im Bild bei der Glättung bevorzugt werden, d. h. die Dämpfung der Bildfrequenzen erfolgt orts- und richtungsinvariant.

- *Kausalität:*

  Die Kausalität fordert für sämtliche auftretende Bildstrukturen in allen Ebenen des Skalenraumes das Vorliegen einer Ursache im Originalbild. Die Kausalität beinhaltet das Kriterium der Nichtverstärkung von Extremwerten beim Übergang von feinen zu gröberen Skalen. Maxima (Minima) in den Grauwerten des Bildes müssen also bei steigendem Skalenparameter t kleiner (größer) werden.

- *Halbgruppeneigenschaft:*

  Die Halbgruppeneigenschaft besagt, dass jedes Bild $L(x,y;t_2)$ im Skalenraum auf einer groben Ebene t_2 auch aus dem Bild $L(x,y;t_1)$ einer feineren Ebene t_1 (wobei $t_2 > t_1$) durch Faltung mit dem Gauß-Kernel $g(x,y;t)$ erzeugt werden kann. Ein Bild einer bestimmten Skalenebene t_2 kann also nicht nur aus dem Originalbild $t=0$ gebildet werden, sondern kann auch aus dem Bild einer anderen, jedoch noch feineren Skalenstufe t_1 durch Faltung erzeugt werden:

$$L(x,y;t_2) = f(x,y) * g(x,y;t_2) = L(x,y;t_1) * g(x,y;t_2 - t_1) \tag{2.11}$$

  Die auf beiden Wegen erzeugten Bilder sind identisch. Diese Eigenschaft des Linearen Skalenraumes wird daher auch Kaskadenglättung genannt. Die Halbgruppeneigenschaft garantiert, dass die Eigenschaften des Skalenraumes zwischen den beiden Skalenebenen eingehalten sind und sämtliche Skalenebenen gleich behandelt werden. Die Halbgruppeneigenschaft skizziert Abbildung 2.11.

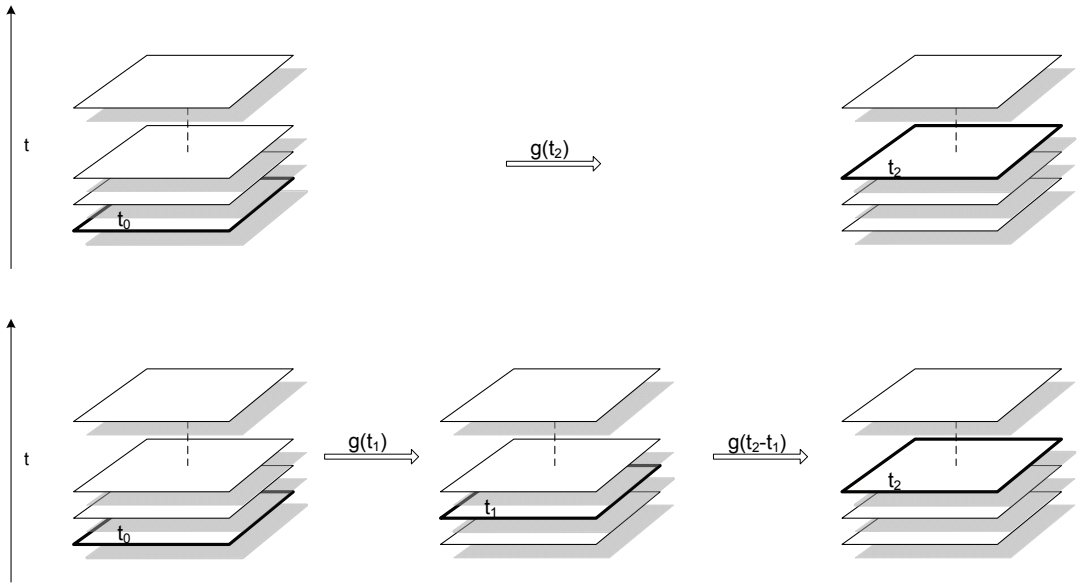


Abbildung 2.11: Halbgruppeneigenschaft des Linearen Skalenraumes: die Faltung des Originalbildes $L(t_0)$ mit dem Gauß-Filter des Skalenparameters t_2 (1.Zeile) ergibt dasselbe Bild im Skalenraum $L(t_2)$ wie die Anwendung des Gauß-Filters des Skalenparameters t_2-t_1 auf das Bild $L(t_1)$ der Skalenebene t_1 (2.Zeile)

Feine Strukturen werden im Linearen Skalenraum mit größerem Skalenparameter stärker unterdrückt. Eine Forderung des Linearen Skalenraumes ist, dass gröbere Ebenen einer Vereinfachung der Bildstruktur entsprechen, die durch Abstraktion des Bildinhalts erzielt wird. Wenn der Skalenparameter t groß genug wird, d. h. das Bild stark genug geglättet wird, verbleibt für jedes Bild nur noch ein einziges Grauwert-Maximum [LOOG et al. 2001], dessen Intensität allerdings sehr klein sein kann.

Eine weitere Eigenschaft des Linearen Skalenraumes ist die Drift von Bildstrukturen. Die Position im Bildraum kann sich in groben Skalenebenen ändern – die Bildstrukturen bewegen sich von ihrem Ursprungsort in der feinen Skala weg. Bei einer starken Glättung kann es weiterhin zu einer Verwischung von Kanten kommen. Aufgrund der starken Dämpfung von hohen Bildfrequenzen erscheinen die Kanten, die durch eine große Intensitätsänderung charakterisiert sind, unscharf, da in den gröberen Skalenebenen nur noch ihre niedrigfrequenten Anteile verbleiben.

Skalenraumereignisse

Beim Übergang zwischen verschiedenen Skalenebenen können in der Bildstruktur so genannte *Skalenraumereignisse* auftreten [KOENDERINK 1984, LINDEBERG 1992]. Diese Veränderungen der Bildstruktur sind diskret in Bezug auf den (kontinuierlichen) Skalenparameter t, d. h. die Änderungen treten plötzlich und komplett zwischen zwei Skalenstufen auf. Bei einem Skalenraumereignis ändert sich auch die Topologie in der Bildstruktur.

Bereits WITKIN [1983] erkannte Ereignisse im Skalenraum für eindimensionale Signale anhand der Kantenposition, d. h. an den Nulldurchgängen der zweiten Ableitung L_{xx}. Abbildung 2.12 illustriert ein Beispiel. Mit steigendem t nähern sich Positionen von Kanten immer schneller an, verschmelzen dann und verschwinden. Witkin interpretierte die Konturen als physikalische Ereignisse, die zur qualitativen Beschreibung der Bildstruktur beitragen.

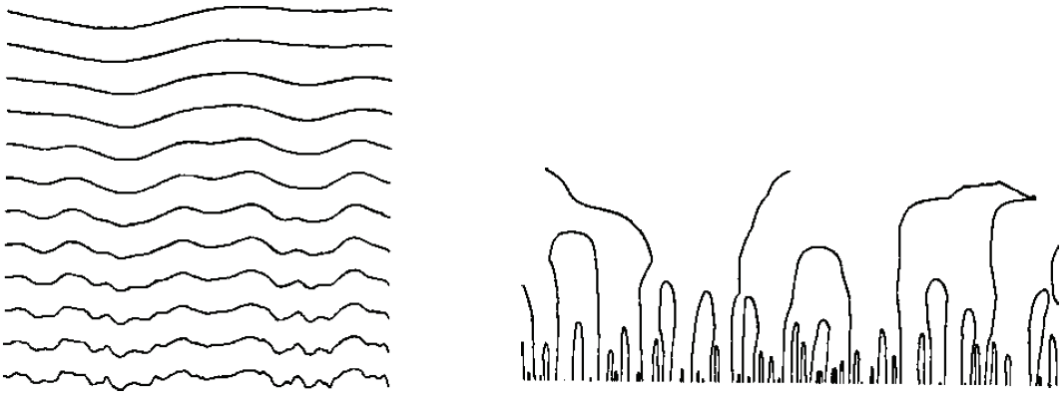


Abbildung 2.12: Nulldurchgänge der zweiten Ableitung (rechts) eines eindimensionalen Signals (links) im Linearen Skalenraum, x-Achse horizontal, t vertikal ansteigend; aus [WITKIN 1983]

Skalenraumereignisse in 2D sind nicht durch einfache Nulldurchgänge wie im eindimensionalen Fall beschreibbar, da die Nulldurchgänge der zweiten Ableitungen L_{xx} und L_{yy} Kurven im (2D-)Bild bilden. Skalenraumereignisse im zweidimensionalen Raum sind jedoch für klar abgegrenzte Objekte handhabbar. Daher werden *Blobs* als homogene Regionen im Bild definiert, die heller bzw. dunkler als ihre Umgebung sind [LINDEBERG 1993]. Das Konzept der Blobs verdeutlicht Abbildung 2.13. Blobs werden gebildet aus einem Paar kritischer Punkte (bzw. Regionen im diskreten Bild) der zweidimensionalen Bildfunktion – einem lokalen Extremum E und einem begrenzenden Sattelpunkt S. Ein Blob $B(E)$ ist daher ein dreidimensionales Objekt mit Ausdehnungen im Bildraum x,y und Grauwert z. Die Ausdehnungen im Bildraum sind durch seine *Basisregion Basis(B)* definiert. Der Grauwert des begrenzenden Sattelpunktes S entspricht dem Basiswert des Blobs $z_{Basis}(B)$. Die Basisregion bildet sich um das Extremum E aus den Bildpunkten mit Grauwerten, die gleich groß oder größer als der Basiswert sind. Das *Blobvolumen Vol(B)* ist definiert durch das Integral der Bildfunktion über die Basisregion. Der *Blobkontrast K(B)* ist gegeben durch die Grauwertdifferenz zwischen Basiswert $z_{Basis}(B)$ und lokalem Extremum E.

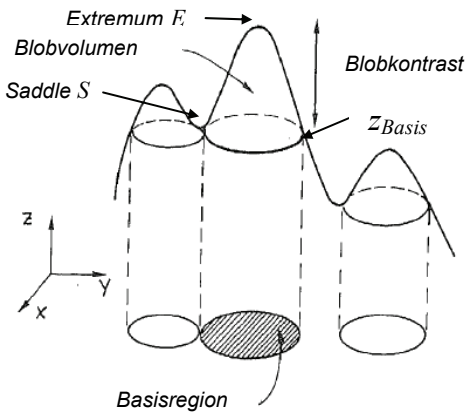


Abbildung 2.13: Definition von Blobs; adaptiert von [LINDEBERG 1993]

LINDEBERG [1994] stellt einen sequentiellen Algorithmus zur Abgrenzung von Blobs in zweidimensionalen diskreten Grauwertbildern vor:

1. Angrenzende Pixel desselben Grauwertes werden zu zusammenhängenden Regionen gruppiert. Ergebnis ist ein Labelbild, das allen Regionen einen eindeutigen Wert zuweist.
2. Alle Regionen werden nach ihren Grauwerten aufsteigend sortiert.
3. Für jede Region wird eine Liste aller angrenzenden Nachbarregionen mit einem höheren Grauwert erzeugt.
4. Die Regionen werden zu Blobs zusammengefasst. Lokale Grauwert-Maxima dienen als Saatpunkte der Blobs, denen nacheinander alle Nachbarregionen mit niedrigerem Grauwert zugeordnet werden. Eine Region wird einem benachbarten Blob zugeschlagen, wenn sämtliche benachbarte Regionen mit höherem Grauwert demselben Blob angehören. Falls die angrenzenden höheren Nachbarregionen einer Region unterschiedlichen Blobs angehören, kann die untersuchte Region nur ein geteilter Sattel dieser Blobs sein und wird als Hintergrund eingestuft. Eine Region wird ebenfalls dem Hintergrund zugeordnet, wenn mindestens eine höhere Nachbarregion bereits als Hintergrund gekennzeichnet wurde.

Sattelpunkte entsprechen nach der Definition dem Basiswert des Blobs und werden vom Algorithmus als nicht zum Blob zugehörig betrachtet. Ein Sattelpunkt ist allerdings mit dem Blob, den er begrenzt, assoziiert. Die abgegrenzten Blobs umfassen daher ausschließlich zusammenhängende Pixel mit einem höheren Grauwert als der Sattelpunkt. Der obige Algorithmus grenzt helle Objekte als Blobs vom Bildhintergrund ab. Alternativ können zu den lokalen Maxima auch die lokalen Minima als Saatpunkte dienen. In diesem Fall stellen die abgegrenzten Blobs dunkle Bildobjekte dar.

Für Blobs werden folgende Skalenraumereignisse unterschieden [LINDEBERG 1994]:

1. *Auslöschung*: Ein Blob verschwindet.
2. *Verschmelzung*: Zwei oder mehrere Blobs verschmelzen zu einem einzigen Blob.
3. *Erzeugung*: Ein neuer Blob erscheint.
4. *Spaltung*: Ein Blob spaltet sich in zwei oder mehrere Blobs.

Im eindimensionalen Raum treten nur Auslöschung und Verschmelzung auf, jedoch keine Erzeugung und Spaltung. Die beiden letztgenannten Ereignisse können im zweidimensionalen Raum, d. h. in Bildern, erscheinen, jedoch nur selten, wie auch experimentell gezeigt werden konnte [KUIJPER & FLORACK 1999]. Bei Erzeugung und Spaltung werden neue Strukturen in gröberen Skalen scheinbar erzeugt – ein Prozess, der auf den ersten Blick widersprüchlich zum Kriterium der Kausalität scheint. Kausalität fordert jedoch nur, dass sämtliche Strukturen in einer gröberen Skala ihre Ursache im Ursprungsbild haben müssen. Es können auch neue Strukturen im geglätteten Bild auftreten, die aufgrund der Glättung der bereits im Ausgangsbild vorhandenen Strukturen entstanden sind. Ein anschauliches Beispiel für solch einen Fall

zeigt LINDEBERG [1990, 1994] (vgl. Abbildung 2.14). Dabei sind zwei Bergkuppen, die unterschiedlich hoch sind, durch einen schmalen Grat miteinander verbunden. Wenn dieser Grat erodiert wird, bis er schließlich verschwindet, verbleiben zwei getrennte Berge. Diese Kuppen können als Grauwerte in der Bildfunktion angesehen werden; die Erosion wird im Linearen Skalenraum durch die Glättung erreicht. Während anfangs nur ein Blob durch einen Extremwert laut Definition für Blobs vorhanden ist, erscheint mit zunehmender Erosion (t) ein zweiter Blob – in dem Beispiel fand demnach eine Spaltung statt.

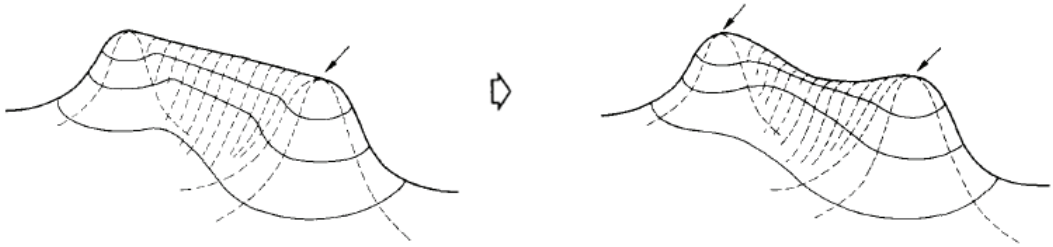


Abbildung 2.14: Spaltung am Beispiel zweier Bergkuppen; adaptiert von [LINDEBERG 1990]

Skalenraumereignisse lassen sich auch vom Standpunkt der *Katastrophentheorie* betrachten. Nach ARNOLD [1992] bezeichnen *Katastrophen* abrupte Änderungen, die durch eine plötzliche Antwort eines Systems auf eine kontinuierliche gleichmäßige Veränderung der externen Bedingungen entsteht. Die Methode der Katastrophentheorie wird sehr vielseitig angewendet, z. B. zur Untersuchung des Herzschlages, für Modelle der Gehirnaktivität oder das Verhalten von Investoren an der Börse [ARNOLD 1992]. Auch das chaotische Verhalten der Ereignisse im Skalenraum in Bezug zu ihrem plötzlichen Auftreten bei kontinuierlicher Veränderung des Skalenparameters t kann in die Sichtweise der Katastrophentheorie eingeordnet werden. Die Extrema der Blobs werden dabei als *kritische Punkte* angesehen, die sich entlang einer *kritischen Kurve* im Skalenraum bewegen. Der steuernde Parameter, der die Katastrophen auslöst, ist dabei der kontinuierliche Skalenparameter t. Ein kritischer Punkt in der Bildfunktion ist ein Extremum (Maximum/Minimum) oder ein Sattelpunkt, an dem der Bildgradient ($\nabla L(x,y;t) = 0$) verschwindet und die Eigenwerte der Hesse-Matrix H nicht negativ sind. Wenn die Determinante der Hesse-Matrix $det\ H \neq 0$ ist, so werden die kritischen Punkte *Morse-kritische Punkte* genannt, da ihre Umgebung nach dem Morse-Lemma durch ein Polynom 2. Grades beschrieben werden kann. Bei $det\ H = 0$ (bei der die Hesse-Matrix also singulär ist) dagegen kann die Umgebung dieser *Nicht-Morse-Punkte* nur über ein Polynom höheren Grades (≥ 3) beschrieben werden. Punkte mit dieser Eigenschaft sind typischerweise isolierte Punkte auf der kritischen Kurve[7]. Ein Nicht-Morse-Punkt kann sich in mehrere Morse-kritische Punkte aufspalten: Dieser Vorgang wird als *Morsifikation* bezeichnet. Der Nicht-Morse-Teil der Bildfunktion, d. h. das Polynom höheren Grades in der Umgebung des Nicht-Morse-Punktes, kann in eine kanonische Form gebracht werden und wird dann *Katastrophenkeim* genannt. Eine Katastrophe entspricht im Kontext der Theorie des Linearen Skalenraumes einem Skalenraumereignis und entsteht immer aus zwei kritischen Punkten mit gegensätzlichen Signaturen der Hesse-Matrix, die sich kurz vor bzw. nach der Katastrophe äußerst schnell gegenseitig annähern oder voneinander entfernen [FLORACK & KUIJPER 2000]. In 1D entsprechen diese zwei kritischen Punkte immer einem Maximum und einem Minimum. Im 2D-Raum dagegen entsprechen diese beiden gegensätzlichen kritischen Punkte einem Extremum (Maximum oder Minimum mit positiver Hesse-Signatur) und einem Sattelpunkt (negative Signatur). Ein Skalenraumereignis kann sogar mit Subpixel-Genauigkeit über die Determinante der Hesse-Matrix und den Vektor der Skalenraumdriftgeschwindigkeit im Bildraum lokalisiert werden [KUIJPER & FLORACK 1999]. Die Skalenebene t für ein bestimmtes Skalenraumereignis kann jedoch nur experimentell durch schrittweise Untersuchung aufeinander folgender Skalenebenen geschätzt werden.

Nach der Katastrophentheorie existieren ausschließlich zwei Arten von Katastrophen (Skalenraumereignissen): Auslöschung und Erzeugung. Dabei werden die Verschmelzung als eine Form der Auslöschung (eines Blobs) und die Spaltung als Erzeugung (eines weiteren Blobs) betrachtet. Genau genommen handelt es sich bei der Verschmelzung um eine Auslöschung eines Blobs, dessen Sattelpunkt mit einem anderen Blob geteilt wird, und bei der Spaltung um eine Erzeu-

[7] Nicht-Morse-Punkte werden in der Literatur auch häufig als *Top Points* bezeichnet.

gung eines neuen Blobs mit geteiltem Sattelpunkt. Grundsätzlich werden generische und nicht-generische Katastrophen unterschieden: Generische Katastrophen sind aus den Daten erklärbar, gewissermaßen vorhersehbar; für nicht-generische Katastrophen gilt dies jedoch nicht. Die einzigen generischen Katastrophen im Skalenraum sind Auslöschung und Erzeugung, wie DAMON [1995] zeigen konnte. Alle anderen Ereignisse können durch Zusammensetzungen dieser beiden Formen beschrieben werden. Aufgrund von numerischen Einschränkungen sowie zu grober Abtastung kann es in der Praxis allerdings manchmal unmöglich sein, diese beiden generischen Ereignisse klar voneinander abzugrenzen. Ereignisse im Skalenraum können unter diesen Umständen als nicht-generische Katastrophen erscheinen. Die Morsifikation einiger nicht-generischer Ereignisse im Skalenraum konnte beispielsweise in KUIJPER & FLORACK [2004] über die Trajektorien der kritischen Punkte modelliert werden.

Deep Structure

Sowohl WITKIN [1983] als auch KOENDERINK [1984] forderten, die Bildstruktur in allen Skalen gleichzeitig zu analysieren (*deep structure*). Durch eine simultane Untersuchung der Bildstruktur im gesamten Skalenraum würde das Bild komplett beschrieben. Der Umgang mit der Deep Structure ist jedoch immer noch ungelöst [FLORACK & KUIJPER 2000]. Die komplette Information über die Bildstruktur ist in den Grauwerten der Bilder in verschiedenen Skalen enthalten, jedoch nur implizit.

Eine Form, Bildstruktur im Linearen Skalenraum explizit zu formulieren und die deep structure zu untersuchen, ist der *Scale-Space Primal Sketch* [LINDEBERG 1993, LINDEBERG 1994]. Blobs dienen als Primitive, die Bildmerkmale sowie ihre Beziehungen zwischen verschiedenen Skalen beschreiben. Dieses Konzept wurde als Basis für die Extraktion signifikanter Bildmerkmale entworfen, die in einem möglichst großen Skalenbereich stabil bleiben. Im Allgemeinen lässt sich für einen Blob, der in einer bestimmten Skalenebene existiert, ein zugehöriger Blob in einer geringfügig feineren bzw. gröberen Skalenebene finden. Durch die Verbindung dieser Blobs zwischen verschiedenen Skalenebenen werden so genannte *Skalenraum-Blobs* gebildet. Diese vierdimensionalen Objekte haben eine Ausdehnung jeweils in Bildraum x,y, Grauwert z und Skala t. Allerdings ist eine einfache Verbindung von Blobs angrenzender Skalenebenen nicht immer möglich, da Skalenraumereignisse der Blobs eine Topologieänderung in der Bildstruktur zwischen diskreten Skalenebenen hervorrufen. Aufgrund der Singularität der Hesse-Matrix, die ein Skalenraumereignis in Richtung t verursacht, ist die Ausdehnung eines Skalenraum-Blobs durch die jeweilige Ebene im Skalenraum, in der das Skalenraumereignis auftritt, begrenzt. Ein Skalenraum-Blob wird daher mit einer ihm zugehörigen Lebenszeit im Skalenraum verknüpft, der durch die Differenz seiner Entstehungsebene t_E und der Skalenebene seines Verschwindens t_V gegeben ist. Um nicht-generische Skalenraumereignisse aufgrund mehrdeutiger Verlinkungen der Grauwert-Blobs zwischen aufeinanderfolgenden Skalenebenen zu vermeiden, empfiehlt Lindeberg einen Algorithmus des adaptiven Samplings, der die Dichte aufeinanderfolgender Skalenebenen so lange verfeinert bis alle Blob-Verbindungen zwischen benachbarten Skalenebenen eindeutig den vier Skalenereignissen zugeordnet werden können. Auf diesem Wege wird eine Repräsentation der Bildstruktur über alle Skalen erzeugt. Diese multiskalige Abbildung der Bildstruktur ist hilfreich für die Detektion signifikanter Bildmerkmale in den am besten geeigneten Skalen.

Ein anderes Verfahren zur Erstellung einer multiskaligen Repräsentation der Bildstruktur im Linearen Skalenraum ist die *Skalenraum-Hierarchie*, die von KUIJPER [2002] bzw. KUIJPER et al. [2003] und KUIJPER & FLORACK [2005] entwickelt wurde. Eine wichtige Rolle spielen Skalenraum-Sattel, an denen sowohl die räumlichen Gradienten $\nabla L(x,y;t)$ als auch die Gradienten in Skalenrichtung $\partial L_t(x,y,t)$ verschwinden ($\nabla L(x,y;t) = 0$ und $\partial L_t(x,y;t) = 0$). Da für Skalenraum-Sattel im Linearen Skalenraum die Diffusionsgleichung gilt, kann $\partial L_t(x,y;t) = 0$ durch $\Delta L(x,y;t) = 0$ ersetzt werden, d. h. der Laplace-Operator mit den zweiten Ableitungen im Bildraum verschwindet. In 1D-Signalen entsprechen Skalenraum-Sattel Katastrophenpunkten; dort findet also eine Auslöschung statt. In 2D-Signalen dagegen müssen Skalenraum-Sattel nicht unbedingt mit Katastrophenpunkten zusammenfallen. Die hierarchische multiskalige Datenstruktur wird aus einzelnen Bereichen des Bildes im Skalenraum aufgebaut. Sie werden Skalenraumsegmente genannt. Die Segmente werden von Mannigfaltigkeiten im Skalenraum mit gleicher Intensität wie die Skalenraum-Sattel eingeschlossen. Ein Skalenraumsegment kann mehrere Extrema enthalten und daher wiederum in Untersegmente unterteilt werden. Aus diesem Zusammenhang wird ein Hierarchie-Baum erzeugt. Die Äste repräsentieren die Segmente; die einzelnen Segmentarme treffen sich an den Knoten des Baumes, die den Katastrophenpunkten im Skalenraum entsprechen. Die Skalenraum-Hierarchie kann als Vorsegmentierung des Bildes ohne a priori Information dienen.

2.2.2.3 Anwendung des Linearen Skalenraumes in der Bildanalyse

Bereits vor der formalen Definition des Linearen Skalenraumes wurde der Gauß-Filter zur gleichmäßigen Glättung der Bilder verwendet, um Fehlereinflüsse wie normalverteiltes (weißes) Rauschen weitgehend vor der Weiterverarbeitung der Bilddaten zu beseitigen. So argumentierte MARR [1982], dass die Reflektanz von Objektoberflächen, die ein Bild wiedergibt, von unterschiedlichen Phänomenen erzeugt werden, die in ihren eigenen (unterschiedlichen) Skalen agieren. Beispielweise bildet sich die Oberfläche eines Flusses aus der Überlagerung von flachem Wasser und zufälligen Störungen, wie Steinen im Wasser und von Wind hervorgerufenen kleinen Wellen. Diese intrinsischen Skalen müssen bei der automatischen Interpretation des Bildes berücksichtigt werden. Daher wurde für den Entwurf von Bildverarbeitungsoperatoren zur Merkmalsextraktion der Gauß-Filter zur Analyse verschiedener Skalenebenen eingebunden. Berühmte Beispiele dafür sind der *Laplacian of Gaussian* (*LoG*) [MARR 1982], der dem mit dem Laplace-Operator gefalteten Gauß-Kernel enstpricht (Δg), vgl. auch Abbildung 2.10 b, und der *DoG* (*Difference of Gaussians*). Diese Operatoren arbeiten auf mehreren Skalenebenen zur Unterscheidung von Kanten und Linien, d. h. Orten mit großer Intensitätsänderung und daher hoher Frequenz, von anderen hochfrequenten Anteilen wie Rauschen. Signifikante Kanten und Linien, die kein Rauschen darstellen, sind in verschiedenen Skalenebenen vorhanden, nicht nur in feinen Skalen, sondern auch in gröberen Skalenebenen. Auch die Merkmalsextraktion über Bildpyramiden folgt einer ähnlichen Idee. Hier wird ebenfalls die Abstraktion in gröberen Pyramidenstufen genutzt, um stabile Bildmerkmale zu detektieren. Oftmals werden Merkmale erst in höheren Pyramidenstufen gesucht und dann anschließend in einer näheren Umgebung in die feinste Stufe propagiert und dort exakt lokalisiert. Auf diesem Weg werden Ungenauigkeiten bei der Detektion vermieden, die sich aus der Drift von Bildstrukturen in gröberen Skalen aufgrund der starken Glättung ergeben. In allen diesen Ansätzen sind jedoch die Analysen auf bestimmte Skalenebenen beschränkt, die vorher festgelegt, d. h. bekannt, sein müssen.

Anwendungen in der Computer Vision

Der Lineare Skalenraum wird in der Bildanalyse meist zur (Vor-)Segmentierung des Eingangsbildes (z. B. mit der Skalenraum-Hierarchie), zur Wahl der besten Skala für die Merkmalsextraktion oder zur Merkmalsextraktion selbst eingesetzt. So können mit dem Scale-Space Primal Sketch durch die Berechnung der höchsten Blob-Stärke im Skalenraum die für eine Extraktion von blob-ähnlichen Bildstrukturen am besten geeigneten Skalenebenen gefunden werden [LINDEBERG 1993]. Die besten Skalen für die Extraktion allgemeiner Bildstrukturen können auch über lokale Extrema in Skalenrichtung der normalisierten Gaußschen Ableitungen gefunden werden [LINDEBERG 1998A]. Hier wird der differenzierte Gauß-Kernel mit derjenigen Größe (und damit Skala) ausgewählt, der am ehesten dem Template der Bildstruktur ähnelt. Alternativ können geeignete Skalenebenen für die Extraktion von Kanten und Linien auch über die *Skalenraumkante* bzw. die *Skalenraumlinie* ausgewählt werden [LINDEBERG 1998B]. Kanten und Linien werden hier als eindimensionale Kurven im dreidimensionalen Skalenraum behandelt. Die Stärke der Kanten und Linien wird über ihre räumlichen Ableitungen in den einzelnen Skalenebenen ermittelt. In der Skalenebene, in der die Kante bzw. Linie am stärksten ist, wird schließlich die Extraktion durchgeführt.

Einer der am häufigsten angewendeten Kantenoperatoren ist der *Canny-Operator*. Er wurde entworfen um möglichst optimal verschiedene Qualitätskriterien an die zu extrahierende Kante zu erfüllen. Eine Untersuchung zeigte, dass der optimale Filter nahezu identische Ergebnisse für reale Bilder liefert wie die erste Ableitung des diskreten Gauß-Filters [CANNY 1986]. Auch der in jüngerer Zeit sehr häufig eingesetzte SIFT[8]-Operator zur Detektion stabiler Interestpunkte nutzt den Linearen Skalenraum [LOWE 2004]. Kandidaten für interessante Punkte werden zuerst über Intensitäts-Extrema aus einem Bereich von 3x3 Bildpunkten in der aktuellen und den beiden angrenzenden Skalenebenen gefunden. Aus diesen Kandidaten werden über die Spur und die Determinante der Hesse-Matrix der DoG-Funktion diejenigen Punkte mit wenig Kontrast oder entlang einer Kante eliminiert. Anschließend werden die dominanten Orientierungen der verbliebenen Punkte aus der Stärke der Gradienten in allen Skalenebenen errechnet. Die so gefundenen Interestpunkte sind sowohl rotations- und skaleninvariant als auch bis zu einem gewissen Maß invariant gegenüber Rauschen, Verdeckung und Beleuchtungsänderung. LOWE konnte die erzeugten SIFT-Merkmale u. a. auch erfolgreich zur Objekterkennung durch Zuordnung der gefundenen Merkmale zu einer bekannten Datenbasis einsetzen.

[8] SIFT: **S**cale-**I**nvariant **F**eature **T**ransform

Ein Beispiel zur Segmentierung durch gradientenbasierte Wasserscheidentransformation im Skalenraum zeigt GAUCH [1999]. Aus den Wasserscheiden und Auffangbecken im Skalenraum, die den Objektkanten und Objektregionen entsprechen, wird eine multiskalige Hierarchie der Objekte aufgebaut. Die Pfade der Minima im Bild, die mit den Auffangbecken assoziiert werden, müssen dazu durch den Skalenraum verfolgt werden. Eine Eltern-Kind-Verbindung in der Hierarchie wird durch die Verschmelzung eines Auffangbeckens mit einem angrenzenden Becken definiert. Auf diesem Wege werden die Skalenraumereignisse sämtlicher Minima über eine vorher festgelegte Reihenfolge der Skalenebenen verfolgt bis der komplette Eltern-Kind-Baum als Hierarchie der Wasserscheidenregionen vollständig ist. Erzeugungen neuer Wasserscheidenregionen werden bei dieser Vorgehensweise allerdings ignoriert. Die Wasserscheiden werden über die Grenzen der Auffangbecken ermittelt. Die individuelle Lebenszeit der Wasserscheiden ist daher durch die Hierarchie der Auffangbecken gegeben. Die Signifikanz eines Objektes im Bild kann dann aus der Lebenszeit der Wasserscheiden im Skalenraum abgeleitet werden. Zur Segmentierung des Bildes werden nur die signifikanten Wasserscheiden mit der größten Lebenszeit berücksichtigt.

Die ersten Ableitungen des Gauß-Filters einer bestimmten Skalenebene t werden als *Local Jet* zusammengefasst, der als die ersten Terme der Taylor-Reihenentwicklung des mit derselben Skala t geglätteten Bildes angesehen werden kann, siehe z. B. [FLORACK et al. 1992]. Der Local Jet stellt wichtige Hinweise zur lokalen Bildstruktur bereit. Diese Eigenschaft kann für die Segmentierung des Bildes genutzt werden. In Dimensionen größer als 1 wird dabei auch die Separierbarkeit des Gauß-Kernels ausgenutzt, der eine effiziente Implementierung der Filteroperation erlaubt. Eine relativ einfache Bildsegmentierung auf einer festen Skalenebene kann beispielsweise mit einem Indikator erreicht werden, der aus einer Kombination verschiedener Ableitungen des Gauß-Filters ermittelt wird [FLORACK et al. 1992]. GRIFFIN & LILLHOLM [2006, 2007] zeigen eine Möglichkeit zur qualitativen Bildbeschreibung über den *Jet Space*, einem Raum, der aus möglichen Kombinationen der einzelnen Jets gebildet wird. Aus der Partitionierung des Jet Space kann die qualitative Beschreibung der lokalen Bildstruktur mit Bildmerkmalen gewonnen werden.

Der Lineare Skalenraum findet in vielen weiteren Bereichen der Computer Vision Anwendung: Z. B. zeigt KUIJPER [2004] eine Verbindung des Linearen Skalenraumes mit der Bildorientierung über Mutual Information. Top Points, an denen Skalenereignisse auftreten, werden ebenfalls als Interestpunkte erfolgreich für die Bildzuordnung verwendet [PLATEL et al. 2005]. BRESSON et al. [2006] beschreiben eine Methode zur Segmentierung mit multiskaligen Snakes, die aus dem gesamten Skalenraum die am besten geeignete Skala zur Segmentierung wählen.

Anwendungen in der Landschaftsökologie

Ziel der Landschaftsökologie ist es, ökologische Prozesse zwischen verschiedenen Landschaften besser verstehen und vorhersagen zu können. Für dieses Ziel wird auch die Interpretation von Satellitenbildern eingesetzt. Oftmals werden verschiedene fraktale Indizes zur Bestimmung der Landschaftsstrukturen und -grenzen genutzt. Die ökologisch relevanten Prozesse agieren jedoch oftmals auf sehr unterschiedlichen Skalen, in dem Sinne, dass verschieden große Areale von diesen Prozessen beeinflusst sein können und sich die Landschaftsstrukturen dementsprechend verändern. Diese Skalenabhängigkeit der Prozesse bewirkt die Ausprägung verschiedener Muster in Bilddaten unterschiedlicher Auflösung und damit der Ergebnisse der verwendeten Indizes. Die Faktoren, die diese Skalenabhängigkeit beeinflussen, sind noch nicht vollständig verstanden [KOJIMA et al. 2006]. Zur besseren Modellierung und Erklärung der Prozess-Muster-Verbindung in Abhängigkeit der Bildauflösung bzw. der beobachteten Skala schlagen HAY et al. [2002] einen multiskaligen Ansatz im Linearen Skalenraum vor. Dazu wird der Scale-Space Primal Sketch (nach LINDEBERG [1994]) über Blob-Detektion mit den Skalenraum-Blobs als Elementen erstellt, die als signifikante landschaftsökologische Objekte angesehen werden. Die letztlich extrahierten Objekte entsprechen der Basisregion des Grauwert-Blobs in der Skalenebene mit dem größten Blob-Volumen.

Anwendungen in der Fernerkundung

Da natürliche Objekte aufgrund ihrer physikalischen Beschaffenheit aus Strukturen unterschiedlicher Skalen zusammengesetzt sind, existieren oftmals viele verschiedene lokale Skalen in einem Bild oder einer Szene. Beispielsweise besteht ein Wald aus Bäumen, ein Baum wiederum aus Blättern und Ästen, usw. Jedes Teilobjekt hat seinen eigenen Skalenbereich. Diese lokalen Skalen sind essentiell für die korrekte Interpretation von Luft- und Satellitenbildern. Eine einzige globale Skalenebene kann jedoch nicht sämtliche Informationen abdecken, die für eine korrekte Interpretation

notwendig sind. Daher werden sehr häufig verschiedene Skalen bzw. Bildauflösungen für die Datenauswertung genutzt. Meist werden Operatoren zur Merkmalsextraktion bzw. Segmentierungsverfahren in unterschiedlichen Bildauflösungen angewendet und die Ergebnisse später kombiniert.

Ein Ansatz zur hierarchischen Objektextraktion ist in der kommerziellen Bildanalyse-Software *eCognition* implementiert. Hier wird zwischen Bildauflösung und Skala unterschieden: Eine Skala wird als Grad der Aggregation und Abstraktion definiert, für den ein Objekt beschrieben werden kann und daher das Auftreten und die Abwesenheit einer Objektklasse bestimmt. Dementsprechend basiert *eCognition*'s multiskalige Segmentierung auf einer Aggregationstechnik: In aufeinanderfolgenden Schritten werden einzelne Pixel zu größeren Objekten verbunden bis der Zuwachs eine über den sog. Skalenparameter vordefinierte Größe überschreitet. Das Zusammenfassen benachbarter Regionen wird nach einem Heterogenitätskriterium beurteilt [BENZ et al. 2004].

LAPTEV et al. [2000] verbinden den Linearen Skalenraum mit Snakes zur Extraktion von Straßen aus Luft- und Satellitenbildern. Die Snakes werden hier zur Kantenextraktion der Fahrbahn genutzt. Zuerst werden die Straßen als Linien in einer groben Skalenebene extrahiert, die anschließend eine Ribbon-Snake zur Kantenextraktion aus der feinen Skala initialisieren. Snakes sind in der Lage durch geometrische Bedingungen örtlich begrenzte Störungen in hochauflösenden Bildern, wie Schatten, Fahrzeuge und überragende Bäume, zu überbrücken. Eine möglichst gute Initialisierung der Snake ist jedoch für ein zuverlässiges Ergebnis eine wichtige Voraussetzung. Diese Initialisierung wird hier unter Ausnutzung des Skalenverhaltens von Straßen erreicht: Die Abstraktion von einem flächenhaften Objekt der Fahrbahn in feinen Skalen zu einer Linie in groben Skalenebenen.

Wie in Abschnitt 2.1.2.2 gezeigt, kann auch die Beschreibung eines kompletten Landschaftsobjektes (Straße) in verschiedenen Bildauflösungen in einem Objektmodell zur Steigerung der Evidenz bei der Objektextraktion eingebunden werden. Dafür müssen die unterschiedlichen Darstellungen des Objektes jedoch vorher bekannt sein. MAYER & STEGER [1998] haben dazu analytische Untersuchungen des Skalenverhaltens von Straßen durchgeführt. Hier wurde das Verhalten des Grauwertprofils für eine und zwei Fahrbahnen mit einem hellen bzw. dunklen Fahrzeug im Linearen Skalenraum untersucht. Außerdem wurde das Ergebnis eines Operators zur Linienextraktion [STEGER 1998A] auf den Profilen im Skalenraum analysiert. Empirische Untersuchungen bestätigen die analytischen Berechnungen. Es konnte sogar eine Vorhersage derjenigen Skalenebene getroffen werden, in der zwei Fahrbahnen als Linien miteinander verschmelzen: Ausgehend vom Modell zur Linienextraktion aus dem Straßenprofil konnte die Standardabweichung σ, d. h. der Grad der nötigen Glättung, für die Verschmelzung analytisch berechnet werden. Diese Berechnung der Skalenebene für die Verschmelzung der Linien wurde für die spezifische Konstellation zweier Fahrbahnen variabler Breite aufgeführt und erfolgte unter der Beobachtung, dass die Verschmelzung mit einem Maximum des mit dem Gauß-Filter gefalteten Profils in Skalenrichtung für diesen konkreten Fall zusammenfiel. Diese analytische Vorhersage wurde jedoch nicht auf allgemeine Fälle erweitert.

Der Lineare Skalenraum wird auch auf andere Datentypen angewendet, wie beispielsweise Synthetic Aperture Radar (SAR). Radardaten sind häufig stark durch Einflüsse wie Rauschen und Speckle geprägt, die eine Analyse dieser Daten erschweren. Die Gaußsche Filterung im Linearen Skalenraum kann aufgrund der Glättung dieser hochfrequenten Bildanteile zur Reduzierung einiger dieser Artefakte genutzt werden. Nicht nur zur Bildvorverarbeitung, sondern auch für weitergehende Auswertungen von SAR-Daten wird der Lineare Skalenraum eingesetzt. BARALDI & PARMIGGIANI [1996] stellen z. B. ein Verfahren zur multiskaligen Segmentierung vor, das mit Gabor-Filtern verschiedener σ arbeitet.

2.3 Diskussion

Semantische Netze als eine Form der Wissensrepräsentation bieten sich für eine automatische Anpassung an andere Bildauflösungen aufgrund ihres Aufbaus an. Die klare Trennung zwischen Objekten und Relationen sowie die explizite Wissensrepräsentation ermöglichen eine automatische Simulation der modellierten Bildszene. Insbesondere unterstützt auch die explizite Modellierung einer Bestandteilshierarchie in Semantischen Netzen eine automatische Abstraktion der Objekte. Außerdem sind die in Semantischen Netzen enthaltenen Informationen über die räumlichen Relationen zwischen den Objektteilen für eine zuverlässige Prädiktion des Skalenverhaltens der Objekte wichtig.

Die Modellgeneralisierung verfolgt ein auf den ersten Blick ähnliches Ziel wie diese Dissertation: Die Anpassung von räumlichen Daten eines feinen Maßstabes an eine geeignete Repräsentation dieser Daten in einem groben Maßstab. Während die Modellgeneralisierung allerdings mit Vektordaten arbeitet und für den Übergang der Daten in eine andere Maßstabsebene vorgegebene Richtlinien verwendet, die sich an der Lesbarkeit und dem Nutzen einer Karte in den beiden Maßstäben orientieren, müssen bei der Anpassung von Objektmodellen für die Bildanalyse an eine niedrigere Bildauflösung die physikalischen Gegebenheiten der Bildaufnahme (Rasterdaten) Beachtung finden. Die Repräsentation von optischen Bilddaten berücksichtigt keine Richtlinien zur Interpretierbarkeit der Bilddaten, sondern wird ausschließlich von physikalischen Prozessen gesteuert. Die im Abschnitt 2.2.1 diskutierte Modellgeneralisierung ist aufgrund der unterschiedlichen Voraussetzungen nicht kompatibel mit der Aufgabe der Anpassung von Bildanalyse-Objektmodellen an eine niedrigere Bildauflösung.

Der Prozess der Bildaufnahme ist gekennzeichnet durch eine Glättung des empfangenen Signals, die aus den Eigenschaften der Optik resultiert, gefolgt von einer Abtastung dieses geglätteten kontinuierlichen Signals zu einem diskreten Bild. Wie der Vergleich verschiedener Skalenräume im Abschnitt 2.2.2.1 zeigt, ist der Lineare Skalenraum der einzig zweckmäßige Skalenraum zur Simulation der Glättung bei der Bildaufnahme aufgrund seiner Erzeugung durch die Gauß-Funktion, die die Tiefpassfilterung in der Optik beim Aufnahmeprozess geeignet approximiert.

Der vorangehende Überblick über die verfügbare Literatur zu dem Thema Linearer Skalenraum hat gezeigt, dass viele intensive Untersuchungen zu dieser Theorie existieren. Keine dieser Untersuchungen kann jedoch eine Methode präsentieren, die in der Lage ist, das Verhalten komplexer Objekte im Skalenraum analytisch exakt zu prädizieren. Es wurden allerdings Verfahren entwickelt, um die Bildstruktur über verschiedene Skalenebenen zu verfolgen. Der Scale-Space Primal Sketch nutzt dabei als Grundelemente Grauwert-Blobs, die über die Skalenebenen miteinander verbunden werden, während die Skalenraum-Hierarchie aus den Pfaden der kritischen Punkte im Skalenraum generiert wird. Für die vorliegende Aufgabe der Simulation von Objekten in verschiedenen Bildauflösungen (siehe Kapitel 3), erscheint das Konzept des Scale-Space Primal Sketch aufgrund der Verwendung von Blobs besser geeignet, da Blobs im Allgemeinen Objekten entsprechen.

Die Anwendung des Linearen Skalenraumes begrenzt sich in der Computer Vision vorwiegend auf Aufgaben der low- und mid-level Bildanalyse, wie Vorverarbeitung, Merkmalsextraktion und Segmentierung. Der Lineare Skalenraum wird auch für die Interpretation von Fernerkundungsdaten angewendet. Vielen dieser Anwendungen, z. B. in der Landschaftsökologie, liegt jedoch kein explizites Objektmodell zugrunde. Es existieren jedoch auch Anwendungen zur Extraktion von Straßen in der Fernerkundung, die eine Objektmodellierung mit dem Linearen Skalenraum verbinden. So integrieren z.B. [MAYER 1998, BAUMGARTNER 2003] die Repräsentation der Straße in verschiedenen Bildauflösungen in ein einziges Objektmodell (als Semantisches Netz). Hier wird die unterschiedliche Erscheinung eines Landschaftsobjektes zur Erhöhung der Evidenz bei der wissensbasierten Straßenextraktion aus optischen Luft- und Satellitenbildern genutzt. Das Skalenverhalten ist jedoch vorher bekannt bzw. analytisch oder empirisch für konkrete Auflösungsebenen bestimmt worden.

Da Landschaftsobjekte in verschiedenen Bildauflösungen unterschiedlich erscheinen, ist die Erstellung geeigneter Objektmodelle für verschiedene Auflösungen bzw. geeigneter Algorithmen zur Objektextraktion notwendig. Die Modelle für die einzelnen Bildauflösungen müssen sich durchaus sehr voneinander unterscheiden. Bisher erfolgt die Erstellung von Objektmodellen in der Bildanalyse meist manuell – selbst wenn dasselbe Objekt in derselben Szene in anderen Bilddaten mit niedrigerer Auflösung modelliert werden soll. Eine automatische Ableitung eines neuen Modells aus einem bereits bestehenden Modell wäre ein wichtiger Fortschritt für die Automatisierung in der Bildanalyse. Dafür müssen die Informationen über das Objekt in die niedrigere Bildauflösung übertragen und entsprechend der Darstellung dieses Objektes in der niedrigeren Bildauflösung aufbereitet werden. Für diese Aufbereitung der Modellinformation ist die Veränderung des Objektes aufgrund der Abstraktion in den verschiedenen Bildauflösungen zu analysieren. Die Entwicklung eines neuen Verfahrens, das die Veränderung des modellierten Objektes in der niedrigeren Bildauflösung automatisch mit Methoden der Theorie des Linearen Skalenraumes analysiert und in ein angepasstes Modell zur Objektextraktion überführt, ist das Thema dieser Dissertation.

3 Ein neues Verfahren zur automatischen auflösungsabhängigen Anpassung von Bildanalyse-Objektmodellen

Die Diskussion in Abschnitt 2.3 hat gezeigt, dass eine Methode zur automatischen Anpassung von vollständigen Objektmodellen an Bilddaten niedrigerer Auflösung bislang nicht verfügbar ist. In diesem Kapitel wird ein neues Verfahren zur automatischen Anpassung von Bildanalyse-Objektmodellen an eine niedrigere Bildauflösung vorgestellt. Zuerst werden die Voraussetzungen bezüglich der gegebenen Bilddaten und Objektmodelle, die für eine erfolgreiche automatische Anpassung nötig sind, erläutert und begründet. Als zweites wird die Strategie des neuen Verfahrens zur Anpassung detailliert beschrieben. Anschließend werden die entwickelten Methoden zur automatischen Anpassung für unterschiedliche Typen der gegebenen Objektmodelle dargestellt. Das Kapitel schließt mit einer Diskussion im Hinblick auf die erzielten Innovationen und Grenzen des neuen Verfahrens.

3.1 Voraussetzungen

Die Formulierung der Objektmodelle als Semantische Netze und die Wahl der Extraktionsstrategie ist nicht standardisiert, wie in 2.1.1 ausgeführt. Dies führt zu einer großen Vielfalt der möglichen Objektmodelle, die in das Verfahren eingehen könnten. Für eine Automatisierung des Verfahrens ist eine Einschränkung dieser Vielfalt nötig. Daher werden in diesem Abschnitt die Voraussetzungen bzgl. des Aufbaus der Objektmodelle und der mit ihnen verbundenen Extraktionsstrategie erläutert, unter denen das Verfahren zur Anpassung entwickelt wurde. Die Eingangsgrößen müssen Randbedingungen erfüllen. Zuerst werden die allgemeinen Randbedingungen zum Aufbau der Semantischen Netze vorgestellt. Als nächstes werden die Attribute der Knoten und die zulässigen Relationen zwischen den Objektteilen, die bei der Anpassung berücksichtigt werden können, definiert. Schließlich wird auch die spezifische Extraktionsstrategie, die hinter dem Semantischen Netz steht, vorgestellt.

Die hier formulierten Voraussetzungen gelten für Semantische Netze. Aufgrund der expliziten Repräsentation des dargestellten Wissens ermöglichen sie eine klare hierarchische Struktur des Objektmodells und erleichtern damit die Abstraktion der beschriebenen Objekte. Die Objektmodelle können jedoch prinzipiell auch in anderen Formen vorliegen. Semantische Netze dienen ausschließlich der Formulierung der Information über die Objekte und werden nicht zur Prädiktion des Skalenverhaltens benötigt. Der Kern des entwickelten Verfahrens zur automatischen Anpassung ist damit nicht auf Semantische Netze beschränkt. Der hier vorgestellte Ansatz ist daher durch wenige Modifikationen, die die Wissensrepräsentation betreffen, auch auf andere Typen von Objektmodellen anwendbar.

Das hier vorgestellte Verfahren zur automatischen Anpassung unterscheidet nicht zwischen verschiedenen Typen von Fernerkundungs-Sensoren (z. B. optische Bilddaten, Radardaten, DGMs). Eine Randbedingung für die Verwendung der mit dem entwickelten Verfahren an eine niedrigere Bildauflösung angepassten Objektmodelle ist daher, dass die Bilddaten in der niedrigeren Zielauflösung ähnliche radiometrische Eigenschaften aufweisen und vom selben Sensortyp aufgenommen wurden wie die in der höheren Ursprungsauflösung.

3.1.1 Allgemeine Randbedingungen an den Aufbau der Semantischen Netze

Aufgrund des Fehlens einer Standardterminologie für den Aufbau Semantischer Netze sind sehr unterschiedliche Möglichkeiten für die Erzeugung und die Zusammensetzung eines Modells für die Extraktion eines bestimmten Objektes erlaubt. Das bedeutet, dass ein Objekt mit sehr unterschiedlichen Semantischen Netzen beschrieben werden kann. Für das Ziel dieser Arbeit – die Entwicklung eines Verfahrens zur automatischen Anpassung von Bildanalyse-Objektmodellen an eine niedrigere Bildauflösung – ist eine Einschränkung dieser Vielzahl möglicher Modelle als Eingangsgröße nötig. Die gegebenen Objektmodelle müssen außerdem die richtige Art und ausreichend Informationen über das zu extrahierende Objekt bereitstellen, um eine eindeutige Vorhersage des Skalenverhaltens zu ermöglichen. Um die Anpassung komplexer Modelle zu vereinfachen, sollte ein gegegebenes Semantisches Netz in Unterstrukturen einteilbar sein. Weiterhin umfassen die Ziele dieser Arbeit nicht nur eine Anpassung der Beschreibung des Objektes in der neuen Bildauflösung, sondern auch eine Adaption bzw. Wechsel der Algorithmen, die zur Extraktion des Objektes aus einem

konkreten Bild der Zielauflösung verwendet werden. Insbesondere sollte die Objektbeschreibung durch das Semantische Netz auch Informationen zu den Extraktionsalgorithmen, die im Bild arbeiten, bereitstellen.

Anpassbare Objektmodelle müssen daher die folgenden allgemeinen *Randbedingungen* erfüllen:

- Das Objektmodell muss das Objekt *vollständig beschreiben* mit allen Charakteristika, die für die Darstellung des Objektes im Bild und für seine Extraktion wichtig sind. Diese Forderung schließt insbesondere auch zusätzliche Attribute mit ein, die in der niedrigeren Zielauflösung für die Extraktion des Objektes relevant sind, auch wenn sie es in der hohen Bildauflösung nicht sind.
- Das Objektmodell muss eine Struktur aufweisen, die es erlaubt, das gesamte Modell automatisch in einzelne *Substrukturen* zu *zerlegen*. Das Skalenverhalten dieser Substrukturen, die möglicherweise aus einzelnen Objektteilen bestehen, sollte automatisch analysiert werden können.
- Die einzelnen Objektteile müssen über geeignete *Operatoren zur Merkmalsextraktion* extrahierbar sein. Die jeweiligen Operatoren müssen im Objektmodell spezifiziert und den Objektteilen zugeordnet sein. Diese Bedingung gilt auch für ein Modell, das nur aus einem einzigen Objektteil besteht.

Die Randbedingungen sind als allgemeine Richtlinien zum Aufbau anpassbarer Semantischer Netze für die automatische Verarbeitung zu verstehen. Trotz Einschränkung der Variabilität bezüglich des Aufbaus anpassbarer Semantischer Netze durch diese Randbedingungen, bleiben die Vorteile einer Modellierung mit Semantischen Netzen erhalten. Insbesondere bleiben die explizite Wissensrepräsentation, die hohe Vernetzbarkeit und die hierarchische Struktur Semantischer Netze unangetastet. Eine explizite Wissensrepräsentation und ein hierarchischer Aufbau sind sogar vorteilhaft für die Zerlegbarkeit der Semantischen Netze in Unterstrukturen zur automatischen Weiterverarbeitung, wie sie bei dem entwickelten Verfahren angewendet und in den Abschnitten 3.2 und 3.3 vorgestellt wird.

3.1.2 Knoten

Die Knoten im Semantischen Netz repräsentieren die Objektteile des modellierten Landschaftsobjektes. Im Folgenden wird als *Objektteil* ein über einen geeigneten Operator zur Merkmalsextraktion extrahierbares Teil des gesamten Landschaftsobjektes verstanden. Jedes Objektteil wird durch genau einen Knoten repräsentiert. Die Attribute der Knoten beschreiben die geometrischen und radiometrischen Eigenschaften der Objektteile, d. h. ihr Aussehen in den Bilddaten. Weiterhin enthalten die Knoten Angaben zum jeweiligen Objekttyp und spezifizieren den Objektnamen.

Für das entwickelte Verfahren der automatischen Anpassung eines Objektmodells, repräsentiert als Semantisches Netz, an eine niedrigere Bildauflösung werden die folgenden *Attribute für die Knoten* definiert:

- *Objektname*: Der Objektname beschreibt das Objektteil in natürlicher Sprache, z. B. Fahrbahn.
- *Objekttyp*: Im entwickelten Verfahren werden linienartige und flächenhafte Objekttypen unterschieden, die die grundlegende Form eines Objektteils widerspiegeln. Punkte als Objekttypen sind nicht berücksichtigt, da diese in der Regel nicht zur Objektmodellierung herangezogen werden. Objektteile desselben Typs zeigen ein ähnliches Verhalten während einer Skalenänderung. Ähnliche Objekttypen können mit ähnlichen Merkmalsoperatoren extrahiert werden. Die Objekttypen lassen sich in zwei Klassen einteilen:
    - *linienartige* Objekttypen: Diese Objekttypen dienen der Beschreibung von langgestreckten, eher schmalen Objekten, die mit Linienoperatoren extrahiert werden können. Zur Klasse der Linien gehören folgende Objekttypen:
        - *Kontinuierliche Linie*: Die Ausprägung der Linie ist durchgängig.
        - *Periodische Linie*: Die Ausprägung der Linie ist mit einer bestimmten Regelmäßigkeit unterbrochen. Periodische Linien können beispielsweise zur Beschreibung von gestrichelten Fahrbahnmarkierungen verwendet werden.

- *flächenhafte* Objekttypen: Flächenhafte Objekte sind meist kompakter als Linien, können allerdings auch langgestreckte Flächen darstellen. Ihre Form ist sehr variabel. Sie werden mit flächenbasierten Merkmalsoperatoren verknüpft. Zur Klasse der flächenhaften Objekttypen gehören:
    - *geometrische Formen*: Geometrische Formen können mathematisch exakt beschrieben werden, z. B. *Rechteck, Dreieck, Ellipse.*
    - *beliebige Muster*: Muster können beliebige Formen annehmen. Sie werden durch ein Template im Rasterformat definiert, das ebenfalls gegeben sein muss.

- *Grauwert*: Dieses Attribut dient der radiometrischen Beschreibung des Objektteils im Bild und gibt die Intensitäten im Bild wieder.
- *Ausdehnung*: Die Ausdehnung bezeichnet die geometrischen Ausmaße des Objektteils in Pixeln im Bild. Für linienartige Objekttypen wird die Breite der Linie angegeben; für flächenhafte Objektteile sollten die Länge und Breite des kleinsten umschließenden Rechtecks angegeben werden.
- *Orientierung*: Die Orientierung gibt die Richtung des Objektteils relativ zur Hauptachse des Gesamtobjektes an und ist daher ausschließlich für Objektmodelle mit beliebig orientierten Objektteilen relevant.
- *Periodizität*: Dieses Attribut ist ausschließlich für periodische Linien relevant. Die Periodizität wird als ein Verhältnis zwischen der Länge der Linienteile l_{Teil} und der Länge der Lücke $l_{Lücke}$ zwischen den Linienteilen definiert:

$$p = \frac{l_{Teil}}{l_{Teil} + l_{Lücke}} \tag{3.1}$$

Die Attribute Grauwert, Ausdehnung, Orientierung und Periodizität können als absolute Werte oder Wertebereiche angegeben werden. Wertebereiche für den Grauwert ermöglichen eine höhere Flexibilität der Modelle für verschiedene Beleuchtungsverhältnisse bei der Bildaufnahme. Bereiche für die Ausdehnung, Orientierung und Periodizität erlauben eine konzeptionelle Beschreibung von Objekten, deren geometrische Eigenschaften gewissen Schwankungen in der Realität unterliegen und sich auch in ihrer Darstellung im Bild widerspiegeln.

Die gegebenen Objektmodelle sollten nicht nur deskriptive Daten, sondern zusätzlich ausführbare Prozesse beinhalten. An die einzelnen Knoten, d. h. Objektteile, werden geeignete *Bildanalyse-Operatoren* angehängt. Ihre Auswahl orientiert sich an den Objekttypen. Dementsprechend sind beispielsweise differentialgeometrische Linienoperatoren, Template-Matching für Muster sowie Operatoren für die geometrischen Formen, basierend auf Flächenwachstumsverfahren, denkbar. Die Parameter der Operatoren sind ebenfalls Eingangsgrößen für das Verfahren zur automatischen Anpassung und sollten an die Extraktion aus der Originalauflösung optimiert sein. Idealerweise sollten die Parameter eines Operators mit den Attributwerten der Knoten des betreffenden Objektteils übereinstimmen. Die oben angegebenen Attribute entsprechen zum großen Teil den Parametern für die genannten Operatortypen. Für andere Operatoren gilt dies jedoch nicht. In diesem Fall ist eine automatische Anpassung der Operatorparameter mit dem entwickelten Verfahren nicht sicher. Für eine größere Auflösungsänderung versagt der Operator in der Zielauflösung mit den Parametern aus der Ursprungsauflösung möglicherweise, obwohl die Extraktion des Objektteils mit optimierten Parametern noch möglich wäre. Nichtsdestotrotz kann der Operator manuell mit den automatisch prädizierten Informationen über das Aussehen des Objektteils in der Zielauflösung angepasst werden. Alternativ können die Operatorparameter auch aus einer zuvor erstellten Liste von Parametern für unterschiedliche Bildauflösungen ausgewählt werden oder durch Methoden des automatischen Lernens (machine learning) automatisch bestimmt werden.

3.1.3 Relationen

In einem Semantischen Netz beschreiben Relationen die Beziehungen zwischen den Objektteilen, die durch die Knoten repräsentiert sind. Wie in Kapitel 2.1.2 dargelegt, werden hierarchische und räumliche Relationen unterschieden. Während hierarchische Relationen die Objektteile in eine Bestandteilshierarchie eingliedern und damit die Objektteile auf

unterschiedlichen hierarchischen Ebenen miteinander verbinden, dienen räumliche Relationen der Beschreibung der Beziehungen zwischen Objektteilen in derselben Hierarchieebene.

In einem automatisch anpassbaren Objektmodell sind folgende *Relationen* zulässig:

- *hierarchische* Relationen:
  - *Part-of*
  - *OptPart-of (Optional part-of)*
  - *Is-a*
  - *Con-of (Concretization-of)*
- *räumliche* Relationen:
  - *adjacent* bzw. *left-of* und *right-of*
  - *parallel*
  - *perpendicular*

Den Relationen können auch Attribute zugewiesen werden, welche die betreffende Relation beschreiben. Ein wichtiges Attribut ist die *Distanz* zwischen den Objektteilen. Sie spezifiziert die Entfernung der über eine räumliche Relation verbundenen Objektteile und wird ebenfalls als Anzahl der Pixel angegeben. Distanzen sollten entlang der Hauptorientierungsrichtungen des Gesamtobjektes angegeben werden. Distanzen haben eine große Bedeutung für die Prädiktion des Skalenverhaltens. Nahe liegende Objekte können sich während der Skalenänderung gegenseitig beeinflussen, wie in Abschnitt 3.2 demonstriert wird. Da der Einfluss von der Entfernung der Objektteile abhängig ist, ist die Kenntnis der Distanz für eine eindeutige Anpassung des Objektmodells an eine andere Bildauflösung entscheidend.

3.1.4 Extraktionsstrategie

Die Extraktion eines Objektes mit wissensbasierter Bildanalyse erfordert nicht nur ein geeignetes Objektmodell, sondern auch eine dazu gehörige Extraktionsstrategie. Diese bestimmt die Nutzung des Wissens, das im Objektmodell formuliert ist. Modell und Strategie müssen aufeinander abgestimmt sein. Strategien zur Objektextraktion können sehr vielfältig sein. Diese Vielfalt muss für eine automatische Verarbeitung wiederum eingeschränkt werden.

Eine weitere Voraussetzung für die automatische Anpassung ist daher eine Annahme über die Extraktionsstrategie, die mit dem Objektmodell verknüpft ist. Da die Extraktionsstrategie zum Objektmodell nicht explizit gegeben ist, wird eine spezifische Extraktionsstrategie angenommen. Die vorausgesetzte Extraktionsstrategie (verdeutlicht in Abbildung 3.1) ist relativ einfach: Zuerst werden alle Objektteile der untersten Ebene durch die jeweiligen Operatoren im Bild extrahiert; anschließend erfolgt eine Zuordnung aller gefundenen Objektteile anhand der gegebenen Relationen untereinander. Wenn passende Merkmale mit den betreffenden Attributen für alle Knoten durch die Operatoren gefunden und sämtliche Bedingungen, die durch die Relationen gegeben sind, erfüllt wurden, wird das Gesamtobjekt als im Bild verifiziert betrachtet. Wenn dagegen mindestens eine Bedingung nicht bestätigt werden konnte, ist die Extraktion des gesamten Objektes fehlgeschlagen. Die gewählte Extraktionsstrategie ist kompatibel zum wissensbasierten System geoAIDA, das für die Verifikation des entwickelten Verfahrens eingesetzt wird (siehe Kapitel 4).

Die Extraktionsstrategie bleibt im entwickelten Verfahren unverändert, d. h. sie gilt auch für das angepasste Objektmodell der niedrigeren Bildauflösung. Sie ist aufgrund ihrer Einfachheit und Allgemeinheit für die meisten Objektmodelle universell anwendbar. Die zugrunde liegende einfache Extraktionsstrategie wird auf Kosten einer optimalen Strategie für eine effiziente Objektextraktion gewählt. Im Zuge einer manuellen Erstellung der Objektmodelle würde möglicherweise eine andere effizientere Strategie bevorzugt. Die hier gewählte Extraktionsstrategie hat jedoch den Vorteil, dass sie in einem automatischen Verfahren verarbeitet werden kann.

Die hier vorgestellte Extraktionsstrategie gilt für sämtliche Typen der Objektmodelle, die durch das entwickelte Verfahren anpassbar sind, mit Ausnahme derer, die lokale Kontextobjekte beinhalten. Die für Kontextobjekte geltende Extraktionsstrategie wird in Abschnitt 3.3.4 ausführlich dargestellt.

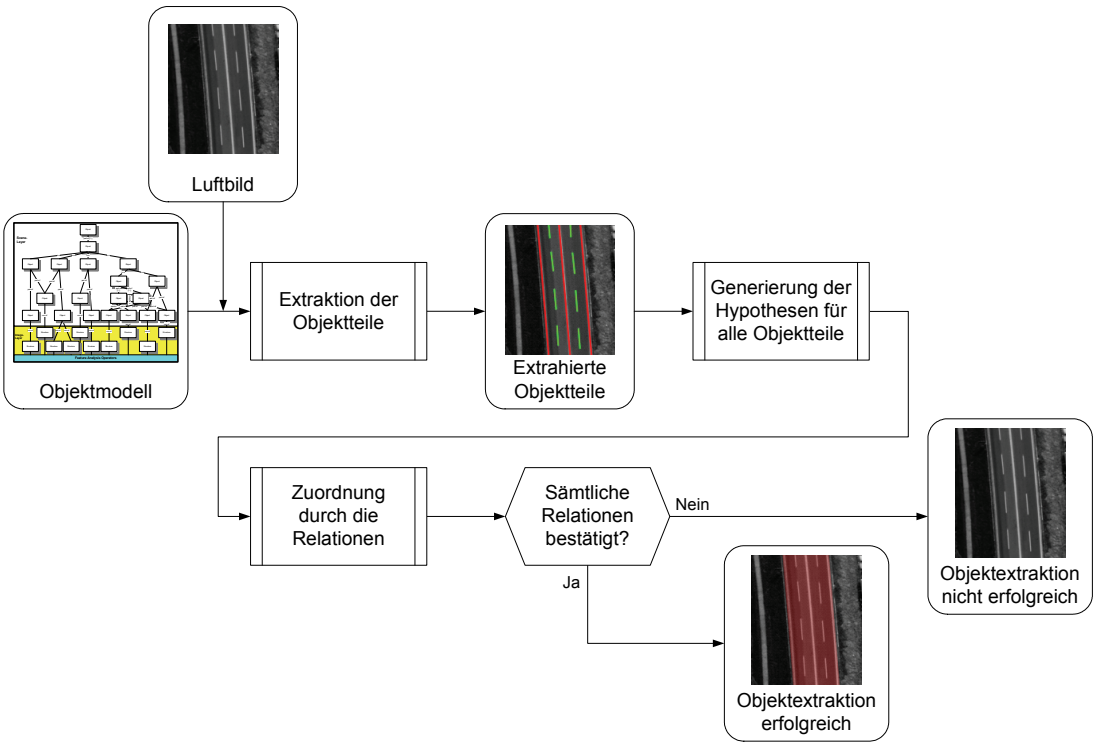


Abbildung 3.1: Extraktionsstrategie: Die Objektextraktion erfolgt in zwei Stufen – Extraktion der einzelnen Objektteile mit gegebenen Attributen und ihre Zuordnung aufgrund der gegebenen Relationen

3.2 Strategie der automatischen Anpassung

Im Gegensatz zur Extraktionsstrategie definiert die in diesem Abschnitt vorgestellte Strategie den Ablauf der einzelnen Schritte des Verfahrens zur automatischen Anpassung des gegebenen Objektmodells an eine niedrigere Zielauflösung. Sie gilt für alle anpassbaren Objektmodelltypen, ausgenommen die Objektmodelle mit Kontextobjekten, für deren Anpassung eine veränderte Strategie angewendet wird. Die entwickelten Methoden der einzelnen Stufen des Anpassungsprozesses werden im folgenden Abschnitt 3.3 präsentiert, da sie sich je nach Typ des Objektmodells unterscheiden.

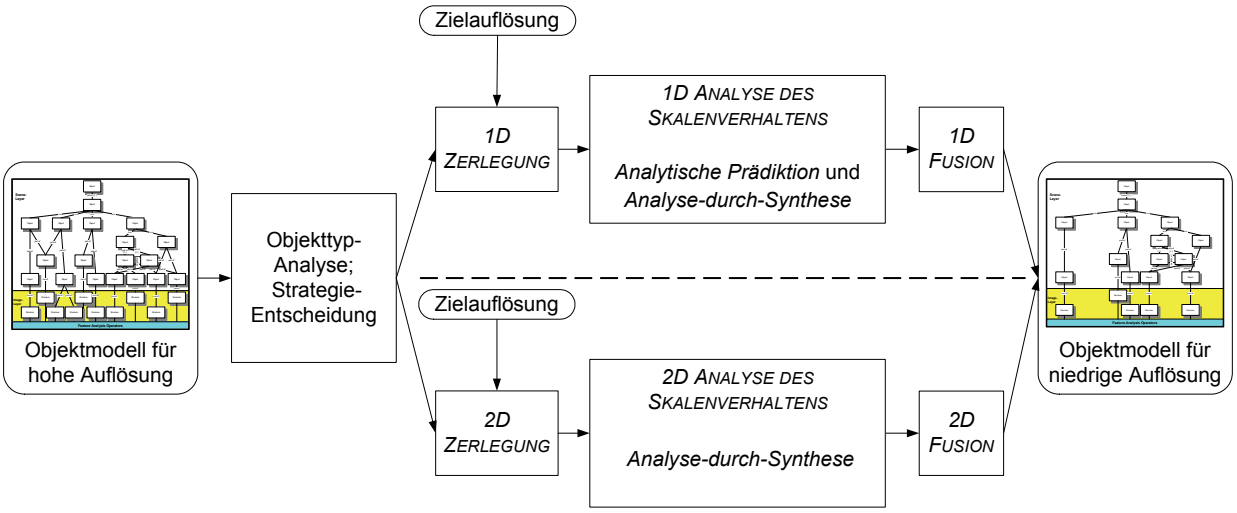


Abbildung 3.2: Strategie des Verfahrens zur automatischen Anpassung von Objektmodellen

Die Strategie zur automatischen Anpassung folgt einem Prozess mit drei Hauptstufen – *Zerlegung*, *Analyse des Skalenverhaltens* und *Fusion*. Abbildung 3.2 skizziert den Prozess der Anpassung mit seinen Teilschritten. In Abhängigkeit der Objekttypen und deren Orientierung im Objektmodell der hohen Auflösung wird zuerst die Entscheidung getroffen, ob die ein- oder die zweidimensionale Analyse-Methode zur Anwendung kommt. Grundsätzlich wird das komplette Objektmodell entweder mit der eindimensionalen oder der zweidimensionalen Strategie angepasst. Tabelle 3.1 gibt einen Überblick über die Art der angewendeten Strategie für jeden Typ der betrachteten Objektmodelle. Sofern das Objektmodell ausschließlich aus linienartigen Objekttypen mit paralleler Orientierung besteht, kann die Prädiktion des Skalenverhaltens auf ein eindimensionales Problem reduziert werden, da die Analyse des Linienprofils in diesem Fall ausreichend ist. Das gilt ebenfalls für die Objektmodelle mit linienartigen Objektteilen variabler Anzahl. Sind jedoch flächenhafte Objekttypen oder linienartige Objekttypen mit beliebiger Orientierung im Objektmodell beschrieben, muss die zweidimensionale Methode zur automatischen Anpassung verwendet werden. Einen Sonderfall stellen Objektmodelle mit lokalen Kontextobjekten dar. Je nach Objekttyp der Kontextobjekte wird für diese Art Objektmodelle eine eindimensionale oder zweidimensionale Anpassung vorgenommen.

Tabelle 3.1: Art der Strategie der automatischen Anpassung in Abhängigkeit des Objektmodelltyps

Objektmodelle mit …	**Eindimensionale Strategie**	**Zweidimensionale Strategie**
parallelen linienartigen Objektteilen	X	
parallelen linienartigen Objektteilen variabler Anzahl	X	
flächenhaften beliebig orientierten Objektteilen		X
lokalen Kontextobjekten	X	X

Zerlegung

Die erste Stufe der automatischen Anpassung ist die *Zerlegung*. In diesem Modul wird das gegebene Objektmodell für die hohe Bildauflösung in Gruppen von Objektteilen zerlegt, deren Aussehen im Bild sich für die gegebene Änderung der Auflösung (bzw. der Zielskala) gegenseitig beeinflussen und deren Skalenverhalten daher gemeinsam analysiert werden muss. Die gegenseitige Beeinflussung benachbarter Objektteile im Zuge der Auflösungsänderung wird im Folgenden als *Interaktion* bezeichnet. Ein Beispiel für Interaktion zweier paralleler Fahrbahnmarkierungen in einem Luftbild demonstriert Abbildung 3.3. Eine Gruppe von Objektteilen, deren Skalenverhalten gemeinsam analysiert wird, wird als *Interaktionsgruppe* bezeichnet. Sie besteht aus mindestens zwei Objektteilen. Das Vorliegen einer Interaktion bei einer bestimmten Zielskala ist von den räumlichen Distanzen der Objektteile untereinander abhängig. Falls die Distanz unter einen bestimmten Betrag fällt, der durch die Zielskala bestimmt ist, kann Interaktion auftreten. Dann muss in der folgenden Stufe des Anpassungsverfahrens das Skalenverhalten dieser Objektteile in einer Interaktionsgruppe gemeinsam analysiert werden.

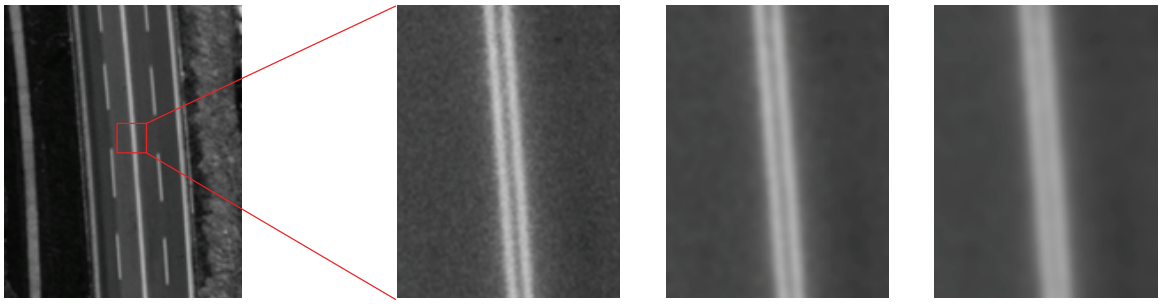


Abbildung 3.3: Interaktion am Beispiel zweier Fahrbahnmarkierungen im Luftbild

Das Prinzip der Zerlegung skizziert Abbildung 3.4. Ausgehend vom gegebenen Objektmodell der hohen Auflösung werden unter Kenntnis der Zielauflösung Interaktionsgruppen gebildet, deren Skalenverhalten in der folgenden Stufe separat analysiert werden.

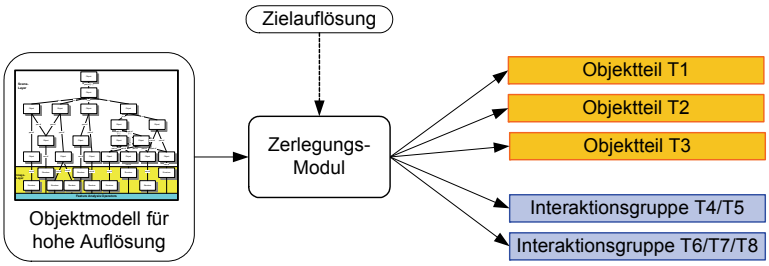


Abbildung 3.4: Prinzip des Zerlegungsmoduls (einzelne Objektteile in orange, Interaktionsgruppen in blau)

Analyse des Skalenverhaltens

In der zweiten Stufe des Verfahrens zur automatischen Anpassung – *Analyse des Skalenverhaltens* – werden das Aussehen der Objektteile und deren Extrahierbarkeit in der niedrigeren Zielauflösung prädiziert. Das Skalenverhalten wird für einzelne Objektteile oder für Interaktionsgruppen analysiert. Die Methode der Prädiktion nutzt u. a. *Analyse-durch-Synthese* zur Simulation des Verhaltens der Objektteile mit Hilfe von synthetischen Bildern bei der gegebenen Skalenänderung. Einen Überblick zum Prinzip dieser Stufe gibt Abbildung 3.5.

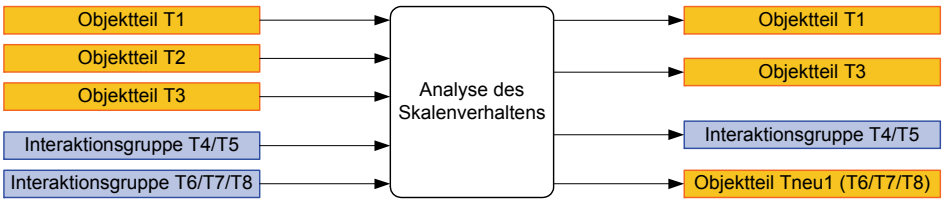


Abbildung 3.5: Prinzip der Analyse des Skalenverhaltens (einzelne Objektteile in orange, Interaktionsgruppen in blau)

Fusion

Die letzte Stufe des Verfahrens zur automatischen Anpassung ist die *Fusion*. In diesem Modul, skizziert in Abbildung 3.6, werden alle an die niedrigere Zielauflösung angepassten Objektteile bzw. Interaktionsgruppen zu einem vollständigen Objektmodell zusammengesetzt. Das komplette Modell ist das Ergebnis der automatischen Anpassung und dient der Objektextraktion aus der Zielauflösung.

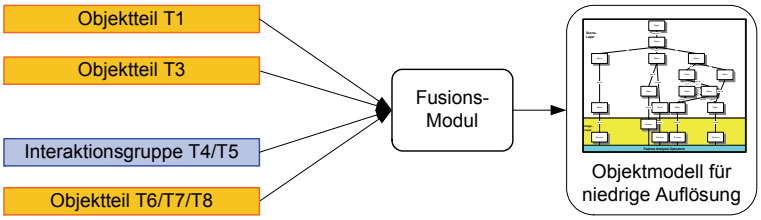


Abbildung 3.6: Prinzip des Fusionsmoduls (einzelne Objektteile in orange, Interaktionsgruppen in blau)

3.3 Methodik

In diesem Abschnitt erfolgt eine detaillierte Erläuterung der drei Hauptstufen des Verfahrens für die verschiedenen Typen der automatisch anpassbaren Objektmodelle. Je nach Typ des Objektmodells werden unterschiedliche Methoden zur Anpassung angewendet, da die Prädiktion des Skalenverhaltens verschiedener Typen unterschiedlich komplex ist. Für einige Typen ist die Prädiktion sogar teilweise analytisch möglich ist. Die allgemeine zweidimensionale Methode zur automatischen Anpassung kann jedoch auch auf die Arten von Objektmodellen angewendet werden, die über die eindimensionale Strategie angepasst werden können.

3.3.1 Objektmodelle mit parallelen linienartigen Objektteilen

Bei Objektmodellen, die ausschließlich aus parallelen linienartigen Objektteilen zusammengesetzt sind, ist für die Prädiktion des Skalenverhaltens die Analyse des Linienprofils ausreichend, da das Linienprofil an jeder beliebigen Stelle dasselbe Skalenverhalten aufweist. Dadurch können diese Objektmodelle als eindimensionale Konstellation der Objektteile betrachtet werden. Der Vorteil eindimensionaler Konstellationen ist eine vereinfachte Analyse des Skalenverhaltens, da im eindimensionalen Skalenraum nur Auslöschung und Verschmelzung als Ereignisse eintreten können. Außerdem sind für einfache eindimensionale Linienprofile teilweise analytische Berechnungen des Skalenverhaltens möglich. Zu den linienartigen Objekttypen zählen die *Kontinuierliche Linie* und die *Periodische Linie*. Die Periodische Linie ist streng genommen kein eindimensionaler Objekttyp. Wenn eine Linie jedoch schmal im Vergleich zur Länge der Linienteile ist, ist der Effekt der Zweidimensionalität im Skalenverhalten vernachlässigbar und eine eindimensionale Betrachtung auch für diesen Objekttyp zulässig.

3.3.1.1 Zerlegung

Aufgabe des Zerlegungsmoduls ist eine zuverlässige Bestimmung der Interaktion zwischen benachbarten Objektteilen und die darauf basierende Erzeugung der einzelnen Interaktionsgruppen. Für die linienartigen Objektteile wird eine stabförmige Linie mit *Linienbreite* $2w$ und *Kontrast* c zur Umgebung als realistisches Grauwertprofil angenommen:

$$f_L(w) = \begin{cases} c, & |x| \le w \\ 0, & |x| > w \end{cases} \tag{3.2}$$

Die Größe der angestrebten Auflösungsänderung bestimmt die Zielskala σ_z, die der Standardabweichung σ des Gauß-Filters der zur Auflösungsänderung korrespondierenden Glättung entspricht. Die Breite des zur Zielskala σ_z korrespondierenden Kernels w_F entspricht dem Einflussgebiet der Gauß-Filterung, die zur Interaktion führt. Falls die Distanz d zweier benachbarter Linien in diesen Einflussbereich fällt, tritt Interaktion bei der Skalenänderung auf. Abbildung 3.7 verdeutlicht diesen Zusammenhang. In der dargestellten Konstellation von Liniendistanz und Filterbreite findet keine Interaktion statt. Bei einer etwas größeren Filterbreite würde jedoch Interaktion auftreten.

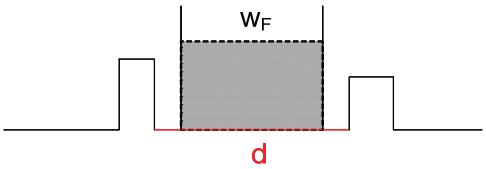


Abbildung 3.7: Einfluss der Breite des Gauß-Filters w_F und der Liniendistanz d auf die Interaktion zweier Linien im Grauwertprofil

Für den eindimensionalen Fall erfolgt für sämtliche Objektteile des Objektmodells (n Linien) die Bestimmung der Interaktion I_{ij} zweier Linien L_i, L_j während der Skalenänderung über den Vergleich der Größe des zur Auflösungsänderung korrespondierenden Gauß-Filters w_F mit der Distanz d_{ij} der betrachteten Linien L_i, L_j:

$$\forall\ i,j \in [0,n] \qquad \text{wenn } d \le w_F \rightarrow I_{ij} \tag{3.3}$$

Die Größe des zur Glättung verwendeten Gauß-Filters wird in Abhängigkeit von der Zielskala σ_z gewählt. Um mindestens 99.7 % des Integrals unter der Gauß-Funktion abzudecken (entsprechend $2 \cdot 3\sigma$), wird ein Filter-Kernel mit folgender Breite generiert:

$$w_F = 6\sigma_z \tag{3.4}$$

Die mit (3.4) berechnete Filterbreite wird gerundet. Falls diese Rundung eine gerade Zahl ergeben sollte, wird die angewendete Filterbreite auf die nächstgrößere ungerade Zahl aufgerundet.

Zerlegung für Distanzbereiche

In Konzeptnetzen zur Objektextraktion werden oftmals keine exakten Distanzen zwischen den Objektteilen, sondern Distanzbereiche angegeben, um eine Anwendbarkeit der Objektmodelle auf eine Vielzahl von Objektszenen zu gewährleisten. Das Verfahren der automatischen Anpassung hält eine Möglichkeit für die Anpassung von Objektmodellen mit Distanzbereichen mit minimalen und maximalen Distanzwerten bereit. Für die Zerlegung bedeuten diese Bereiche, dass eine Interaktion möglicherweise für die minimale Distanz auftritt, während für die maximale Distanz keine Interaktion vorliegt. Zur Lösung dieses Widerspruchs werden die minimale und maximale Attributkonfiguration wie zwei Objektmodelle gehandhabt. Dazu werden Interaktionsgruppen für die minimalen und die maximalen Attributwerte, einschließlich der Distanzen, gebildet. Anschließend wird das Skalenverhalten jeweils für die minimale und maximale Konfiguration untersucht. Das Ergebnis der automatischen Anpassung sind dann u. U. zwei Modelle für die Zielauflösung.

3.3.1.2 Analyse des Skalenverhaltens

In dieser Stufe des Verfahrens zur automatischen Anpassung wird für alle Interaktionsgruppen, die in der vorhergehenden Stufe der Zerlegung gebildet wurden, das Aussehen in der Zielauflösung separat voneinander prädiziert. Diese Aufgabe kann analytisch oder empirisch gelöst werden. Der hier vorgestellte Ansatz zur Analyse des Skalenverhaltens für parallele linienartige Objektteile verwendet eine Kombination aus beiden Herangehensweisen. Soweit möglich, wird der analytische Ansatz bevorzugt. Für Einzellinien können teilweise der Eintritt von Skalenraumereignissen und ein Teil der Attribute in der Zielauflösung analytisch prädiziert werden, während das Skalenverhalten der Interaktionsgruppen und das Verhalten der Merkmalsoperatoren in der Zielauflösung empirisch über *Analyse-durch-Synthese* untersucht werden müssen. Synthetische Bilder werden hier zur Simulation des Skalenverhaltens der Objektteile generiert. Abbildung 3.8 beschreibt schematisch das Konzept des Prozesses der Analyse-durch-Synthese.

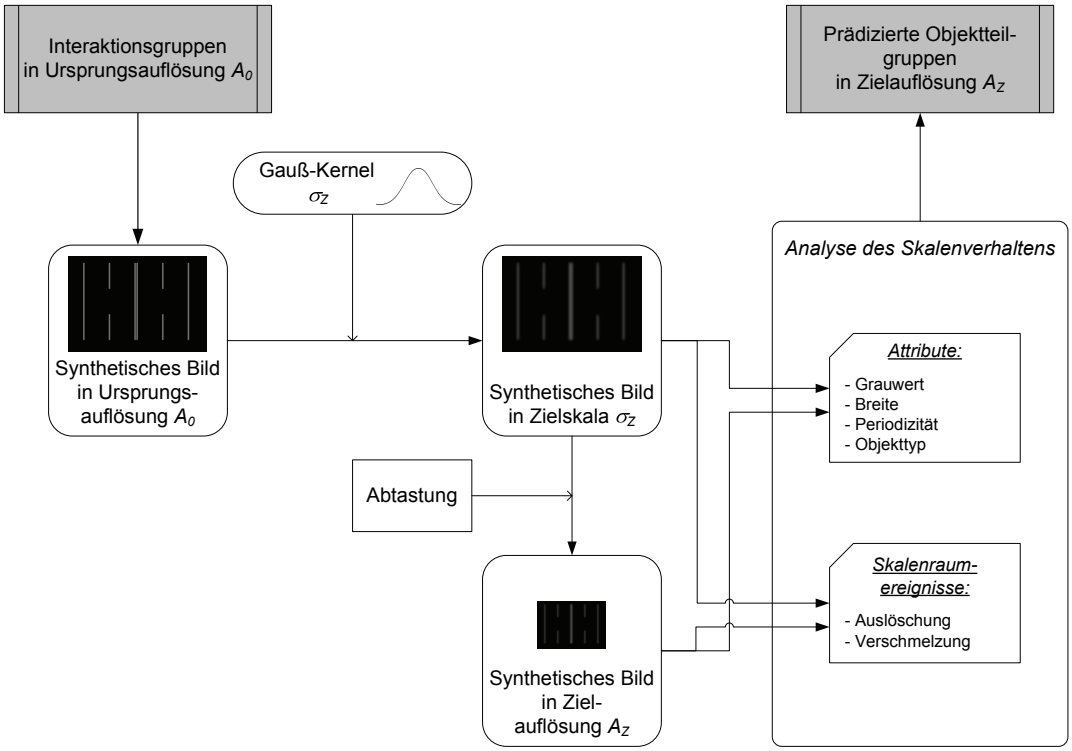


Abbildung 3.8: Konzept der Analyse-durch-Synthese für eindimensionale Objektmodelle

Für jede Interaktionsgruppe wird ein synthetisches Bild aus den Informationen über das Aussehen der Objektteile erzeugt, die von den Knoten und Relationen bzw. ihrer Attribute aus dem gegebenen Semantischen Netz bereitgestellt werden. Dieses Anfangsbild in der Ursprungsauflösung A_0 wird auf die Ebene der Zielskala σ_z durch Faltung mit dem korrespondierenden Gauß-Kernel transferiert. Unter Verwendung einer bilinearen Interpolation wird das Bild der Zielskala anschließend in die Zielauflösung A_z überführt. Im Bild der Zielskala $I_{\sigma z}$ werden der Eintritt sicherer Skalenraum-

ereignisse und die Attributwerte der Objektteile bestimmt. Wenn kein sicheres Skalenraumereignis prädiziert wurde, kann anschließend im Bild der Zielauflösung I_{Az} mit den angepassten Operatoren zur Merkmalsextraktion der Eintritt effektiver Skalenraumereignisse überprüft werden. Die Prädiktion der sicheren Skalenraumereignisse und der meisten Attributwerte ist im Bild der Zielskala $I_{\sigma z}$ sinnvoller als im Bild der Zielauflösung, da die Höhe der Grauwerte im abgetasteten Bild der Zielauflösung I_{Az} von der Wahl der Abtaststellen abhängt. Aus dem Bild der Zielskala $I_{\sigma z}$ können verschiedene Bilder der Zielauflösung I_{Az} je nach Wahl der Abstatstellen abgeleitet werden, die sich in ihren Grauwerten unterscheiden. Um eine sichere Prädiktion der Skalenraumereignisse und umfassende Prädiktion der Attributwerte vornehmen zu können, müssten sämtliche sich unterscheidende Bilder der Zielauflösung analysiert werden. Aus dem Bild der Zielskala $I_{\sigma z}$ dagegen können die sicheren Skalenraumereignisse und die minimalen und maximalen Attributwerte der Zielauflösung direkt ermittelt werden. Zur Vereinfachung der Analyse wird die Prädiktion der sicheren Skalenraumereignisse und der meisten Attributwerte daher im Bild der Zielskala $I_{\sigma z}$ durchgeführt. Das Ergebnis der Analyse des Skalenverhaltens ist eine an die Zielauflösung angepasste Einzellinie bzw. Gruppe von interagierenden Objektteilen. Das Skalenverhalten wird für alle aus der Zerlegung hervorgegangenen Interaktionsgruppen bzw. einzelnen Objektteile separat analysiert. Die Menge aller angepassten Objektteilgruppen bildet die Eingangsgröße für die letzte Stufe des Verfahrens, die Fusion. Die entwickelte Methode für die Prädiktion des Skalenverhaltens wird auf den nächsten Seiten ausführlich erläutert.

Bei der Analyse des Skalenverhaltens der Interaktionsgruppen wird zuerst das Auftreten von Skalenraumereignissen untersucht. Diese beeinflussen die Anzahl der resultierenden Objektteile und damit die Struktur des Objektmodells in der Zielauflösung. Im Anschluss werden für die verbliebenen Objektteile in der Zielauflösung die Attributwerte prädiziert.

Prädiktion von Skalenraumereignissen

Bei Skalenänderung verändern nicht nur Objekte ihr Aussehen, sondern auch die Topologie der Bildstruktur kann sich durch diskrete Ereignisse im Skalenraum ändern (vgl. Abschnitt 2.2.2.2). Die Änderung der Topologie bedeutet eine Veränderung der Anzahl der Objekte beim Übergang in eine gröbere Skalenebene. Diese Veränderung wird vom vorgestellten Verfahren prädiziert. Bei Objektmodellen mit parallelen linienartigen Objektteilen können ausschließlich zwei Typen von Skalenraumereignissen auftreten: Verschmelzung mehrerer Objekte und Auslöschung eines Objektes. Das Auftreten dieser beiden Ereignisse wird nacheinander für die gegebene Konstellation und Zielauflösung untersucht.

Verschmelzung:

In einer bestimmten Skalenebene beginnen benachbarte Linien miteinander zu interagieren, wie für die Zerlegung in Abschnitt 3.3.1.1 beschrieben. Diese Skalenebene wird durch den Abstand der Linien voneinander bestimmt. Mit steigendem Skalenparameter werden die Linienprofile immer breiter und nähern sich immer weiter an bis schließlich eine Verschmelzung eintritt und nur noch eine einzige Linie übrig bleibt. Da offensichtlich nur in der Zielskala miteinander interagierende Linien von einer Verschmelzung betroffen sein können, wird dieses Ereignis ausschließlich für Interaktionsgruppen untersucht. Ein Beispiel für eine Verschmelzung zweier Linien illustriert Abbildung 3.9. In der Ursprungsskala σ_0 bestehen im Grauwertprofil zwei getrennte Maxima, die ein einziges Minimum einschließen. Mit größerem σ werden die Linienprofile breiter und flacher bis schließlich das Minimum verschwindet und nur noch ein einziges Maximum verbleibt, das die *sichere Verschmelzung* der Linien signalisiert.

Die Entwicklung des Niveaus der Verschmelzung im Skalenraum kann in drei Zonen unterteilt werden. In der ersten Zone sind die Objekte klar voneinander getrennt. Zwischen dem *Interaktionspunkt* σ_I, d. h. der Skalenebene, in der die Interaktion der benachbarten Linien beginnt (siehe Abschnitt 3.2), und dem *Punkt der sicheren Verschmelzung* σ_{SV}, ab dem nur noch ein einziges Grauwert-Maximum verbleibt, liegt die *Unsicherheitszone*. Diese Zone stellt einen Bereich im Skalenraum dar, in dem die benachbarten Linien ihr Aussehen gegenseitig beeinflussen, jedoch noch nicht vollständig verschmolzen sind. In der dritten Zone sind beide Linien komplett miteinander verschmolzen und haben eine gemeinsame neue Einzellinie gebildet. Die betreffenden Verschmelzungszonen mit ihren verschiedenen Niveaus der Verschmelzung sind ebenfalls in Abbildung 3.9 dargestellt. Obwohl das Beispiel in Abbildung 3.9 zwei benachbarte Linien mit derselben Linienbreite und -intensität beschreibt, ist das Modul zur Analyse des Skalenverhaltens fähig, Linien mit beliebiger Breite und Kontrast zu verarbeiten.

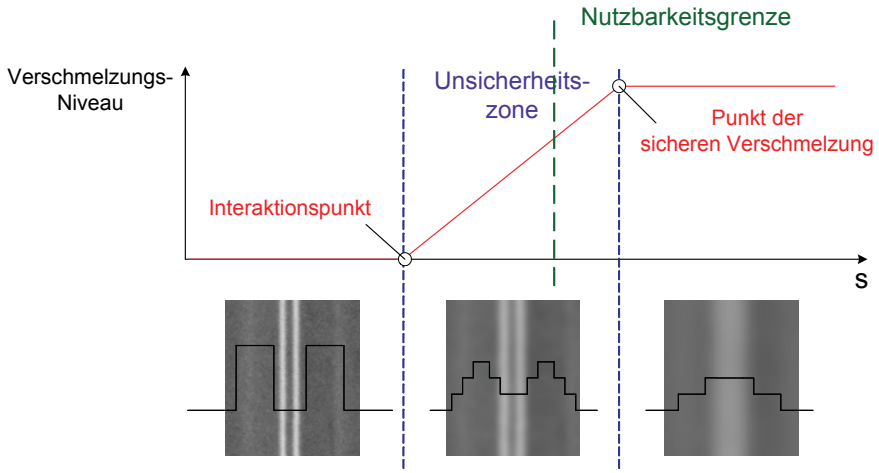


Abbildung 3.9: Verschmelzung zweier Linien im Grauwertprofil mit Einteilung der Verschmelzungszonen und Nutzbarkeitsgrenze des Merkmalsoperators (grün gestrichelt)

Die Extraktion der einzelnen Objektteile aus den Bildern wird von Merkmalsoperatoren ausgeführt, die zusammen mit der Beschreibung der Objektteile im Modell gegeben sind. Die Eigenschaften des jeweiligen Operators bestimmen den Erfolg der Extraktion der Objektteile und damit auch die Anzahl der verbleibenden Objektteile im Objektmodell der Zielauflösung. Bevor Interaktion stattfindet, d. h. in der ersten Verschmelzungszone, kann ein Operator zwei separate Linien extrahieren. In der letzten Zone dagegen, nach der sicheren Verschmelzung, kann nur noch eine einzelne Linie extrahiert werden. In der Unsicherheitszone hängt die Anzahl der extrahierten Linien jedoch von den Eigenschaften des jeweiligen Merkmalsoperators ab. Jeder Operator hat seine eigene *Nutzbarkeitsgrenze* σ_{EV} im Skalenraum (dargestellt durch die grün gestrichelte Linie in Abbildung 3.9). In einer gröberen Skalenebene detektiert der Operator nur noch eine Linie trotz der, möglicherweise, noch vorhandenen zwei Linien mit schwachem Kontrast. An der Nutzbarkeitsgrenze findet daher eine *effektive Verschmelzung* aufgrund der Leistungsfähigkeit des verwendeten Operators statt. Die Nutzbarkeitsgrenze kann durch empirische Analyse gefunden werden: Dazu wird der Merkmalsoperator mit automatisch angepassten Parametern (siehe Abschnitt zur Anpassung der Operatorparameter) auf das synthetische Bild I_{Az} angewendet, das die Liniengruppe mit ihren Attributen in der Zielauflösung simuliert, wie in Abbildung 3.8 skizziert. Zur Simulation eines möglichst realistischen Bildes kann weißes Rauschen hinzugefügt werden, z. B. mit einer Amplitude von 2 bis 3 Grauwertstufen. Das Ergebnis des Operators im synthetischen Bild drückt seine Fähigkeit zur separaten Extrahierbarkeit der Linien aus und bestimmt daher die Anzahl der Objektteile im resultierenden Objektmodell der Zielauflösung. Mit der empirischen Analyse ist die Methode zur Analyse des Skalenverhaltens sehr flexibel bezüglich unterschiedlicher Operatoren, da die Wahl des verwendeten Operators im gegebenen Objektmodell nicht eingeschränkt ist.

Zur Bestimmung, ob die Zielskala in den Unsicherheitsbereich fällt, wird das Grauwertprofil der Linien aus der Interaktionsgruppe in der Zielskala untersucht. Dazu wird der Interaktionspunkt σ_I berechnet und ein Test auf die sichere Verschmelzung für die gegebene Interaktionsgruppe durchgeführt. Der Test sucht nach der Existenz eines Minimums zwischen den Linien im Grauwertprofil. Kann ein Minimum nachgewiesen werden und ist die Zielskala σ_z größer als der Interaktionspunkt σ_I, befindet sich die Zielskala in der Unsicherheitszone und die Interaktionsgruppe wird auf effektive Verschmelzung durch Anwendung des Operators auf das synthetische Bild untersucht. Ist die Zielskala kleiner als der Interaktionspunkt, befindet sie sich in der ersten Verschmelzungszone, in der keine Verschmelzung stattfinden kann. Falls kein Minimum existiert, ist eine Verschmelzung der beiden Linien eingetreten und die Anzahl der verbleibenden Objektteile im Objektmodell der Zielauflösung muss reduziert werden.

Auslöschung:

Das Ergebnis der Faltung des stabförmigen Linienprofils $f_L(x)$ mit der Breite $2w$ und dem Kontrast c mit der eindimensionalen Gauß-Funktion $g(x,\sigma)$ ist gegeben durch:

$$r_L(x,\sigma,w,c) = f_L(x) * g(x,\sigma) \tag{3.5}$$

Im kontinuierlichen Raum wird eine Linie durch Faltung mit der Gauß-Funktion steigender σ breiter und flacher. Die Linienmitte wird jedoch niemals vollständig verschwinden, da die Amplitude der Gauß-Funktion für $\sigma \rightarrow \infty$ zwar sehr klein, aber niemals 0 wird. Der Grauwert an der Linienmitte x_M in der Zielskala σ_z im kontinuierlichen Raum kann mit dem normalisierten Faltungsintegral bestimmt werden:

$$r_b(x_M, \sigma, w, c) = \frac{c}{\sqrt{2\pi}\sigma} \int_{x_M - w}^{x_M + w} e^{\frac{t^2}{2\sigma^2}} dt \tag{3.6}$$

Auch im diskreten Raum wird das Linienprofil durch die Faltung mit immer größer werdenden Gauß-Kerneln breiter und flacher, aber die Linie wird bei einer bestimmten Filtergröße verschwinden. Aufgrund von Abtasteffekten wird der Grauwert der Linienmitte in einer diskreten Pixelmatrix der Zielauflösung gleich oder niedriger sein als im kontinuierlichen Raum, wie in Gleichung (3.6) dargestellt. Der resultierende Grauwert im diskreten Raum liegt in einem Bereich Δr, der von der Abtaststelle, von der Abtastweite sd, d. h. Pixelgröße, sowie von der Breite w und dem Kontrast c der Linie abhängig ist. Abbildung 3.10 stellt die Beziehung zwischen der Wahl der Abtaststelle und dem Bereich der resultierenden Grauwerte bildlich dar.

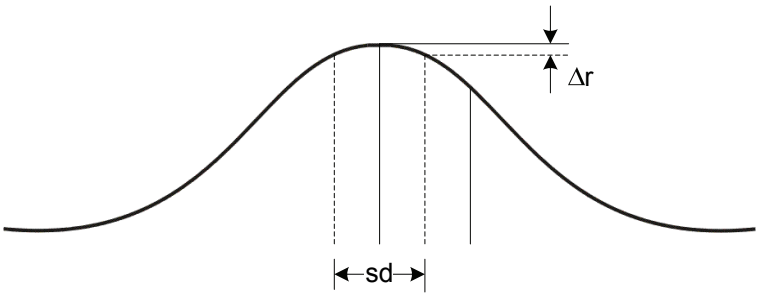


Abbildung 3.10: Entstehung des Grauwertbereiches Δr der Linienmitte in der Zielauflösung in Abhängigkeit von der Abtaststelle (mit sd: Abtastweite)

Die Grenzen des Grauwertbereiches Δr der Linienmitte im diskreten Bild sind durch die jeweiligen Grauwerte an zwei Abtaststellen im kontinuierlichen Raum bestimmt. Die obere Grenze r_{ub} ist durch die Abtaststelle in der Linienmitte gegeben, während sich die untere Grenze r_{lb} aus der Abtaststelle entlang des Linienprofils versetzt um die halbe Abtastweite $sd/2$ ergibt. Die obere (3.7) und untere Grauwertgrenze (3.8) für die Linienmitte berechnen sich im diskreten Raum daher mit:

$$r_{oG}(x, \sigma, w, c) = \frac{c}{\sqrt{2\pi}\sigma} \int_{-w}^{+w} e^{\frac{t^2}{2\sigma^2}} dt \tag{3.7}$$

$$r_{uG}(x, \sigma, w, c) = \frac{c}{\sqrt{2\pi}\sigma} \int_{\frac{sd}{2} - w}^{\frac{sd}{2} + w} e^{\frac{t^2}{2\sigma^2}} dt \tag{3.8}$$

Die Auslöschung einer Linie im diskreten Bild der Zielauflösung kann mit obigen Gleichungen durch numerische Berechnung der Faltung des Linienprofils mit dem Gauß-Filter der korrespondierenden Zielskala an verschiedenen Abtaststellen prädiziert werden. So lange sämtliche Grauwerte größer als eine Quantisierungsstufe des Bildes sind, existiert (zumindest theoretisch) die Linie. Erst wenn der Grauwert unter diese Schwelle fällt, ist die Linie sicher verschwunden. Da der maximal mögliche Grauwert jedoch dem Grauwert im kontinuierlichen Skalenraum entspricht, der immer größer als 0 ist, existiert keine sichere Auslöschung für Linien. Viele Operatoren können jedoch für kleine Grauwertkontraste an der Extraktion scheitern. Obwohl die Linie in diesem Fall in den Bilddaten existiert, tritt eine *effektive Auslöschung* auf. Sie ist vom jeweiligen Operator bzw. seinen Parametereinstellungen abhängig. Häufig werden Parameter in Relation zum erwarteten Signal-Rausch-Verhältnis der Bilddaten gewählt. Zur Abschätzung, ob der zur untersuchten Linie

gegebene Operator an der Linienextraktion aus der Zielauflösung scheitert, und damit eine effektive Auslöschung verursacht, wird der betreffende Operator in Zweifelsfällen mit an die Zielauflösung automatisch angepassten Parametern auf das synthetische Bild I_{Az} angewendet, das die Linie bzw. die Interaktionsgruppe in der Zielauflösung simuliert. Dazu wird ein Schwellwert r_{eA} des resultierenden Grauwertes aus der Faltung eingeführt, unter dem die Linie auf effektive Auslöschung in der Zielauflösung getestet wird. Oberhalb dieses Schwellwertes sollte ein Operator in jedem Fall in der Lage sein, die Linie zu extrahieren. Die Höhe des Schwellwertes sollte großzügig gewählt werden, um die Fehlerquote bei der Bestimmung der effektiven Auslöschung möglichst gering zu halten. Zur Berücksichtigung des Einflusses von Rauschen sollte der Schwellwert r_{eA} entsprechend der erwarteten Rauschamplitude höher gewählt werden.

Falls eine Auslöschung prädiziert wird, darf die betroffene Linie in dem angepassten Objektmodell der Zielauflösung nicht mehr als Objektteil verbleiben, so dass die Anzahl der verbleibenden Objektteile in der Zielauflösung reduziert ist.

Prädiktion von Attributen

Nach der Analyse der Skalenraumereignisse ist die Anzahl der Objektteile, und damit der Knoten im angepassten Objektmodell, in der Zielauflösung bekannt. Allerdings müssen sich für die Mehrheit der Anpassungsprozesse auch die Attribute der Knoten ändern, die die Eigenschaften der Objektteile in den Bilddaten der Zielauflösung wiedergeben. Die Werte der Attribute ändern sich oftmals, auch wenn keine Skalenraumereignisse aufgetreten sind. Die Attributwerte in der Zielauflösung werden für alle Objektteile jeder Interaktionsgruppe bzw. Einzellinien prädiziert, die nach der Analyse der Skalenraumereignisse in der Zielauflösung verblieben sind.

Grauwert:

Das Attribut *Grauwert* in der Zielauflösung kann für Einzellinien analytisch ermittelt werden. Der Grauwert ist durch die Summe der Intensität des hierarchisch höher liegenden Knotens (*Part-of*-Relation) und dem Grauwertkontrast an der Linienmitte gegeben. Für diskrete abgetastete Bilder liegt der resultierende Grauwert der Linie in einem Bereich, der durch die gewählte Abtaststelle bestimmt wird, wie für Auslöschungen bereits beschrieben (siehe Gleichungen 3.7 und 3.8). Für Interaktionsgruppen wird der maximale Grauwert r_{IGmax} im Bild der Zielskala $I_{\sigma z}$ an der Position der Linienmitte x_M bestimmt. Der minimale Grauwert für Linien in Interaktionsgruppen r_{IGmin} entspricht dem Grauwert im Linienprofil im geglätteten Bild $I_{\sigma z}$ an der Abtaststelle, die um die halbe Abtastweite *sd* zur Linienmitte versetzt ist:

$$r_{IG\max} = r\left(x_M\right) \tag{3.9}$$

$$r_{IG\min} = r\left(x_M + \frac{sd}{2}\right) \tag{3.10}$$

Ausdehnung:

Die *Ausdehnung* wird für parallele linienartige Objekttypen durch die Breite der Linie ausgedrückt, die der Distanz zwischen ihren Kanten entspricht. Die Positionen der Kanten werden in der Analyse-durch-Synthese über die Wendepunkte WP_{Breite} des Linienprofils im synthetischen Bild bestimmt, das die Linie oder Interaktionsgruppe simuliert. Die Wendepunkte entsprechen den Maxima im Gradientenbild der Zielskala $G_{\sigma z}$, da der Gradient senkrecht zur Linie seinen größten absoluten Wert an der Position der Wendepunkte hat. Unter Berücksichtigung des Auflösungsfaktors F_A, der der Abtastweite im Bild der Zielskala zur Überführung in die Zielauflösung entspricht (siehe 2.2.2.1), wird anschließend die Breite der Linie in der Zielauflösung berechnet.

Periodizität:

Der Wert des Attributes *Periodizität p* kann sich für periodische Linien ändern, d. h. für Linien mit $p<1$, wenn nachfolgende Linienteile l_{Teil} derselben periodischen Linie miteinander interagieren. Diese Interaktion kann analog zur Bestimmung der Interaktion zwischen benachbarten Linien im Zerlegungs-Modul über den Vergleich der Lückenlänge $l_{Lücke}$ zwischen den Linienteilen l_{Teil} und der Breite des korrespondierenden Gauß-Filters w_f bestimmt werden. Falls eine Interaktion zwischen den Linienstücken vorliegt, wird die Veränderung der Lückenlänge ähnlich wie die Linienbreite aus dem Gradientenbild bestimmt. Hier wird jedoch der größte Gradientenwert entlang der periodischen Linie verwen-

det, der dem Wendepunkt $WP_{periodisch}$ entspricht. Von diesem Wert kann das Verhältnis der Länge der Linienstücke l_{Teil} zu deren Lücken $l_{Lücke}$, d. h. die Periodizität der Linie, abgeleitet werden. Abbildung 3.11 verdeutlicht schematisch die Ermittlung der Periodizität im Gradientenbild.

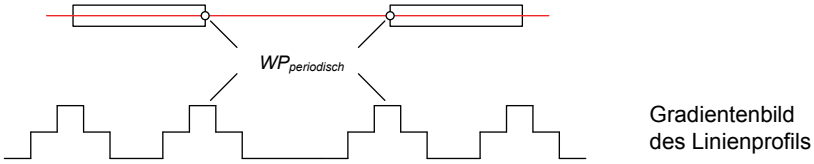


Abbildung 3.11: Ermittlung der Periodizität durch die Wendepunkte $WP_{periodisch}$ – oben: zwei Linienteile einer periodischen Linie, deren Lücke durch die Wendepunkte entlang der Linie begrenzt ist; unten: Maxima im Gradientenbild des Linienprofils entlang der Linienrichtung

Objekttyp:

Auch die *Objekttypen* können sich zwischen den Auflösungsstufen ändern. Objekttypänderungen sind stark mit Skalenraumereignissen verknüpft. So kann sich eine periodische Linie mit einer verhältnismäßig kleinen Lücke zwischen langen Linienstücken bei starker Glättung in eine kontinuierliche Linie wandeln, wenn die Linienstücke miteinander verschmelzen. Falls sich der Objekttyp ändert, muss ein anderer Merkmalsoperator den ursprünglichen Operator im angepassten Objektmodell der Zielauflösung ersetzen, denn der Operator muss dem jeweiligen Objekttyp entsprechen. Die Extraktion einer periodischen Linie benötigt einen anderen Operator-Typ als die einer kontinuierlichen Linie, auch wenn die Operatoren Ähnlichkeiten untereinander aufweisen. Die definierten linienartigen Objekttypen zeigen ein unterschiedliches Skalenverhalten. Für Einzellinien, die keine Interaktion zu Nachbarlinien aufweisen, ist in Tabelle 3.2 das mögliche Skalenverhalten aufgeführt:

Tabelle 3.2: Objekttypänderung für Einzellinien

Objekttyp in Ursprungsauflösung	**Objekttyp in Zielauflösung**
Kontinuierliche Linie (KL)	KL
Periodische Linie (PL)	PL oder KL

Für Interaktionsgruppen sind zusätzlich die Kombinationen der interagierenden Objekttypen zu berücksichtigen. Beispielsweise können zwei dicht nebeneinander liegende kontinuierliche Linien in einer bestimmten Skalenebene miteinander verschmelzen. Die folgenden Objekttypänderungen können für sämtliche Kombinationen der linienartigen Objekttypen in Interaktionsgruppen bei einer Skalenänderung auftreten:

Tabelle 3.3: Objekttypänderung für mögliche Kombinationen der Linientypen in Interaktionsgruppen

Kombinationen	Kontinuierliche Linie	Periodische Linie
Kontinuierliche Linie	KL	KL
Periodische Linie	KL	PL oder KL

Die obigen beiden Tabellen zeigen, dass eine Objekttypänderung für linienartige Objektteile ausschließlich bei periodischen Linien auftreten kann. Falls eine Verschmelzung einer periodischen Linie mit einer kontinuierlichen Linie im vorhergehenden Verfahrensschritt zur Analyse von Skalenraumereignissen prädiziert worden ist, entspricht der resultierende Objekttyp immer einer kontinuierlichen Linie. Falls periodische Linien miteinander verschmolzen sind, kann der Objekttyp unverändert geblieben oder eine kontinuierliche Linie entstanden sein. Dasselbe gilt für einzelne periodische Linien, die keine Interaktion mit Nachbarlinien ausüben. Zur Bestimmung, ob sich der Objekttyp zu einer kontinuierlichen Linie geändert hat, wird der prädizierte Wert der Periodizität in der Zielauflösung hinzugezogen. Falls die Periodizität im Vergleich zur Ursprungsauflösung stark gestiegen ist, kann von einer starken Interaktion zwischen den Linienteilen ausgegangen werden. Ob wirklich eine effektive Verschmelzung der Linienteile detektierbar ist, ist wiederum vom zugehörigen Operator zur Extraktion der periodischen Linie abhängig. Um auf eine effektive Verschmelzung der Linienteile zu schließen, die eine Objekttypänderung in der Zielauflösung nach sich zieht, wird daher (ähnlich wie bei

der Prädiktion von Verschmelzungen zwischen Nachbarlinien) für einen hohen Wert der Periodizität im synthetischen Bild der betroffenen periodischen Linie der zugehörige Operator angewendet.

Wertebereiche der Attribute

Die Attributwerte der Objektteile können als Bereiche mit minimalem und maximalem Wert im Objektmodell der Ursprungsauflösung gegeben sein. Da vor allem der Kontrast der Linien und die Distanz zwischen den Linien maßgeblich für das Skalenverhalten sind, können sich je nach Auflösungsänderung und Größe der Wertebereiche nicht nur verschiedene Attributwerte in der Zielauflösung ergeben, sondern auch unterschiedliche Skalenraumereignisse eintreten. Da Skalenraumereignisse eine Änderung der Struktur der Objektmodelle erfordern, sind verschiedene Objektmodelle als Ergebnis der Anpassung von Objektmodellen mit Wertebereichen möglich.

Für Objektmodelle mit Wertebereichen wird das Skalenverhalten für den minimalen und den maximalen Wert getrennt voneinander analysiert. Falls die Struktur der beiden angepassten Modelle in der Zielauflösung identisch ist, besteht das Ergebnis der Anpassung aus einem einzigen Objektmodell. Für alle Attribute, für die Wertebereiche angegeben sind, werden diese in der Zielauflösung für den minimalen und maximalen Wert ermittelt. Damit ergibt sich für diese Attribute in der Zielauflösung wiederum ein Wertebereich.

Falls unterschiedliche Skalenraumereignisse für die minimale und maximale Konfiguration prädiziert werden, resultieren mehrere Objektmodelle mit unterschiedlicher Struktur als Ergebnis der Anpassung. Die Konsequenzen dieser Mehrdeutigkeit des Ergebnisses werden in Abschnitt 3.4 diskutiert. In diesem Fall besteht das angepasste Konzeptnetz aus verschiedenen Varianten, die die möglichen Repräsentationen des Objektes in der Zielauflösung widerspiegeln. Als Attributbereiche werden jeweils die angepassten Werte für die minimale und maximale Konstellation übernommen, obwohl nicht der gesamte prädizierte Attributbereich für alle Objektmodelle gültig ist. Für eine präzisere Prädiktion der Attributbereiche der einzelnen Objektmodelle ist eine genaue Bestimmung des Attributwertes nötig, für den das Skalenraumereignis eintritt. Das kann durch schrittweise Untersuchung der Attribute erreicht werden. Die Größe der dafür gewählten Schritte bestimmt die Genauigkeit der Prädiktion angepasster Attributwerte.

Anpassung der Operatorparameter

Den einzelnen Objektteilen zugehörige Merkmalsoperatoren mit ihren an die Ursprungsauflösung optimierten Parametern liefern aufgrund unterschiedlicher Darstellungen der Objekte in anderen Bildauflösungen unterschiedliche Ergebnisse. Üblicherweise nimmt der Erfolg der Operatoren in niedrigeren Bildauflösungen ab, wie z. B. Untersuchungen einiger häufig angewendeter Kanten- und Linienoperatoren zeigen [PAKZAD & HELLER 2004]. Sofern die Operatoren flexibel genug sind, kann eine Anpassung der Parameter an die Zielauflösung bessere Ergebnisse liefern und die Nutzbarkeit des Operators im Skalenraum verlängern. Die prädizierten Attribute können als angepasste Operatorparameter in der Zielauflösung dienen. Operatoren, deren Parameter mit den Attributen übereinstimmen, die für die einzelnen Objektteile spezifiziert sind, können im Anpassungsprozess an die Zielauflösung optimiert werden. Dagegen erreichen Operatoren, deren Parameter nicht flexibel genug sind oder nicht den jeweiligen Attributen entsprechen, möglicherweise keine ausreichenden Extraktionsergebnisse aus der Zielauflösung. Obwohl eine Linie in der Zielauflösung noch vorhanden ist und gut extrahierbar wäre, sind diese unflexiblen Operatoren dazu nicht fähig. In diesem Fall müssen sie durch einen anderen geeigneten Operator ersetzt werden.

3.3.1.3 Fusion

In der letzten Stufe des Anpassungsverfahrens werden alle Einzellinien und Interaktionsgruppen, deren Aussehen in der Zielauflösung in vorherigen Schritten prädiziert wurde, zu einem kompletten Objektmodell zusammengeführt. Dieses angepasste Modell berücksichtigt sämtliche Skalenraumereignisse und enthält angepasste Attribute und Merkmalsoperatoren. Beim Auftreten von Skalenraumereignissen während der Auflösungsänderung kann für Objektmodelle mit linienartigen parallelen Objektteilen die Anzahl ihrer Teile in der Zielauflösung nur sinken, da ausschließlich Auslöschung und Verschmelzung auftreten können. Falls eine Auslöschung eingetreten ist, erscheint die betroffene Linie nicht mehr im Modell der Zielauflösung. Sämtliche Relationen zu anderen Knoten werden gelöscht. Bei einer Verschmelzung von Nachbarlinien wird die Anzahl der Linien ebenfalls reduziert. Für die verschmolzenen Linien wird ein

neuer Knoten generiert. Die Knoten und die Relationen zwischen den verschmolzenen Linien verschwinden. Der Typ der hierarchischen und räumlichen Relationen aller anderen Linien, die nicht von Skalenraumereignissen betroffen sind, bleibt unverändert.

Distanzen

Die Distanzen als Attribute der räumlichen Relationen zwischen den Linien d_{LiLj} in der Zielauflösung müssen angepasst werden, wenn sich die Breite der Linien verändert, sich die Linienpositionen aufgrund einer Verschmelzung verschoben haben oder eine Auslöschung aufgetreten ist. Bei der Auslöschung einer Nachbarlinie entfallen alle Distanzen zu der betroffenen Linie und es wird geprüft, ob stattdessen eine räumliche Relation mit einer Distanzangabe zu einer anderen Linie infrage kommt. In diesem Fall wird die Distanz zur nun nächstgelegenen Linie ermittelt. Die Distanzen zwischen den angepassten Linien in der Zielauflösung können durch die Auflösungsänderung $F_A=A_z/A_0$ berechnet werden:

$$d_{AzLiLj} = \frac{1}{F_A}\left(d_{A0LiLj} - \tfrac{1}{2}\Delta w_{\sigma zLi} - \tfrac{1}{2}\Delta w_{\sigma zLj} + \Delta t_{LiLj}\right) \tag{3.11}$$

mit d_{AzLiLj}: Distanz der Linie i und Linie j in der Zielauflösung A_z

d_{A0LiLj}: Distanz der Linie i und Linie j in der Ursprungsauflösung A_0

$\Delta w_{\sigma zLi,j}$: Veränderung der Breite der Linie i bzw. j von A_0 zu σ_z

Δt_{LiLj}: Differenz der Translation der Positionen von Linie i und Linie j in σ_z

Die Distanzen müssen für fast alle durchgeführten Anpassungen geändert werden, da sich die Breite der Linie schon bei kleinen Auflösungsänderungen verändert. Die Werte der Translationen von Linien werden dagegen erst bei sehr großen Auflösungsänderungen signifikant, weil Verschiebungen sich auf die Position der Linie erst bei sehr großer Glättung deutlich auswirken. Bei Verschmelzungen wird sich die Position der neu gebildeten Linie allerdings zwischen denen der verschmolzenen Linien in der Ursprungsauflösung befinden und damit in der Regel auch eine signifikante Translation auftreten. Weitere Ausführungen zur entwickelten Methode für Objektmodelle mit parallelen linienartigen Objektteilen mit Beispielen zur Anpassung enthalten HELLER & PAKZAD [2005A] und [2005B].

3.3.2 Objektmodelle mit parallelen linienartigen Objektteilen variabler Anzahl

Eine variable Anzahl gleichartiger Objektteile mit denselben Eigenschaften kann für die Modellierung von Landschaftsobjekten sinnvoll sein, die aus regelmäßigen, sich wiederholenden Teilen bestehen, z. B. Parkanlagen oder Häuserreihen. Durch das Hinzufügen einer Variabilität der exakten Anzahl dieser gleichartigen Objektteile sind die adaptierbaren Objektmodelle weniger in Bezug auf einen bestimmten Typ des Objektes beschränkt. Sie sind damit für eine Extraktion von ähnlichen Landschaftsobjekten flexibler einsetzbar. Für Straßen können beispielsweise mehrere Fahrbahnen in demselben Modell beschrieben werden, ohne dass vorheriges Wissen über die genaue Anzahl der Fahrbahnen erforderlich ist. Daher können verschiedene Straßentypen mit einem einzigen Modell beschrieben werden. Die variable Anzahl kann im Semantischen Netz über eine *dynamische Relation* realisiert werden. Diese kann sowohl hierarchische als auch räumliche Relationen zwischen Objektteilen im Konzeptnetz beschreiben. Im Instanzennetz jedoch, das sich auf spezifische Bildszenen bezieht, ist die exakte Anzahl der variablen Objektteile bekannt und kann sich für unterschiedliche Bildszenen unterscheiden. Die Anzahl der variablen Objektteile kann zur Erhöhung der Extraktionssicherheit auf einen bestimmten minimalen und maximalen Wert begrenzt werden. Auch für die Prädiktion des Skalenverhaltens dieser Objektmodelle sollte die maximale Anzahl der variablen Objektteile bekannt sein (siehe Abschnitt 3.3.2.2). Ein einfaches Beispiel einer möglichen dynamischen Relation für Fahrbahnen illustriert Abbildung 3.12. Die Fahrbahnen sind hier durch parallele Fahrbahnmarkierungen als periodische Linien modelliert. Die Objektteile, die durch die dynamische Relation beschrieben sind, werden im Folgenden als *dynamische Gruppe* bezeichnet.

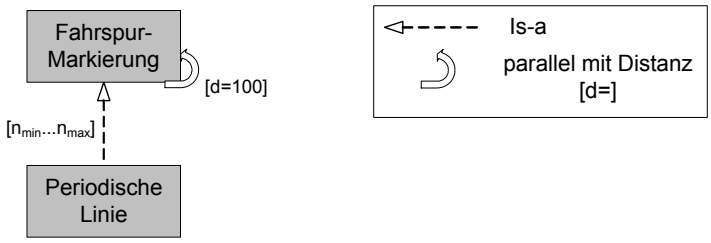


Abbildung 3.12: Dynamische Relation zur Modellierung einer variablen Anzahl von Fahrbahnmarkierungen

Die in diesem Abschnitt und in HEUWOLD&PAKZAD [2006] bereits in Teilen präsentierte Methode zur Anpassung von Objektmodellen mit einer variablen Anzahl paralleler linienartiger Objektteile erweitert das im vorigen Abschnitt vorgestellte Verfahren zur Anpassung von Objektmodellen mit einer konstanten Anzahl Objektteile. Die wesentlichen Grundzüge der Methode bleiben erhalten. Die Besonderheiten, die sich aufgrund der unbekannten Anzahl der Objektteile ergeben, werden detailliert ausgeführt.

3.3.2.1 Zerlegung

Das gegebene Objektmodell für die Extraktion aus hoch aufgelösten Bilddaten wird in dieser Stufe des Anpassungsprozesses in Einzellinien und Interaktionsgruppen zerlegt. Aufgrund ihrer räumlichen Nähe können auch Objektteile der dynamischen Gruppe während der Auflösungsänderung miteinander oder mit anderen benachbarten Objektteilen interagieren. Die Objektteile der dynamischen Gruppe werden als eigenständige Objektteile in der Zerlegung behandelt. Für die Ermittlung von Interaktion innerhalb dieser Gruppe ist eine einmalige Bestimmung der Interaktion ausreichend, da alle angrenzenden Objektteile nach der Definition in der dynamischen Gruppe dieselbe Distanz d_{dyn} zueinander haben, (siehe das Beispiel in Abbildung 3.13). Die Bestimmung der Interaktion wird analog zur Zerlegung bei eindimensionalen Objektmodellen ohne variable Anzahl der Objektteile (Abschnitt 3.3.1.1) durch den Vergleich der Distanz d_{dyn} mit der Breite w_F des zur Zielauflösung korrespondierenden Gauß-Filters durchgeführt. Falls keine Interaktion zwischen den Linien innerhalb der dynamischen Gruppe für die gegebene Zielauflösung auftritt, werden alle Linien als Einzellinien ausgewiesen. Wenn sie dagegen interagieren, bilden ausnahmslos alle Linien der dynamischen Gruppe eine Interaktionsgruppe, deren Skalenverhalten dementsprechend gemeinsam in der folgenden Stufe des Anpassungsverfahrens analysiert werden muss.

Abbildung 3.13: Dynamische Gruppe mit n_{min}=2 und n_{max}=4 Objektteilen und ihre Distanzen d_{dyn}

Für die anderen Objektteile des Objektmodells wird die gesamte dynamische Gruppe in der Zerlegung wie ein einzelnes Objektteil behandelt. Für angrenzende Objektteile muss eine mögliche Interaktion berücksichtigt werden. Falls diese vorliegt, müssen die äußeren Linien (interagierenden) der dynamischen Gruppe in die Interaktionsgruppe aufgenommen werden. Wenn auch Interaktion innerhalb der dynamischen Gruppe vorliegt, bilden die außerhalb liegenden angrenzenden Linien zusammen mit allen Linien dieser Gruppe eine Interaktionsgruppe. Wenn keine Interaktion innerhalb der dynamischen Gruppe auftritt, wird diese aufgespaltet. In diesem Fall bilden nur die äußeren Linien eine Interaktionsgruppe mit den außerhalb der dynamischen Gruppe liegenden Nachbarlinien.

3.3.2.2 Analyse des Skalenverhaltens

Die Analyse des Skalenverhaltens prädiziert das Verhalten der Objektteile während der Auflösungsänderung. Für jede Interaktionsgruppe und jedes nicht interagierende Objektteil wird das Aussehen in der groben Zielauflösung unter Berücksichtigung von Skalenraumereignissen prädiziert.

Prädiktion von Skalenraumereignissen

Für dynamische Gruppen mit parallelen linienartigen Objektteilen sind ausschließlich Auslöschung und Verschmelzung möglich. Wenn die Objektteile in der dynamischen Gruppe eine Interaktiongruppe ohne angrenzende Linien von außerhalb bilden, treten Skalenraumereignisse für alle Linien gemeinsam ein. Das gleichzeitige Auftreten der Ereignisse im Skalenraum begründet sich mit den identischen Eigenschaften der variablen Objektteile, die durch die dynamische Relation beschrieben sind. Diese Eigenschaften werden über die Attribute bezüglich Grauwert, Ausdehnung und Distanz definiert, die grundsätzlich die kritischen Parameter für den Eintritt von Skalenraumereignissen sind.

Verschmelzung:

Eine Verschmelzung in der dynamischen Gruppe kann eintreten, wenn Interaktion vorliegt. Der mögliche Fall einer Verschmelzung muss daher lediglich für Interaktionsgruppen untersucht werden. Die Anzahl der resultierenden Linien bei einer Verschmelzung ist von der tatsächlichen Anzahl der beteiligten Objektteile abhängig. Ein einfaches Beispiel zur Glättung einer dynamischen Gruppe im diskreten Linearen Skalenraum verdeutlicht Abbildung 3.14, das die Grauwertprofile einer dynamischen Gruppe mit drei bzw. vier beteiligten Linien in der Ursprungsauflösung σ_0 und in der Zielskala σ_z=3 mit einer Filterbreite w_F darstellt, die genau drei dynamischen Linien mit dem Abstand d_{dyn} entspricht.

Abbildung 3.14: Linienprofil in σ_0=0 (grau) und in der Zielskala σ_z=3 (blau) mit resultierendem Grauwert r_i (rot) für drei und vier interagierende Objektteile der dynamischen Gruppe

Für beide Konstellationen in Abbildung 3.14 (mit drei und vier Linien) tritt Verschmelzung auf. Während bei nur drei Linien in der dynamischen Gruppe ein einziges Grauwert-Maximum entsteht (und damit eine einzige Linie), verbleiben nach der Glättung bei der Konstellation mit vier Linien zwei Grauwert-Maxima und damit zwei Linien. Dieses unterschiedliche Verhalten kann auftreten, wenn der verwendete Filter nicht die komplette dynamische Gruppe überdeckt. Bei der Konstellation mit drei Linien entspricht die Filterbreite genau der Gesamtbreite der Linienkonstellation. Nicht nur die benachbarten, sondern auch die äußersten Linien in der dynamischen Gruppe interagieren miteinander. Bei der Konstellation mit vier Linien hingegen ist die Breite des Filters kleiner als die Gesamtbreite der Gruppe und die äußeren Linien interagieren nicht miteinander. Für die Prädiktion von Verschmelzungen muss daher zwischen zwei Fällen unterschieden werden – in Abhängigkeit des Verhältnisses von der maximalen Gesamtbreite w_{dyn_max} der dynamischen Gruppe zur Filterbreite w_F in der Zielskala:

1. $w_{dyn_max} \leq w_F$:

   Die maximale Gesamtbreite der Linienkonstellation ist höchstens so groß wie die verwendete Filterbreite. Damit überdeckt der Filter alle beteiligten Linien und alle Linien interagieren miteinander. Ob eine Verschmelzung vorliegt, kann mit nur zwei der Linien aus der dynamischen Gruppe bestimmt werden. Das Ergebnis dieser Untersuchung ist auf die gesamte Gruppe übertragbar. Wenn die beiden beispielhaft untersuchten Linien eine Verschmelzung ausüben, verschmelzen alle Linien zu einer einzigen.

2. $w_{dyn_max} > w_F$:

   Die maximale Gesamtbreite der Linienkonstellation übersteigt die Filterbreite. Der verwendete Filter überdeckt nur einen Teil der Linien. Die äußeren Linien interagieren nicht miteinander. Die Anzahl der resultierenden Linien bei einer Verschmelzung ist von der tatsächlichen, aber für die dynamische Gruppe unbekannten, Anzahl der beteiligten Linien abhängig. Sie kann somit bei einer unbekannten Linienzahl nicht eindeutig prädiziert werden. Allgemein gilt für die Anzahl der resultierenden Linien n_z in der Zielskala bei einer Verschmelzung:

$$n_z = n - n_F + 1 \tag{3.12}$$

mit n: Anzahl der Linien der dynamischen Gruppe

n_F: Anzahl der vom Filter überdeckten Linien

Die Berechnung der resultierenden Linien einer dynamischen Gruppe aus einer Verschmelzung wird für deren minimale Anzahl n_{min} und deren maximale Anzahl n_{max} durchgeführt. Das Ergebnis in der Zielauflösung ist wiederum eine dynamische Gruppe mit einer möglicherweise reduzierten minimalen und maximalen Anzahl Linien.

Auslöschung:

Wenn die Zerlegung der dynamischen Gruppe in Einzellinien resultierte, kann ausschließlich Auslöschung als mögliches Skalenraumereignis eintreten. Da exakt dieselben Bedingungen für alle Einzellinien gelten, ist die Untersuchung einer einzigen Linie als Beispiel für die gesamte dynamische Gruppe ausreichend. Die Untersuchung bezüglich einer möglichen Auslöschung erfolgt analog zu Abschnitt 3.3.1.2 für Objektmodelle mit einer bekannten Anzahl linienartiger Objektteile.

Im Allgemeinen tritt eine Auslöschung ein, wenn der zugehörige Operator bei der Extraktion des Objektteils aufgrund zu geringen Kontrastes versagt. Der Kontrast der Interaktionsgruppe in der gröberen Zielskala wird aufgrund der Interaktion von der tatsächlichen Anzahl der beteiligten Objektteile beeinflusst. Für relativ zur Skalenänderung kleine Gruppenbreiten w_{dyn} kann der zur Glättung verwendete Gauß-Filter mehr als eine Linie überlagern. Der resultierende Grauwert wird dann unter dem Einfluss mehrerer interagierender Linien erzeugt. Abbildung 3.15 illustriert die Abhängigkeit des resultierenden Grauwertes in der Zielskala σ_z von der konkreten Anzahl der beteiligten Objektteile beispielhaft für zwei und drei beteiligte Linien der dynamischen Gruppe.

Abbildung 3.15: Linienprofil in σ_0 (grau) und in der Zielskala σ_z=3 (blau) mit resultierendem Grauwert r_i (rot) für zwei und drei interagierende Objektteile

Auch bei der Prädiktion von Auslöschungen in dynamischen Gruppen muss daher zwischen zwei Fällen unterschieden werden – in Abhängigkeit des Verhältnisses von der maximalen Gesamtbreite w_{dyn_max} der dynamischen Gruppe zur Filterbreite w_F in der Zielskala:

1. $w_{dyn_max} \leq w_F$:

   Falls die maximale Gesamtbreite der Linienkonstellation w_{dyn_max} kleiner ist als die verwendete Filterbreite w_F, ist das Ereignis der Auslöschung für eine Interaktionsgruppe von einem zusätzlichen Parameter abhängig: der exakten Anzahl der an der Interaktion beteiligten variablen Objektteile. Je mehr Objektteile interagieren, desto mehr relativ hohe Grauwerte können zum resultierenden Wert beitragen. Dann kann der Grauwert der Interaktionsgruppe größer werden und damit wird später im Skalenraum eine Auslöschung eintreten. Eine Auslöschung ist deshalb für eine variable Anzahl beteiligter Objektteile aus der dynamischen Gruppe nicht eindeutig und zuverlässig prädizierbar.

2. $w_{dyn_max} > w_F$:

   Bei einer kleineren Filterbreite als der maximalen Gesamtbreite der Linienkonstellation kann der resultierende Grauwert in der Zielskala zuverlässig ermittelt werden (vgl. Abbildung 3.14). Eine eindeutige Prädiktion der Auslöschung ist möglich für alle Linienkonstellationen, die breiter sind als der zur Zielskala korrespondierende Glättungsfilter.

Für ein Landschaftsobjekt ist es meist möglich, adäquate Grenzen für die Anzahl der Objektteile in der dynamischen Gruppe anzugeben. Das obige Beispiel der Fahrbahnen könnte z. B. eine realistische Anzahl von ein bis vier Fahrbahnen ohne bauliche Trennung umfassen. Die Angabe von minimalen und maximalen Grenzen der Anzahl der variablen Objektteile ermöglicht die Ableitung einer übersichtlichen Menge der resultierenden angepassten Objektmodelle. Zur Gewährleistung einer automatischen Anpassung von variablen und damit flexiblen Objektmodellen sollte die Anzahl der Objektteile im gegebenen Modell begrenzt werden. Diese Forderung stellt eine zusätzliche Randbedingung bzgl. des Aufbaus der anpassbaren Objektmodelle mit einer variablen Anzahl von Objektteilen dar.

Der Eintritt eines Skalenraumereignisses verändert die Struktur des adaptierten Objektmodells bezüglich der erwarteten Präsenz eines oder mehrerer Objektteile. Das Ergebnis der Anpassung muss daher mehr als ein mögliches Objektmodell (als Konzeptnetz) für die Zielauflösung umfassen. Die Konsequenzen daraus werden in Abschnitt 3.4 diskutiert.

Prädiktion von Attributen

Die Werte der Attribute in den Knoten des Semantischen Netzes beschreiben das Aussehen der Objektteile im Bild. Für Einzellinien der dynamischen Gruppe, die keine Interaktion ausüben, können die Attribute beispielhaft für eine Linie nach der Methode in Abschnitt 3.3.1.2 bestimmt werden. Da die Linien dieselben Eigenschaften besitzen und daher dasselbe Skalenverhalten aufweisen, sind die Ergebnisse auf die gesamte dynamische Gruppe übertragbar.

Für eine Interaktionsgruppe werden dagegen der Grauwertkontrast und die Linienbreite auch von der konkreten Anzahl der beteiligten Objektteile in der dynamischen Gruppe beeinflusst, wie für die Prädiktion der Skalenraumereignisse erläutert. Aufgrund der Variabilität der Anzahl der Objektteile werden die prädizierten Attributwerte nicht exakt sein, sondern einen Bereich umfassen, der entsprechend im Konzeptnetz der Zielauflösung ausgewiesen werden muss. Wenn die Anzahl der minimalen und maximalen Objektteile jedoch begrenzt ist, werden die für die Zielauflösung prädizierten Grauwerte und Linienbreiten auch hier in einem überschaubaren Bereich liegen. Falls trotz Interaktion in der dynamischen Gruppe keine Verschmelzung eintritt, können der Grauwert und die Linienbreite mit derselben Methode wie für Einzellinien prädiziert werden.

3.3.2.3 Fusion

Die Fusion fügt alle in der voran gegangenen Stufe prädizierten Objektteile zu einem oder mehreren vollständigen Objektmodell(en) in der Zielauflösung zusammen. Falls keine Skalenraumereignisse in der dynamischen Gruppe eingetreten sind, wird diese bei der Fusion wie ein einzelnes Objektteil behandelt und die Struktur des Objektmodells bleibt erhalten. Wenn Skalenraumereignisse aufgetreten sind, verändert sich dagegen die Struktur im Vergleich zur Ursprungsauflösung. In diesem Fall muss sich die Anzahl der Objektteile in der Zielauflösung ändern. Für Objektmodelle mit parallelen linienartigen Objektteilen variabler Anzahl kann sich die Anzahl der Objektteile nur reduzieren, da ausschließlich Auslöschung und Verschmelzung auftreten können.

Falls eine Verschmelzung der dynamischen Gruppe zu einer einzigen Linie eingetreten ist, d. h. $w_{dyn_max} \leq w_F$, muss die Verschiebung der Position der verschmolzenen Objektteile für die Anpassung der Distanzen in der Stufe der Fusion berücksichtigt werden. Diese Translation beeinflusst die Distanz des durch die Verschmelzung entstandenen Objektteils zu seinen Nachbarlinien. Die Distanz $d_{AzLgiLj}$ der verschmolzenen dynamischen Gruppe Lg_i zu einer benachbarten Linie L_j in der Zielauflösung A_z kann aus der dynamischen Distanz d_{dyn} in der Ursprungsauflösung A_0 und der Anzahl n der beteiligten Objektteile in der dynamischen Gruppe berechnet werden:

$$d_{AzLgiLj} = \frac{1}{F_A}\left(d_{A0LgiLj} - \tfrac{1}{2}\Delta w_{\sigma zLgi} - \tfrac{1}{2}\Delta w_{\sigma zLj} + \frac{(n-1)\cdot d_{dyn}}{2} \right) \tag{3.13}$$

für $n>1$

mit $\Delta w_{\sigma zLgi}$: Veränderung der Breite der dynamischen Gruppe i von A_0 zu σ_z

$\Delta w_{\sigma zLi}$: Veränderung der Breite der Linie i von A_0 zu σ_z

F_A: Auflösungsänderung $F_A = A_z/A_0$.

Wenn die Linien der dynamischen Gruppe zu mehreren Linien verschmolzen sind, ist deren Anzahl von der Größe des zur Glättung genutzten Gauß-Filters abhängig, also davon, wie viele Linien in der dynamischen Gruppe vom Filter überdeckt werden. Die Distanz $d_{AzLgiLj}$ der äußersten verschmolzenen Linie der dynamischen Gruppe zu den benachbarten Linien außerhalb der dynamischen Gruppe berechnet sich dann folgendermaßen:

$$d_{AzLgiLj} = \frac{1}{F_A}\left(d_{A0LgiLj} - \tfrac{1}{2}\Delta w_{\sigma zLgi} - \tfrac{1}{2}\Delta w_{\sigma zLj} + \frac{(n_F - 1)\cdot d_{dyn}}{2} \right) \tag{3.14}$$

für $n_F > 1$.

Die dynamischen Distanzen d_{dyn} in der Zielauflösung A_z innerhalb der dynamischen Gruppe, in der Verschmelzung auftritt, werden im Gradientenbild der Zielskala $G_{\sigma z}$ aus den Abständen zwischen den Kanten der neu entstandenen Linien ermittelt. Dabei werden die Maxima des Linienprofils im Gradientenbild als Kantenpositionen betrachtet.

Die Fusion wird analog zu Abschnitt 3.3.1.3 wie für Objektmodelle ohne variabler Anzahl Objektteile jeweils für die minimale und maximale Anzahl der Objektteile in der dynamischen Gruppe durchgeführt. Eine variable Anzahl von Objektteilen kann zu sehr unterschiedlichem Skalenverhalten für die untere und obere Grenze führen. Für eine sehr unterschiedliche Anzahl von Objektteilen entstehen daher häufig mehrere Modelle als Ergebnis der automatischen Anpassung.

3.3.3 Objektmodelle mit flächenhaften beliebig orientierten Objektteilen

Ein Objektmodell der hier behandelten Art kann sowohl flächenhafte Objekttypen, linienartige Objekttypen mit beliebiger Orientierung sowie eine Kombination beider enthalten. Mit flächenhaften beliebig orientierten Objektteilen können sehr viel komplexere Objekte als mit parallelen Linien modelliert werden, die nur einen begrenzten Bereich der Landschaftsobjekte abdecken können. Objektteile mit beliebiger Orientierung ermöglichen beispielsweise die Modellierung von Straßen in komplexen Stadtgebieten oder in Kreuzungsbereichen, während flächenhafte Objekttypen zur Beschreibung zusätzlicher Markierungen der Straße, z. B. Richtungspfeile, genutzt werden können.

Objektmodelle mit flächenhaften und/oder beliebig orientierten Objektteilen müssen mit der zweidimensionalen Strategie angepasst werden, da die Prädiktion des Skalenverhaltens nicht auf eine Untersuchung des Grauwertprofils (wie bei parallelen linienartigen Objektteilen) beschränkt werden kann. Flächenhafte Objektteile unterliegen im Zuge der Auflösungsänderung Veränderungen in zwei Dimensionen. Obwohl das Verfahren für die zweidimensionale Anpassung ebenfalls wie in der eindimensionalen Strategie in drei Stufen gegliedert ist, unterscheiden sich die Methoden stark, vor allem die Analyse des Skalenverhaltens. Insbesondere können im zweidimensionalen Raum zwei zusätzliche Skalenraumereignisse auftreten, die für die Prädiktion des Skalenverhaltens der Objektteile entsprechend berücksichtigt werden müssen.

3.3.3.1 Zerlegung

Interaktion zwischen Objektteilen in der Zielauflösung ist von der Distanz der Objektteile und der Größe des zur Zielskala σ_z korrespondierenden Gauß-Filters abhängig. Da die Abstände zwischen beliebig orientierten Objektteilen nicht konstant sind, ist ein einfacher Vergleich wie im eindimensionalen Fall zur Bestimmung von Interaktion nicht ausreichend. Als Basis zur Zerlegung der Objektteile in Interaktionsgruppen und nicht interagierende einzelne Objektteile wird stattdessen eine *Interaktionszone* eingeführt. Dabei handelt es sich um einen Bereich um jedes Objektteil, dessen Größe der Filtergröße w_F des diskreten Gauß-Kernels entspricht. Falls sich die individuellen Interaktionszonen von zwei Objektteilen an mindestens einer Stelle überschneiden, liegt Interaktion vor und die zugehörigen Objektteile werden einer gemeinsamen Interaktionsgruppe zugewiesen. Andernfalls werden die Objektteile separat als nicht-interagierende einzelne Objektteile analysiert. Abbildung 3.16 illustriert das Konzept der Interaktionszonen. Zur Bestimmung der Interaktion wird die im Objektmodell gegebene Konstellation der Objektteile in einem synthetischen Bild simuliert und mit Bildverarbeitungsalgorithmen ermittelt, ob sich die Interaktionszonen schneiden.

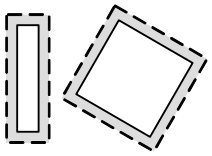


Abbildung 3.16: Interaktionszonen (grau, gestrichelte Grenze) um zwei Objektteile (weiß, durchgezogene Grenze)

Wertebereiche für Attribute

Falls Wertebereiche für die Attribute im Objektmodell der hohen Auflösung gegeben sind, wird ähnlich wie in der eindimensionalen Strategie der automatischen Anpassung verfahren. Für die minimale und maximale Attributkonfiguration werden jeweils eigene Interaktionsgruppen erzeugt und ihr Skalenverhalten wird separat analysiert. Ein großer Bereich möglicher Werte für die gegenseitige Orientierung benachbarter Objektteile impliziert eine Mehrdeutigkeit des Skalenverhaltens. Ob Interaktion zwischen zwei beliebig orientierten Objektteilen eintritt, ist von ihrer jeweiligen Orientierung abhängig. Im Gegensatz zu Distanzen muss Interaktion nicht unbedingt bei der Konstellation mit den minimalen Werten des zwischen den Objektteilen eingeschlossenen Winkels auftreten. Um Fälle der maximalen und minimalen Interaktion abzufangen, müssen die Konstellationen betrachtet werden, in denen aufgrund der variablen Orientierung die kürzeste und maximale Distanz zwischen den Objektteilen entsteht. Dazu wird um die einzelnen Objektteile jeweils ein *Interaktionskreis* mit dem Durchmesser des größten Abstandes zwischen zwei Ecken des Objektteils gezogen. An der Stelle, an denen sich die Kreise am nächsten sind, tritt potentiell maximale Interaktion auf. Die Verbindung der Interaktionskreise an dieser Stelle entspricht der Orientierung der Achsen des größten Abstandes für beide Objektteile. Die Konstellation mit der geringstmöglichen Interaktion entsteht bei paralleler Orientierung der größten Halbachse der Objektteile. Das Skalenverhalten dieser beiden Konstellationen wird in der anschließenden Stufe des Anpassungsverfahrens untersucht.

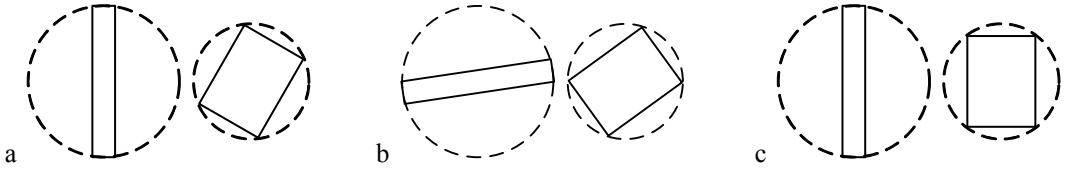


Abbildung 3.17: Interaktionskreise (gestrichelte Grenze) um zwei beliebig orientierte Objektteile (weiß, durchgezogene Grenze): Ausgangssituation (a), maximale Interaktion (b), minimale Interaktion (c)

3.3.3.2 Analyse des Skalenverhaltens

Eine analytische Prädiktion des Skalenverhaltens ist für Objekte mit flächenhaften beliebig orientierten Objektteilen in der Regel nicht möglich [HEUWOLD et al. 2007, 2008A]. Die Analyse des Skalenverhaltens für diese komplexen Objekte wird daher empirisch in einem Analyse-durch-Synthese-Prozess durchgeführt, der die Objektteile in synthetischen Bildern der Ursprungsauflösung A_0, der Zielskala σ_z und der Zielauflösung A_z untersucht. Für die Analyse wird das Konzept der Grauwert-Blobs nach LINDEBERG [1994, 1993] genutzt (im Folgenden vereinfacht als *Blobs* bezeichnet). Die Analyse berücksichtigt die möglichen Skalenraumereignisse, die bei der Auflösungsänderung auftreten können. Das Ergebnis ist eine Beschreibung der Objektteile in der gegebenen Zielauflösung mittels der Attribute der Knoten und Relationen. Das Konzept des Prozesses der Analyse-durch-Synthese ist in Abbildung 3.18 dargestellt.

In der Analyse des Skalenverhaltens wird als erstes für jede Interaktionsgruppe aus den Angaben zu den Objektteilen in den Knoten des gegebenen Semantischen Netzes der Ursprungsauflösung ein synthetisches Bild I_0 erstellt, das die Interaktionsgruppe in der Ursprungsauflösung A_0 simuliert. Im nächsten Schritt wird das Ursprungsbild I_0 in das Bild $I_{\sigma z}$ der Zielskala σ_z durch Faltung mit dem entsprechenden diskreten Gauß-Filter überführt. Daraus wird nachfolgend das Bild I_{Az} der Zielauflösung A_z durch Abtastung und bilineare Interpolation berechnet. Die Prädiktion von sicheren Skalenraumereignissen und der Attributwerte wird im Bild der Zielskala $I_{\sigma z}$ durchgeführt. Der Eintritt der effektiven Skalenraumereignisse wird dagegen im Bild der Zielauflösung I_{Az} untersucht. Der Grund ist, dass die Prädiktion des gesam-

ten Attributwertebereiches und der sicheren Skalenraumereignisse in der Zielskala im Gegensatz zur Untersuchung in der Zielauflösung einfacher ist (siehe auch 3.3.1.2).

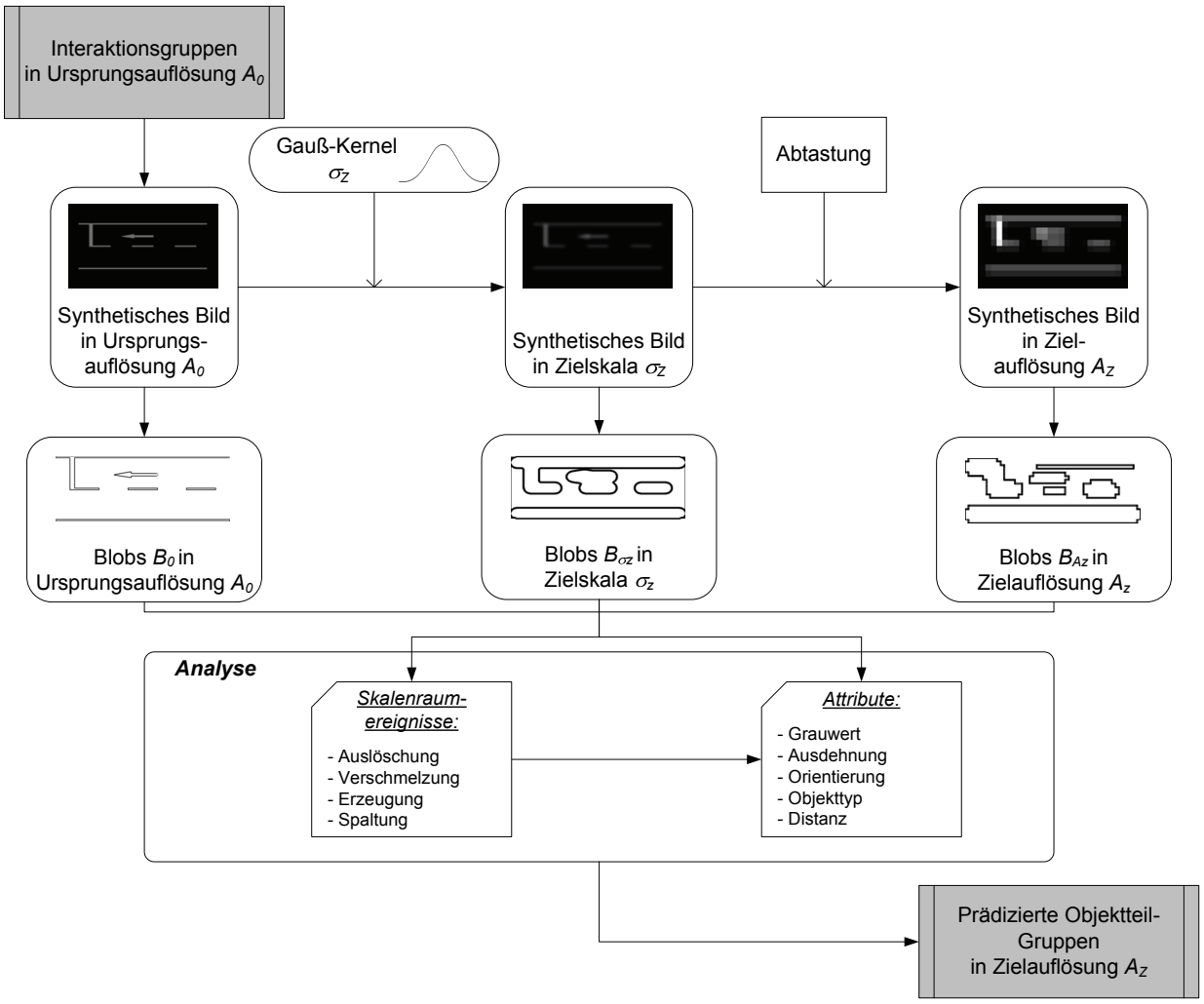


Abbildung 3.18: Konzept der Analyse-durch-Synthese für zweidimensionale Objektmodelle

Prädiktion von Skalenraumereignissen

Die Analyse der Skalenraumereignisse muss die vier möglichen Ereignisse im zweidimensionalen diskreten Skalenraum berücksichtigen – Auslöschung, Verschmelzung, Erzeugung und Spaltung. Zur Bestimmung der Ereignisse wird das Konzept der *Blobs* angewendet. Diese entsprechen im Allgemeinen den einzelnen Objektteilen der Interaktionsgruppen. Die Abgrenzung der Blobs wird auf der Basis des sequentiellen Algorithmus zur Blob-Detektion nach LINDEBERG [1994] durchgeführt (siehe Kapitel 2.2.2.2). Die Grenzen der Blobs werden in diesem Algorithmus über Wendepunkte in der Funktion der Bildintensitäten bestimmt, die in der Regel den Kanten der einzelnen Objektteile entsprechen. Zur Vermeidung von Fehlern bei der Abgrenzung wird im verwendeten Algorithmus jedoch zusätzlich nach dem Finden einer Hintergrundregion, die an einen noch wachsenden Blob angrenzt, das zukünftige Wachstum dieses Blobs angehalten und die Hintergrundregion als Sattelregion mit diesem Blob assoziiert. Damit wird sichergestellt, dass die Basisregion des Blobs nicht mehr mit anderen angrenzenden Regionen, die einen niedrigeren Grauwert als die Hintergrundregion aufweisen, erweitert wird. Ergebnis sind die Positionen der Basisregionen *Basis(B)* sowie der Extrema *E* und begrenzenden Sattelpunkte *S* der Blobs in Ursprungs- und Zielskala sowie in der Zielauflösung. Eine Konstruktion des kompletten Scale-space primal sketch, der aus den Blobs der einzelnen Skalenebenen gebildet wird, über den gesamten Skalenraum ist für eine Analyse des Skalenverhaltens bei einer vorgegebenen spezifischen Auflösungsänderung nicht erforderlich. Stattdessen ist für das Ziel der Arbeit – die Anpassung von Objektmodellen an eine bestimmte niedrigere Bildauflösung – die Analyse der Anzahl und des Aussehens der in der Zielskala bzw. Zielauflösung verbliebenen Blobs ausreichend.

Zuerst werden die Blobs sowohl im synthetischen Bild L_0 der Ursprungsskala als auch im Bild $L_{\sigma z}$ der Zielskala detektiert. Anschließend werden die Blobs zwischen der Ursprungs- und Zielskala durch den Vergleich der Blob-Eigenschaften und -Positionen in beiden Auflösungen einander zugeordnet bzw. miteinander verbunden. Die Basisregion der Ursprungsauflösung wird dazu als Suchraum in der Zielskala genutzt. Hierbei wird angenommen, dass das Extremum in der Zielskala nicht aus der korrespondierenden Basisregion in der Ursprungsauflösung herausdriftet. Nach LINDEBERG [1992] agiert die Basisregion als grobe Grenze für die Drift der lokalen Extrema: Obwohl die Driftgeschwindigkeit der lokalen Extrema in der Nähe eines Skalenraumereignisses gegen ∞ streben kann, definiert die Grauwert-Blob-Basisregion einen natürlichen räumlichen Bereich für die Suche des Blobs in der nächsten Skalenebene.

Für die Verbindung der Blobs zwischen den Skalenebenen wird davon ausgegangen, dass die meisten Blobs keinem Skalenraumereignis beim Skalenübergang unterliegen. Daher wird zuerst versucht eine sogenannte *gewöhnliche Verbindung* für alle Blobs zwischen der Ursprungsskala und der Zielskala herzustellen. Basierend auf der Menge der Basisregionen aus m Blobs in der Ursprungsskala $Basis_0$ und der Menge der Basisregionen aus n Blobs in der Zielskala $Basis_{\sigma z}$ wird im neuen Verfahren zur automatischen Anpassung gefordert, dass eine gewöhnliche Verbindung die folgende Bedingung erfüllt (siehe auch HEUWOLD et al. [2007, 2008A]):

Gewöhnliche Verbindung:

Ein spezifischer Blob in der Ursprungsauflösung hat einen und nur einen korrespondierenden Blob in der Zielskala. Die Basisregion dieses Blobs B_i in der Ursprungsauflösung $Basis_0(B_i)$ schneidet nur die Basisregion des Blobs B_q in der Zielskala $Basis_{\sigma z}(B_q)$. Keine Basisregion eines anderen Blobs in der Zielskala darf $Basis_0(B_i)$ schneiden:

$$\exists^{=1}\, q\ (Basis_0(B_i) \cap Basis_{\sigma z}(B_q) \neq 0)\ ,\ i \in \{1 \ldots m\},\ q \in \{1 \ldots n\} \tag{3.15}$$

Falls nicht für alle Blobs eine gewöhnliche Verbindung hergestellt werden kann, müssen bei der gegebenen Skalenänderung Skalenraumereignisse aufgetreten sein. Der Typ der eingetretenen Ereignisse kann automatisch abgeleitet werden. Die im Bild der Zielskala auftretenden Blob-Ereignisse sind *sichere Skalenraumereignisse*, für die im entwickelten Verfahren die folgenden Postulate festgelegt werden:

Auslöschung:

Ein spezifischer Blob in der Ursprungsauflösung hat keinen korrespondierenden Blob in der Zielskala. Die Menge der Basisregionen in der Zielskala $Basis_{\sigma z}$ ist an der Stelle der Basisregion des Blobs in der Ursprungsauflösung $Basis_0(B_i)$ leer:

$$\exists\ i\ (Basis_0(B_i) \cap Basis_{\sigma z} = 0)\ ,\ i \in \{1 \ldots m\} \tag{3.16}$$

Verschmelzung:

Zwei oder mehr Blobs in der Ursprungsauflösung haben ein und denselben korrespondierenden Blob in der Zielskala. Die Basisregionen von mindestens zwei Blobs in der Ursprungsauflösung $Basis_0(B_i)$ und $Basis_0(B_j)$ schneiden die Basisregion desselben Blobs in der Zielskala $Basis_{\sigma z}$:

$$\exists\ q\ ((Basis_0(B_i) \cap Basis_{\sigma z}(B_q) \neq 0) \wedge (Basis_0(B_j) \cap Basis_{\sigma z}(B_q) \neq 0))\ ,\ i,j \in \{1 \ldots m\},\ q \in \{1 \ldots n\},\ i \neq j \tag{3.17}$$

Erzeugung:

Ein spezifischer Blob in der Zielskala hat keinen korrespondierenden Blob in der Ursprungsauflösung. Die Menge der Basisregionen in der Ursprungsauflösung $Basis_0$ ist leer an der Stelle der Basisregion des Blobs in der Zielskala $Basis_{\sigma z}(B_q)$:

$$\exists\ q\ ((Basis_0 \cap Basis_{\sigma z}(B_q) = 0),\ q \in \{1 \ldots n\} \tag{3.18}$$

Spaltung:

Ein spezifischer Blob in der Ursprungsauflösung hat zwei oder mehr korrespondierende Blobs in der Zielskala. Die Basisregion eines Blobs in der Ursprungsauflösung $Basis_0(B_i)$ schneidet die Basisregionen $Basis_{\sigma z}(B_q)$ und $Basis_{\sigma z}(B_s)$ von mindestens zwei Blobs in der Zielskala:

$$\exists\, q\, \exists\, s\, ((Basis_0(B_i) \cap Basis_{\sigma z}(B_q) \neq 0) \wedge (Basis_0(B_i) \cap Basis_{\sigma z}(B_s) \neq 0))\,,\; i \in \{1 \ldots m\},\; q,s \in \{1 \ldots n\},\; q \neq s \quad (3.19)$$

Grauwert-Blobs können aus mehreren individuellen Objektteilen bestehen, die direkt aneinander grenzen. Daher muss die Anzahl der Blobs sowohl in der Ursprungs- als auch in der Zielskala nicht exakt mit der Anzahl der Objektteile der jeweiligen Bildauflösung übereinstimmen. Da allerdings die Zusammensetzung der Blobs in der Ursprungsauflösung aus dem Synthetisierungsprozess bekannt ist, können die individuellen Objektteile, die einen gemeinsamen Blob in der Zielskala bilden, voneinander separiert werden. In den meisten Fällen entspricht ein individueller Blob jedoch einem einzelnen Objektteil, da die meisten Objektteile nicht direkt aneinander grenzen.

Bei der Analyse von Skalenraumereignissen von Objektteilen, die in der Ursprungsauflösung bereits einen gemeinsamen Blob mit direkt aneinander grenzenden anderen Objektteilen bilden, können jedoch fälschlicherweise Auslöschungen prädiziert werden. Der verwendete Algorithmus zur Blob-Detektion grenzt die Blobs in Abhängigkeit der Basiswerte der angrenzenden Blobs ab. Wenn die Sattelregionen der Blobs bei relativ hohen Grauwerten liegen, werden Objekte mit niedrigeren Grauwerten in der Umgebung als Hintergrund eingestuft und damit komplett ignoriert. Direkt aneinander angrenzende Objektteile, die in der Ursprungsauflösung bereits durch einen gemeinsamen Blob repräsentiert sind, können möglicherweise in gröberen Skalen durch die Blob-Detektion nicht korrekt abgegrenzt werden und es können Auslöschungen prädiziert werden, obwohl das Objektteil weiterhin existiert und extrahierbar ist. Daher wird für die Objektteile, für die aufgrund der Veränderungen der Basisregion ihres Blobs Auslöschungen prädiziert werden, dieses Skalenraumereignis mit einem unabhängigen Test geprüft. Dazu wird das Skalenverhalten des Objektteils, für das eine Auslöschung aus dem Blob-Konzept heraus prädiziert wurde, ungeachtet seiner vorliegenden Interaktion zum angrenzenden Objektteil nochmals separat analysiert. Mit den in der separaten Analyse angepassten Operatorparametern wird in dem synthetischen Bild der Zielauflösung I_{Az} der gesamten Interaktionsgruppe auf eine effektive Auslöschung des Objektteils getestet. Falls die Auslöschung bestätigt werden sollte, wird das Objektteil aus dem angepassten Objektmodell entfernt. Sollte die Auslöschung jedoch nicht bestätigt werden, wird das betreffende Objektteil mit den in der separaten Analyse prädizierten Attributen in die anschließende Verfahrensstufe, die Fusion, übernommen.

Andererseits kann auch ein Objektteil aus mehreren Blobs bestehen. Dies kann für jene Objektteile auftreten, deren Operatoren nicht mit differentialgeometrischen Eigenschaften arbeiten, wie beispielsweise Template-Matching für *Muster*. Deren Extraktionsmodell entspricht nicht dem Konzept der Abgrenzung der Blobs über Wendepunkte in der Funktion der Bildintensitäten. In diesem Fall können Spaltungen in der Zielskala auftreten, deren Berücksichtigung in der Zielauflösung für die Extraktion dieser Objektteile nicht sinnvoll wäre. Durch Spaltungen wird die Größe der Objektteile und damit der Mustermatrix reduziert. Den Matching-Operatoren steht bei kleineren Mustermatrizen jedoch weniger Evidenz für die Objektextraktion zur Verfügung, so dass diese daher schlechtere Extraktionsergebnisse liefern. Für Muster, die in der Ursprungsauflösung mit einem Template assoziiert sind, werden deshalb Erzeugungen oder Spaltungen innerhalb der Ausdehnung des Musters für die Zielauflösung ignoriert. Diese Objektteile können allerdings auch eine Verschmelzung mit anderen benachbarten Objektteilen eingehen und ausgelöscht werden. Diese beiden Skalenraumereignisse sind bedeutend für die Extraktion des Musters und müssen berücksichtigt werden.

Im Anschluss an die Prädiktion der sicheren Skalenraumereignisse kann eine Untersuchung von *effektiven Skalenraumereignissen* in der Zielauflösung nötig sein. Wurde für Objektteile in der Zielauflösung ein niedriger Kontrast unterhalb der kritischen Grauwertschwelle r_{eA} prädiziert, werden im synthetischen Bild der Zielauflösung L_{Az} die korrespondierenden Operatoren mit den angepassten Parametern auf die betreffenden Objektteile angewendet. Bei Erfolg des Operators ist kein effektives Skalenraumereignis eingetreten. Versagt der Operator jedoch, wird eine effektive Auslöschung bzw. Verschmelzung für die Zielauflösung prädiziert. Wurde eine Spaltung in der Zielskala prädiziert, wird der Operator des aufgespalteten Objektteils auf das Bild der Zielauflösung angewendet. Kann der Operator mehrere Objektteile extrahieren, wird eine effektive Spaltung bestätigt und ein zusätzliches Objektteil im Modell der Zielauflösung eingefügt. Falls der Operator dagegen nicht in der Lage ist, die prädizierte Anzahl Objektteile zu extrahieren, kann die Spaltung nicht bestätigt werden und im Modell der Zielauflösung verbleibt ein einziges Objektteil. Bei einer prädizierten Erzeugung in der Zielskala werden die Operatoren von sämtlichen anderen Objektteilen auf dem Bild des erzeugten Objektteils in der Zielauflösung getestet. Bei Erfolg von mindestens einem Operator kann dem erzeugten Objektteil in der Zielauflösung ein Operator zugewiesen werden. Sollte jedoch keiner der Operatoren ein Ergebnis liefern, muss für

das betreffende Objektteil im Modell der Zielauflösung ein geeigneter Merkmalsoperator aus einer zuvor erstellten Liste von Operatoren ausgewählt werden.

Prädiktion von Attributen

Die Attribute in den Knoten eines Semantischen Netzes beschreiben das Aussehen eines Objektteils im Bild. Die Attributwerte der Objektteile in der niedrigeren Bildauflösung werden daher in den synthetischen Bildern der Zielskala $I_{\sigma z}$ und der Zielauflösung I_{Az} analysiert. Für jedes Objektteil werden aus der Basisregion der Blobs die folgenden Attribute abgeleitet:

Ausdehnung:

Die Basisregion des Blobs in der Zielskala $Basis_{\sigma z}$ *(B)*, die von seinem Sattelpunkt $S_{\sigma z}$ begrenzt ist, dient als Abschätzung für die Ausdehnung des betreffenden Objektteils in der Zielauflösung. Dabei ist der Auflösungsfaktor F_A zu berücksichtigen. Die Werte der Ausdehnungen werden entlang der beiden Hauptorientierungen des Objektteils bestimmt. Bei direkt aneinander grenzenden Objektteilen, die in einem Blob zusammengefasst sind, werden die Ausdehnungen in der Zielauflösung proportional zu den in der Ursprungsauflösung gegebenen Ausdehnungen mit dem Auflösungsfaktor F_A errechnet. Bei Objektteilen, die über mehrere Blobs repräsentiert sind, wird deren kleinstes umschließendes Rechteck für die Bestimmung der Ausdehnung herangezogen.

Das Ausdehnungsverhältnis von Mustern, deren Operatoren mit Template-Matching arbeiten, wird als unverändert in der Zielauflösung angenommen und daher unabhängig von der Basisregion des zugehörigen Blobs ermittelt. Die Ausdehnung in der Zielauflösung wird mit dem Auflösungsfaktor F_A aus den Werten der Ausdehnung in der Ursprungsauflösung berechnet.

Grauwert:

Der maximale Grauwert wird für flächenhafte Objektteile innerhalb seiner zuvor ermittelten Ausdehnungen bestimmt. Der höchste Grauwert in der Zielauflösung kann nicht größer als der Grauwert des Extremums in der Zielskala $E_{\sigma z}$ sein, der daher als maximaler Grauwert des Objektteils in der Zielauflösung gewählt wird. Der minimale Grauwert kann jedoch aufgrund der Abtastung kleiner sein als der kleinste Wert innerhalb der Basisregion $Basis_{\sigma z}$. Die Abtaststelle kann maximal um die halbe Abtastweite *sd*/2 außerhalb der Basisregion in der *Abtastzone* liegen. Zur Ermittlung des minimalen Grauwertes wird daher um die Basisregion des Blobs in der Zielskala eine Pufferzone dieses Abtastversatzes *sd*/2 erzeugt. Der kleinste Grauwert innerhalb dieser Pufferzone entspricht dem minimalen Grauwert des Objektteils in der Zielauflösung. Abbildung 3.19 verdeutlicht die Ermittlung des minimalen und maximalen Grauwertes. Für linienartige Objekttypen dagegen werden der minimale und der maximale Grauwert analog zu 3.3.1.2 bestimmt.

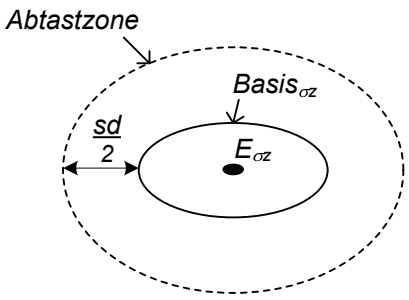


Abbildung 3.19: Ermittlung des Grauwertbereiches Δr eines flächenhaften Objektteils in der Zielauflösung in Abhängigkeit von der Abtaststelle (mit sd: Abtastweite) aus einem Blob in der Zielskala mit Extremum $E_{\sigma z}$, Basisregion $Basis_{\sigma z}$ und Abtastzone

Orientierung:

Da die Orientierung eines Objektteils hinsichtlich der Auflösungsänderung invariant ist, muss sie nur bestimmt werden, wenn Objektteile in der Zielauflösung miteinander verschmolzen sind oder durch Erzeugung bzw. Spaltung generiert wurden. Für diese Aufgabe wird die Hauptorientierung der den Blob umspannenden Bounding Box ermittelt.

Objekttyp:

Der Objekttyp kann sich für Primitive mit exakt geometrischen Formen aufgrund von Skalenraumereignissen und signifikanten Formveränderungen im Zuge einer größeren Auflösungsänderung ändern. Dann muss das betreffende Objektteil durch ein Muster repräsentiert werden, das willkürliche Formen beschreiben kann. Miteinander verschmolzene Muster führen immer zu neuen Mustern.

Anpassung der Operatorparameter

Im Anschluss werden mit den prädizierten Attributwerten die *Parameter der Operatoren* an die Zielauflösung angepasst. Die angepassten Operatoren werden dann bei Bedarf zur Prädiktion effektiver Skalenraumereignisse eingesetzt. Falls eine effektive Verschmelzung auftritt, werden die Attributwerte des durch die Verschmelzung entstandenen Objektteils auf Grundlage der neu enstandenen größeren Basisregion prädiziert. Für eine effektive Spaltung werden die Attributwerte für die kleineren Basisregionen ermittelt.

Für die zu Mustern gehörenden Merkmalsoperatoren können die neuen Templates in der Zielauflösung direkt aus dem Prozess der Analyse-durch-Synthese geliefert werden. Deren Größe wird mit dem Auflösungsfaktor F_A aus den gegebenen Templates in der Ursprungsauflösung A_0 berechnet. Die angepassten Templates werden aus dem synthetischen Bild in der Zielauflösung um den Schwerpunkt der Basisregion des zum Objektteil korrespondierenden Blobs entnommen. Bei Verschmelzungen von Mustern mit benachbarten Objektteilen entsteht wieder ein Muster. Innerhalb der ermittelten Ausdehnungen des neu entstandenen Objektteils wird aus dem synthetischen Bild der Zielauflösung I_{Az} das neue Template für das Muster entnommen.

3.3.3.3 Fusion

Alle in der Zielauflösung existierenden Objektteile einschließlich ihrer Attribute werden in der Stufe der Fusion zu einem vollständigen Objektmodell zusammengefügt. Sind für die minimale und maximale Attributkonfiguration dieselben Skalenraumereignisse eingetreten, ist die Struktur der Objektmodelle für beide Konfigurationen in der Zielauflösung gleich und das Ergebnis der automatischen Anpassung kann in einem einzigen angepassten Objektmodell repräsentiert werden. Die minimalen und maximalen Werte der Attribute aus beiden Konfigurationen bilden den Wertebereich der prädizierten Attribute in der Zielauflösung. Wurden dagegen unterschiedliche Skalenraumereignisse in beiden Attributkonfigurationen prädiziert, müssen diese Unterschiede im Skalenverhalten in der Struktur der Objektmodelle repräsentiert werden, so dass zwei verschiedene Objektmodelle das Ergebnis der Anpassung sind.

Die hierarchischen Relationen zwischen den Objektteilen bleiben unberührt, so lange keine Skalenraumereignisse aufgetreten sind. Im Falle einer Auslöschung wird das betreffende Objektteil einschließlich seiner Relationen entfernt. Für verschmolzene Objektteile verbleibt eine einzelne *Part-of*-Relation. Dagegen müssen für neu generierte Objektteile, die durch Erzeugung und Spaltung entstanden sind, zusätzliche *Part-of*-Relationen in die unterste Hierarchieebene des angepassten Objektmodells eingefügt werden. Der Typ der räumlichen Relationen (parallel oder senkrecht) bleibt unverändert. Allerdings müssen die Distanzen zwischen den Objektteilen an deren veränderte Ausdehnungen angepasst werden. Die angepassten Werte werden aus den kürzesten Distanzen zwischen den Basisregionen der Blobs in der Zielskala $Basis_{\sigma z}$ entlang der beiden Hauptorientierungen des Gesamtobjektes abgeleitet. Die Distanzen $d_{\sigma zOiOj}$ zwischen Objektteil i und j in der Zielskala werden unter Berücksichtigung des Faktors der Auflösungsänderung $F_A=A_z/A_0$ in die Distanzen der Zielauflösung d_{AzOiOj} umgerechnet:

$$d_{AzOiOj} = 1/F_A \cdot d_{\sigma zOiOj} \tag{3.20}$$

3.3.4 Objektmodelle mit lokalen Kontextobjekten

Der Extraktionsprozess eines Landschaftsobjektes von Interesse (im Folgenden *Primärobjekt* genannt) kann durch *lokale Kontextobjekte* stark beeinträchtigt werden, d. h. durch andere nahe gelegene Objekte in der Umgebung, die das Primärobjekt teilweise oder vollständig verdecken. Beispielweise sind Straßen häufig von Fahrzeugen, Bäumen, Gebäuden oder deren Schatten überdeckt. Unter dem Einfluss lokaler Kontextobjekte kann die Extraktion des Primärobjektes versagen, da einige Objektteile nicht oder nicht komplett extrahiert werden können. Daher ist eine Modellierung des Landschaftsobjektes, die sich ausschließlich auf seine inhärenten Merkmale beschränkt, in komplexen Gebieten, wie z. B. Innenstadträumen, für eine erfolgreiche Extraktion oftmals nicht ausreichend. Die Berücksichtigung von möglicherweise vorhandenen lokalen Kontextobjekten kann jedoch die Extraktionsergebnisse deutlich verbessern (siehe Abschnitt 2.1.2.2). Das entwickelte Verfahren zur automatischen Anpassung aus dem vorigen Abschnitt wird daher in Bezug auf lokale Kontextobjekte erweitert, die als optionale Objektteile in die Modellierung integriert werden können.

Die lokalen Kontextobjekte werden hier ebenfalls wie das Primärobjekt über Semantische Netze modelliert, deren Attribute ihr Aussehen in den Bildern beschreiben und denen spezifische Merkmalsoperatoren zugeordnet sind. Die Kontextobjekte sind als eigenständige Objekte über lokale Kontextrelationen mit dem Objektmodell des Primärobjektes verbunden. Während von den Objektmodellen, die bisher behandelt wurden, sämtliche Teile des Landschaftsobjektes sowie ihre Positionen zueinander bekannt waren, gilt dies nicht für lokale Kontextobjekte. Weder ihre Existenz noch ihre Position ist im Vorfeld bekannt und als Vorinformation verfügbar. Insbesondere die Position lokaler Kontextobjekte kann extrem variabel sein und kann, wenn überhaupt, nur sehr grob im Konzeptnetz angegeben werden. Diese Variabilität steht im Widerspruch zu einer eindeutigen Prädiktion des Skalenverhaltens und muss entsprechend in der Methode der Anpassung berücksichtigt werden.

Im Folgenden werden ein Konzept und eine Methode für die automatische Anpassung von Objektmodellen, die lokale Kontextobjekte berücksichtigen, an eine niedrigere Bildauflösung eingeführt und detailliert erläutert. Zuerst werden verschiedene Möglichkeiten für das Konzept der automatischen Anpassung diskutiert, die zum entwickelten Ansatz führen. Anschließend werden die Methode der Anpassung und ihre Voraussetzungen hinsichtlich der Modellierung des lokalen Kontextes und seiner Integration in die Extraktionsstrategie beschrieben.

3.3.4.1 Konzept der Anpassung

Die Position lokaler Kontextobjekte ist im Vorfeld normalerweise unbekannt. Straßen beispielsweise können von Bäumen an deren Rändern flankiert sein. Es ist in der Regel jedoch nicht bekannt, wo und ob überhaupt Bäume vorkommen. Fahrzeuge können sich sogar auf allen Teilen der Straße befinden. Während im Instanzennetz einer bestimmten Szene die genaue Position des Kontextobjektes bekannt ist, muss das Konzeptnetz aufgrund der Variabilität der Position sehr flexibel sein. Das Skalenverhalten von Objekten bei Auflösungsänderung unterscheidet sich erheblich für verschiedene relative Objektpositionen. Angenommen, es tritt für eine bestimmte Auflösungsänderung F_A ein Skalenraumereignis für die Objekte A und B ein, die relativ dicht beieinander liegen, dann kann für eine andere Objektkonstellation mit größerer Distanz zwischen A und B ein anderes oder kein Skalenraumereignis für dieselbe Auflösungsänderung F_A auftreten. Das Skalenverhalten der einzelnen Objektkonstellationen wird also stark von der relativen Position der einzelnen Objekte zueinander beeinflusst. Diese direkte Abhängigkeit erschwert eine eindeutige Prädiktion des gemeinsamen Skalenverhaltens von lokalen Kontextobjekten mit dem Primärobjekt. Das Ergebnis der automatischen Anpassung ist eine Reihe von verschiedenen Konzeptnetzen in der Zielauflösung – im Extremfall je ein Konzeptnetz für jede mögliche Konstellation. Daher kann für eine komplette Analyse des Skalenverhaltens eine sehr große Anzahl angepasster Objektmodelle in der Zielauflösung nicht ausgeschlossen werden, die die Ergebnisse für reale Anwendungen unbrauchbar macht.

Trotz dieser Schwierigkeiten sind u. a. folgende Ansätze zur automatischen Prädiktion des Skalenverhaltens von Objektmodellen mit lokalen Kontextobjekten denkbar [HEUWOLD et al. 2008B]:

1. Ansatz der Konstellationen mit den höchsten Wahrscheinlichkeiten:

   Die Analyse des gemeinsamen Skalenverhaltens der lokalen Kontextobjekte zusammen mit dem Primärobjekt wird nur für einige (wenige) Konstellationen durchgeführt, die am wahrscheinlichsten auftreten. Beispielsweise könnten für Fahrzeuge als lokale Kontextobjekte auf einer Straße im Kreuzungsbereich zwei Konstellationen betrachtet werden – Fahrzeugpositionen in der Kreuzungsmitte und an der Haltelinie. Der Vorteil dieser Vorgehensweise ist eine eindeutige Prädiktion des Skalenverhaltens für die untersuchten Konstellationen unter Berücksichtigung des umgebenden Primärobjektes. Die Variabilität der Position der lokalen Kontextobjekte ist jedoch in der Anpassung nur teilweise berücksichtigt. Dementsprechend spiegeln die angepassten Objektmodelle nicht das gesamte Potenzial der Variabilität wider. Diese Beschränkung der Position stellt einen Widerspruch zur eigentlichen Charakteristik lokaler Kontextobjekte dar, da ihr Auftreten und die daraus resultierenden Auswirkungen auf das Primärobjekt natürlicherweise extrem variieren können.

2. Statistischer Ansatz:

   Alle Konstellationen werden mit den jeweiligen Wahrscheinlichkeiten für ihr Auftreten analysiert. Für die Position der einzelnen lokalen Kontextobjekte in der Ursprungsauflösung können Wahrscheinlichkeitsdichtefunktionen aus der Erfahrung heraus oder durch statistische Untersuchungen aus Trainingsdaten modelliert werden. Das Skalenverhalten für jede Konstellation wird dann mit der jeweiligen Wahrscheinlichkeit ihres Auftretens prädiziert. Das Ergebnis ist eine Menge von Konzeptnetzen in der Zielauflösung mit statistischer Information über deren jeweilige Wahrscheinlichkeit.

3. Separater Ansatz:

   Das Skalenverhalten des Primärobjektes und der lokalen Kontextobjekte wird separat analysiert. Mit diesem Ansatz bleibt die Variabilität der Kontextobjekte auf Kosten einer exakten gemeinsamen Prädiktion des Skalenverhaltens erhalten. Die Relation des lokalen Kontextes zum Primärobjekt wird dabei vernachlässigt. Allerdings müssen die meisten Bildanalyse-Operatoren zur Extraktion lokaler Kontextobjekte aufgrund der Variabilität der Position in sehr verschiedenartigen Umgebungen arbeiten. Dadurch können die Operatoren unmöglich sämtliche Details des Primärobjektes für die Extraktion berücksichtigen. Sie gehen von einem relativ homogenen Bildhintergrund aus. Das gilt sowohl für eine hohe als auch für mittlere und niedrigere Bildauflösungen.

Das Verfahren der Anpassung von Objektmodellen, die lokale Kontextobjekte integrieren, nutzt den separaten Ansatz. Für eine realistische Objektextraktion ist der Erhalt der Flexibilität der Kontextrelation im Modell ein entscheidender Aspekt. In den meisten Fällen führt eine separate Anpassung der Objektmodelle für das Primärobjekt und die lokalen Kontextobjekte zu einem ausreichenden Ergebnis der Extraktion mit einem angepassten Objektmodell, da die meisten Bildanalyse-Operatoren von einem homogenen Bildhintergrund ausgehen und die Variabilität des Primärobjektes nicht berücksichtigen. Der statistische Ansatz kann zwar prinzipiell eine genauere Prädiktion des Skalenverhaltens ermöglichen und gleichermaßen die Variabilität der Kontextobjekte mit einbeziehen. Eine Analyse des Skalenverhaltens mit Wahrscheinlichkeitsdichten erfordert jedoch einen erheblich größeren Aufwand der Modellierung, da die Wahrscheinlichkeitsdichten für das Primärobjekt und die einzelnen Kontextobjekte erfasst werden müssen.

3.3.4.2 Kontextmodellierung

Im entwickelten Anpassungsansatz werden das Primärobjekt und die lokalen Kontextobjekte als getrennte Objekte betrachtet: Die lokalen Kontextobjekte werden nicht in ein einziges gemeinsames Objektmodell für das Primärobjekt und die Kontextobjekte eingebunden, sondern es werden separate Modelle für das Primärobjekt und für jedes lokale Kontextobjekt vorausgesetzt.

Für die Einbindung der lokalen Kontextobjekte in das Objektmodell für das Primärobjekt werden die folgenden *Kontextrelationen* für eine automatische Anpassung definiert:

- *Optional part-of*: Diese Relation beschreibt die Verbindung zwischen Primärobjekt und dem lokalen Kontextobjekt. Das lokale Kontextobjekt kann, muss aber nicht, vorhanden sein. Für eine automatische Anpassung ei-

nes Objektmodells, das lokale Kontextobjekte integriert, ist ausschließlich diese Relation für die Verbindung von Primärobjekt und Kontextobjekten zulässig.

Zusätzlich zur hierarchischen Kontextrelation *Optional part-of* kann noch der Typ der Störung des Primärobjektes im Bild über die folgenden Relationen angegeben werden:

- *occlusion:* Das lokale Kontextobjekt verdeckt ein mit dieser Relation verbundenes Objektteil des Primärobjektes.
- *shadow:* Das lokale Kontextobjekt wirft einen Schatten auf ein mit dieser Relation verbundenes Objektteil des Primärobjektes.

In der Modellierung ist auch die Berücksichtigung von mehreren lokalen Kontextobjekten möglich. Bei mehreren Kontextobjekten kann deren Priorität im Extraktionsprozess des Primärobjektes durch ein Ablaufschema für die einzelnen Kontextobjekte definiert werden. Durch diese Vorgehensweise wird eine Reihenfolge für den Aufruf der Modelle der Kontextobjekte zur Verbesserung der Extraktionsergebnisse in einem konkreten Bild festgelegt.

3.3.4.3 Extraktionsstrategie

Die im Anpassungsverfahren zugrunde gelegte Extraktionsstrategie eines Landschaftsobjektes unter Berücksichtigung von lokalem Kontext folgt einem zweistufigen Prozess, der in Abbildung 3.20 dargestellt ist [HEUWOLD et al. 2008B].

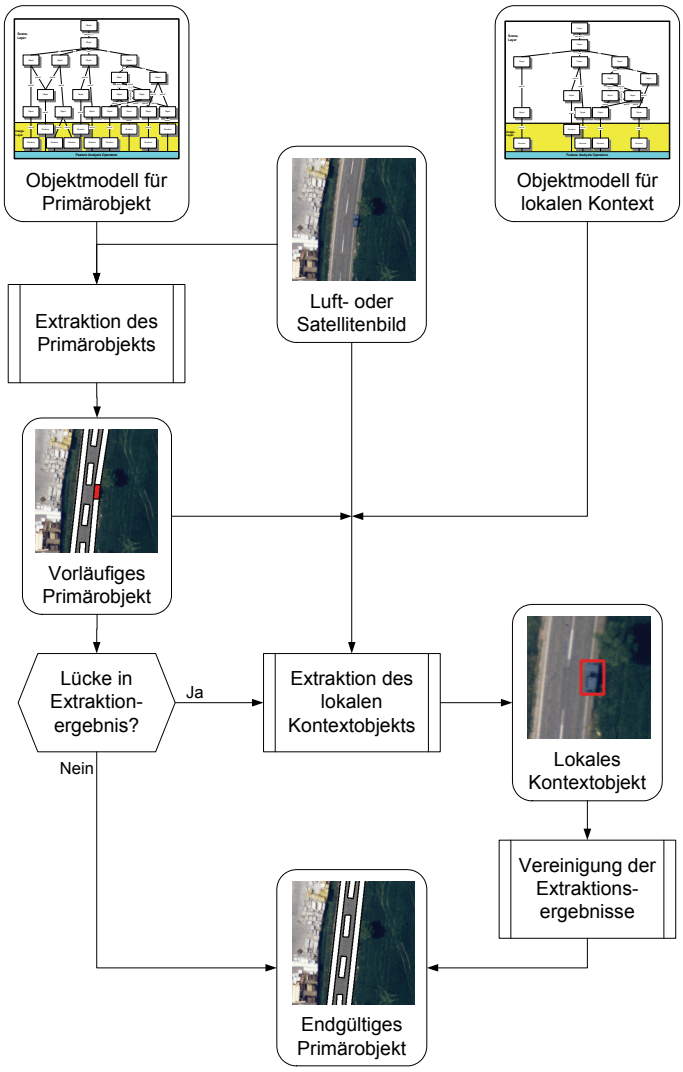


Abbildung 3.20: Extraktionsstrategie mit lokalen Kontextobjekten

Im ersten Schritt wird die Extraktion des Primärobjektes mit dem zugehörigen Objektmodell ohne Berücksichtigung der lokalen Kontextobjekte durchgeführt. Falls diese Extraktion teilweise fehlschlägt und eine Lücke im Extraktionsergebnis auftaucht, wird in einem zweiten Schritt versucht, die Lücke durch das Vorliegen eines Kontextobjektes zu erklären, das eine Störung der Bildinformation hervorruft. Daher wird eine Extraktion des lokalen Kontextobjektes unter Nutzung des gegebenen Objektmodells in dem Bereich des Bildes durchgeführt, in dem die Extraktionslücke aufgetreten ist. Kann ein lokales Kontextobjekt in diesem Bereich extrahiert werden, wird die Extraktion des Primärobjektes als vollständig angesehen und seine Existenz bestätigt. Das teilweise Versagen der Extraktion kann durch lokalen Kontext erklärt werden. Wenn dagegen kein Kontextobjekt verifiziert werden kann, verbleibt die Extraktion des Primärobjektes unvollständig. Lokale Kontextobjekte werden hier für das Schließen von Lücken im Extraktionsergebnis des Primärobjektes verwendet. Die einzige zulässige hierarchische Kontextrelation zwischen Primärobjekt und lokalem Kontextobjekt ist *Optional part-of*.

Eine einfache Modellierung der lokalen Kontextobjekte führt in vielen Fällen zu ausreichenden Extraktionsergebnissen für das Primärobjekt, da die Extraktion der Kontextobjekte in einer zuvor ausgewählten Bildregion durchgeführt wird. Detaillierte Modelle der Kontextobjekte können jedoch ebenso in die Extraktionsstrategie aufgenommen werden.

3.3.4.4 Methode der Anpassung

Die automatische Anpassung an die Zielauflösung wird für die Objektmodelle des Primärobjektes und die lokalen Kontextobjekte unabhängig voneinander durchgeführt. Für die Anpassung wird je nach Typ des Objektmodells mit Ausnahme kleiner Besonderheiten die Methode aus Abschnitt 3.3.1 für eindimensionale Objektmodelle bzw. Abschnitt 3.3.3 für zweidimensionale Objektmodelle angewendet.

Zerlegung

Die Zerlegung wird für das Objektmodell des Primärobjektes und die Modelle der einzelnen Kontextobjekte getrennt voneinander nach der in Abschnitt 3.3.1.1 bzw. Abschnitt 3.3.3.1 beschriebenen Methode durchgeführt.

Analyse des Skalenverhaltens

Das Skalenverhalten des Primärobjektes und der lokalen Kontextobjekte wird separat mit der Methode der eindimensionalen Objektmodelle bzw. mit Hilfe von Analyse-durch-Synthese für zweidimensionale Objekte prädiziert. Diese Analyse umfasst sowohl die Prädiktion aller vier möglichen Skalenraumereignisse, die während einer Auflösungsänderung auftreten können, als auch der Attribute, die das Aussehen der Objektteile im Bild beschreiben. Die für die Zielauflösung prädizierten Attribute dienen als angepasste Parameter der Operatoren. Merkmalsoperatoren für lokale Kontextobjekte nutzen üblicherweise Eigenschaften der Form und der Radiometrie – üblicherweise aus Farbbildern – für die Auswahl der jeweiligen Objekte und deren Abgrenzung zu anderen Objekten. Die Verwendung von Farbbildern erfordert die Analyse des Skalenverhaltens in mehreren Kanälen. Für jeden Kanal muss ein eigenes synthetisches Bild analysiert werden. Häufig wird für lokale Kontextobjekte aufgrund ihrer Vielseitigkeit jedoch keine Einschränkung bezüglich möglicher Farbwerte in den Attributen angegeben, d. h. das Kontextobjekt kann beliebige Farben haben. Da die Kontextobjekte mit dem Primärobjekt über die Kontextrelation *Optional part-of* verbunden sind, ist eine Simulation der Szene mit maximaler Attributkonfiguration für einen einzigen Farbkanal bzw. für ein einziges synthetisches Bild in der Analyse-durch-Synthese für die Prädiktion der Auslöschung des Kontextobjektes ausreichend. In welchem Kanal die maximale Konfiguration auftritt, ist irrelevant, weil die maximale Konfiguration mit der längsten Existenz des Objektteils im Skalenraum einhergeht. Ein Objekt(teil) mit maximaler Konfiguration kann noch in einer bestimmten gröberen Auflösung existieren, während ein Objekt mit geringerem Kontrast und geringerer Ausdehnung möglicherweise nicht mehr extrahierbar ist.

Fusion

Solange das lokale Kontextobjekt in der Zielauflösung noch extrahierbar ist, bleibt der Typ der Kontextrelation zwischen dem Primärobjekt und dem lokalen Kontextobjekt unverändert. Wenn das lokale Kontextobjekt aufgrund einer Auslöschung nicht mehr extrahiert werden kann, wird auch die betreffende Kontextrelation gelöscht.

3.4 Diskussion

Mit dem in diesem Kapitel vorgestellten Verfahren ist erstmals eine automatische Anpassung von Bildanalyse-Objektmodellen, die für eine bestimmte Bildauflösung erstellt wurden, an eine niedrigere Bildauflösung möglich. Die Modelle der Landschaftsobjekte sind dabei in Form Semantischer Netze als Konzeptnetze repräsentiert. Bei der Anpassung werden mögliche *Skalenraumereignisse* der Objektteile bei der Auflösungsänderung berücksichtigt, die zu erheblichen strukturellen Veränderungen im angepassten Objektmodell führen können. Des Weiteren wird das Aussehen der einzelnen Objektteile bezüglich *geometrischer und radiometrischer Eigenschaften der Objektteile* bzw. ihrer Beziehungen im Bild der niedrigeren Auflösung prädiziert. Mit diesen angepassten Eigenschaften der Objektteile können auch die *Operatoren zur Merkmalsextraktion* an die neue Bildauflösung angepasst werden. Damit wird eine direkte Verbindung zu ausführenden Prozessen auf dem Bild im Verfahren der automatischen Anpassung gewährleistet.

Die automatische Anpassung mit dem entwickelten Verfahren ist nur für Bilddaten in der Zielauflösung gültig, die vom selben Sensortyp wie die Bilddaten der hohen Auflösung stammen, für die das gegebene Objektmodell erstellt wurde. Für eine Anpassung der Objektmodelle an niedriger aufgelöste Bilddaten anderer Sensortypen müssen mit Hilfe eines Sensormodells zusätzlich bei der Prädiktion des Aussehens der Objektteile in der Zielauflösung die individuellen radiometrischen und geometrischen Unterschiede der einzelnen Sensoren berücksichtigt werden. In Anbetracht der verfügbaren Menge verschiedener Luft- und Satellitenbildsensoren erscheint diese Möglichkeit für die Zielsetzung dieser Arbeit nicht zweckmäßig. Das entwickelte Verfahren zur Anpassung kann bei Bedarf für eine begrenzte Anzahl verschiedener Sensortypen nach dem vorliegenden Konzept erweitert werden.

Im vorgestellten Verfahren der automatischen Anpassung werden *Semantische Netze* als Repräsentationsform des Objektmodells vorausgesetzt. Semantische Netze dienen dabei jedoch ausschließlich als Träger von Informationen über die beschriebenen Landschaftsobjekte. Für den Kern des Verfahrens – der Prädiktion des Skalenverhaltens – sind dagegen keine Semantischen Netze erforderlich. Die Beschreibung des Objektes kann ebenso gut in anderer Form repräsentiert werden. In diesem Fall müsste die Objektinformation aus einer anderen Beschreibungsform ausgelesen und für die automatische Analyse entsprechend aufbereitet werden.

Zur Prädiktion des Skalenverhaltens wird das Aussehen des modellierten Landschaftsobjektes für die gegebene Auflösungsänderung mit einem synthetischen Bild simuliert. In einer *Analyse-durch-Synthese* werden über eine Simulation der Objektszene die Veränderungen des Objektes zwischen der hohen und der niedrigeren Bildauflösung automatisch analysiert und die Ergebnisse für ein angepasstes Objektmodell anschließend aufbereitet. Synthetische Daten bieten den Vorteil, das Skalenverhalten für jede mögliche zulässige Objektbeschreibung analysieren und das Objektmodell dementsprechend anpassen zu können. Bei einer Analyse von realen Bilddaten dagegen besteht eine Abhängigkeit von den zuvor ausgewählten Szenen. Damit ist eine Berücksichtigung aller möglichen Konstellationen von Objektteilen nicht möglich und die Anpassung wäre auf die zuvor untersuchten Beispiele beschränkt.

Aufgrund verschiedener Objekttypen (linienartig und flächenhaft), für die eine Beschreibung zulässig ist, und der Integration möglicher lokaler Kontextobjekte in den Extraktionsprozess, ist eine Anpassung flexibler Objektmodelle möglich. Der Prädiktion des Skalenverhaltens von Objektmodellen sind jedoch Grenzen bei zu starker Variabilität des Aussehens der modellierten Objekte gesetzt. Bei variabler Geometrie, Radiometrie und Topologie der einzelnen Objektteile ist das Skalenverhalten der Objekte im Rahmen der Auflösungsänderung nicht eindeutig prädizierbar und es kann nicht immer eine eindeutige Lösung in Form eines einzigen angepassten Objektmodells erzielt werden. Trotzdem ist eine Flexibilität der anpassbaren Objektmodelle auf Kosten der Eindeutigkeit der Anpassung des gegebenen Modells möglich. Bei variablen Attributwerten und variabler Anzahl gleichartiger Objektteile werden mehrere Objektmodelle für die Zielauflösung als Ergebnis der automatischen Anpassung prädiziert. Eine zu starke Variabilität würde eine Explosion der Zahl möglicher Modelle der Zielauflösung hervorrufen, die nicht nur Probleme beim Umgang bei einer Extraktion in der Zielauflösung verursachen, sondern auch eine Explosion der Rechenzeit für die automatische Anpassung verursachen. Einige Modelle (z. B. bis zu vier) sind als Ergebnis jedoch noch in angemessener Zeit berechenbar und bleiben übersichtlich.

Die vorgestellten Methoden zur automatischen Anpassung wurden am Beispiel Straßen entwickelt. Landschaftsobjekte, die für ihre Modellierung andere Objekttypen benötigen, sind mit den präsentierten Methoden nicht anpassbar. Die Methoden zur Anpassung müssten dann auf diese Objekttypen erweitert werden. Andere linienhafte Objekte, z.B. Flüsse oder Bahntrassen, sind mit den entwickelten Methoden möglicherweise durch wenige Veränderungen anpassbar. Objektmodelle für Gebäude beispielsweise müssen dagegen sehr flexibel sein, da das Aussehen von Häusern extrem variabel ist. Ihre Formen folgen nicht bzw. nur eingeschränkt einer Konvention und ihrer Funktionalität wie bei Straßen. Außerdem kann die Dachfarbe sehr unterschiedlich ausgeprägt sein. Wie bereits ausgeführt, bewirkt extreme Variabilität der Attribute der Objektmodelle jedoch ein mehrdeutiges Skalenverhalten, so dass eine Explosion der Ergebnismenge resultieren kann.

Für die Anpassung von Objektmodellen, die *lokale Kontextobjekte* berücksichtigen, wurde eine separate Anpassung der Objektmodelle für das Primärobjekt und für die Kontextobjekte gewählt. Obwohl diese Strategie keine exakte Prädiktion des gemeinsamen Skalenverhaltens von lokalem Kontext mit dem Primärobjekt erzielt, entspricht dieses Konzept der üblichen Vorgehensweise der Bildanalyse-Operatoren für lokale Kontextobjekte (vgl. beispielsweise MÜLLER et al. [2003], HINZ [2004], ERIKSON & OLOFSSON [2005]). Aufgrund der extremen Variabilität der Umgebung für lokale Kontextobjekte fokussiert deren Extraktion ausschließlich auf die eigenen charakteristischen Merkmale der lokalen Kontextobjekte, die sie von ihrer Umgebung im Allgemeinen hervorheben. Die getrennte Anpassung liefert ausreichend gute Ergebnisse, solange die Umgebung des Kontextobjektes relativ homogen ist und die Interaktion mit anderen Objekten gering bleibt. Bei zu starker Auflösungsänderung in einem heterogenen urbanen Gebiet wird diese Annahme allerdings verletzt und Fehler in der Prädiktion des Skalenverhaltens und damit der exakten Anpassung der Objektmodelle können auftreten. Eine interessante Alternative zum hier gewählten separaten Ansatz ist ein statistischer Ansatz, der das Skalenverhalten des Primärobjektes und der lokalen Kontextobjekte gemeinsam mit den Wahrscheinlichkeitsdichten für ihr Auftreten analysiert. Dazu müssten allerdings umfassende Untersuchungen bzgl. des stochastischen Rahmens vorgenommen werden. Außerdem wäre ein erheblicher Mehraufwand der Modellierung wie bei den meisten statistischen Ansätzen nötig, da die Wahrscheinlichkeitsdichten für das Primärobjekt sowie für sämtliche Kontextobjekte vorliegen müssen. Der Vorteil wäre die exakte Prädiktion des Skalenverhaltens unter Berücksichtigung der Variabilität der Kontextobjekte.

Wie bedeutend die Einschränkungen aufgrund der formulierten Randbedingungen und des gewählten Konzeptes des entwickelten Verfahrens sind, soll im nächsten Kapitel anhand von realen Beispielen analysiert werden.

4 Bewertung des neuen Verfahrens

Die Ziele dieses Kapitels sind die Demonstration und Bewertung der entwickelten Methoden zur automatischen Anpassung von Bildanalyse-Objektmodellen an niedrigere Bildauflösungen. Untersucht werden soll vor allem die Qualität des Verfahrens der automatischen Anpassung. Dazu muss der Erfolg der angepassten Modelle bei der Objektextraktion bewertet werden. Weiterhin soll eine Aussage zur Übertragbarkeit des Verfahrens getroffen werden können: Wo sind die Grenzen des Verfahrens? Welche Modelle sind anpassbar? Welche nicht?

Die Analyse des entwickelten Verfahrens erfolgt an realen Beispielen für Objektmodelle und Bilddaten. Zuerst werden Beispiele für die verschiedenen Methoden der automatischen Anpassung, die in Abschnitt 3.3 vorgestellt wurden, gezeigt. Zur Demonstration des Prozesses der automatischen Anpassung wird eine Anpassung von drei verschiedenen Objektmodellen an jeweils drei verschiedene Bildauflösungen durchgeführt. Das erste der untersuchten Objektmodelle ist vom Typ linienartiger paralleler Objektteile. Das zweite Beispiel besteht aus linienartigen und flächenhaften Objektteilen und das dritte Modell integriert zusätzlich lokale Kontextobjekte. Die hohe Bildauflösung, für die die gegebenen Modelle erstellt wurden, ist für alle Beispiele 0.03 m/Pixel. Die Zielauflösungen reichen bis zu 5.00 m/Pixel und zeigen repräsentativ die Möglichkeiten und Probleme sowie die Grenzen der automatischen Anpassung auf. Anschließend wird zur umfangreichen Bewertung der Leistungsfähigkeit des entwickelten Verfahrens eine Verifikation der Methoden durchgeführt. Dazu werden die zuvor präsentierten Beispiele mit den gegebenen Modellen und den automatisch angepassten Modellen auf reale Bilddaten angewendet. Der Vergleich der Extraktionsergebnisse aus verschiedenen Bildauflösungen wird zur Bewertung der entwickelten Methoden herangezogen.

Das Kapitel schließt mit einer Diskussion der Ergebnisse, aus denen Schlussfolgerungen über die Qualität und die Grenzen des neuen Verfahrens abgeleitet werden.

4.1 Beispiele zur automatischen Anpassung von Objektmodellen

Als Anwendungsobjekt für die Bewertung des entwickelten Verfahrens zur automatischen Anpassung wurden Straßen ausgewählt. Die Bewertung wird anhand von repräsentativen Beispielen für die Typen der Objektmodelle durchgeführt, für die verschiedene Methoden zur automatischen Anpassung in Kapitel 3 vorgestellt wurden. Das Beispielobjekt Straße wurde gewählt, da die Modellierung von Straßen mit Semantischen Netzen gut erforscht ist (siehe Abschnitt 2.1.2.2) und die bestehende Vielseitigkeit der verschiedenen Straßentypen und mit ihnen verbundenen lokalen Kontextobjekten eine umfassende Bewertung des neuen Verfahrens ermöglicht. Dazu wird in diesem Abschnitt je ein Beispiel für eine einfache Straße mit parallelen linienartigen Objektteilen (Beispiel 1), für ein komplexes Straßensegment im Kreuzungsbereich mit linienartigen und flächenhaften Objektteilen (Beispiel 2) und für eine Straße mit einem Fahrzeug als lokales Kontextobjekt (Beispiel 3) vorgestellt. Die drei Objektmodelle wurden für eine Bildauflösung von 0.03 m/Pixel manuell erstellt und werden automatisch an jeweils drei niedrigere Bildauflösungen angepasst. Die Ergebnisse der Zwischenschritte für alle drei Stufen der automatischen Anpassung werden detailliert erläutert und die angepassten Objektmodelle gezeigt.

Die Berechnung der Zielskala σ_z, d.h. der nötigen Glättung vor der Auflösungsänderung, erfolgt in Abhängigkeit des Faktors F_A zwischen Zielauflösung A_z und Ursprungsauflösung A_0 mit $F_A=A_0/A_z$. Wie in Kapitel 2 erläutert, sollte die Gauß-Funktion zur Vermeidung von Abtastfehlern an der gewählten Grenzfrequenz k genügend klein geworden sein. Abbildung 4.1 illustriert beispielhaft die Gauß-Funktionen für σ=2 und σ=3 im Frequenzraum. Wenn als Vielfaches d von σ_{Gauss} im Frequenzraum $d=\pi$ gewählt (siehe Gleichung 2.3) wird, sind die Werte der beiden Funktionen an den zugehörigen Grenzfrequenzen 1/2 bzw. 1/3 bereits sehr klein, so dass die Dämpfung der Bildfrequenzen an dieser Stelle fast vollständig gegeben ist. Damit gilt folgendes Verhältnis zwischen Auflösungsänderung F_A und Zielskala σ_z:

$$\sigma_z = F_A = \frac{A_0}{A_z} \tag{4.1}$$

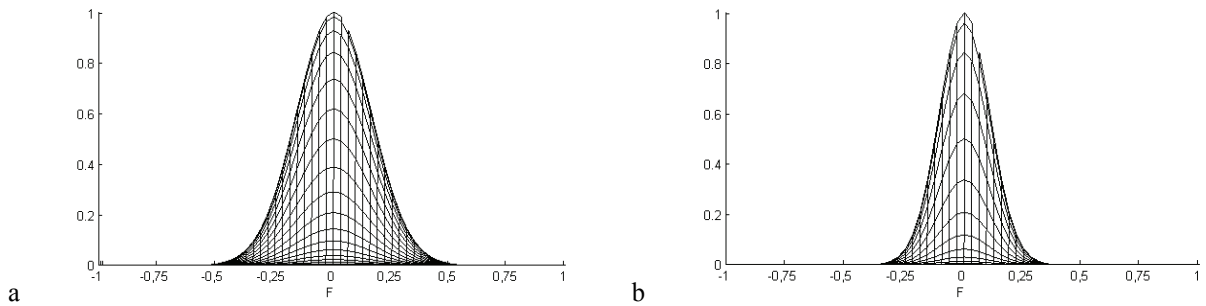


Abbildung 4.1: Gauß-Funktion im Frequenzraum für σ=2 (a) und σ=3 (b)

JÄHNE [1992] konnte zeigen, dass der Binomialfilter B^{16} der am besten geeignete Filter für die Erzeugung der nächsten Bildpyramidenstufe ist, für deren Faktor der Auflösungsänderung F_A=2 gilt. Der Binomialfilter B^{16} approximiert eine Gauß-Funktion mit σ=2. Dieses Ergebnis steht damit im Einklang mit Gleichung 4.1.

4.1.1 Beispiel 1 – Einfache Straße

In diesem Abschnitt wird ein Beispiel zur Anpassung eines Objektmodells gezeigt, das ausschließlich aus parallelen linienartigen Objektteilen besteht. Die Attribute der einzelnen Objektteile und die Distanzen sind als Wertebereiche der Anzahl der Pixel im Bild gegeben. Für die automatische Anpassung wird die in Abschnitt 3.3.1 vorgestellte Methode auf ein Objektmodell für die Extraktion einer vierspurigen Straße angewendet. Die Anpassung wird für drei niedrigere Bildauflösungen (0.10 m, 1.00 m und 5.00 m) demonstriert.

4.1.1.1 Objektmodell für die hohe Bildauflösung

Das Semantische Netz für die Extraktion einer vierspurigen Straße mit einer Doppellinie aus einem panchromatischen Luftbild von ~0.03 m/Pixel Bildauflösung wurde manuell erstellt und erfüllt die in Abschnitt 3.1 formulierten Bedingungen an den Aufbau automatisch anpassbarer Objektmodelle. Das Objektmodell für die hohe Auflösung ist in Abbildung 4.2 dargestellt. Die synthetischen Bilder zur Simulation der modellierten Szene zeigen Abbildung 4.7 a+b.

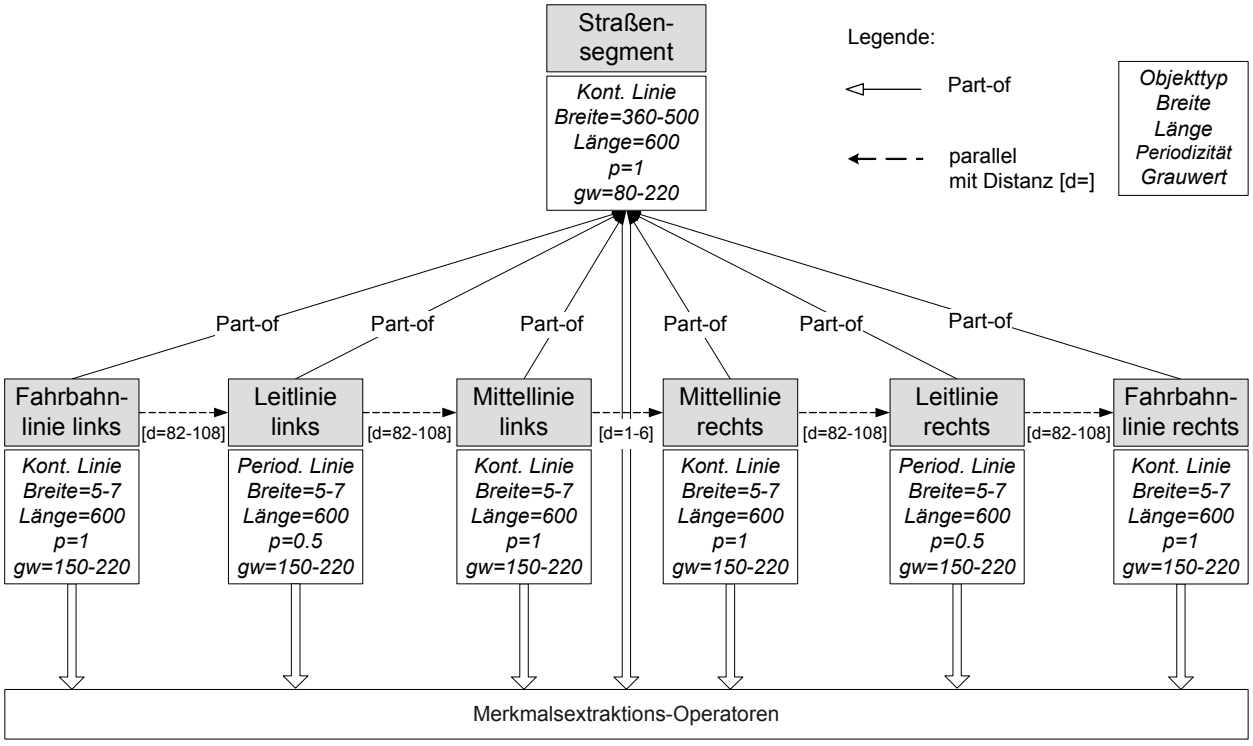


Abbildung 4.2: Objektmodell für eine einfache Straße in 0.03 m/Pixel

Das Semantische Netz des Beispiels enthält ein Straßensegment und die Fahrbahnmarkierungen als Objektteile der Straße, die als eigenständige Knoten repräsentiert sind. Der hierarchisch höchste Knoten *Straßensegment* ist als kontinuierliche Linie modelliert. Die Fahrbahnmarkierungen liegen in der unteren Hierarchiestufe und sind Teil des Straßensegments. Die *Fahrbahnlinien* begrenzen die Fahrbahn auf beiden Seiten als durchgezogene Markierungen und sind daher als kontinuierliche Linien modelliert. *Leitlinien*, die die einzelnen Fahrspuren mit gleicher Fahrtrichtung voneinander abgrenzen, erscheinen als gestrichelte Linien im Bild und werden daher als periodische Linie beschrieben. Die *Mittellinien* links und rechts liegen sehr nah beieinander, erscheinen als durchgezogene Doppellinie im Bild und werden daher ebenfalls als kontinuierliche Linien im Modell repräsentiert. Sämtliche Knoten des Objektmodells beinhalten nicht nur Informationen über den jeweiligen Objekttyp, sondern auch Werte der Attribute *Ausdehnung*, *Grauwert* und *Periodizität*. Jeder Knoten hat eine Verbindung zu einem spezifischen, für die Extraktion des repräsentierten Objektteils aus der gegebenen Bildauflösung geeigneten, Merkmalsextraktions-Operator. Für sämtliche kontinuierliche Linien wird derselbe Operator zur Merkmalsextraktion genutzt (siehe Abschnitt 4.2). Sowohl der Operator für die kontinuierlichen Linien als auch der für die periodischen Linien basiert auf dem Linienoperator nach STEGER [1998A], der für Fahrbahnmarkierungen optimiert wurde.

Die Angabe der räumlichen Relationen und der zugehörigen Distanzen zwischen den Objektteilen derselben Hierarchiestufe ist für den automatischen Anpassungsprozess entscheidend. Die Distanz zwischen den Fahrbahnmarkierungen entspricht im obigen Beispiel der Breite einer einzelnen Fahrspur, mit Ausnahme der beiden Mittellinien, die einen kürzeren Abstand zueinander haben.

4.1.1.2 Automatische Anpassung

Zielauflösung 0.1 m

Die erste gewählte Zielauflösung von A_{z1}=0.1 m entspricht einer Auflösungsänderung F_{A1}=3. Sie erfordert nach (4.1) eine Glättung mit einem Gauß-Filter mit σ_{z1}=3. Für den Filter wird nach Gleichung (3.4) eine Breite von 19 Pixeln gewählt.

In der Stufe der *Zerlegung* wird für jedes benachbarte Linienpaar in der niedrigsten Hierarchieebene untersucht, ob Interaktion in der Zielskala besteht. Falls Interaktion vorliegt, werden die betroffenen Linien in Interaktionsgruppen zusammengefasst. In den Interaktionsgruppen wird das Skalenverhalten der einzelnen Linien in der folgenden Stufe des Verfahrens dann gemeinsam analysiert. Falls keine Interaktion vorliegt, werden die Linien als einzelne Objektteile prozessiert. Da die Distanzen als Wertebereiche vorliegen, müssen sowohl die maximalen als auch die minimalen Distanzen untersucht werden. Für die erste Zielauflösung A_{z1}=0.1m besteht für die maximale und minimale Distanz Interaktion zwischen den beiden Mittellinien, die im Vergleich zur Größe des korrespondierenden Gauß-Kernels eine kleinere Distanz zueinander haben. Alle anderen benachbarten Linien haben Distanzen, die die Größe des Gauß-Kernels überschreiten und üben daher keine Interaktion in der Zielskala aus. Das gegebene Objektmodell der Ursprungsauflösung A_0=0.03 m wird daher jeweils für die minimale und maximale Attributkonfiguration in vier einzeln zu analysierende Linien und eine Interaktionsgruppe, bestehend aus den beiden Mittellinien, zerlegt. Da beide Leitlinien und die Fahrbahnlinien am linken und rechten Fahrbahnrand identische Attribute aufweisen, besitzen sie dasselbe Skalenverhalten. Daher wird hier die Untersuchung nur einer Leitlinie und einer Fahrbahnlinie demonstriert. In der folgenden Stufe des Verfahrens, der Analyse des Skalenverhaltens, wird daher jeweils das Skalenverhalten für zwei einzelne Linien und eine Interaktionsgruppe prädiziert.

Die *Analyse des Skalenverhaltens* wird für die Interaktionsgruppe vollständig über Analyse-durch-Synthese und für die Einzellinien zum Teil mit analytischen Berechnungen durchgeführt. Im Prozess der Analyse-durch-Synthese wird zur Simulation der zu untersuchenden Interaktionsgruppe je ein synthetisches Bild mit den maximalen und den minimalen Attributwerten aus den Informationen über die Objektteile im Objektmodell der Ursprungsauflösung A_0 erzeugt. Durch Überführung der synthetischen Bilder in die Zielskala σ_z und die Zielauflösung A_z können die Eigenschaften der Objektteile in der Zielskala bzw. der Zielauflösung untersucht werden. Die synthetischen Bilder der Interaktionsgruppe der beiden Mittellinien für die minimale und maximale Attributkonfiguration in der Ursprungsauflösung, der Zielskala und

der Zielauflösung zeigt Abbildung 4.3. Im Folgenden sind zur besseren Visualisierung die Grauwerte in den Abbildungen sämtlicher synthetischer Bilder invertiert dargestellt.

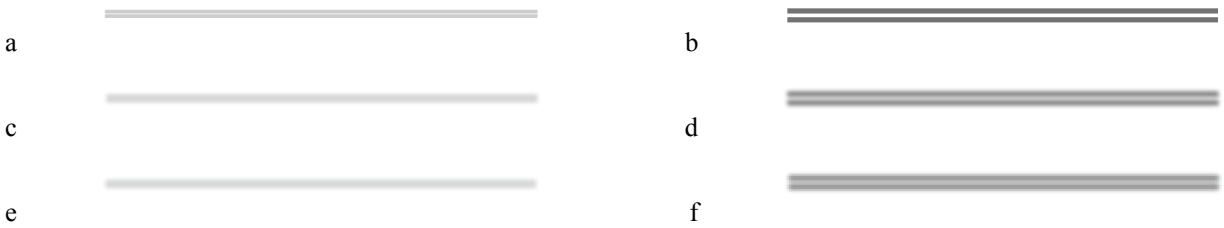


Abbildung 4.3: Synthetische Bilder der Interaktionsgruppe der minimalen (linke Spalte) und maximalen Konfiguration (rechte Spalte) in der Ursprungsauflösung A_0 (a+b), der Zielskala σ_{z1} (c+d) und der Zielauflösung A_{z1} (e+f)

Zuerst wird die Interaktionsgruppe der minimalen und maximalen Konfiguration auf Skalenraumereignisse in der Zielskala untersucht. Dazu werden die Grauwert-Maxima in den Linienprofilen in den synthetischen Bildern der Zielskala ermittelt und mit der Anzahl der Linien in der Ursprungsauflösung verglichen. Für die minimale Konfiguration ist im vorliegenden Beispiel mit zwei interagierenden Linien eine Verschmelzung als sicheres Skalenraumereignis eingetreten, da nur noch ein einziges Maximum im Linienprofil in der Zielskala existiert. Dagegen sind in der maximalen Konfiguration mit der größeren Distanz zwischen den beiden Mittellinien auch in der Zielskala zwei Grauwert-Maxima existent. Damit ist für die maximale Konfiguration keine sichere Verschmelzung eingetreten. Allerdings kann auch hier effektiv eine Verschmelzung oder Auslöschung auftreten, falls die zugeordneten Linienoperatoren aufgrund niedrigerer Kontraste im Bild der Zielauflösung nicht in der Lage sind, zwei Linien zu extrahieren. Bei niedrigem Kontrast werden die Operatoren mit den angepassten Parametern auf das synthetische Bild der Zielauflösung I_{Az} angewendet und das Ergebnis geprüft. Dieser Test auf effektive Skalenraumereignisse kann jedoch erst nach der Anpassung der Attributwerte erfolgen, die als neue Parameterwerte für die Operatoren dienen. Für die Einzellinien wird keine Verschmelzung geprüft, da ohne Interaktion keine Verschmelzung auftreten kann.

Nachdem die sicheren Skalenraumereignisse prädiziert wurden, werden die Attributwerte der Objektteile in der Zielauflösung ermittelt. Der Bereich der möglichen Grauwerte der Linie kann über die analytische Berechnung der Faltungsintegrale an der Position der Linienmitte und an der um eine halbe Pixelweite von der Linienmitte versetzten Abtaststelle berechnet werden (siehe Abbildung 3.10). Dazu wird für jede Einzellinie der Grauwert der Linienmitte in der Zielskala nach Gleichung (3.7) berechnet. Die Pixelweite entspricht dabei dem Faktor der Auflösungsänderung F_{A1}=3. Die Breite der Linie in der Zielauflösung wird über die Distanz der Kanten berechnet, die über die Maxima des Linienprofils im Gradientenbild der Zielskala ermittelt werden.

Die synthetischen Bilder der kontinuierlichen und periodischen Linien, die in der Analyse-durch-Synthese genutzt werden, zeigen Abbildung 4.4 und Abbildung 4.5. Für die beiden periodischen Linien erfolgt auch eine Prädiktion der Periodizität in der Zielauflösung. Ähnlich zur Linienbreite wird dazu das Gradientenbild der simulierten Linie in der Zielskala genutzt. Anders als bei der Linienbreite werden für die Bestimmung der Periodizität allerdings die Maxima im Grauwertprofil entlang der Linienrichtung ausgewertet.

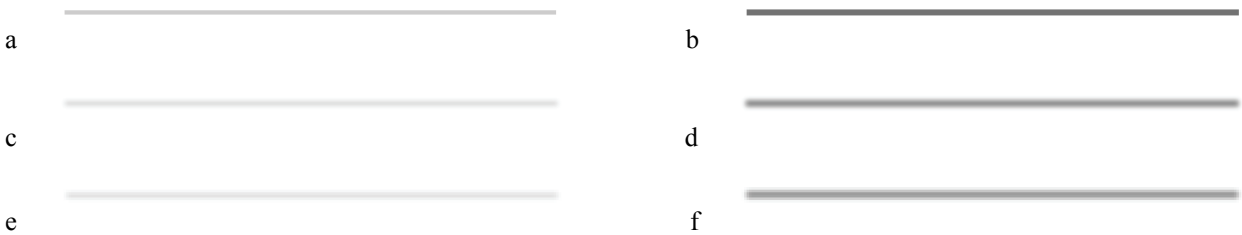


Abbildung 4.4: Synthetische Bilder der kontinuierlichen Einzellinien für die minimale (linke Spalte) und maximale Konfiguration (rechte Spalte) in der Ursprungsauflösung A_0 (a+b), der Zielskala σ_{z1} (c+d) und der Zielauflösung A_{z1} (e+f)

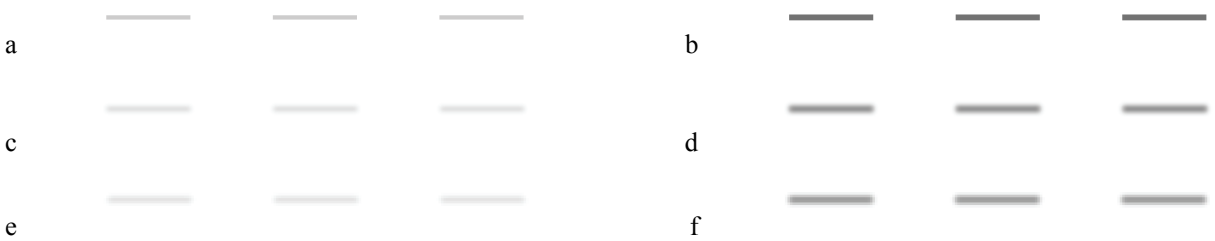
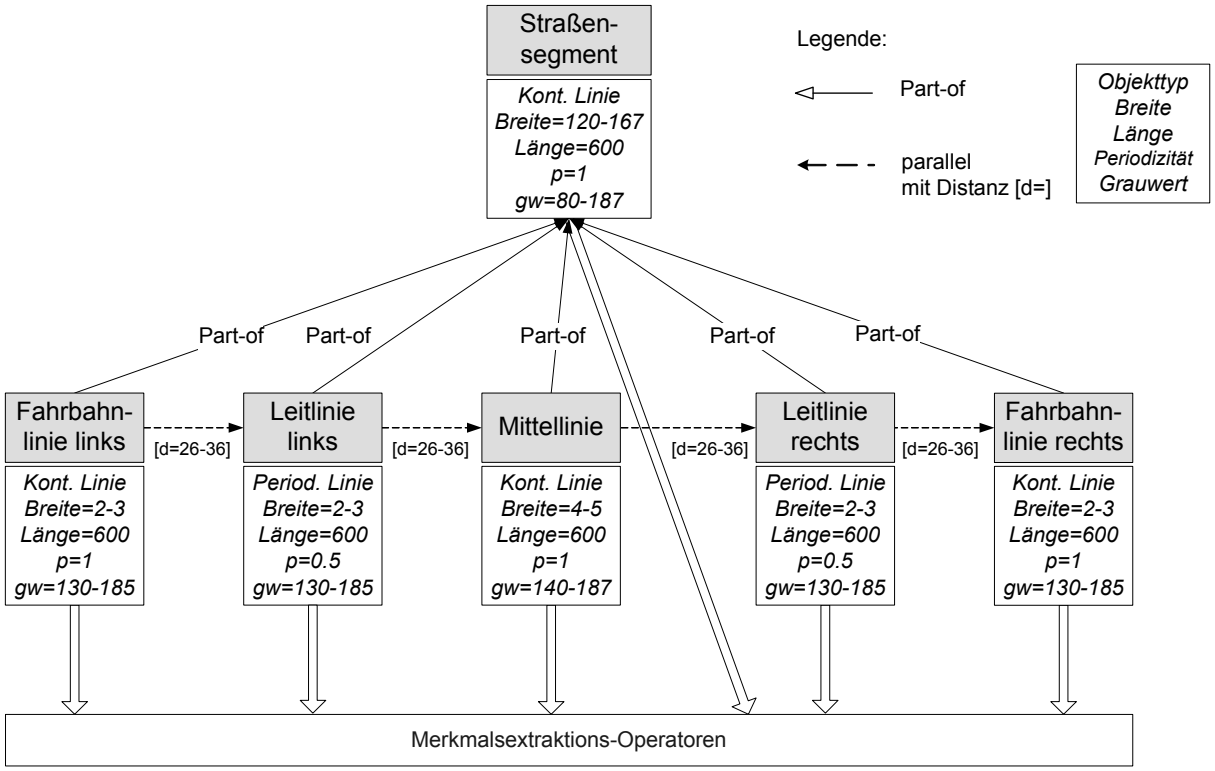


Abbildung 4.5: Synthetische Bilder der periodischen Einzellinien für die minimale (linke Spalte) und maximale Konfiguration (rechte Spalte) in der Ursprungsauflösung A_0 (a+b), der Zielskala σ_{z1} (c+d) und der Zielauflösung A_{z1} (e+f)

Mit den prädizierten Attributen können anschließend die Parameter der Linienoperatoren an die Zielauflösung A_{z1}=0.1m angepasst werden. Für den in den Linienoperatoren eingesetzten Steger-Operator ist eine automatische Anpassung der Parameter bei Kenntnis der Linienbreite und des Kontrastes zur Umgebung in der Zielauflösung möglich. Die anzupassenden Parameter entsprechen der Linienbreite sowie einer minimalen und maximalen Kontrastschwelle, die über die Faltung der Linienfunktion mit der zweiten Ableitung der Gauß-Funktion mit dem entsprechenden σ berechnet werden können [STEGER 1998A]. Diese maximale Kontrastschwelle wird mit dem prädizierten minimalen Grauwert berechnet. Für die minimale Kontrastschwelle wird das Verhältnis (0.4) zur maximalen Kontrastschwelle aus den gegebenen Parametern in der hohen Bildauflösung beibehalten.

Mit den angepassten Operatoren kann nun, falls nötig, ein Test auf effektive Skalenraumereignisse durchgeführt werden. Falls der berechnete Grauwert der Einzellinie bzw. Interaktionsgruppe in der Zielskala sehr klein ist, ist auch der Kontrast der Linie zur Umgebung schwach. Der Linienoperator kann bei der Extraktion bei schwachem Kontrast versagen und somit eine effektive Auslöschung bewirken. Bei schwachem Kontrast wird der Linienoperator mit angepassten Parametern auf das synthetische Bild der Zielauflösung angewendet und auf effektive Auslöschung getestet. Als Grauwertschwelle, unter der ein Test auf effektive Auslöschung durchgeführt wird, wird r_{eA}=15 gewählt. Oberhalb dieses Kontrastes sollte ein Linienoperator eine Linie sicher extrahieren können. Für die Interaktionsgruppe beträgt der prädizierte Kontrast zwischen den Maxima und dem von den beiden Linien eingeschlossenen Minimum in der Zielskala für die maximale Konfiguration 31 Graustufen. Damit wird kein Test einer effektiven Verschmelzung durchgeführt und die beiden Linien bleiben in der maximalen Konfiguration erhalten. Sowohl für die kontinuierliche als auch für die periodische Einzellinie beträgt der minimale Grauwertkontrast 28, so dass hier ebenfalls kein Test auf effektive Auslöschung durchgeführt wird und sämtliche Einzellinien in der minimalen und maximalen Konfiguration in der Zielauflösung erhalten bleiben.

Da unterschiedliche Skalenraumereignisse für die Interaktionsgruppe in der minimalen und maximalen Konfiguration in der Analyse des Skalenverhaltens prädiziert wurden, werden in der Stufe der *Fusion* zwei komplette Objektmodelle für die Zielauflösung A_{z1} von 0.1 m erzeugt (Abbildung 4.6). Das erste Objektmodell hat die Struktur der minimalen Konfiguration mit der Verschmelzung der beiden Mittellinien zu einer gemeinsamen Linie. Die Anzahl der Objektteile in der Zielauflösung ist damit geringer als in der Ursprungsauflösung. Dagegen behält das andere Objektmodell die Struktur aus der Ursprungsauflösung mit den gegebenen Objektteilen, da keine Skalenraumereignisse in der maximalen Konfiguration eingetreten sind. Die Wertebereiche der Attribute für beide Objektmodelle entsprechen den prädizierten Werten der minimalen und maximalen Konfiguration.



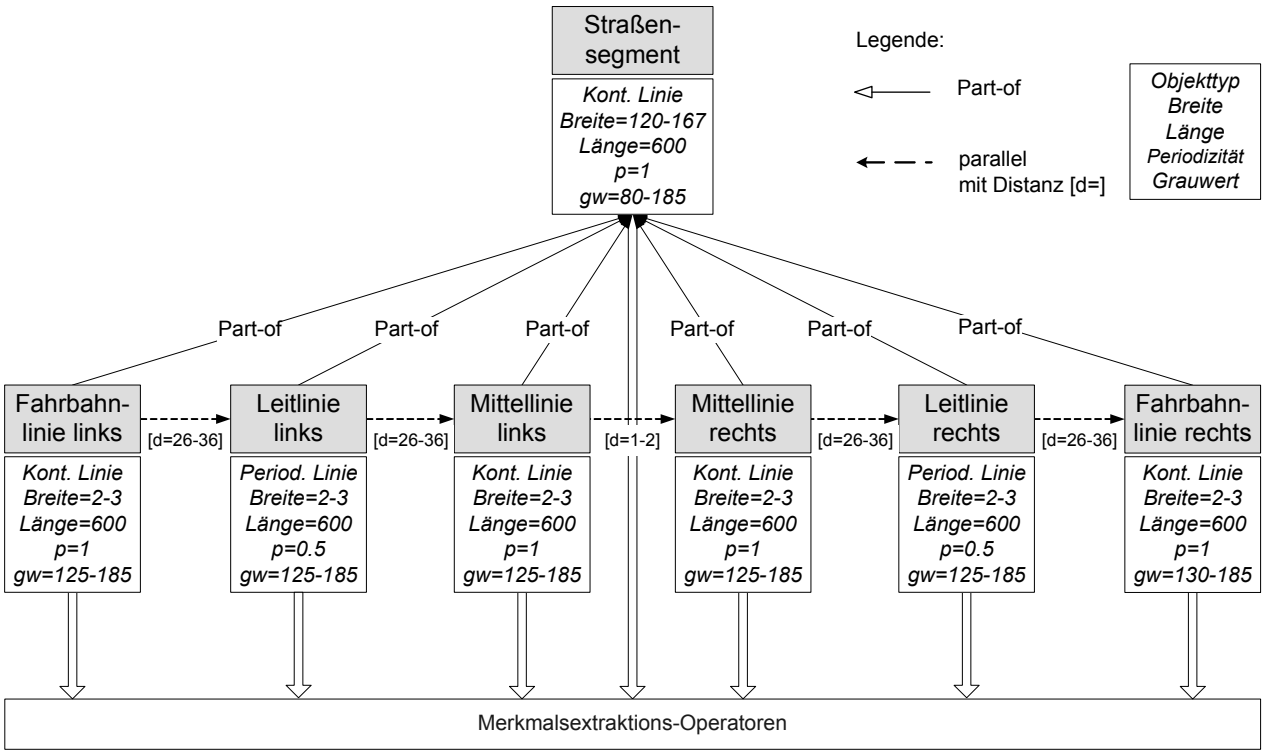


Abbildung 4.6: Automatisch angepasste Objektmodelle für die Zielauflösung A_{z1}=0.1m

Zielauflösung 1.00 m

Für die Zielauflösung A_{z2}=1.00 m ist ausgehend von der Ursprungsauflösung A_0=0.03 m eine Glättung mit einem Gauß-Filter von σ_{z2}=30 nötig. Der korrespondierende Filter-Kernel hat eine Breite von 181 Pixeln. Da in der Ursprungsauflösung der Abstand zwischen benachbarten Linien jeweils 82 bis 108 Pixel beträgt, besteht Interaktion zwischen allen benachbarten Linien. In der *Zerlegung* wird daher eine einzige Interaktionsgruppe jeweils für die maximale und die minimale Konfiguration erzeugt, die sämtliche Linien des Modells beinhaltet.

Da die Zerlegung keine Einzellinien ergeben hat, muss die *Analyse des Skalenverhaltens* vollständig in Analyse-durch-Synthese durchgeführt werden. Die dafür benötigten synthetischen Bilder der minimalen und maximalen Attributkonfiguration in der Ursprungsauflösung A_0, die geglätteten Bilder in der Zielskala σ_{z2} und die abgetasteten Bilder in der Zielauflösung A_{z2} illustriert Abbildung 4.7. Die minimale und die maximale Konfiguration zeigen ein unterschiedliches Skalenverhalten für diese Zielskala. Die beiden Mittellinien verschmelzen in beiden Konfigurationen aufgrund ihrer geringen Distanz zueinander. In der minimalen Konfiguration verschmelzen jedoch außerdem die benachbarten gestrichelten Leitlinien mit der Mittellinie, während in der maximalen Konfiguration die Leitlinien ein eigenes Maximum im Grauwertprofil behalten und damit keine sichere Verschmelzung eingehen. Ihr Grauwert-Kontrast zum Hintergrund ist jedoch sehr klein (12), so dass im Anschluss an die Anpassung der Operatorparameter die Leitlinien auf effektive Auslöschung untersucht werden. Dasselbe gilt für den Grauwertkontrast der Fahrbahnbegrenzungslinien sowohl in der minimalen als auch in der maximalen Konfiguration, der 3 bzw. 13 Graustufen beträgt.

Für die Prädiktion der Attributwerte in der Zielauflösung A_{z2}=1.00 m werden die Grauwerte in der Linienmitte sowie um eine halbe Pixelweite von der Linienmitte senkrecht zur Linienrichtung versetzt aus dem Bild der Zielskala $I_{\sigma z}$ abgegriffen. Die Pixelweite entspricht dabei dem Faktor der Auflösungsänderung F_{A2}=30. Die Linienbreite wird aus der Distanz der Maxima des Linienprofils im Gradientenbild der Zielskala ermittelt. Der bei der Verschmelzung der beiden kontinuierlichen Mittellinien mit den benachbarten gestrichelten Leitlinien entstandene Objekttyp der neuen Linie kann nach Tabelle 3.3 nur *Kontinuierliche Linie* sein.

Mit den für die Zielauflösung prädizierten Attributwerten können die Parameter der Operatoren für die Extraktion der Linien aus der Zielauflösung angepasst werden. Mit den angepassten Operatoren können nun effektive Skalenraumereignisse untersucht werden. Da der kleinste prädizierte Grauwertkontrast in der Zielauflösung für alle verbliebenen Linien unterhalb der Schwelle für effektive Auslöschung r_{eA}=15 liegt, müssen alle angepassten Operatoren auf das synthetische Bild der Zielauflösung angewendet werden. In der minimalen Attributkonfiguration wird für die beiden Fahrbahnbegrenzungslinien eine effektive Auslöschung festgestellt, da der Linienoperator nicht in der Lage ist, diese Linien mit einem Grauwertkontrast von 3 im Bild der Zielauflösung zu extrahieren. In der maximalen Konfiguration dagegen, in der die Fahrbahnbegrenzungslinien einen Kontrast von 13 aufweisen, bleiben diese Linien erhalten. Der Operator für die gestrichelten Leitlinien versagt jedoch, so dass für die Leitlinien eine effektive Auslöschung in der maximalen Konfiguration prädiziert wird, obwohl keine sichere Auslöschung vorliegt.

In der *Fusion* werden zwei angepasste Objektmodelle für die Zielauflösung A_{z2}=1.00 m erzeugt, um die unterschiedliche Struktur der Modelle, aufgrund der verschiedenen Skalenraumereignisse, für die minimale und maximale Konfiguration wiederzugeben. Die angepassten Distanzen zwischen den verbliebenen Linien werden nach Formel (3.9) berechnet. Für die Linien, die einer Verschmelzung unterliegen, müssen hierbei auch die Translationen ihrer Positionen Δt_{LiLj} berücksichtigt werden. Diese Positions-Verschiebungen werden aus den Positionen der Grauwertprofil-Maxima der Linien in der Ursprungsauflösung und dem Maximum im Profil der durch die Verschmelzung entstandenen Linie ermittelt. Die aus der automatischen Anpassung resultierenden Objektmodelle sind in Abbildung 4.8 dargestellt.

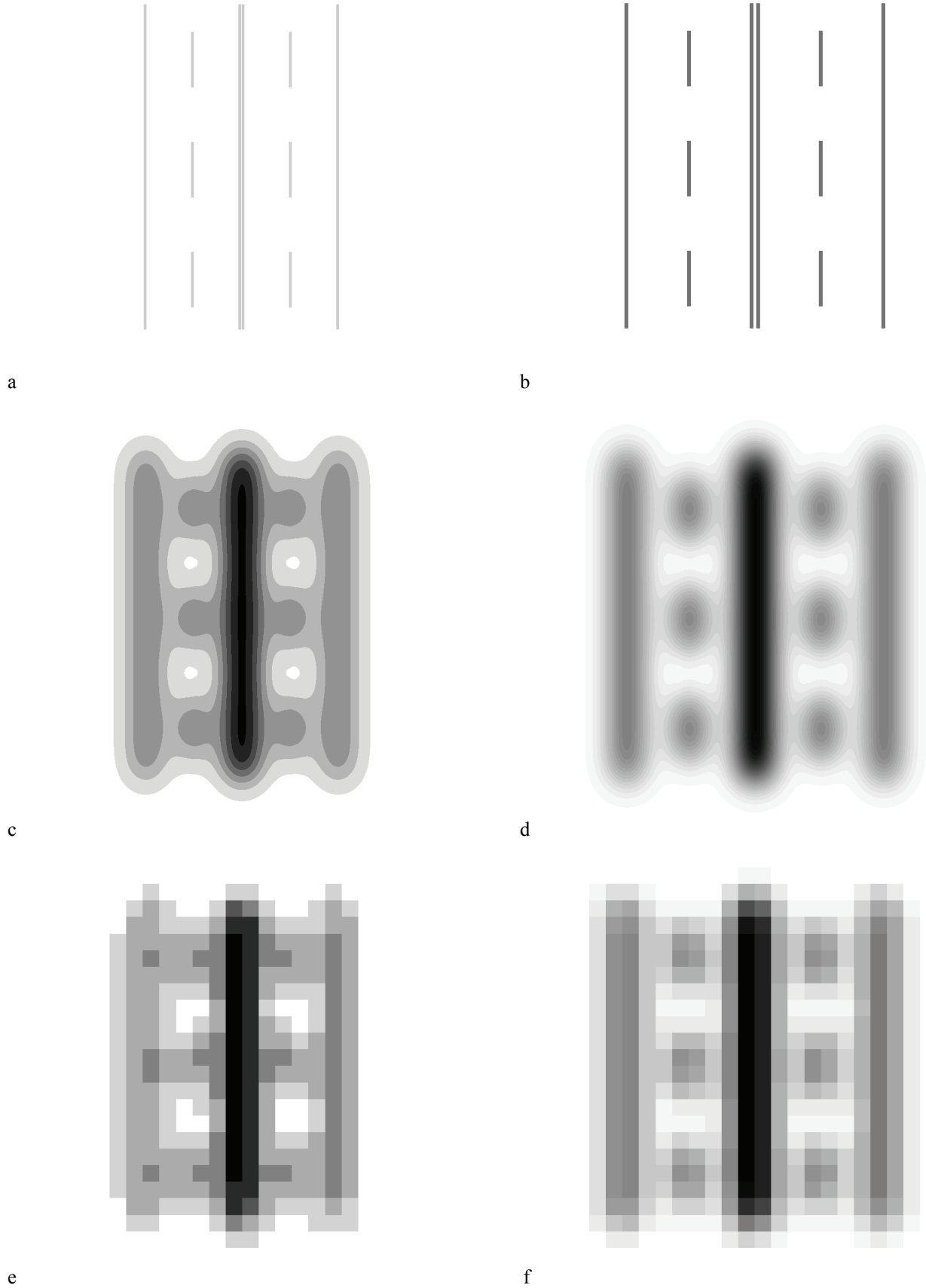
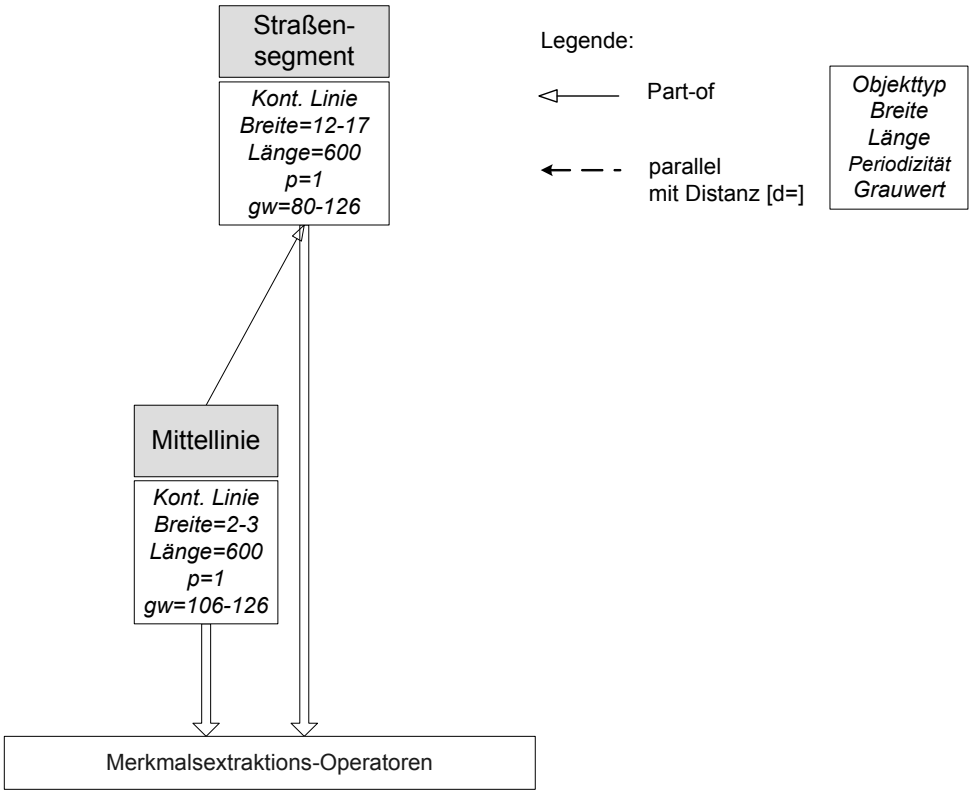


Abbildung 4.7: Synthetische Bilder der Interaktionsgruppe der minimalen (linke Spalte) und maximalen Konfiguration (rechte Spalte) in der Ursprungsauflösung A_0 (a+b), der Zielskala σ_{z2} (c+d) und der Zielauflösung A_{z2} (e+f) (zur besseren Visualisierung sind c,d,e,f kontrastverstärkt)



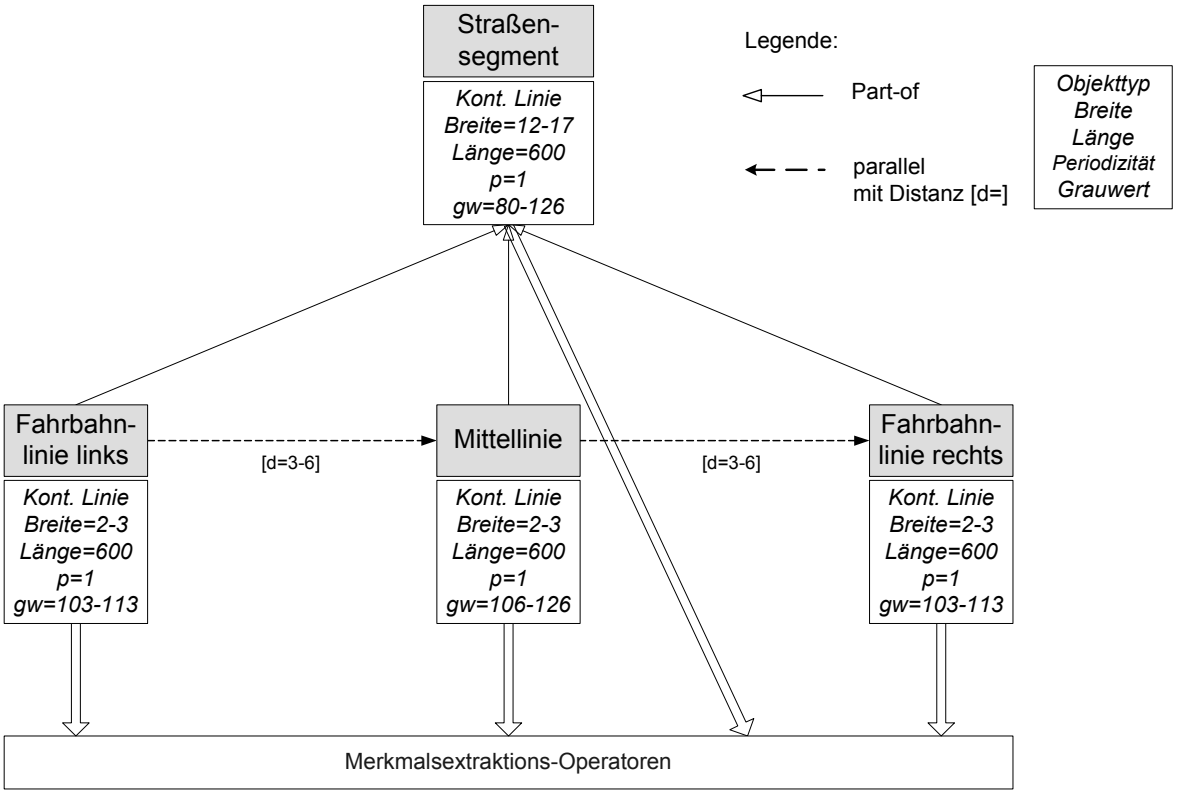


Abbildung 4.8: Automatisch angepasste Objektmodelle für die Zielauflösung A_{z2}=1.00m

Zielauflösung 5.00 m

Die Zielauflösung von A_{z3}=5.00 m bedeutet eine sehr große Auflösungsänderung im Vergleich zur gegebenen Ursprungsauflösung A_0=0.03 m. Die dafür benötigte Glättung entspricht einer Zielskala von σ_{z3}=150, für die ein 901 Pixel breiter Gauß-Filter gewählt wird. Da die Breite des Gauß-Filters die Distanzen aller benachbarten Linien bei weitem übersteigt, resultiert die *Zerlegung* des Objektmodells in eine einzige Interaktionsgruppe, deren Skalenverhalten über Analyse-durch-Synthese untersucht wird. Die synthetischen Bilder der Ursprungsauflösung A_0, der Zielskala σ_{z3} und der Zielauflösung A_{z3} zeigt Abbildung 4.9.

Sowohl in der minimalen als auch in der maximalen Konfiguration wird eine sichere Verschmelzung aller Linien der untersten hierarchischen Ebene prädiziert, da nur noch ein einziges Maximum im Grauwertprofil der Zielskala existiert. Der Kontrast dieser verschmolzenen Linie ist jedoch sehr klein (2 bzw. 8), so dass zusätzlich mit dem angepassten Linienoperator eine effektive Auslöschung in der Zielauflösung untersucht und in beiden Konfigurationen bestätigt wird. Damit entfallen sämtliche Linien der hierarchischen Ebene und können für die Extraktion des Straßensegmentes aus der Zielauflösung nicht mehr genutzt werden. Die einzig verbliebene Evidenz in der Zielauflösung besteht daher in dem Element der hierarchisch höher gelegenen Stufe des ursprünglichen Objektmodells – dem Straßensegment, d.h. der Fahrbahn selbst.

Die Skalenraumereignisse in der minimalen und maximalen Attributkonfiguration sind identisch. Damit resultiert ein einziges Objektmodell aus der automatischen Anpassung, dargestellt in Abbildung 4.10.

4.1.2 Beispiel 2 – Kreuzungsarm

Das zweite Beispiel zur Demonstration der entwickelten Methoden zur automatischen Anpassung modelliert einen komplexen Kreuzungsarm. Das in Abbildung 4.11 dargestellte Objektmodell entspricht dem Typ mit linienartigen und flächenhaften Objektteilen, zu dessen automatischer Anpassung die in 3.3.3 erläuterten Methoden entwickelt wurden. Die synthetischen Bilder in Abbildung 4.20 a+d illustrieren den modellierten Kreuzungsarm. Ein Kreuzungsbereich wird aus mehreren Kreuzungsarmen gebildet, die an Straßensegmenten wie der einfachen Straße aus dem Beispiel 1 an den Kreuzungsbereich angrenzen und in Richtung Kreuzung komplexer aufgebaut sind. Die automatische Anpassung wird für die drei Zielauflösungen 0.10 m, 0.20 m und 1.00 m demonstriert.

4.1.2.1 Objektmodell für die hohe Bildauflösung

Das anzupassende Objektmodell dient der Extraktion eines Kreuzungsarmes mit zwei durchgezogenen Fahrbahnlinien, einer Haltelinie und zwei Richtungspfeilen (links und geradeaus) als Objektteile. Der Objekttyp der beiden Fahrbahnlinien (Mittellinie und Leitlinie) ist Kontinuierliche Linie. Die Haltelinie und die Richtungspfeile sind dagegen über flächenhafte Objekttypen modelliert. Während die Richtungspfeile aufgrund ihrer individuellen Form über Muster repräsentiert sind, wird die Haltelinie als ein Rechteck beschrieben. Die Haltelinie wiederum besteht aus fünf einzelnen untergeordneten rechteckigen Objektteilen. Die meisten Attribute der Knoten sind als Wertebereiche modelliert. Das Modell enthält Informationen zum Objekttyp, sowie Ausdehnung (Breite und Länge) und Grauwert der einzelnen Objektteile im Bild. Die relative zweidimensionale Position der Objektteile zueinander wird über Distanzen angegeben (parallel und senkrecht zur Hauptrichtung des Kreuzungsarmes). Die Attribute der beiden Fahrbahnlinien entsprechen denen der korrespondierenden Objektteile aus dem Objektmodell für die einfache Straße (Abbildung 4.2) im Beispiel 1.

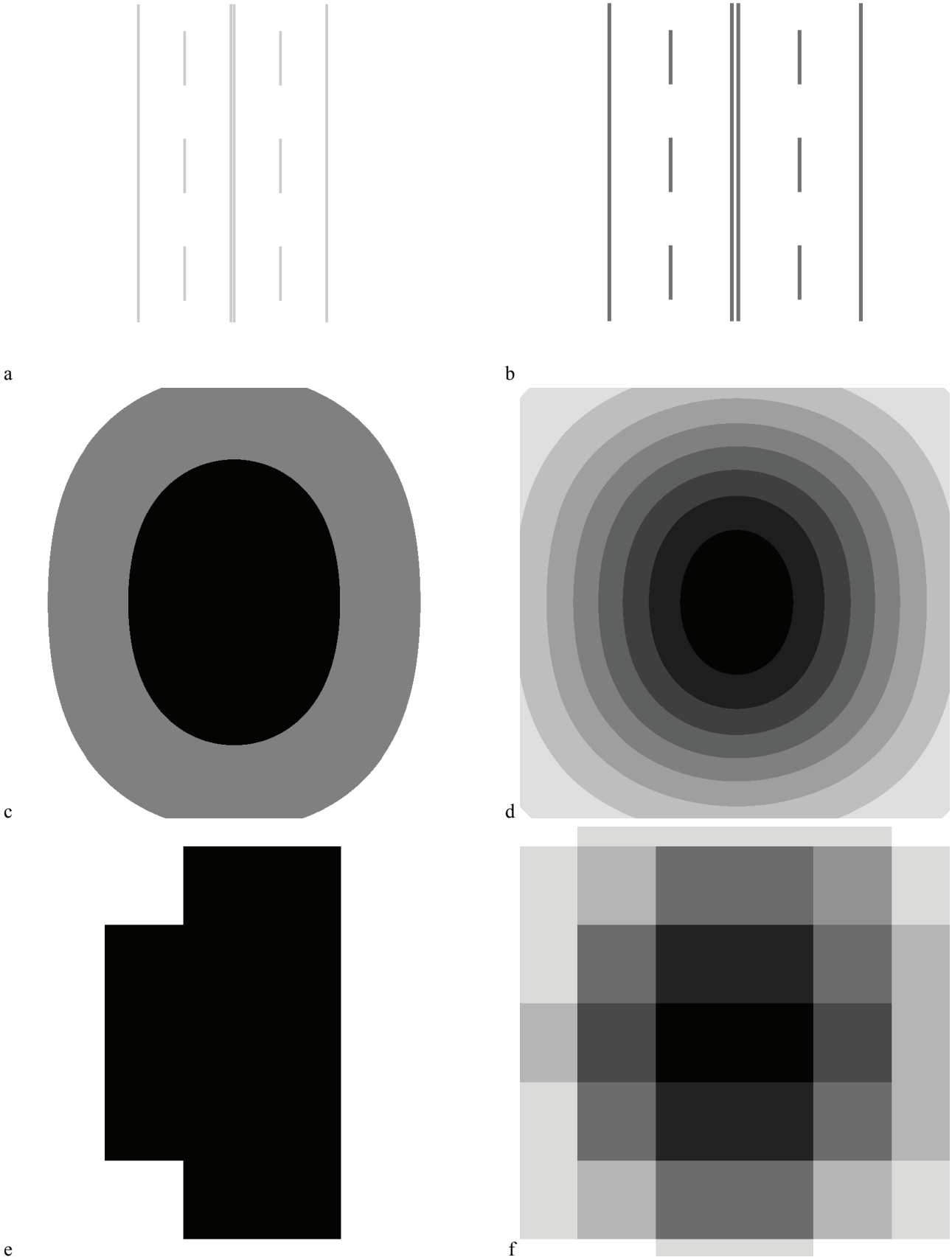


Abbildung 4.9: Synthetische Bilder der Interaktionsgruppe der minimalen (linke Spalte) und maximalen Konfiguration (rechte Spalte) in der Ursprungsauflösung A_0 (a+b), der Zielskala σ_{z2} (c+d) und der Zielauflösung A_{z2} (e+f) (zur besseren Visualisierung sind c,d,e,f kontrastverstärkt)

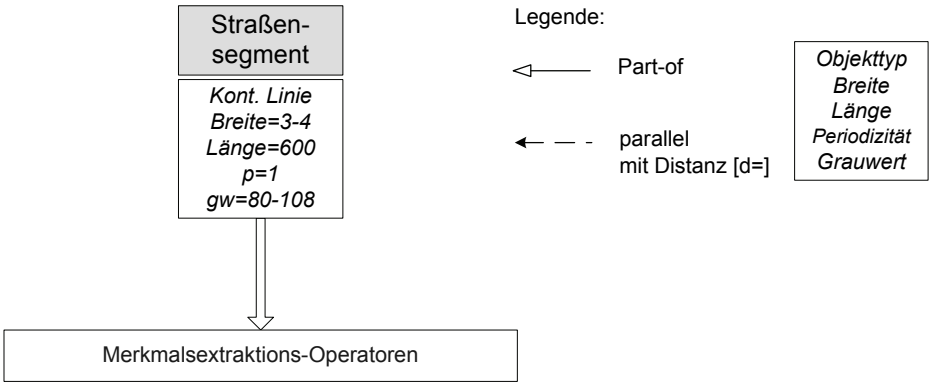


Abbildung 4.10: Automatisch angepasstes Objektmodell für die Zielauflösung A_{z3}=5.00m

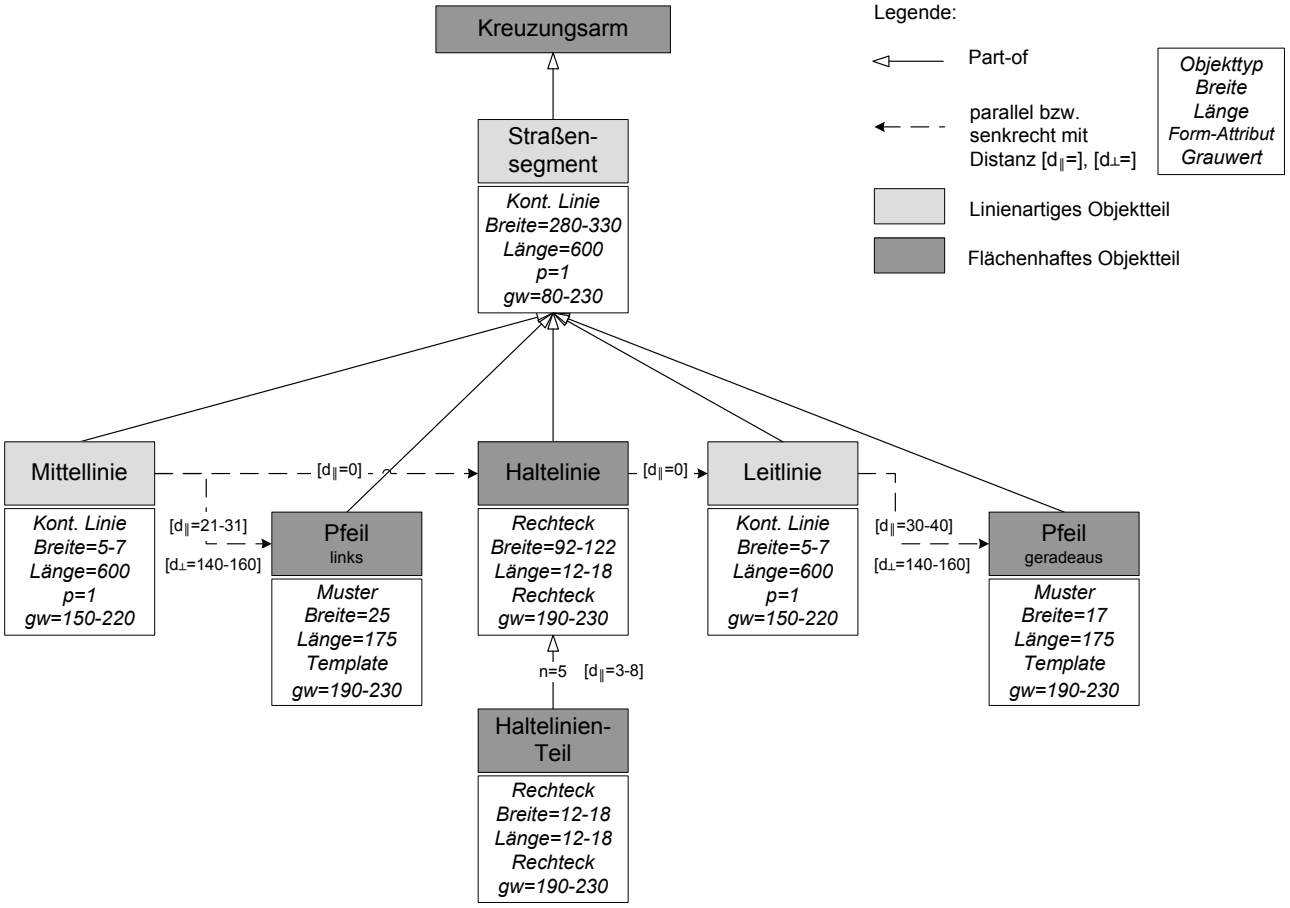


Abbildung 4.11: Objektmodell für ein komplexes Straßensegment (Kreuzungsarm) in 0.03 m/Pixel

Den Objektteilen des Semantischen Netzes ist jeweils ein geeigneter Merkmalsoperator zugeordnet. Unterschiedliche Objekttypen sind mit unterschiedlichen Operatoren verknüpft. Die linienartigen Objektteile nutzen differentialgeometrische Linienoperatoren, vgl. das Objektmodell für Einfache Straße in 4.1.1.1. Für die fünf Teile der Haltelinie sind keine eigenständigen Operatoren zugeordnet. Stattdessen wird die gesamte Haltelinie vom geometrischen Typ *Rechteck* mit einem Operator extrahiert, der auf dem Flächenwachstums-Verfahren basiert. Die einzelnen Teile der Haltelinie werden zuerst über Flächenwachstum extrahiert und anschließend über ihre geometrischen Formen und Distanzen zueinander selektiert. Die Modellierung der einzelnen Teile der Haltelinie ermöglicht im Zuge des Anpassungsverfahrens eine automatische Anpassung der Operatorparameter für die Haltelinie. Die flächenhaften Muster werden mit Template-Matching über bereitgestellte Musterbilder extrahiert [SCHRAMM 2005]. Da für die Richtungspfeile Musterbilder gegeben sind, wird deren Ausdehnung im Objektmodell nicht über Wertebereiche, sondern über feste Attributwerte, beschrieben.

4.1.2.2 Automatische Anpassung

Zielauflösung 0.1 m

Als erste Zielauflösung wird A_{z1}=0.1 m entsprechend einer Auflösungsänderung F_{A1}=3 gewählt. Es ist eine Glättung mit einem Gauß-Filter mit σ_{z1}=3 und einer Breite von 19 Pixeln nötig.

Da die Attribute als Wertebereiche gegeben sind, wird jeweils die Konfiguration mit den minimalen und den maximalen Attributwerten untersucht. Die *Zerlegung* bildet Gruppen von Objektteilen, die während der Auflösungsänderung interagieren und deren Skalenverhalten daher gemeinsam in einer Interaktionsgruppe analysiert werden muss. Für das komplexe Objektmodell wird die zweidimensionale Methode der Zerlegung angewendet, die Interaktionszonen um die einzelnen Objektteile bildet. Sowohl für die minimale als auch die maximale Konfiguration wird von der Haltelinie mit ihren fünf Teilen zusammen mit der Mittellinie und Leitlinie, aufgrund ihrer räumlichen Nähe, eine Interaktionsgruppe gebildet. Die beiden Richtungspfeile sind keiner Interaktion unterworfen und werden in der folgenden Stufe des Verfahrens einzeln prozessiert. Für die beiden Pfeile wird keine minimale und maximale Attributkonfiguration unterschieden, da für diese feste Attributwerte und Musterbilder gegeben sind.

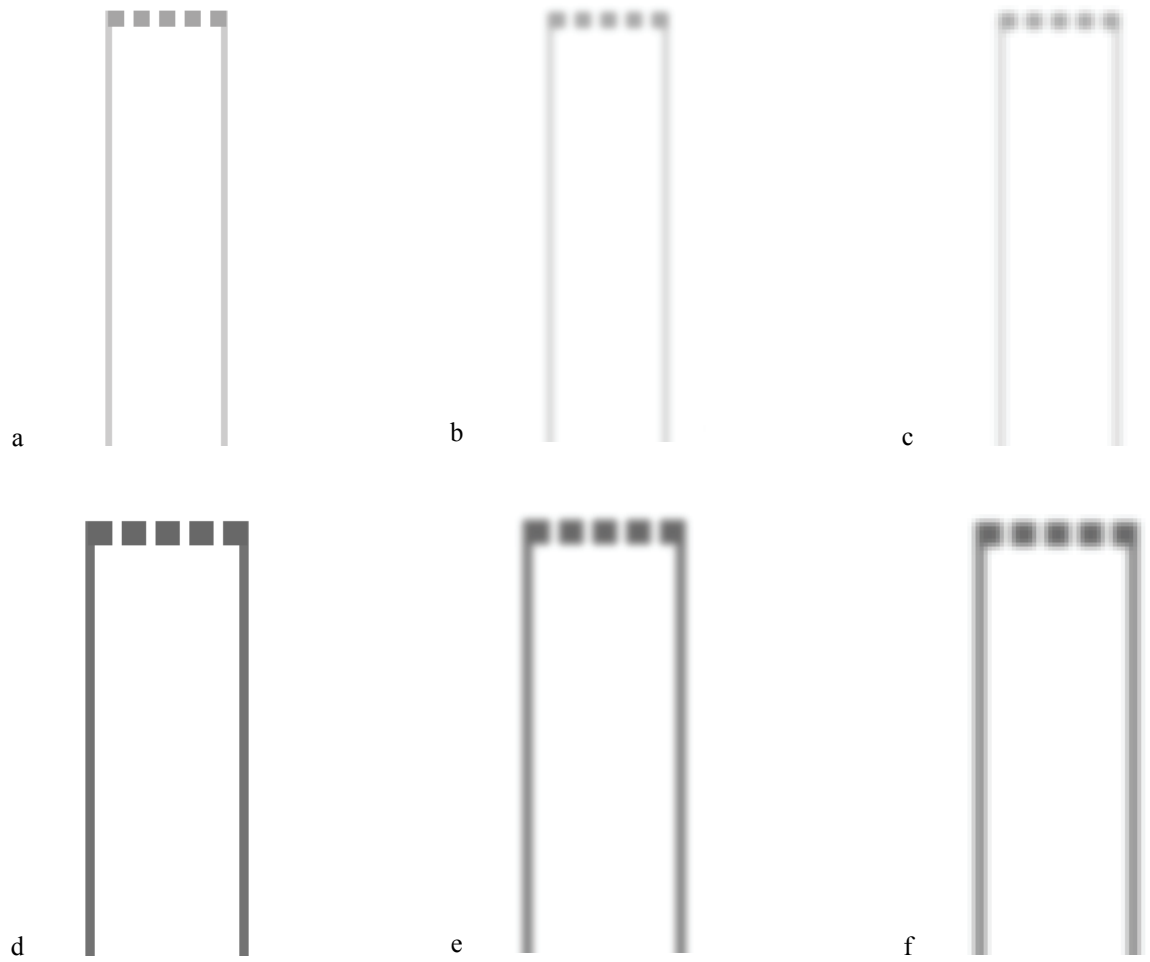


Abbildung 4.12: Synthetische Bilder für die minimale (oben) und maximale Konfiguration (unten) der Interaktionsgruppe in der Ursprungsauflösung A_0 (a+d), Zielskala σ_{z1} (b+e) und Zielauflösung A_{z1} (c+f) (Bilder sind im unteren Bereich gekürzt)

Die *Analyse des Skalenverhaltens* wird für zweidimensionale Objektmodelle vollständig über Analyse-durch-Synthese durchgeführt im Gegensatz zur eindimensionalen Methode, die teilweise auch die analytische Prädiktion nutzt. Aus den Attributwerten der Knoten und Relationen werden synthetische Bilder generiert, die in die Zielskala σ_{z1} und Zielauflösung A_{z1} überführt werden. Die erzeugten synthetischen Bilder in der Ursprungsauflösung I_0 und der Zielskala $I_{\sigma z1}$ für

die Interaktionsgruppe zeigt Abbildung 4.12. Die synthetischen Bilder zur Simulation der Richtungspfeile sind in Abbildung 4.14 dargestellt. Zur Prädiktion von Skalenraumereignissen und der Attributwerte in der Zielauflösung wird für jede Interaktionsgruppe eine Blob-Detektion in den synthetischen Bildern der Ursprungsauflösung I_0 und der Zielskala $I_{\sigma z1}$ durchgeführt. Abbildung 4.13 illustriert die Ergebnisse der Blob-Detektion (Basisregionen und Extrema) der Interaktionsgruppe in der minimalen und maximalen Konfiguration. Die Ergebnisse für die Blob-Detektion der Richtungspfeile sind ebenfalls in Abbildung 4.14 dargestellt.

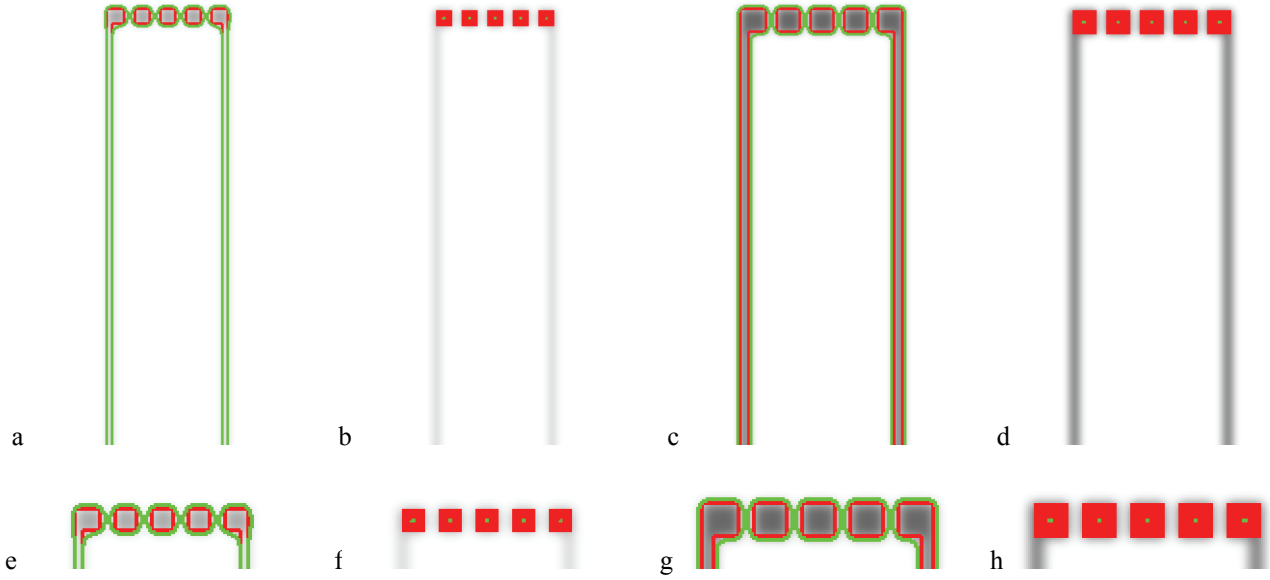


Abbildung 4.13: Ergebnisse der Blob-Detektion für die minimale (a+b) und maximale Konfiguration (c+d) der Interaktionsgruppe in der Ursprungsauflösung A_0 (rot) und der Zielskala σ_{z1} (grün) auf dem Bild der Zielskala $I_{\sigma z1}$: Basisregionen (a+c) und Extrema der Blobs (b+d) sowie Vergrößerungen in Bereich der Haltelinie (e-h)

Für die Interaktionsgruppe wurden in der minimalen und maximalen Konfiguration unterschiedliche Skalenraumereignisse prädiziert. Damit resultieren aus der automatischen Anpassung in der Zielauflösung zwei Objektmodelle, die je nach tatsächlichen Attributwerten im Bild unterschiedliche Strukturen des Objektes wiedergeben. Während für die minimale Konfiguration die sichere Auslöschung der beiden kontinuierlichen Fahrbahnlinien vom Algorithmus prädiziert wird, bleiben alle Objektteile der Interaktionsgruppe in der maximalen Konfiguration erhalten. Bei der Abgrenzung des Blobs mit dem verwendeten Algorithmus (siehe Abschnitt 3.3.3.2) werden die beiden kontinuierlichen Linien in der minimalen Konfiguration als Hintergrund eingeordnet, da ihr Grauwert niedriger als der Basiswert der fünf Teile der Haltelinie ist. Damit existiert an der Stelle der Fahrbahnlinien in der Ursprungsauflösung keine Blob-Basisregion mehr in der Zielskala. In der maximalen Konfiguration dagegen ist der Grauwert der kontinuierlichen Fahrbahnlinien höher und der Basiswert der zugehörigen Blobs kleiner als der Grauwert der Fahrbahnlinien. Dementsprechend werden diese in der maximalen Konfiguration den Blobs der angrenzenden Haltelinien-Teile zugeschlagen. Trotzdem können die kontinuierlichen Linien mit dem zugeordneten Linienoperator noch in der minimalen Konfiguration extrahiert werden (wie in Abschnitt 4.1.1.2 am Beispiel der Einfachen Straße gezeigt). Würde in dem Beispiel die Haltelinie nicht gestückelt, sondern zusammenhängend sein, würde der Algorithmus der Blob-Detektion die Fahrbahnlinien auch in der minimalen Konfiguration dem Blob der angrenzenden Haltelinie zuordnen. An diesem Beispiel wird ein Defizit des Blob-Konzeptes deutlich. Diese Auslöschung wurde auf einer Interaktionsgruppe prädiziert, die bereits in der Ursprungsauflösung als ein gemeinsamer Blob abgegrenzt wurde. Um dieses Defizit zu beheben, werden die betroffenen Objektteile, für die eine Auslöschung durch Blob-Analyse prädiziert wurde, in einer separaten Analyse untersucht (wie in Abschnitt 3.3.3.2 erläutert). Für die beiden kontinuierlichen Fahrbahnlinien können die Ergebnisse dieser separaten Analyse Abschnitt 4.1.1.2 entnommen werden, da sie denen für die äußeren Fahrbahnlinien entsprechen. Die zugehörigen Linienoperatoren werden anschließend mit den angepassten Parametern auf die gesamte Interaktionsgruppe angewendet. Da die Linienextraktion erfolgreich ist, bleiben auch für die minimale Konfiguration in A_{z1}=0.1 m die Fahr-

bahnlinien erhalten. Die Attributwerte der Interaktionsgruppe werden aus dem Bild der Zielskala $I_{\sigma z1}$ abgeleitet (vgl. Abschnitt 3.3.3.2).

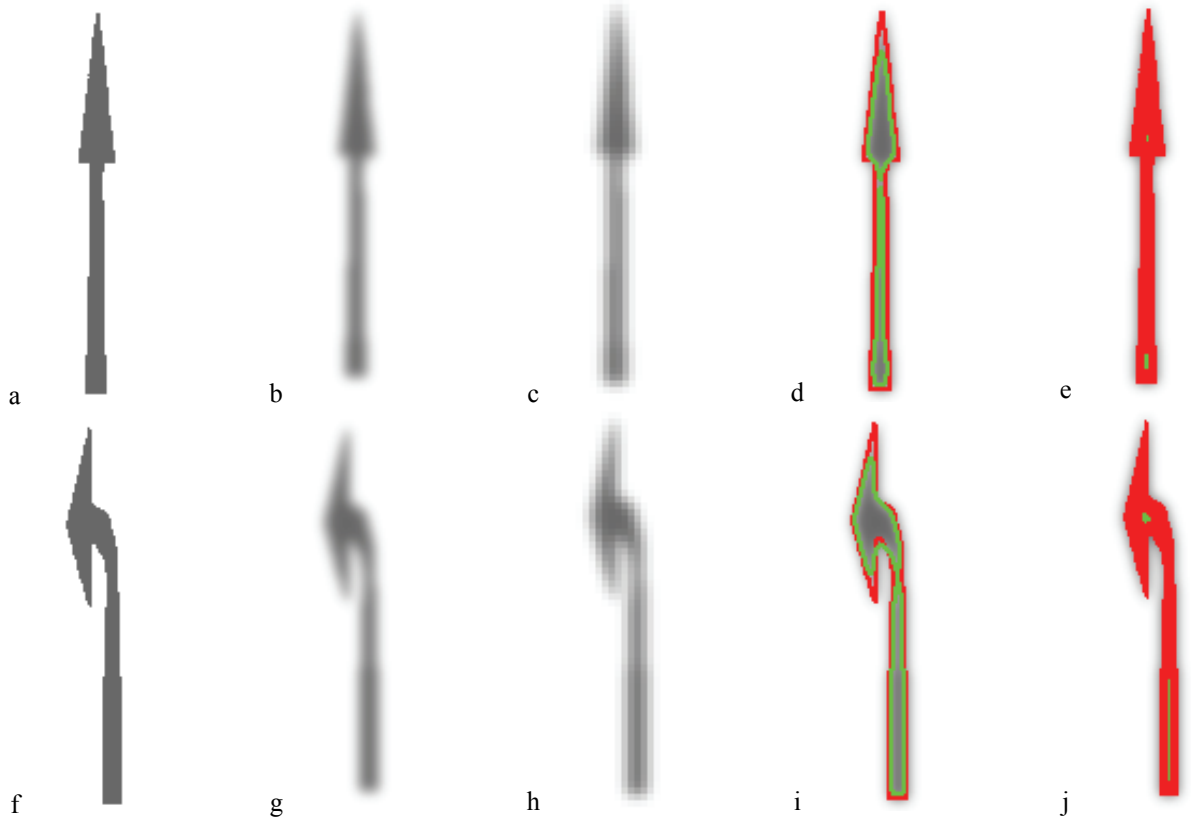


Abbildung 4.14: Synthetische Bilder für die Richtungspfeile „geradeaus“ (oben) und „links“ (unten) in der Ursprungsauflösung A_0 (a+f), Zielskala σ_{z1} (b+g) und Zielauflösung A_{z1} (c+h) sowie die Ergebnisse der Blob-Detektion (Basisregionen d+i und Extrema e+j)

Für die Richtungspfeile, die als Muster modelliert sind, wurde in der Blob-Analyse jeweils eine Spaltung prädiziert. Die schmalste Stelle der Pfeile wird durch die Glättung der Grauwerte dünner, so dass zwei Extrema innerhalb der gegebenen Muster für die Richtungspfeile entstehen. An der schmalsten Stelle, die am stärksten ausgedünnt wird, entsteht der geteilte Sattel der Blobs (vgl. Abbildung 4.14 d+e). Obwohl diese Spaltung nach der Blob-Definition tatsächlich vorliegt, ist sie für die Extraktion der Richtungspfeile im Zusammenhang mit dem gegeben Merkmalsoperator (Template-Matching) nicht sinnvoll. Bei Mustern wird die aus den Basisregionen der Blobs prädizierte Spaltung nicht berücksichtigt, wie bereits für die Analyse des Skalenverhaltens für zweidimensionale Objektmodelle in Abschnitt 3.3.3.2 erläutert. Das Verhältnis der Ausdehnungen der Richtungspfeile, die aus den Templates gegeben sind, bleibt in der Zielauflösung erhalten. Die maximalen Grauwerte der Richtungspfeile werden an den Stellen der Extrema der Blobs ermittelt (bei einer Spaltung im Muster wird der höchste Grauwert von mehreren entstandenen Blob-Extrema gewählt). Der minimale Grauwert wird aus der Abtastzone um die Basisregionen der Blobs im Bild der Zielskala $I_{\sigma z1}$ ermittelt.

Die Anpassung der Operatorparameter an die Zielauflösung erfolgt mit den prädizierten Attributwerten. Für die Haltelinie wurde ein geringer Grauwertkontrast prädiziert (1–50), der unterhalb des Schwellwertes für effektive Skalenraumereignisse r_{eA}=15 liegen kann. Daher wird der zugehörige Operator auf das synthetische Bild der Interaktionsgruppe in der Zielauflösung I_{Az1}, zu der die Haltelinie gehört, angewendet. Die Ergebnisse sowohl für die minimale als auch für die maximale Konfiguration bestätigen kein effektives Skalenraumereignis. Die Haltelinie verbleibt daher mit fünf Teilen in den Objektmodellen der Zielauflösung. Für die Richtungspfeile werden die Templates für ihre Extraktion aus der niedrigeren Zielauflösung aus dem synthetischen Bild der Zielauflösung I_{Az1} abgeleitet. Um eine für die Extraktion aus realen Bilddaten mit Template-Matching geeignete Mustermatrix zu erhalten, wird zu den synthetischen Bildern I_{Az1} der Richtungspfeile in der Zielauflösung der Grauwert des hierarchisch höheren Knotens des Straßensegmentes addiert.

In der *Fusion* wird ein komplettes Objektmodell aus den Interaktionsgruppen und den Pfeilen für die minimale und maximale Konfiguration erstellt. Die Distanzen zwischen den Objektteilen werden aus den Entfernungen der Basisregionen der Blobs in der Zielskala berechnet. Das aus der automatischen Anpassung resultierende Objektmodell für die Zielauflösung A_{z1} ist in Abbildung 4.15 dargestellt.

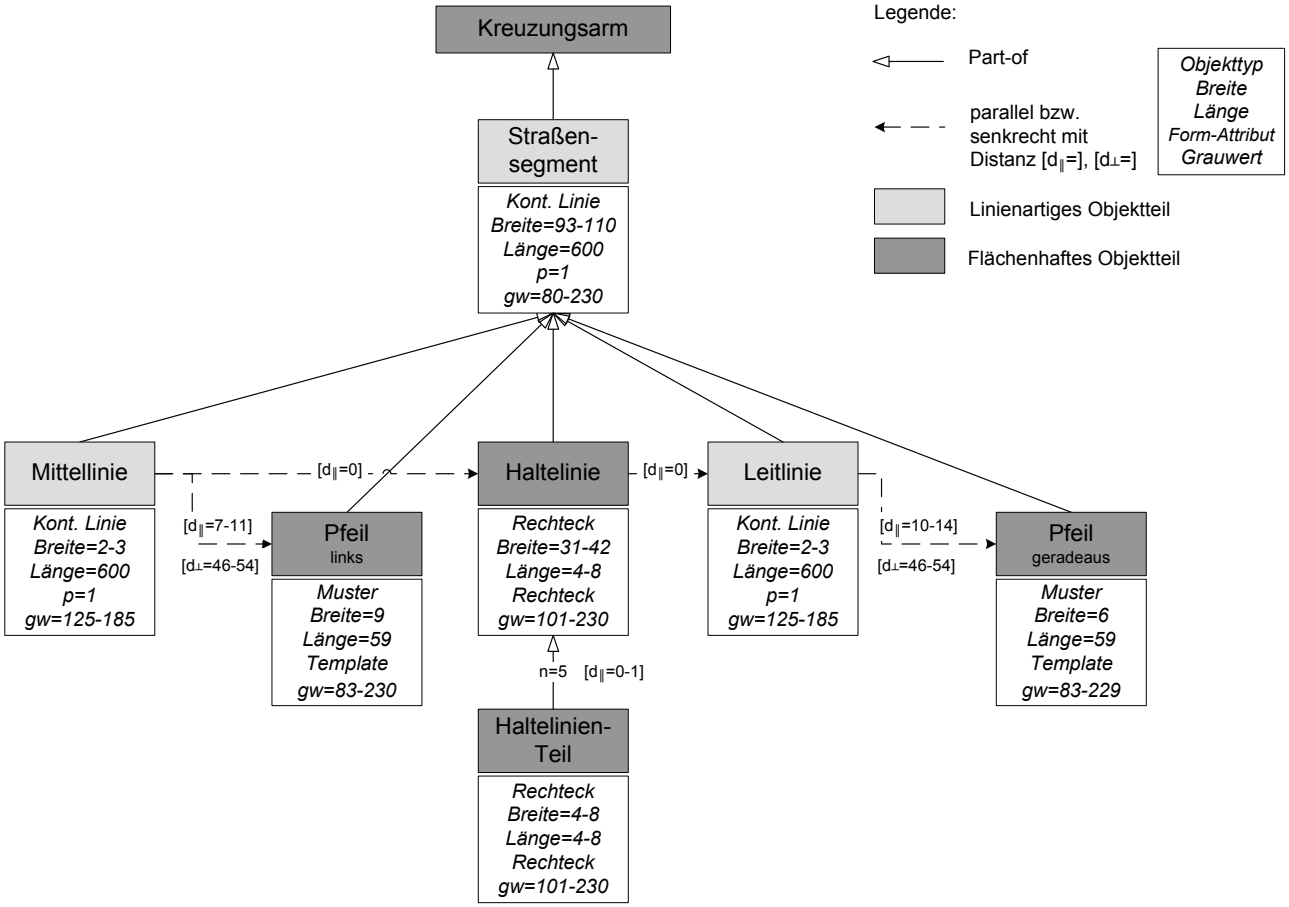


Abbildung 4.15: Automatisch angepasstes Objektmodell für die Zielauflösung A_{z1}=0.1m

Zielauflösung 0.2 m

Die zweite gewählte Zielauflösung A_{z2} entspricht einer Auflösungsänderung F_{A1}=6 und damit einer Zielskala σ_{z2}=6. Der zugehörige Gauß-Filter wird mit einer Breite von 37 Pixeln generiert.

In der *Zerlegung* wird für die minimale Attributkonfiguration aus allen Objektteilen des Kreuzungsarmes eine Interaktionsgruppe gebildet, da alle Objektteile interagieren. In der maximalen Konfiguration interagieren alle Objektteile mit Ausnahme des Richtungspfeils „geradeaus“. Das Skalenverhalten dieses Richtungspfeils wird einzeln analysiert, die anderen Objektteile bilden eine Interaktionsgruppe. Die synthetischen Bilder der Interaktionsgruppen und des einzelnen Pfeils, die für die Analyse des Skalenverhaltens gebildet werden, zeigen Abbildung 4.16 und Abbildung 4.18 a-c.

Zur *Analyse des Skalenverhaltens* wird die Blob-Detektion auf den synthetischen Bildern der Ursprungsauflösung I_{A0} und der Zielskala $I_{\sigma z2}$ durchgeführt. Die Basisregionen und Extrema der Blobs der Interaktionsgruppen minimaler und maximaler Konfiguration in der Ursprungs- und Zielskala illustriert Abbildung 4.17. Die Ergebnisse der Blob-Detektion des einzelnen Richtungspfeils aus der maximalen Konfiguration zeigt Abbildung 4.18 d+e. Wie bereits in der Zielauflösung A_{z1}=0.1m bleiben auch in der zweiten Zielauflösung A_{z2} die einzelnen Teile der Haltelinie erhalten und verschmelzen nicht, während die Fahrbahnlinien in der Zielskala nicht mehr in den Blobs enthalten sind. Auch hier werden diese Linien getrennt analysiert und der Linienoperator mit den angepassten Parametern auf das Bild der Interaktionsgruppe in der Zielauflösung angewendet, um auf effektive Auslöschung zu testen. In dem Test wurde keine Auslöschung bestätigt, da die beiden Linien extrahiert werden konnten. Für den Pfeil wurde ebenfalls wie in A_{z1} eine Spaltung mit dem

Blob-Analyse prädiziert, die jedoch wiederum im angepassten Modell ignoriert wird, da der Pfeil von Objekttyp Muster ist, für dessen Extraktion die Berücksichtigung einer Spaltung nicht erfolgt (vgl. Abschnitt 3.3.3.2).

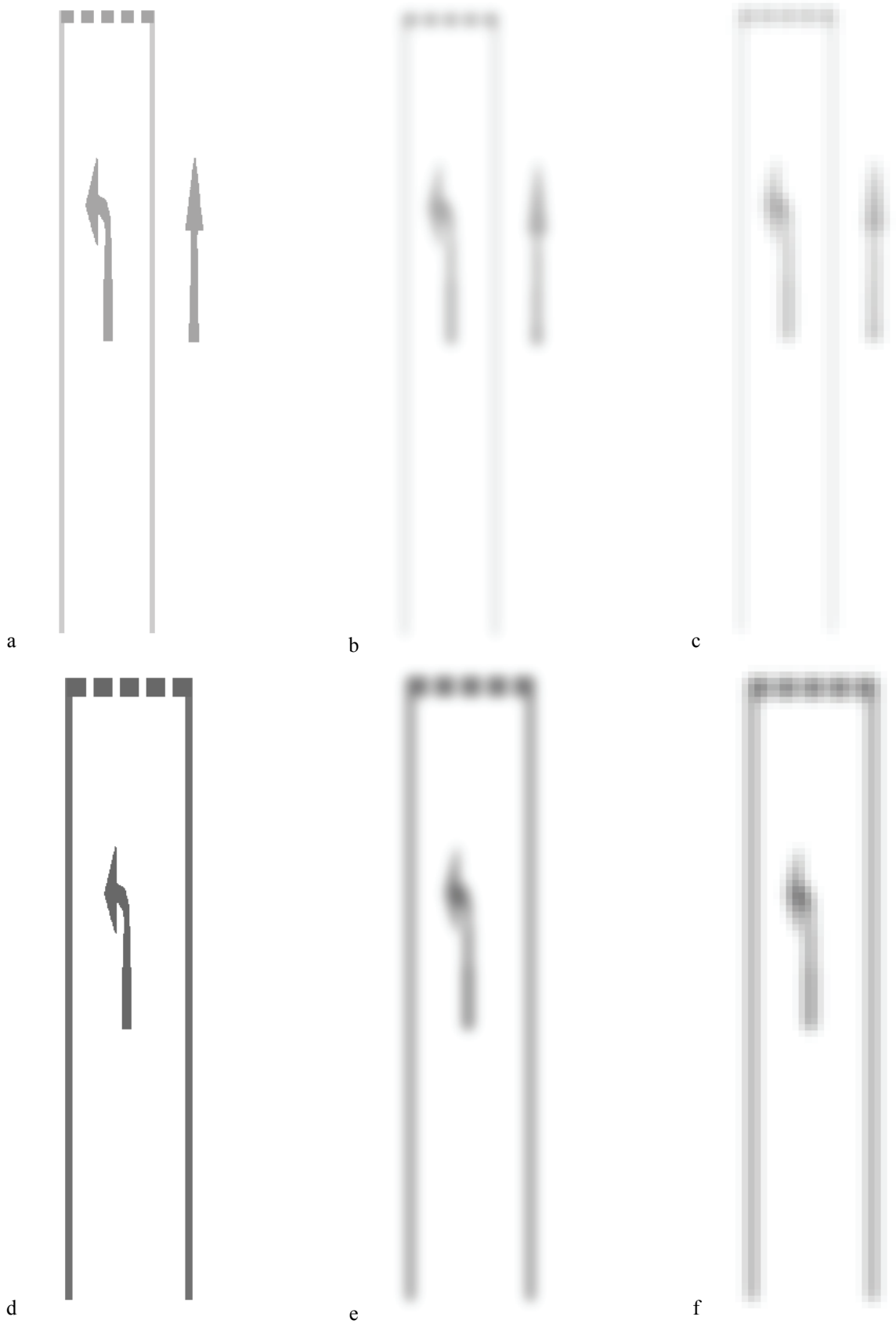


Abbildung 4.16: Synthetische Bilder für die minimale (oben) und maximale Konfiguration (unten) der Interaktionsgruppe in der Ursprungsauflösung A_0 (a+d), Zielskala σ_{z2} (b+e) und Zielauflösung A_{z2} (c+f)

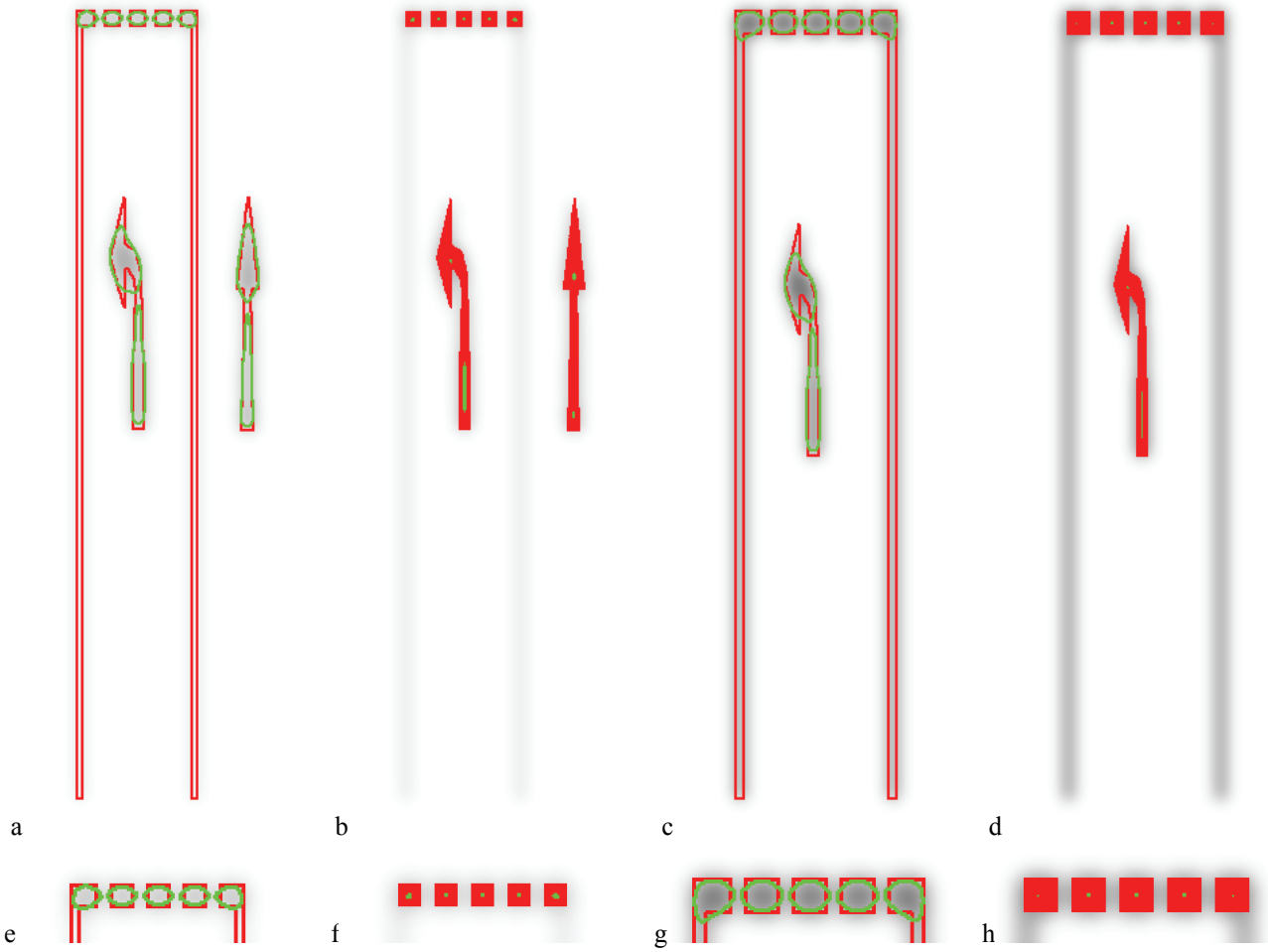


Abbildung 4.17: Ergebnisse der Blob-Detektion für die minimale (a+b) und maximale Konfiguration (c+d) der Interaktionsgruppe in der Ursprungsauflösung A_0 (rot) und der Zielskala σ_{z2} (grün) auf dem Bild der Zielskala $I_{\sigma z2}$: Basisregionen (a+c) und Extrema der Blobs (b+d) sowie Vergrößerungen in Bereich der Haltelinie (e-h)

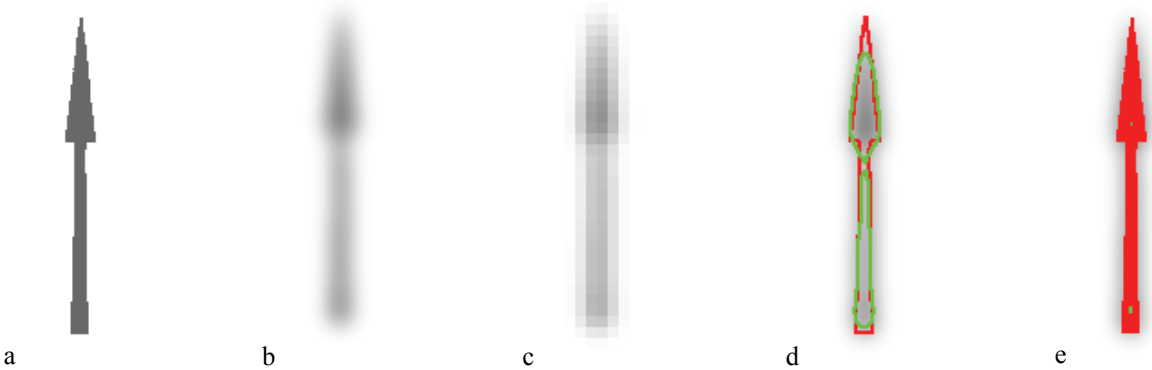


Abbildung 4.18: Synthetische Bilder des Richtungspfeils „geradeaus" für die maximale Konfiguration in der Ursprungsauflösung A_0 (a), Zielskala σ_{z2} (b) und Zielauflösung A_{z2} (c) sowie die Ergebnisse der Blob-Detektion (Basisregionen d und Extrema e)

In der *Fusion* wird ein komplettes Objektmodell für die Zielauflösung A_{z2} erstellt, da in der minimalen und maximalen Konfiguration dieselben Skalenraumereignisse aufgetreten sind. Das angepasste Objektmodell ist in Abbildung 4.19 dargestellt.

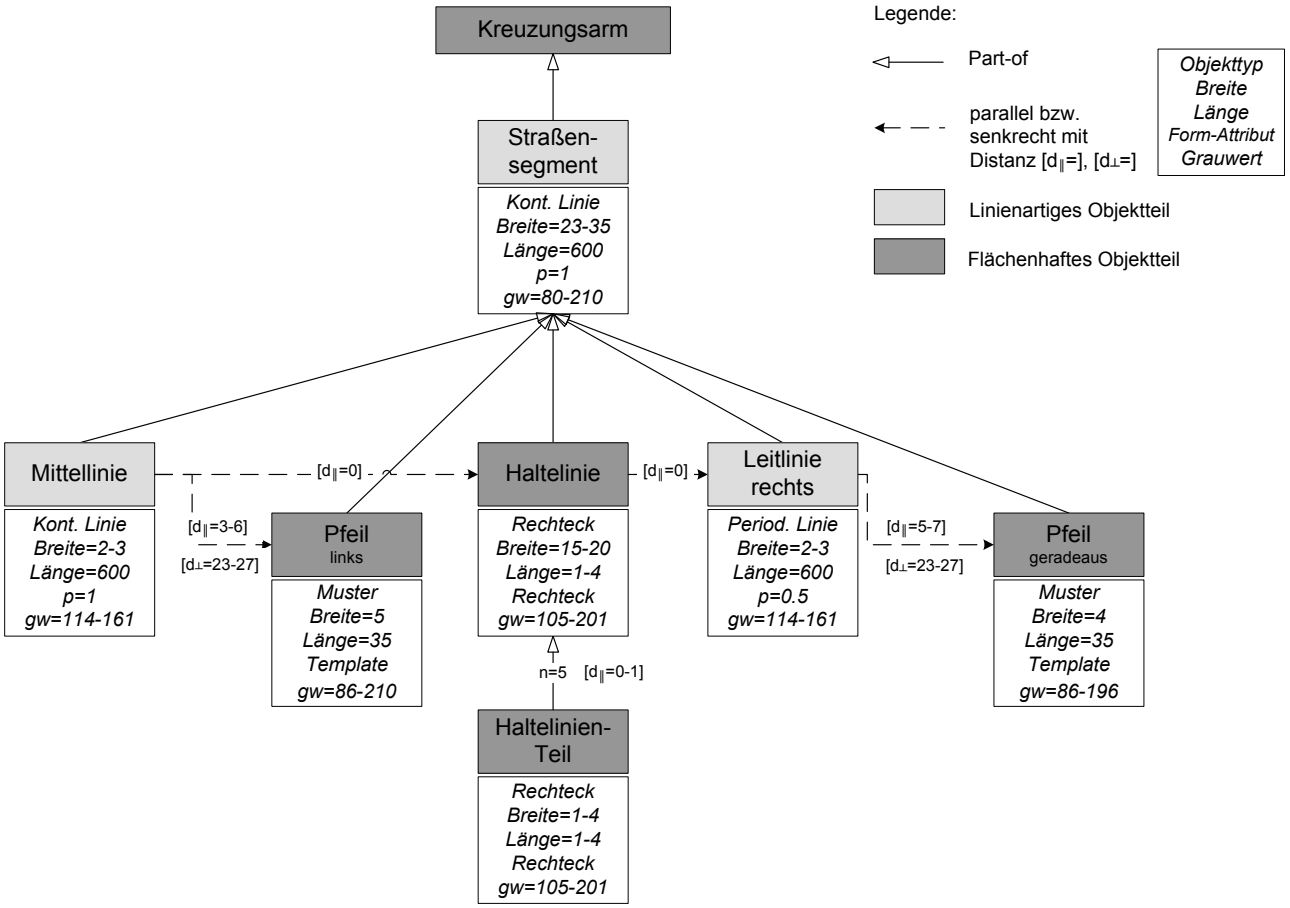


Abbildung 4.19: Automatisch angepasstes Objektmodell für die Zielauflösung A_{z2}=0.2 m

Zielauflösung 1.00 m

Die Zielauflösung A_{z3}=1.00 m entspricht einer Auflösungsänderung F_{A3}=30. Der korrespondierende Gauß-Filter mit σ_{z3}=30 hat eine Filterbreite von 181 Pixel.

In der *Zerlegung* kann weder in der minimalen noch der maximalen Attributkonfiguration in getrennte Interaktionsgruppen zerlegt werden. Aufgrund der großen Glättung und der damit verbundenen Größe des Gauß-Filters interagieren alle Objektteile. Daher entsteht in der Zerlegung jeweils eine Interaktionsgruppe mit sämtlichen Objektteilen für die minimale und die maximale Konfiguration. Die synthetischen Bilder der Interaktionsgruppe minimaler und maximaler Konfiguration sind in Abbildung 4.20 dargestellt.

Die *Analyse des Skalenverhaltens* prädiziert verschiedene Skalenraumereignisse für die minimale und maximale Attributkonfiguration. Die fünf Teile der Haltelinie verschmelzen in beiden Konfigurationen zu einer einzigen Linie. Die Haltelinie wird nur noch über einen einzigen Blob in der Zielskala abgegrenzt (vgl. Abbildung 4.21). Die Attribute der neu entstandenen Haltelinie werden aus den Eigenschaften des verbliebenen Blobs abgeleitet. Wie bereits in der Zielauflösung A_{z2} werden für die kontinuierlichen Fahrbahnlinien keine Blobs detektiert. Die Fahrbahnlinien werden einer separaten Analyse unterzogen, deren Ergebnisse denen der äußeren Fahrbahnbegrenzungslinien in Abschnitt 4.1.1.2 entsprechen. Für die Fahrbahnlinien der minimalen Konfiguration wird daher eine effektive Auslöschung prädiziert. Bei maximaler Attributkonfiguration sind die Linien extrahierbar und verbleiben daher in der Zielauflösung. Damit resultieren zwei unterschiedliche Objektmodelle aus der Anpassung. Die Blobs der Richtungspfeile unterliegen keinen Skalenraumereignissen. Ihren Ausdehnungen haben sich jedoch aufgrund der starken Glättung erheblich verändert. Die Ausdehnungen von Objektteilen, die über Muster repräsentiert sind, werden jedoch nach Abschnitt 3.3.3.2 nicht aus den Ausdehnungen der Blobs ermittelt, sondern mit dem jeweiligen Faktor der Auflösungsänderung F_A an die Zielauflösung angepasst.

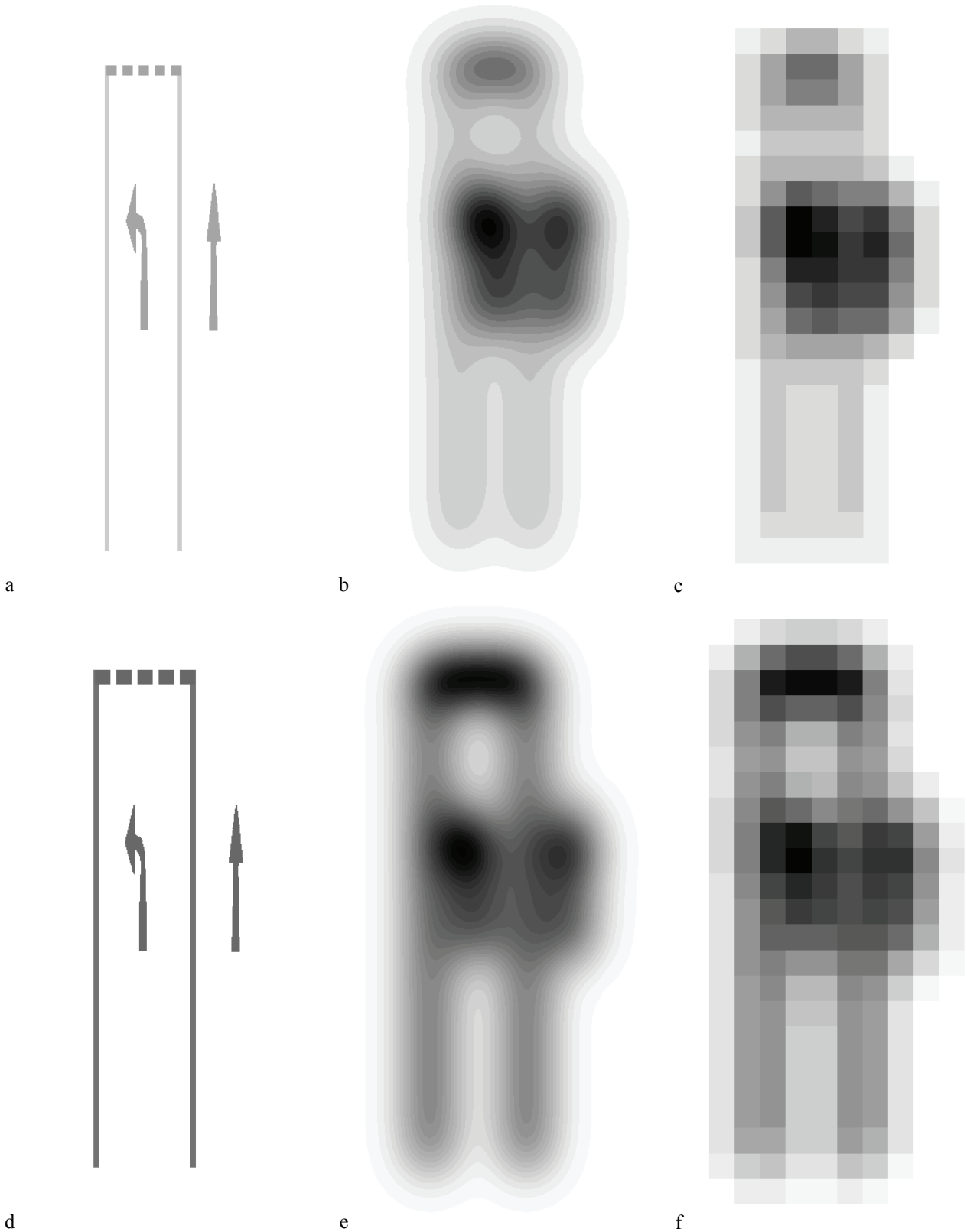


Abbildung 4.20: Synthetische Bilder für die minimale (oben) und maximale Konfiguration (unten) der Interaktionsgruppe in der Ursprungsauflösung A_0 (a+d), Zielskala σ_{z3} (b+e) und Zielauflösung A_{z3} (c+f) (zur besseren Visualisierung sind b,c,e,f kontrastverstärkt)

Die *Fusion* fügt alle verbliebenen Objektteile in der Zielauflösung zu zwei kompletten Objektmodellen zusammen. Da für die linienartigen Objektteile unterschiedliche Skalenraumereignisse für die minimale und maximale Konfiguration eingetretenen sind, werden zwei angepasste Objektmodelle für die Zielauflösung A_{z3}=1.00m erstellt (s. Abbildung 4.22).

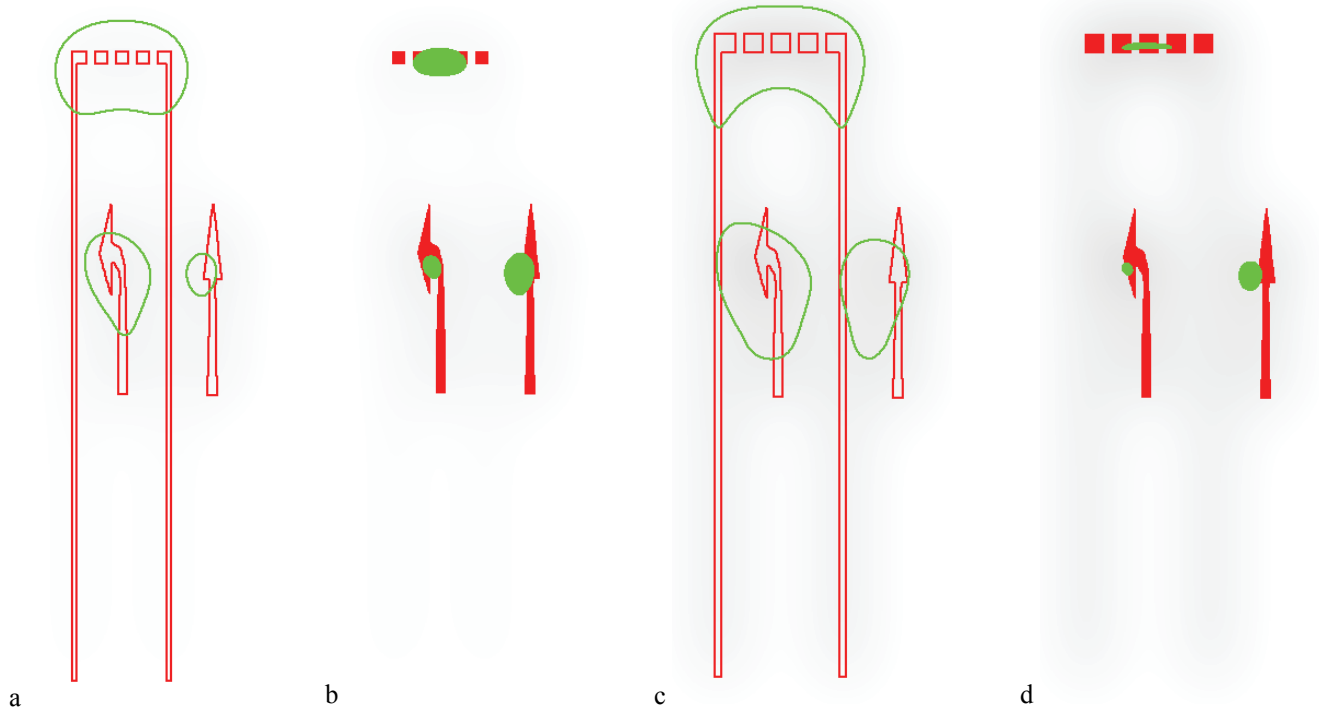


Abbildung 4.21: Ergebnisse der Blob-Detektion für die minimale (a+b) und maximale Konfiguration (c+d) der Interaktionsgruppe in der Ursprungsauflösung A_0 (rot) und der Zielskala σ_{z3} (grün) auf dem Bild der Zielskala $I_{\sigma z3}$: Basisregionen (a+c) und Extrema der Blobs (b+d)

4.1.3 Beispiel 3 – Straße mit Fahrzeug als lokales Kontextobjekt

Dieser Abschnitt zeigt ein Beispiel zur Anpassung eines Objektmodells zur Extraktion einer Straße, das Fahrzeuge als mögliche lokale Kontextobjekte integriert. Das Objektmodell des Primärobjektes Straße entspricht dem der einfachen Straße aus Beispiel 1 in einer Ursprungsauflösung A_0 von 0.03 m. Die Anpassung wird für die Zielauflösungen A_{z1}=0.1 m, A_{z1}=0.2 m und A_{z1}=1.00 m demonstriert.

4.1.3.1 Objektmodell für die hohe Bildauflösung

Das Objektmodell der Ursprungsauflösung A_0 = 0.03 m für das Primärobjekt Straße und das lokale Kontextobjekt Fahrzeug (hier als Lkw) mit den lokalen Kontextrelationen ist in Abbildung 4.23 gegeben. Um eine flexible Modellierung des lokalen Kontextes zu erreichen, kann die Position des Fahrzeugs innerhalb der Fahrbahn variieren. Die Verbindung zum lokalen Kontextobjekt wird über die Relationen *optional occlusion* oder *optional shadow* hergestellt. Da sämtliche Objektteile des Primärobjektes eine Verbindung zum lokalen Kontext aufweisen, können alle Objektteile des Primärobjektes von einer Verdeckung oder einem Schattenwurf durch das Fahrzeug betroffen sein.

Das Objektmodell des Fahrzeugs in der Ursprungsauflösung (dargestellt in Abbildung 4.24) ist relativ einfach und weniger detailliert als beispielsweise das Modell für Fahrzeuge in HINZ [2004]. Trotz seiner Einfachheit sollte es jedoch für eine Demonstration des Anpassungskonzeptes ausreichen. Abbildung 4.29 a+f zeigt synthetische Bilder für die Simulation des modellierten Fahrzeugs. In diesem Beispiel ist ein größeres Fahrzeug modelliert, das aus zwei Objektteilen für *Front* und *Anhänger* besteht, die beide als Rechtecke und damit als flächenhafte Objekttypen beschrieben sind. Die Merkmalsoperatoren der Fahrzeugteile nutzen Flächenwachstum zur Segmentierung. Geeignete Kandidaten für die Fahrzeugteile werden anschließend aus der Segmentierung aufgrund ihrer Formmerkmale (Konvexität, Kompaktheit, Exzentrizität, Fläche) und radiometrischer Eigenschaften ausgewählt (siehe Abschnitt 4.2).

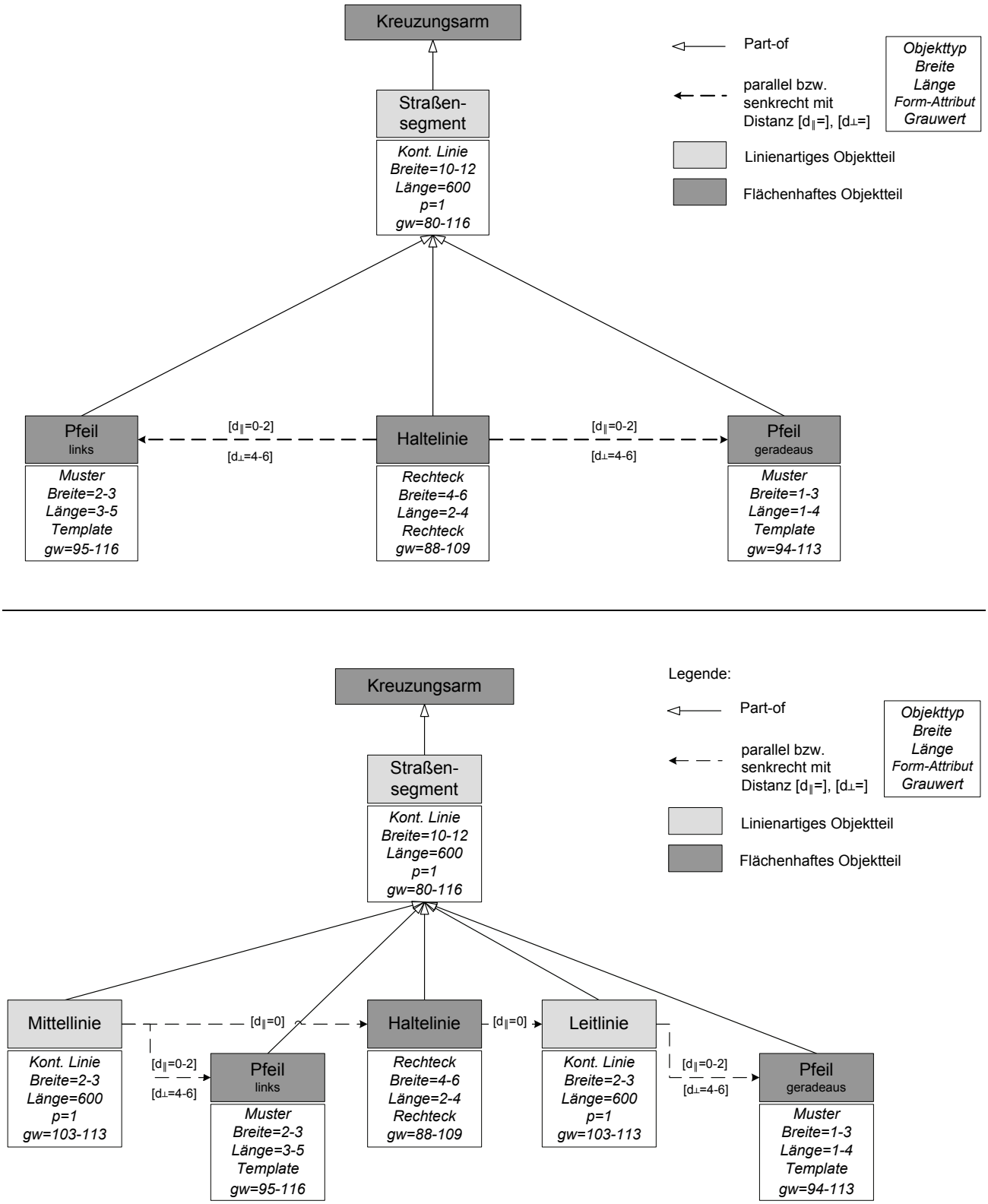


Abbildung 4.22: Automatisch angepasste Objektmodelle für die Zielauflösung A_{z3}=1.00 m

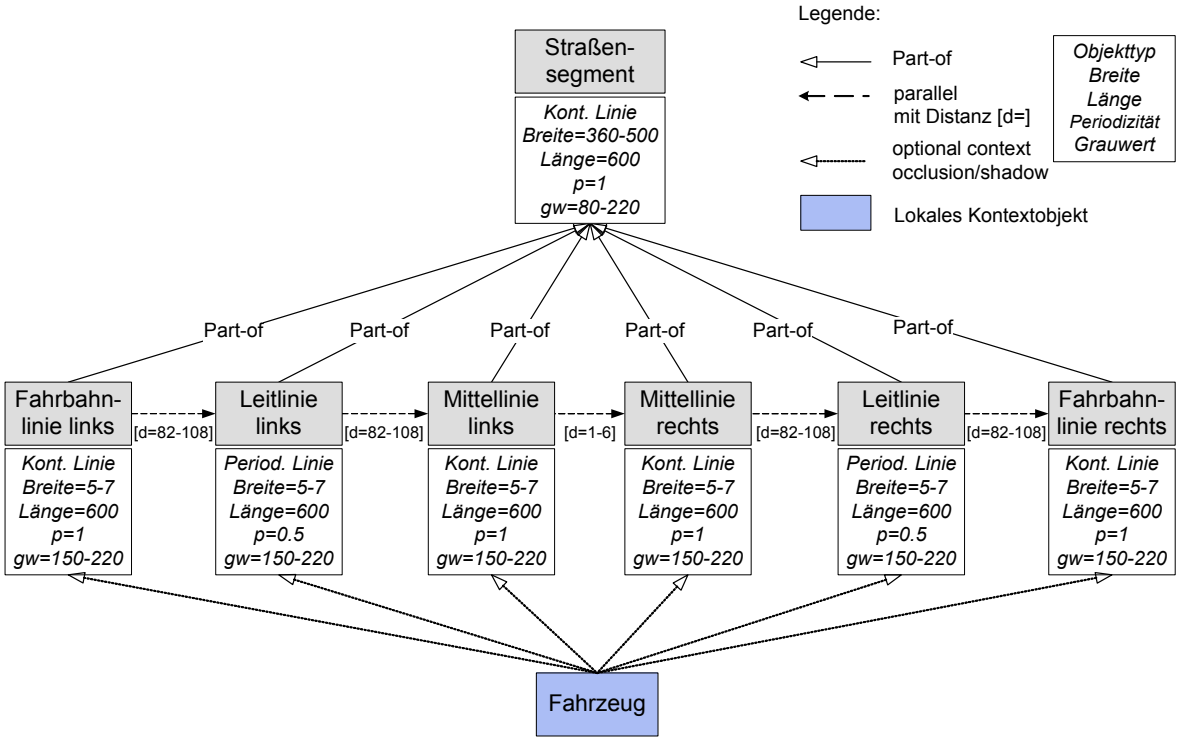


Abbildung 4.23: Objektmodell für eine Straße mit lokalem Kontext (Fahrzeug) in 0.03 m/Pixel

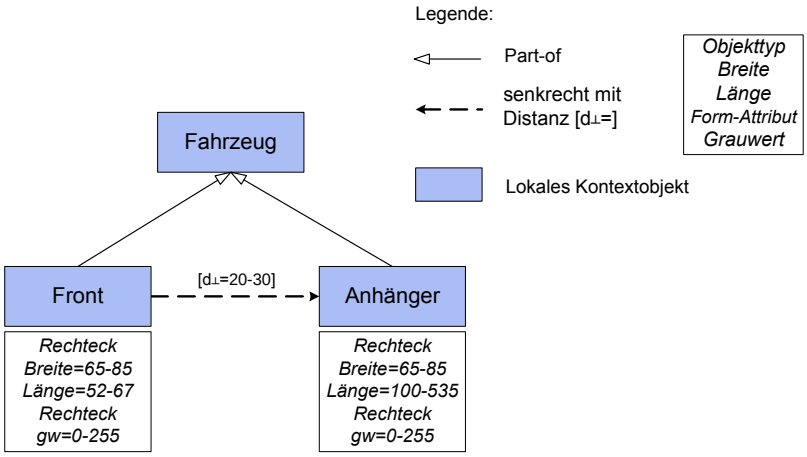


Abbildung 4.24: Objektmodell für ein Fahrzeug als lokales Kontextobjekt in 0.03 m/Pixel

4.1.3.2 Automatische Anpassung

Nach Abschnitt 3.3.3 wird die automatische Anpassung von Objektmodellen, die lokale Kontextobjekte berücksichtigen, für das Objektmodell des Primärobjektes und für das Objektmodell des lokalen Kontextobjektes separat durchgeführt. In diesem Abschnitt wird die automatische Anpassung des Objektmodells für das Fahrzeug als lokales Kontextobjekt demonstriert.

Zielauflösung 0.1 m

Die erste Zielauflösung A_{z1}=0.1 m entspricht einer Auflösungsänderung F_{A1}=3. Es ist eine Glättung mit einem Gauß-Filter mit σ_{z1}=3 und einer Breite von 19 Pixeln nötig. Die Ergebnisse und Anpassung für das Primärobjekt Straße an diese Bildauflösung wurden bereits im Beispiel 1 in Abschnitt 4.1.1 gezeigt.

In der *Zerlegung* wird weder für die minimale noch für die maximale Konfiguration des Fahrzeugs Interaktion festgestellt, so dass das Skalenverhalten beider Objektteile jeweils getrennt untersucht werden kann. In der *Analyse des Skalenverhaltens* wird daher ein synthetisches Bild für jedes Objektteil erzeugt, das in die Zielskala und die Zielauflösung überführt wird. Zur Prädiktion des Skalenverhaltens der Fahrzeugteile wird die minimale und maximale Attributkonfiguration mit dem maximal möglichen Grauwert untersucht, da das ursprüngliche Objektmodell keine Beschränkungen hinsichtlich des Grauwertes angibt, d. h. für ein Fahrzeug sind alle Werte von 0–255 möglich. Da das Fahrzeug als optionales Objekt modelliert ist, wird der optimistischste Fall mit maximalem Kontrast simuliert, der zu der längsten Lebenszeit des Kontextobjektes im Skalenraum führt. Die synthetischen Bilder der minimalen und maximalen Konfiguration für die Fahrzeug-Front und den Anhänger zeigen Abbildung 4.25 + 4.26 a-c+f-h.

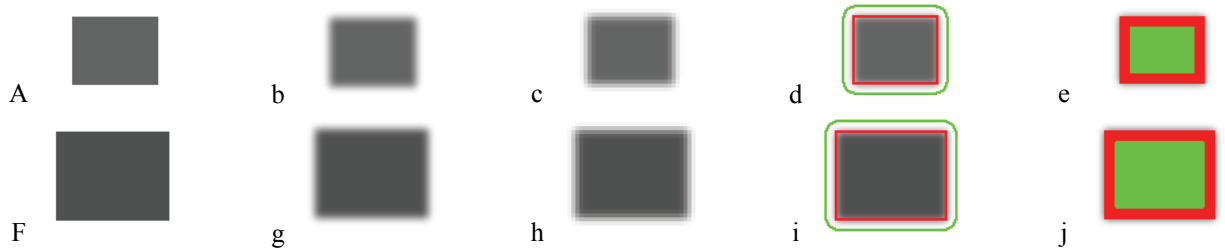


Abbildung 4.25: Synthetische Bilder der Fahrzeug-Front für die minimale (oben) und die maximale Konfiguration (unten) in der Ursprungsauflösung A_0 (a+f), Zielskala σ_{z1} (b+g) und Zielauflösung A_{z1} (c+h) sowie die Ergebnisse der Blob-Detektion (Basisregionen d+i und Extrema e+j)

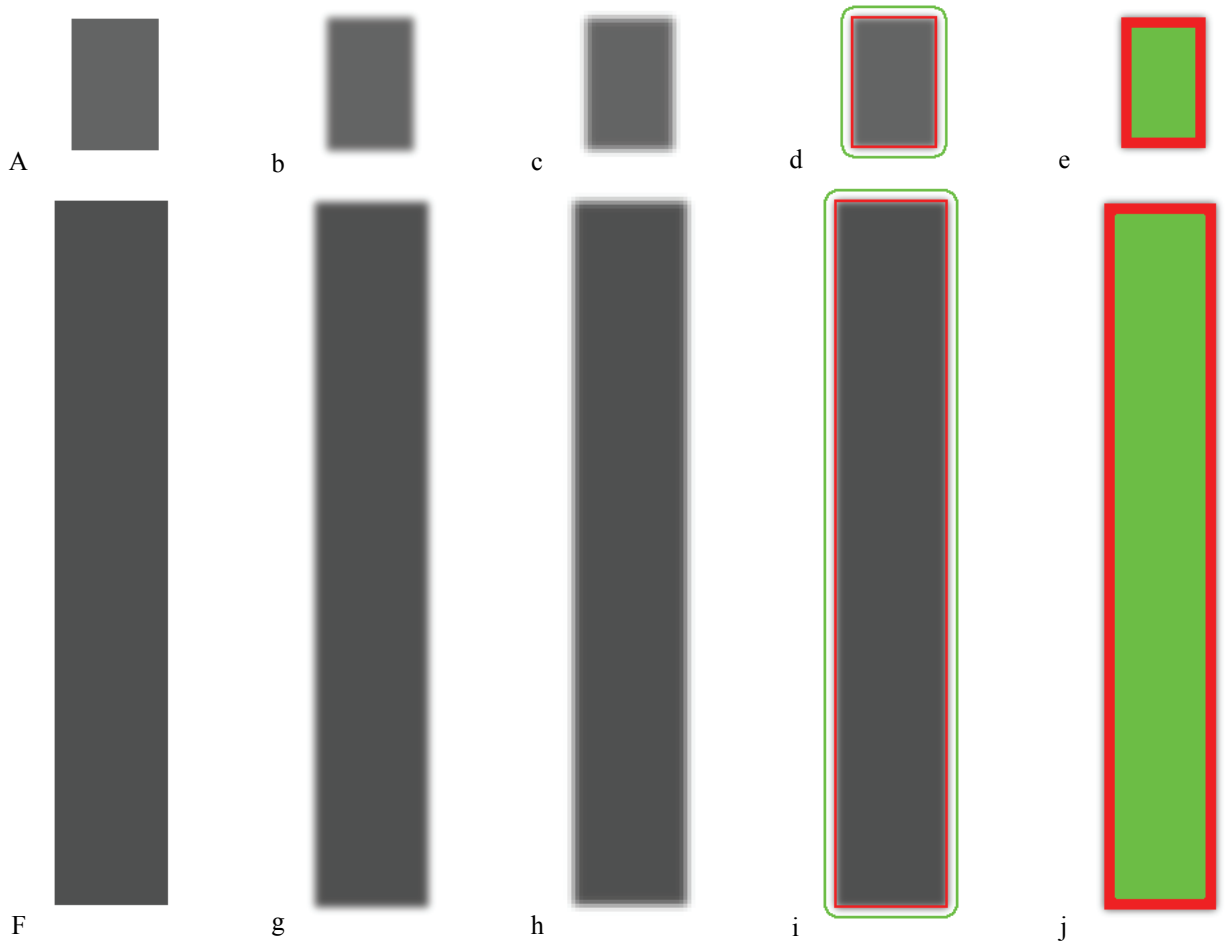


Abbildung 4.26: Synthetische Bilder des Fahrzeug-Anhängers für die minimale (oben) und maximale Konfiguration (unten) in der Ursprungsauflösung A_0 (a+b), Zielskala σ_{z1} (c+d) und Zielauflösung A_{z1} (e+f) sowie die Ergebnisse der Blob-Detektion (Basisregionen d+i und Extrema e+j)

Da keine Interaktion für die Zielauflösung A_{z1}=0.1 m zwischen den Objektteilen stattfindet, können auch keine Verschmelzungen bei der vorliegenden Auflösungsänderung eingetreten sein. Es treten auch keine anderen Skalenraumereignisse für beide Objektteile ein, wie aus der Blob-Detektion hervorgeht. Die Attribute in der Zielauflösung werden aus den Basisregionen und Extrema der Blobs in der Zielskala ermittelt. Die Ergebnisse der Blob-Detektion mit den Extrema und Basisregionen in der Ursprungs- und Zielskala zeigen Abbildung 4.25 + 4.26 d,e,i,j.

Da keine Skalenraumereignisse prädiziert wurden, werden in der *Fusion* alle Objektteile mit den angepassten Attributen wieder zu einem gemeinsamen Objektmodell für das Fahrzeug zusammengesetzt. Das angepasste Fahrzeug-Objektmodell für die Zielauflösung A_{z1}=0.1 m ist in Abbildung 4.27 dargestellt.

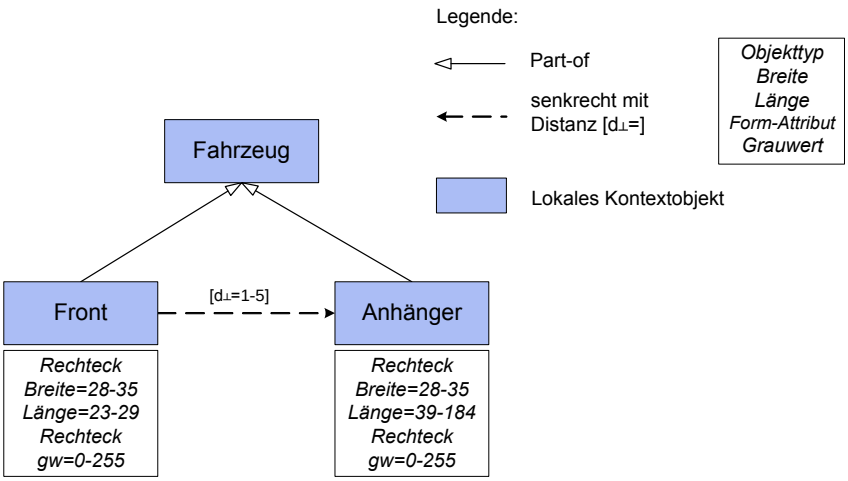


Abbildung 4.27: Automatisch angepasstes Objektmodell für ein Fahrzeug in der Zielauflösung A_{z1}=0.1 m

Zielauflösung 0.2 m

Für die zweite Zielauflösung A_{z2}=0.2 m (entsprechend einer Auflösungsänderung F_{A2}=6) ist eine Glättung mit einem Gauß-Filter mit σ_{z2}=6 und einer Breite von 37 Pixeln nötig. Für das Objektmodell des Primärobjektes Straße wurde die automatische Anpassung mit der eindimensionalen Strategie und Methode durchgeführt (siehe Beispiel 1). Das an die Zielauflösung A_{z2}=0.2 m angepasste Objektmodell für eine einfache Straße ist in Abbildung 4.28 dargestellt.

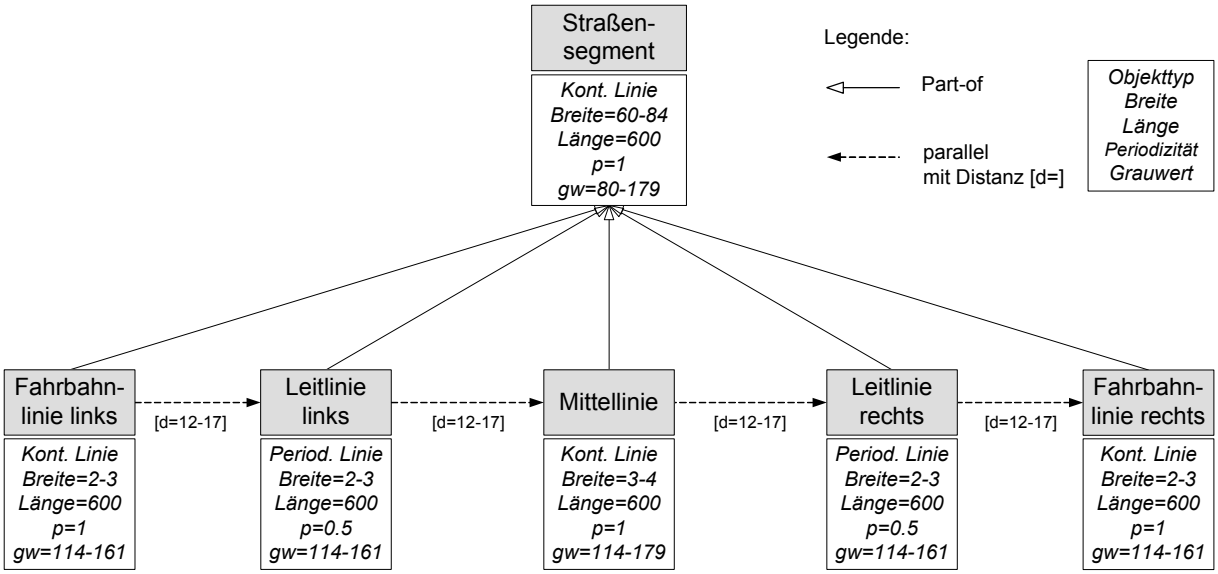


Abbildung 4.28: Automatisch angepasstes Objektmodell für einfache Straße in der Zielauflösung A_{z2}=0.2 m

Für diese Auflösungsänderung findet sowohl in der minimalen als auch in der maximalen Konfiguration des lokalen Kontextobjektes Fahrzeug Interaktion statt. Das Skalenverhalten muss daher für beide Objektteile gemeinsam in einer Interaktionsgruppe untersucht werden, die in der *Zerlegung* erzeugt wird.

Die synthetischen Bilder, die die Interaktionsgruppe zur *Analyse des Skalenverhaltens* simulieren, sind in Abbildung 4.29 a-c+f-h dargestellt. Aus den Ergebnissen der Blob-Detektion in der Ursprungsauflösung A_0 und der Zielskala σ_{z2}=6 wurden keine Skalenraumereignisse prädiziert (Abbildung 4.29 d,e,i,j). Es existieren jeweils zwei Blobs mit einer gemeinsamen Sattelregion in der Ursprungs- und Zielauflösung. Jeder Blob entspricht einem Objektteil. Die Anzahl der Objektteile bleibt in der Zielauflösung trotz Interaktion zwischen den beiden Objektteilen unverändert.

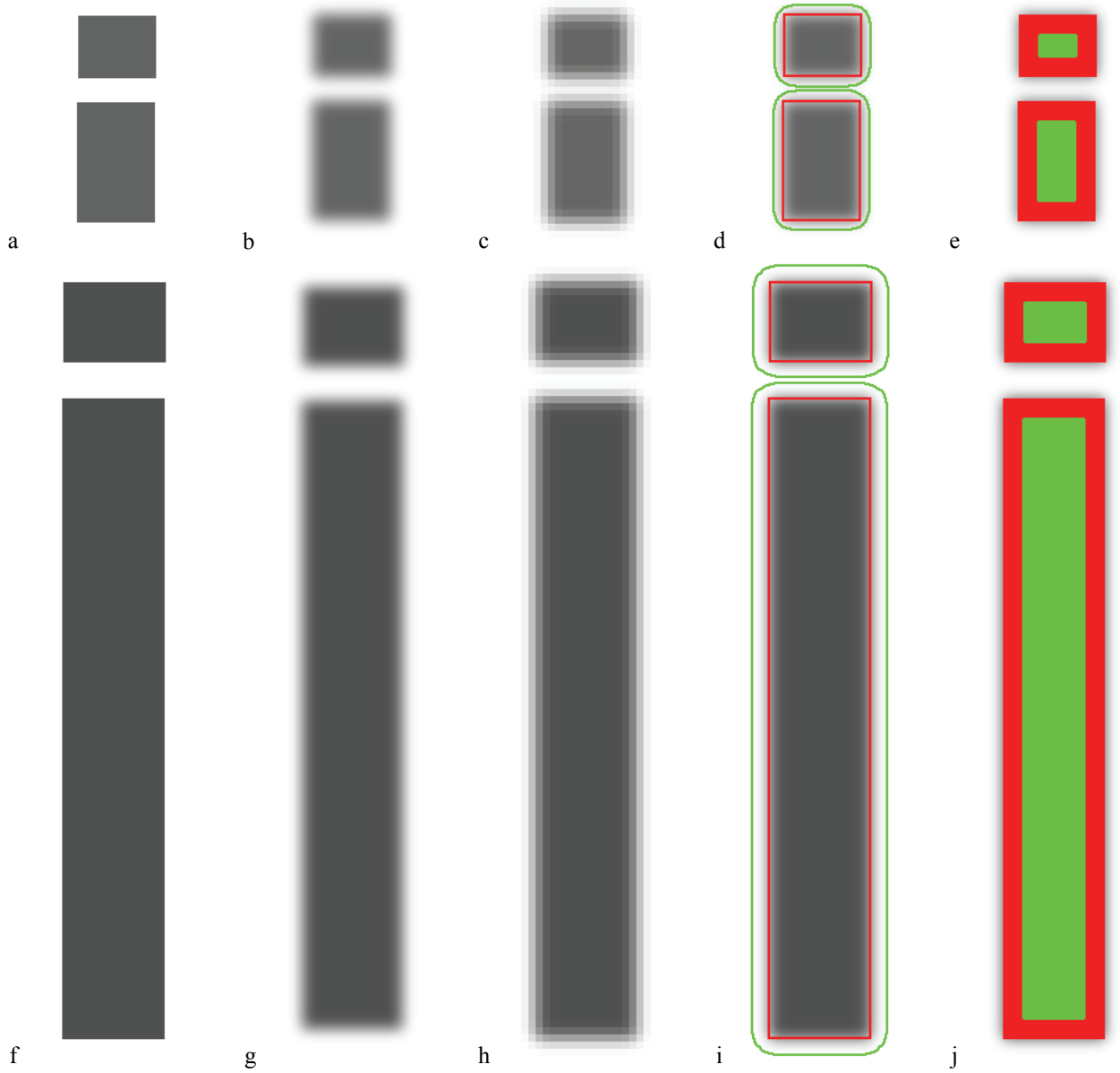


Abbildung 4.29: Synthetische Bilder für die minimale (oben) und die maximale Konfiguration (unten) der Interaktionsgruppe in der Ursprungsauflösung A_0 (a+f), Zielskala σ_{z2} (b+g) und Zielauflösung A_{z2} (c+h) sowie die Ergebnisse der Blob-Detektion (Basisregionen d+i und Extrema e+j)

Da für diese Auflösungsänderung keine Skalenraumereignisse stattfinden, bleibt die Struktur des gegebenen Objektmodells in der Zielauflösung erhalten. In der *Fusion* wird daher ein Objektmodell mit den angepassten Attributen und Operatoren beider Objektteile des Fahrzeugs erstellt. Das an die Zielauflösung A_{z2}=0.2 m angepasste Objektmodell für das lokale Kontextobjekt Fahrzeug zeigt Abbildung 4.30.

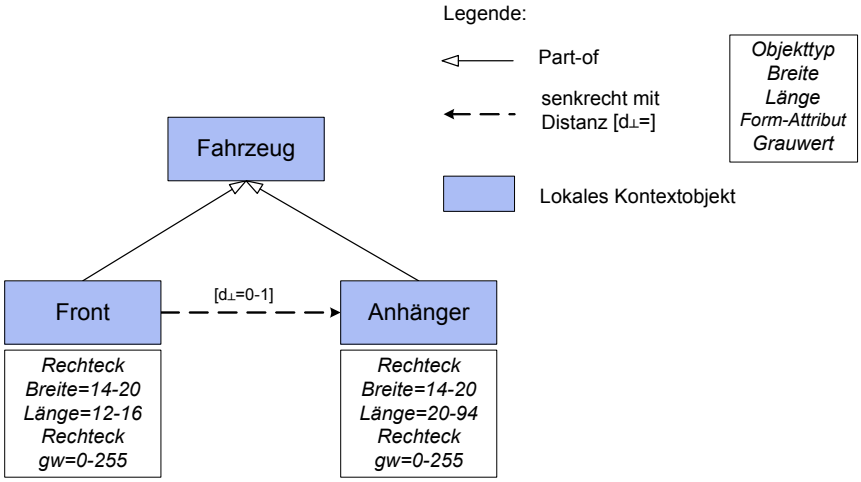


Abbildung 4.30: Automatisch angepasstes Objektmodell für ein Fahrzeug in der Zielauflösung A_{z2}=0.2 m

Zielauflösung 1.00 m

Die dritte gewählte Zielauflösung A_{z3}=1.00 m bewirkt eine Auflösungsänderung von F_{A3}=30. Der zur Zielskala σ_{z3}=30 korrespondierende Gauß-Filter hat eine Filterbreite von 181 Pixel. Die mit der eindimensionalen Methode angepassten Objektmodelle des Primärobjektes Straße für die Zielauflösung A_{z3} sind in Abbildung 4.8 dargestellt.

Die Filterbreite des zur Glättung nötigen Gauß-Filters bewirkt für das Kontextobjekt Fahrzeug eine Interaktion zwischen den beiden Objektteilen. In der *Zerlegung* wird daher für das Fahrzeug wieder eine einzige Interaktionsgruppe für die minimale und maximale Konfiguration erzeugt. In der *Analyse des Skalenverhaltens* werden die synthetischen Bilder untersucht, die diese Interaktionsgruppe simulieren (Abbildung 4.31). Die Interaktionsgruppe wird auf Skalenraumereignisse mit Hilfe der in den synthetischen Bildern in der Ursprungsauflösung I_0 und in der Zielskala $I_{\sigma z3}$ abgegrenzten Blobs analysiert. Für die minimale Konfiguration tritt eine sichere Verschmelzung auf. Es existiert nur noch ein einziger Blob in der Zielskala (siehe Abbildung 4.31). Für die maximale Konfiguration dagegen wird kein sicheres Skalenraumereignis prädiziert, da in der Zielskala zwei getrennte Blobs verbleiben, wie Abbildung 4.31 illustriert. Aufgrund des prädizierten geringen Kontrastes der beiden Objektteile zum Hintergrund wird der Operator für das Fahrzeug mit angepassten Parametern auf das synthetische Bild der Zielauflösung I_{Az3} der maximalen Konfiguration angewendet, um auf effektive Skalenraumereignisse zu testen. Der angepasste Operator ist für die Zielauflösung A_{z3} nicht in der Lage, zwei getrennte Objektteile zu extrahieren, sondern nur ein einziges zusammenhängendes Objektteil. Damit ist für die maximale Konfiguration eine effektive Verschmelzung eingetreten und sowohl die Anpassung der minimalen als auch der maximalen Konfiguration resultiert in ein einziges verbliebenes Objektteil. Das automatisch an die Zielauflösung A_{z3}=1.00 m angepasste Objektmodell für das Fahrzeug ist in Abbildung 4.32 dargestellt.

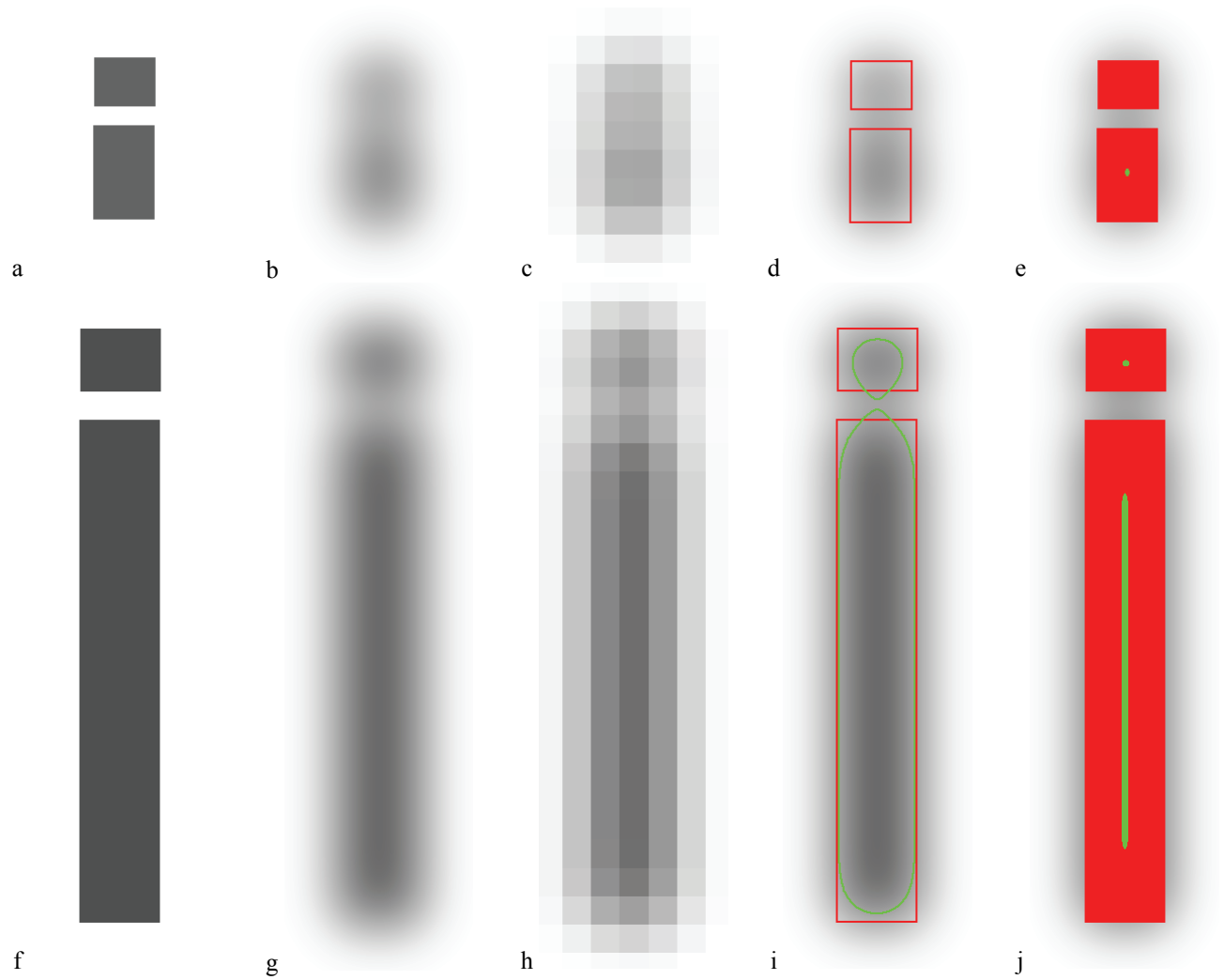


Abbildung 4.31: Synthetische Bilder für die minimale (oben) und maximale Konfiguration (unten) der Interaktionsgruppe in der Ursprungsauflösung A_0 (a+f), Zielskala σ_{z3} (b+g) und Zielauflösung A_{z3} (c+h) sowie die Ergebnisse der Blob-Detektion (Basisregionen d+i und Extrema e+j)

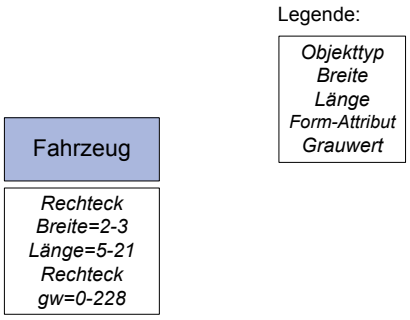


Abbildung 4.32: Automatisch angepasstes Objektmodell für ein Fahrzeug in der Zielauflösung A_{z3}=1.00 m

4.2 Implementierung der Beispielsysteme

Die gegebenen als auch die automatisch angepassten Objektmodelle der vorgestellten Beispiele wurden größtenteils im wissensbasierten Bildinterpretationssystem *geoAIDA* für eine Objektextraktion implementiert (siehe Kapitel 2.1.2.1). Die erzielten Extraktionsergebnisse wurden für die Verifikation des entwickelten Verfahrens genutzt, die im folgenden Abschnitt 4.3 dokumentiert ist.

Die vorausgesetzte Extraktionsstrategie unterteilt den Extraktionsprozess in zwei Schritte: die Extraktion der Kandidaten für die Objektteile sowie deren Bewertung bzw. Gruppierung zu einem Gesamtobjekt (vgl. Abschnitt 3.1.4). Für die Extraktion der Objektteile wurden die folgenden Operatoren zur Merkmalsextraktion verwendet, die den Top-down-Operatoren in geoAIDA entsprechen:

Fahrbahnbegrenzung: Dieser Operator extrahiert linienartige durchgezogene Fahrbahnmarkierungen als helle Linien mit vorgegebenen Eigenschaften bzgl. Linienbreite und Kontrast zum Hintergrund. Der Kontrast wird als Hysterese-Schwellwert angegeben und entspricht der zweiten Ableitung des stabförmigen Linienprofils mit der Gauß-Funktion, deren σ sich proportional zur Linienbreite verhält. Die Linienextraktion basiert auf dem Linienoperator nach STEGER [1998A], gefolgt von einer Bewertung und Fusion der Linien nach WIEDEMANN [2002]. Der Algorithmus von Wiedemann wurde auf die speziellen Bedingungen der Extraktion von Fahrbahnmarkierungen aus hoch auflösenden optischen Bilddaten angepasst [SCHRAMM 2005]. Die Ausgabe des Operators umfasst die Linien im Bild mit ihren Attributen bzgl. der Linienbreite und des mittleren Grauwertes in der Linienmitte.

Leitlinie: Leitlinien werden als linienartige gestrichelte Fahrbahnmarkierung extrahiert. Der Operator arbeitet ähnlich zu dem für die Fahrbahnbegrenzung. Zur Bewertung der extrahierten Linienstücke wird zusätzlich der regelmäßige Abstand zwischen ihnen herangezogen. Ergebnis des Operators sind die Linien im Bild mit ihren Attributen bzgl. der Linienbreite, des mittleren Grauwertes in der Linienmitte sowie der Periodizität der extrahierten periodischen Linien [SCHRAMM 2005].

Haltelinie: Die Haltelinie wird als eine Kombination mehrerer nahe liegender flächenhafter Rechtecke basierend auf Flächenwachstum extrahiert. Die extrahierten Haltelinienteile werden nach ihrer Größe, Form (Kompaktheit) und mittlerem Grauwert ausgewählt. Eingabeparameter des Operators sind Anzahl, Fläche und Kompaktheit der Rechtecke sowie deren mittlerer Grauwert als minimale und maximale Werte.

Richtungspfeile: Der Operator zur Extraktion von Richtungspfeilen auf der Fahrbahn detektiert flächenhafte Muster für verschiedene Pfeil-Typen basierend auf Template-Matching. Die Templates werden als Musterbild für den gesuchten Pfeil bereitgestellt. Zuerst wird das Bild mit dem Flächenwachstumsverfahren segmentiert, um Kandidaten für die Pfeilregionen zu erhalten. Die Übereinstimmung des Templates mit den ermittelten Pfeilkandidaten wird über die Kreuzkorrelationsmethode bewertet. Die Ausgabe des Operators umfasst die im Bild extrahierten Pfeile mit ihren Attributen (mittlerer Grauwert des gefundenen Musters, Größe des kleinsten umschließenden Rechtecks um den extrahierten Pfeil sowie der erreichte Wert des Korrelationskoeffizienten zum Musterbild).

Fahrzeuge: Fahrzeuge sind nach dem Modell in flächenhafte Rechtecke unterteilt. Zuerst wird das Eingabebild mit dem Flächenwachstumsverfahren segmentiert. Anschließend werden aus den Segmenten die Kandidaten der Fahrzeugteile nach ihren Formmerkmalen (Kompaktheit, Konvexität, Exzentrizität), Größe (Fläche) und radiometrischen Eigenschaften ausgewählt, die dem Operator als Eingabeparameter vorher angegeben werden. Die Kandidaten werden nach ihren Abständen und topologischen Relationen bewertet. Der Operator gibt die extrahierten Fahrzeuge im Eingabebild mit ihren Attributen (mittlerer Grauwert und die Größe des kleinsten umschließenden Rechtecks) aus.

Der Bottom-Up-Operator nimmt die Gruppierung bzw. Bewertung der von den Top-down-Operatoren extrahierten Objektteile wahr. Die Auswahl der Objektteile mit den korrekten Eigenschaften, nach Vorgabe der Attribute und Relationen, resultiert in dem Extraktionsergebnis des modellierten Landschaftsobjektes aus einem gegebenen Bild. Nur wenn alle Bedingungen aus den Relationen erfüllt sind, wird die Extraktion des Objektes (Straße im Beispielsystem) als erfolgreich betrachtet.

4.3 Verifikation

Um eine Bewertung und Analyse des entwickelten Verfahrens zur automatischen Anpassung von Bildanalyse-Objektmodellen vornehmen zu können, wurde eine Verifikation durchgeführt. Die Verifikation soll nicht nur eine Aussage darüber treffen, ob eine Extraktion mit den automatisch angepassten Objektmodellen überhaupt erfolgreich ist, sondern auch die Qualität der Anpassung der Objektmodelle an eine niedrigere Bildauflösung bewerten.

4.3.1 Methode der Verifikation

Zur Analyse einer neu entwickelten Methode sollten die erzielten Ergebnisse anhand einer geeigneten Referenz empirisch bewertet werden. Das Ergebnis des in dieser Arbeit entwickelten Verfahrens zur automatischen Anpassung sind Objektmodelle für die Extraktion eines Landschaftsobjektes aus einer bestimmten Bildauflösung.

Aufgabe der Objektmodelle ist die automatische Extraktion des modellierten Objektes aus geeigneten Bilddaten. Daher dient in der Methode der Verifikation das mit dem gegebenen Objektmodell erreichte Extraktionsergebnis des Landschaftsobjektes aus Bilddaten der hohen Auflösung als Referenz zur Bewertung der Qualität des Verfahrens. Mit dieser Referenz wird das mit dem angepassten Objektmodell erzielte Extraktionsergebnis aus der niedrigeren Bildauflösung verglichen. Das Konzept der angewendeten Verifikation zeigt Abbildung 4.33 (siehe auch [HEUWOLD 2006]).

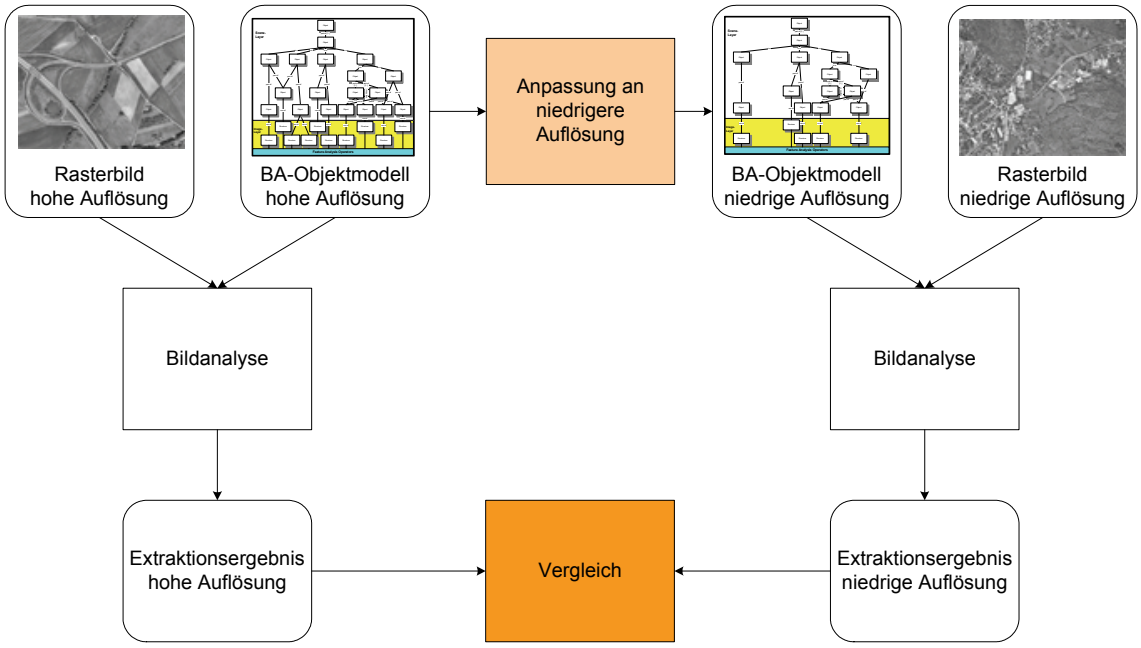


Abbildung 4.33: Konzept der Verifikation

Sowohl mit dem originalen Objektmodell für eine hohe Bildauflösung als auch mit dem automatisch angepassten Objektmodell für die niedrigere Bildauflösung wird eine Objektextraktion aus Bilddaten der jeweiligen Auflösung durchgeführt. Um eine Vergleichbarkeit der Extraktionsergebnisse zu gewährleisten, wird das Bild der niedrigeren Auflösung direkt aus dem Bildmaterial der hohen Auflösung abgeleitet. Zu diesem Zweck wird das hochauflösende Bild erst einer Filterung mit einem geeigneten Gauß-Kernel unterzogen. Das geglättete Bild wird anschließend abgetastet und mit Hilfe der bilinearen Interpolation in die niedrigere Bildauflösung überführt. Die aus beiden Bildauflösungen erzielten Extraktionsergebnisse werden dann miteinander verglichen.

Um einen detaillierten Einblick in die potenziellen Schwächen des automatischen Anpassungsverfahrens zu erhalten, bezieht sich die Verifikation auf die einzelnen Objektteile des Landschaftsobjektes. Da sich die Struktur des Objektmodells im Prozess der auflösungsabhängigen Anpassung aufgrund von Skalenraumereignissen ändern kann, ist ein direkter Vergleich der Extraktionsergebnisse für die Objektteile nicht möglich. Daher müssen bei dem Vergleich der erzielten Extraktionsergebnisse aus beiden Bildauflösungen die Unterschiede eingehend analysiert werden. Grundsätzlich

kann ein Unterschied in den Extraktionsergebnissen drei Ursachen haben: Die erste Ursache ist das Auftreten eines oder mehrerer Skalenraumereignisse, die dazu führen, dass betroffene Objektteile in der niedrigeren Auflösung nicht mehr extrahiert werden können, oder zusätzliche Objektteile extrahiert werden (bei Erzeugungen und Spaltungen). Diese Ursache kann leicht bestätigt werden, wenn das gegebene und das angepasste Objektmodell unterschiedliche Strukturen aufweisen, da sich die Skalenraumereignisse im Modell der niedrigeren Auflösung widerspiegeln sollten. Ein weiterer Grund für Extraktionsunterschiede kann die inkonsistente Leistungsfähigkeit der Merkmalsoperatoren der Objektteile in unterschiedlichen Bildauflösungen sein. Die Extraktion aus den niedrig aufgelösten Bilddaten könnte schlechtere oder sogar bessere Ergebnisse erreichen, falls zuvor bei den Verfahren für die hohe Bildauflösung nicht die niedrigere Bildauflösung mit berücksichtigt worden ist. Drittens kann die automatische Anpassung des Objektmodells falsch sein. In diesem Fall muss die Methode der automatischen Anpassung korrigiert werden.

Für den Vergleich der mit den automatisch angepassten Objektmodellen erzielten Extraktionsergebnisse dienen *Vollständigkeit, Korrektheit* und *Qualität*, die wie folgt definiert sind [McGlone & Shufelt 1994], [Heipke et al. 1998]:

$$Vollständigkeit = \frac{True\ Positives}{True\ Positives + False\ Negatives} \tag{4.2}$$

$$Korrektheit = \frac{True\ Positives}{True\ Positives + False\ Positives} \tag{4.3}$$

$$Qualität = \frac{True\ Positives}{True\ Positives + False\ Positives + False\ Negatives} \tag{4.4}$$

mit *True Positives*: Länge der korrekt extrahierten Objekte

False Negatives: Länge der nicht extrahierten, aber vorhandenen Objekte

False Positives: Länge der extrahierten, aber nicht existierenden Objekte

Das Ergebnis der Extraktion aus der hohen Bildauflösung dient wie oben erwähnt als Referenzdatensatz, d.h. dieses Extraktionsergebnis repräsentiert 100% für Vollständigkeit, Korrektheit und Qualität. Durch Vergleich der Ergebnisse der Extraktion, die mit dem automatisch angepassten Objektmodell aus der korrespondierenden Bildauflösung erzielt wurden, mit den Referenzdaten, kann die Qualität des entwickelten Verfahrens zur automatischen Anpassung einschließlich der Performance der Merkmalsextraktoren bewertet werden. Üblicherweise werden die o. g. Qualitätsmaße Vollständigkeit, Korrektheit und Qualität im Bereich der automatischen Objektextraktion zur Bewertung der Extraktionsgüte einer Objektklasse in Bezug zu einer manuell erstellten Referenz verwendet. Im Gegensatz dazu wird in der hier durchgeführten Verifikation die Güte des gegebenen Objektmodells nicht bewertet.

Bei dem hier durchgeführten Vergleich der Extraktionsergebnisse aus verschiedenen Bildauflösungen kann ein besonderes Phänomen auftreten: Die Extraktionsergebnisse aus der niedrigeren Bildauflösung können, wie bereits angesprochen, besser sein als die aus der hohen Bildauflösung erzielten Resultate, denn die Ergebnisse einer automatischen Extraktion sind auch aus der hohen Bildauflösung, etwa im Vergleich zu einer manuell erstellten Referenz, nicht perfekt. Störungen im Bildmaterial führen häufig zum Versagen eines Operators in einigen Bereichen. In einer niedrigeren Bildauflösung können kleinere Störungen im Bild, die von der Verdeckung durch kleinere Objekte oder deren Schatten resultierten, ausgelöscht bzw. geglättet sein und keinen störenden Einfluss mehr auf die Extraktion desselben Objektteils nehmen. Dann würde der Vergleich zum Extraktionsergebnis der hohen Bildauflösung diese zusätzlich extrahierten Stellen als False Positives einordnen und demnach zu einer verringerten Korrektheit führen. Da diese Extraktionsergebnisse jedoch mit der Bildinformation übereinstimmen und aus der mangelhaften Performance des Operators in der hohen Bildauflösung herrühren, werden diese zusätzlich extrahierten Bereiche im hier durchgeführten Vergleich der Verifikation als True Positives bewertet. Die Bewertung erfolgte manuell.

4.3.2 Bilddaten

Für den Prozess der Verifikation werden hoch auflösende Grauwertbilder einer Vorstadtregion nahe Hannover verwendet. Die Bilder wurden im Maßstab 1:3300 mit einer RMK 30/23 aufgenommen und auf 0.03 m Bildauflösung mit 10µm Pixelgröße digitalisiert. Abbildung 4.34 bis Abbildung 4.37 zeigen die für die Verifikation genutzten Testbilder. Für die Verifikation der Methode zur Anpassung von Objektmodellen mit linienartigen, parallelen Objektteilen (Beispiel 1) wurden drei jeweils 5000x5000 Pixel umfassende Testbilder (Bild 1–3) verwendet, die sowohl gerade als auch kurvige Straßen zeigen. In den Bildern treten teilweise Störungen auf, die die Extraktion behindern: z. B. Schatten von Bäumen und Gebäuden, Fahrzeuge sowie eine nicht dauerhafte Baustellenmarkierung, die nicht in dem gegebenen Objektmodell aus Beispiel 1 enthalten ist.

Für die Verifikation der entwickelten Methode zur automatischen Anpassung von Objektmodellen mit flächenhaften Objektteilen wurde ein Testbild (Bild 4) ausgewählt, das in der hohen Bildauflösung (0.03 m/Pixel) aus 1500x3000 Pixel besteht. Das gegebene Objektmodell für den Kreuzungsbereich und die angepassten Objektmodelle aus Beispiel 2 wurden auf Bild 4 angewendet.

Die Verifikation für die automatische Anpassung von Objektmodellen, die lokale Kontextobjekte integrieren, wurde mit einem Testbild (Bild 5) durchgeführt. Das Testbild umfasst 3000x4500 Pixel und zeigt Fahrzeuge, die die Extraktion der Straßenmarkierungen behindern.

Abbildung 4.34: Testbild 1 in 0.03 m/Pixel

Abbildung 4.35: Testbild 2 in 0.03 m/Pixel

Abbildung 4.36: Testbild 3 in 0.03 m/Pixel

Bild 4

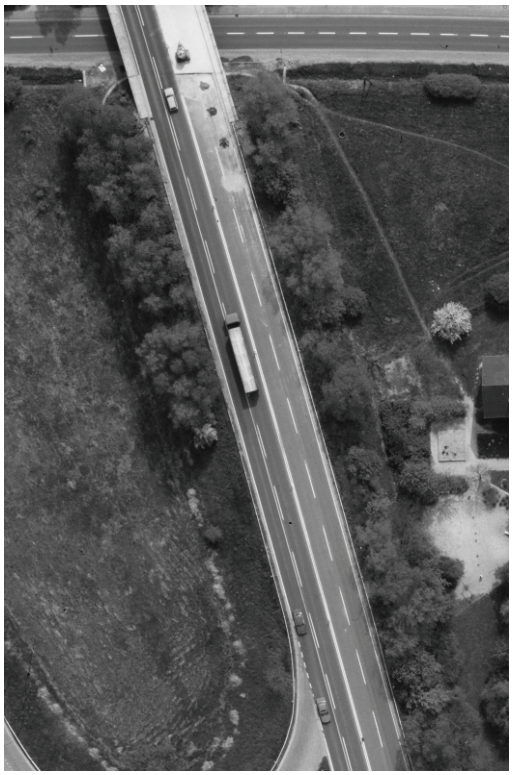


Bild 5

Abbildung 4.37: Testbilder 4 und 5 in 0.03 m/Pixel

4.3.3 Ergebnisse

In diesem Abschnitt werden die Ergebnisse der Verifikation für die in Abschnitt 4.1 präsentierten Beispiele der automatischen Anpassung dargestellt. Die erzielten Ergebnisse werden für jedes Beispiel diskutiert. Soweit möglich, werden Erklärungen für Probleme der Extraktion gegeben.

4.3.3.1 Beispiel 1 – Einfache Straße

Extraktionsergebnis Referenz 0.03m/Pixel

Das mit dem gegebenen Objektmodell (Abbildung 4.) in den Testbildern 1–3 aus der Ursprungsauflösung A_0=0.03 m/Pixel erzielte Extraktionsergebnis dient als Referenz für die Verifikation des entwickelten Verfahrens zur automatischen Anpassung. Das Ergebnis der Extraktion der Objektteile (linienartige Fahrbahnmarkierungen) mit den im Objektmodell für die Einfache Straße angegebenen Bildanalyse-Operatoren zeigt Abbildung 4.38. Die Extraktion aller Fahrbahnmarkierungen war aus den ausgewählten Testbildern erfolgreich. An einigen Stellen in den Bildern war die Extraktion der Fahrbahnbegrenzungslinien jedoch aufgrund von überdeckenden Fahrzeugen (Bild 1 und 2) oder Schattenwurf durch angrenzende Baumreihen (Bild 3) gestört. Auch ein Teil der Leitlinien konnte im unteren Drittel von Bild 2 nur unvollständig extrahiert werden. Sämtliche extrahierten Markierungen wurden zu einem Objekt Einfache Straße durch die Bottom-Up-Operatoren anhand der vom Modell der hohen Bildauflösung vorgegebenen räumlichen Relationen gruppiert. Trotz kleinerer Bildstörungen konnte aus allen drei Testbildern der modellierte Straßentyp fast vollständig extrahiert werden. Die Straßen ohne Fahrbahnmarkierungen in den Testbildern wurden korrekterweise nicht extrahiert, da ihre Eigenschaften nicht dem Modell entsprechen.

Bild 1

Bild 2

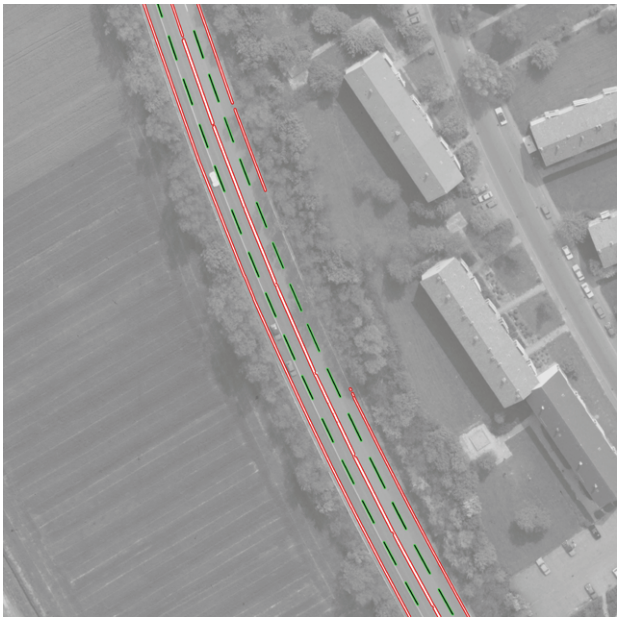


Bild 3

Abbildung 4.38: Extraktionsergebnisse für Bild 1–3 in 0.03 m/Pixel – Referenz (weiß/rot: Fahrbahnlinien und Mittellinien, grün/schwarz: Leitlinien)

Extraktionsergebnisse Einfache Straße A_{z1}=0.10 m/Pixel

Für die Zielauflösung A_{z1}=0.10 m resultierten zwei Objektmodelle aus der automatischen Anpassung, da in der minimalen und maximalen Attributkonfiguration verschiedene Skalenraumereignisse auftraten. Dementsprechend wurden in der Verifikation beide angepassten Objektmodelle zur Extraktion aus den Testbildern angewendet. Das Objektmodell, das aus der maximalen Konfiguration resultierte, enthält weiterhin die beiden Mittellinien aus dem Modell der Ursprungsauflösung. Im Objektmodell der minimalen Konfiguration dagegen sind die zwei Mittellinien zu einer einzigen Mittellinie verschmolzen, für deren Merkmalsoperator (ein Linienoperator) andere Parameter prädiziert wurden. Aus den ausgewählten Testbildern konnten keine zwei eigenständigen Mittellinien extrahiert werden, so dass das aus der maximalen Konfiguration resultierende Objektmodell nicht verifiziert werden konnte. Die Ergebnisse der Extraktion der einzelnen Objektteile für beide Objektmodelle aus den Bildern 1–3 zeigt Abbildung 4.39.

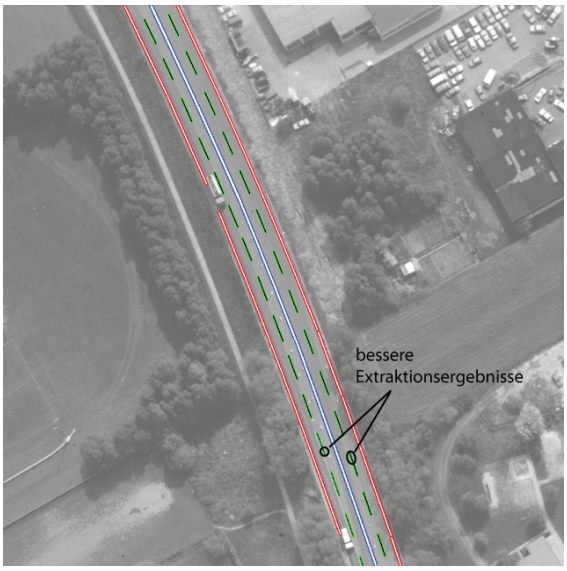


Bild 1 – Objektmodell 1

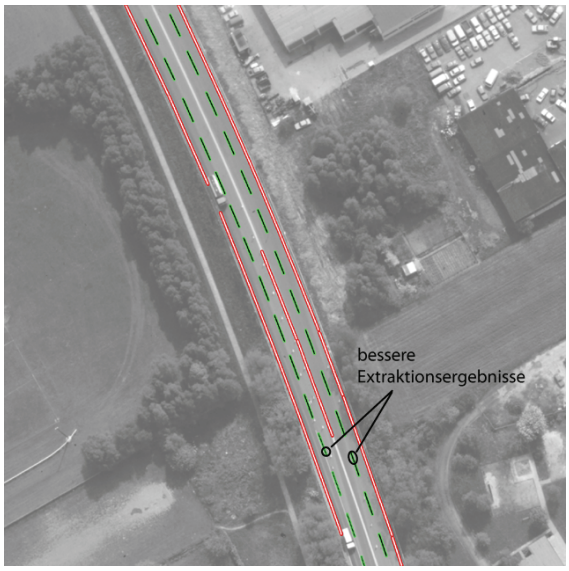


Bild 1 – Objektmodell 2

Bild 2 – Objektmodell 1

Bild 2 – Objektmodell 2

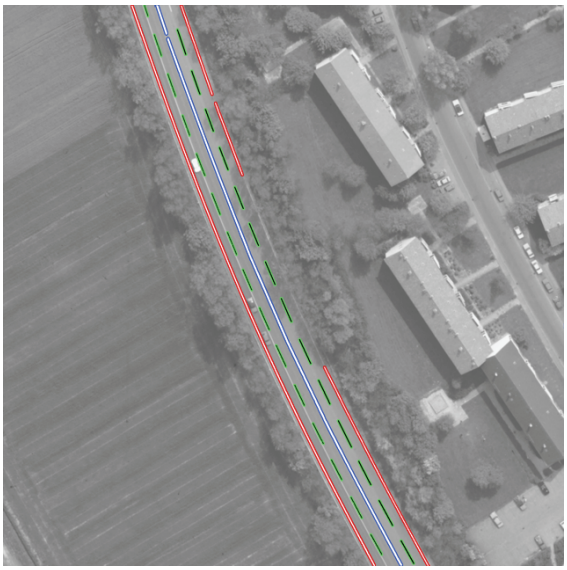


Bild 3 – Objektmodell 1

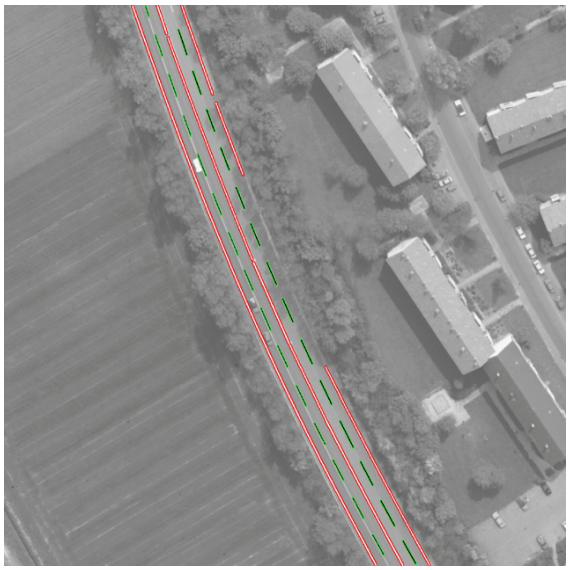


Bild 3 – Objektmodell 2

Abbildung 4.39: Extraktionsergebnisse für Bild 1–3 in 0.1m /Pixel (weiß/rot: Fahrbahnlinien, weiß/blau: Mittellinie, grün/schwarz: Leitlinien)

Im Bild 1 tritt für die Extraktion der Leitlinien im unteren Bildbereich der Effekt der besseren Extrahierbarkeit aus einer niedrigeren Bildauflösung auf (in Abbildung 4.39 gesondert gekennzeichnet). In der hohen Bildauflösung erscheint ein Teil der Leitlinien durch Störungen auf der Fahrbahn unterbrochen, so dass die Referenz an zwei Stellen zu kurze Teilstücke der Leitlinien aufweist. In der Zielauflösung 0.1 m/Pixel sind die kleinen Störungen auf der Fahrbahn geglättet, so dass eine komplette Extraktion der betreffenden Leitlinienteile mit dem angepassten Operator möglich ist. Die Vollständigkeit der Leitlinie beträgt im Vergleich zur Referenz 100%. An den zusätzlich extrahierten Stellen treten im Vergleich zur Referenz streng genommen False Positives auf, die jedoch in der durchgeführten Verifikation als korrekt und damit trotz ihrer Abweichung von der Referenz als True Positives angesehen werden (siehe Abschnitt 4.3.1).

Extraktionsergebnisse Einfache Straße A_{z2}=1.00 m/Pixel

Für die zweite Zielauflösung A_{z2}=1.00 m wurden aufgrund unterschiedlicher Skalenraumereignisse in der minimalen und maximalen Attributkonfiguration wieder zwei Objektmodelle prädiziert (dargestellt in Abbildung 4.8). Für beide Konfigurationen wurden die Auslöschungen der Leitlinien links und rechts prädiziert. Für das Modell 1 konnten im Bild 2 keine Objektteile extrahiert werden, so dass die Vollständigkeit 0 % beträgt. Obwohl Linien an der passenden Stelle im Bild mit dem angepassten Linienoperator für die Mittellinie extrahiert werden konnten, stimmten die Eigenschaften (insbesondere der Grauwert) dieser Linien nicht mit den prädizierten Attributwerten in der Zielauflösung überein. Sie wurden daher im Bottom-Up-Schritt der Bildanalyse nicht als gültige Objektteile anerkannt. Die Extraktionsergebnisse der einzelnen Objektteile der beiden angepassten Objektmodelle auf den drei Testbildern zeigt Abbildung 4.40.

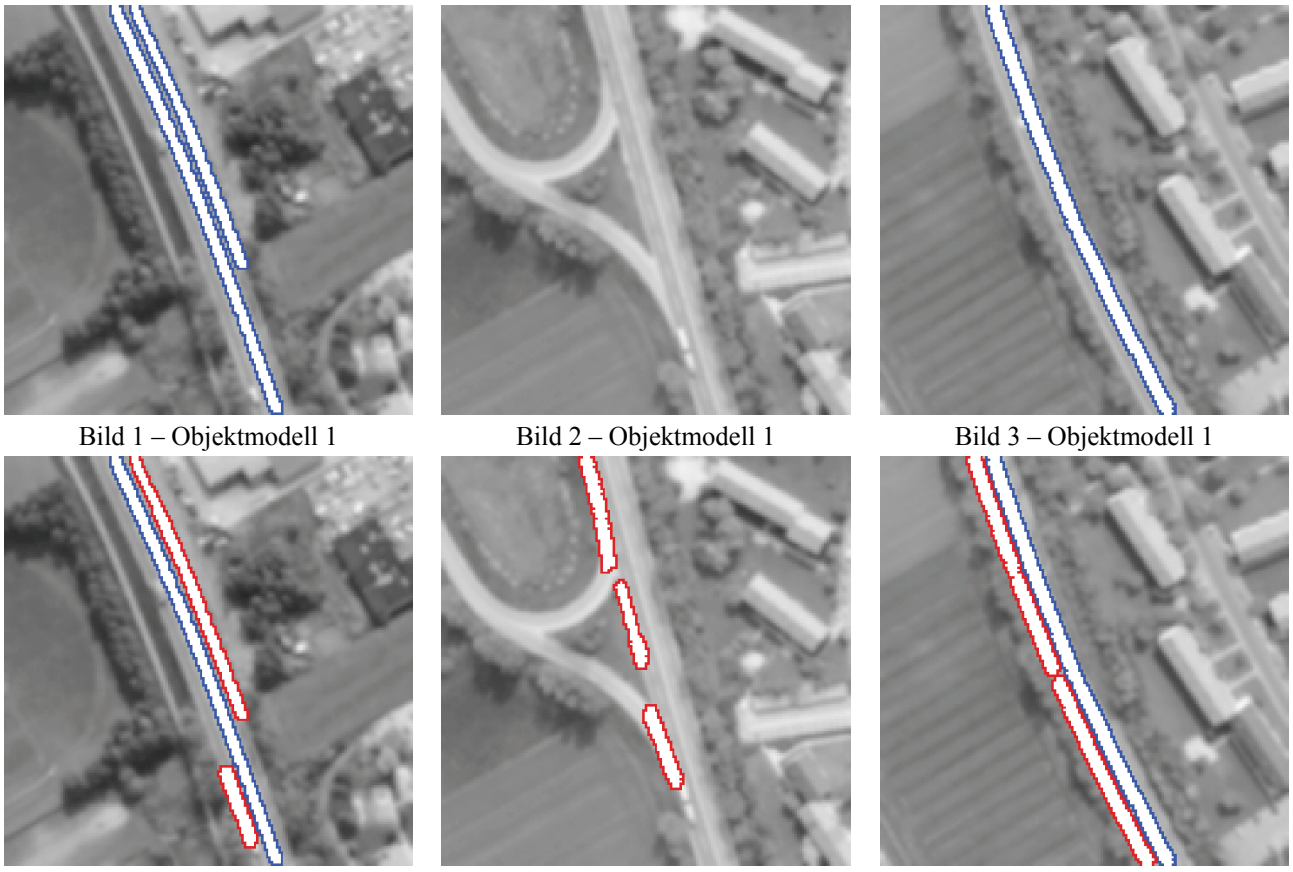


Abbildung 4.40: Extraktionsergebnisse für Bild 1–3 in 1.00 m/Pixel (weiß/rot: Fahrbahnlinien, weiß/blau: Mittellinie)

Extraktionsergebnisse Einfache Straße A_{z3}=5.00 m/Pixel

Die dritte Zielauflösung A_{z3}=5.00 m/Pixel für das Beispiel 1 stellt eine sehr große Auflösungsänderung (Auflösungsfaktor F_A=150) im Vergleich zur Ursprungsauflösung dar. Das automatisch angepasste Objektmodell (Abbildung 4.9) für diese Bildauflösung besteht nur noch aus einem einzigen Objektteil auf der obersten Hierarchieebene – der Fahrbahn des Straßensegments selbst. Aufgrund dieser geringen Evidenz ist eine relativ hohe Zahl von False Positives im Extrak-

tionsergebnis aus den Testbildern (vgl. Abbildung 4.41) zu erwarten, da die Verwechslungsgefahr mit anderen lang gestreckten Objekten mit ähnlichen radiometrischen Eigenschaften und ähnlicher Linienbreite sehr groß ist. Auch die Vollständigkeit der extrahierten Objekte ist niedrig, da die Umgebung des gesuchten Objektes bei dieser großen Auflösungsänderung einen extrem starken Einfluss auf die Extraktion nimmt. Dieser Effekt führt insbesondere im Bild 2 dazu, dass kein einziges Objekt extrahiert werden kann (d. h. die Vollständigkeit beträgt 0 %). In diesem Bild wurden die Fahrbahnen zwar vom angepassten Linienoperator extrahiert, jedoch in der Bewertung aufgrund ihrer Eigenschaften abgelehnt, da sämtliche Linien einen höheren Grauwert als im Modell (80–108) prädiziert aufwiesen. Es ist anzunehmen, dass die relativ hellen Gebäude in der Nähe der Straße bei der Überführung in die niedrige Bildauflösung zu erhöhten Intensitätswerten im Straßenbereich beitrugen.

Bild 1 Bild 2 Bild 3

Abbildung 4.41: Extraktionsergebnisse für Bild 1–3 in 5.00 m /Pixel (weiß/blau: Fahrbahn)

Qualitätsmaße

Aus den Extraktionsergebnissen, die mit den angepassten Objektmodellen in den drei Zielauflösungen erzielt wurden, werden die Maße Vollständigkeit, Korrektheit und Qualität abgeleitet. Als Referenz (100 % Vollständigkeit, Korrektheit und Qualität) dient das Extraktionsergebnis aus der Ursprungsauflösung A_0=0.03 m/Pixel. Die ermittelten Werte für Vollständigkeit, Korrektheit und Qualität in den drei Zielauflösungen im Vergleich zur Referenz zeigt Tabelle 4.1. Korrektheit und Qualität in Bild 2 für 5.00 m/Pixel und 1.00 m/Pixel wurden nicht berechnet, da keine True Positives für dieses Bild vorliegen. Die Ergebnisse der Qualitätsmaße werden für sämtliche Beispiele in Abschnitt 4.4 diskutiert.

Tabelle 4.1: Qualitätsmaße der Extraktionsergebnisse der Objektteile für Beispiel 1

		0.10 m/Pixel		**1.00 m/Pixel**		**5.00 m/Pixel**
		Modell 1	**Modell 2**	**Modell 1**	**Modell 2**	
Bild 1	Vollständigkeit	100 %	83 %	100 %	63%	23 %
	Korrektheit	100 %	100 %	61%	100%	32 %
	Qualität	100 %	83 %	61 %	63 %	15 %
Bild 2	Vollständigkeit	92 %	80 %	0 %	20 %	0 %
	Korrektheit	100 %	100 %	–	79 %	–
	Qualität	92 %	80 %	–	19 %	–
Bild 3	Vollständigkeit	100 %	100 %	100 %	76 %	100 %
	Korrektheit	100 %	100 %	100 %	100 %	80 %
	Qualität	100 %	100 %	100 %	76 %	80 %
Bild 1-3 gesamt	Vollständigkeit	97 %	87 %	67 %	51 %	41 %
	Korrektheit	100 %	100 %	77 %	97 %	62 %
	Qualität	97 %	87 %	56 %	50 %	33 %

4.3.3.2 Beispiel 2 – Kreuzungsarm

Extraktionsergebnis Referenz 0.03 m/Pixel

Die Referenz für die Verifikation des Beispiel 2 (Straßenarm im Kreuzungsbereich) entspricht dem Extraktionsergebnis des gegebenen Objektmodells aus A_0=0.03 m/Pixel (Abbildung 4.11). Das Ergebnis der Extraktion des modellierten Kreuzungsarmes mit den einzelnen Objektteilen aus Testbild 4 zeigt Abbildung 4.42. Alle Objektteile des Modells konnten im Testbild 4 verifiziert werden.

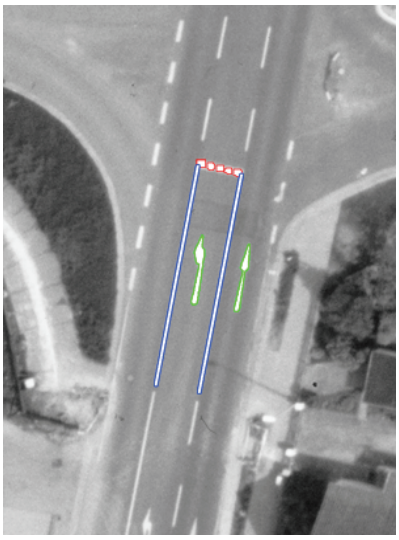


Abbildung 4.42: Extraktionsergebnisse für Bild 4 in 0.03 m/Pixel – Referenz (weiß/rot: Haltelinie, weiß/blau: Mittellinie und Leitlinie, weiß/grün: Richtungspfeile)

Extraktionsergebnisse Kreuzungsarm A_{z1}=0.10 m/Pixel

Die Extraktionsergebnisse der Operatoren sind in Abbildung 4.43 dargestellt.

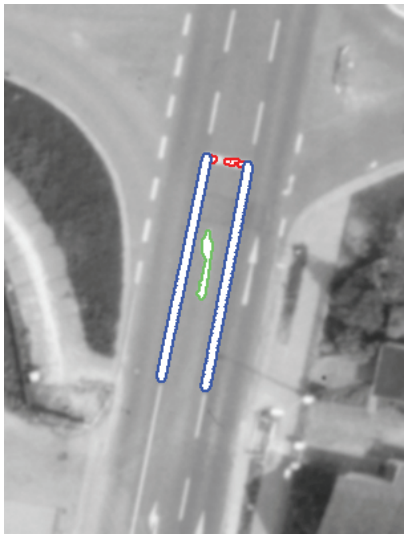


Abbildung 4.43: Extraktionsergebnisse für Bild 4 in 0.10 m/Pixel (weiß/rot: Haltelinie, weiß/blau: Mittellinie und Leitlinie, weiß/grün: Richtungspfeile)

Die Einzelteile der Haltelinie konnten mit dem angepassten Operator im Bild der Zielauflösung 0.1 m/Pixel erfolgreich extrahiert werden. Als Eingabeparameter für den Größenbereich der Haltelinienteile und deren Abstand wurden die automatisch angepassten Attributwerte verwendet. Da der Parameter der Grauwert-Toleranz für das Flächenwachstumsverfahren keinem der prädizierten Attribute entspricht somit keine Informationen über sein Skalenverhalten vorliegen, wurde dieser Wert dagegen festgehalten, d. h. für ihn wurde der Wert der Grauwert-Toleranz des Operators der

Ursprungsauflösung (2) beibehalten. Der Richtungspfeil „links“ wurde mit dem angepassten Template aus der automatischen Anpassung ebenfalls erfolgreich extrahiert. Der Richtungspfeil „geradeaus“ wurde dagegen nicht erkannt. Da somit ein Objektteil bei der Bewertung in der Extraktion fehlt, wird auch das gesamte Objekt Kreuzungsarm nicht verifiziert.

Extraktionsergebnisse Kreuzungsarm A_{z2}=0.20 m/Pixel

Wie in der ersten Zielauflösung konnten auch in der zweiten Zielauflösung 0.20 m/Pixel sämtliche Objektteile korrekt extrahiert werden, mit Ausnahme des Richtungspfeils „geradeaus“ (Abbildung 4.44). Das Objekt Kreuzungsarm wurde demnach auch in dieser Zielauflösung nicht verifiziert. Die automatisch angepassten Operatoren waren jedoch zum größten Teil erfolgreich.

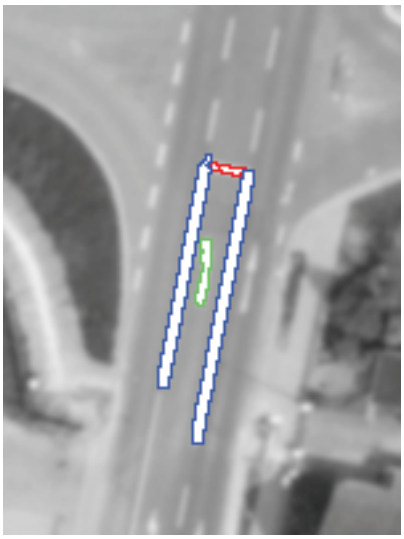


Abbildung 4.44: Extraktionsergebnisse für Bild 4 in 0.20 m/Pixel (weiß/rot: Haltelinie, weiß/blau: Mittellinie und Leitlinie, weiß/grün: Richtungspfeile)

Extraktionsergebnisse Kreuzungsarm A_{z3}=1.00 m/Pixel

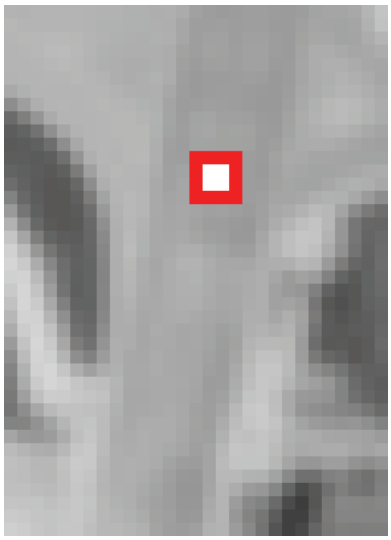


Abbildung 4.45: Extraktionsergebnisse für Bild 4 in 1.00 m/Pixel (weiß/rot: Haltelinie)

Das Verfahren zur automatischen Anpassung ergab für die Zielauflösung von 1.00 m/Pixel zwei Objektmodelle, die das mögliche Erscheinungsbild des Objektes Kreuzungsarm widerspiegeln. Die beiden Modelle unterscheiden sich in der Anzahl ihrer Objektteile. Während die Haltelinie sowie die beiden Richtungspfeile Objekteile in beiden Modellen sind, beinhaltet nur das Objektmodell, das aus der maximalen Attributkonfiguration resultiert, die Mittellinie und die Leitlinie. Die Teilstücke der Haltelinie sind in beiden Objektmodellen zu einer einzelnen Linie verschmolzen, die mit dem

angepassten Operator aus dem Testbild für den Kreuzungsbereich extrahiert werden konnte. Auf dem verwendeten Testbild konnten jedoch weder die beiden linienartigen Objektteile (Mittellinie und Leitlinie) noch die Richtungspfeile erkannt werden (Abbildung 4.45). Die Musterbilder für das Template-Matching der Richtungspfeile konnten zwar in der Analyse-durch-Synthese aus den synthetischen Bildern der Zielauflösung entnommen werden, besitzen jedoch kaum noch Evidenz. Es ist auch visuell in dieser groben Zielauflösung im Testbild nicht mehr möglich, die Richtungspfeile zu erkennen. Eine Extraktion der Richtungspfeile aus dieser Bildauflösung ist daher unmöglich. Für die Richtungspfeile hätte daher in der automatischen Anpassung eine effektive Auslöschung prädiziert werden müssen.

Qualitätsmaße

Die Qualitätsmaße für die erzielten Extraktionsergebnisse der Objektteile aus den drei Zielauflösungen des Testbildes 4 für den Kreuzungsarm sind in Tabelle 4.2 zusammengefasst. Aufgrund der großen Auflösungsänderung in der niedrigsten Zielauflösung A_{z3}=1.00 m/Pixel im Vergleich zur Größe der Objektteile, ist die erreichte Vollständigkeit in dieser Auflösung sehr gering. Dagegen ist die Korrektheit in allen Zielauflösungen sehr hoch. Dieses gute Ergebnis ist insbesondere auf die charakteristischen Formen der Objektteile des Kreuzungsarmes zurückzuführen, die sich deutlich von anderen Objekten im Bild unterscheiden.

Tabelle 4.2: Qualitätsmaße der Extraktionsergebnisse der Objektteile für Beispiel 2

		0.10 m/Pixel	**0.20 m/Pixel**	**1.00 m/Pixel**	
				Modell 1	**Modell 2**
Bild 4	Vollständigkeit	89 %	89 %	8 %	28 %
	Korrektheit	99 %	98 %	100 %	100 %
	Qualität	89 %	89 %	8 %	28 %

4.3.3.3 Beispiel 3 – Lokale Kontextobjekte

Extraktionsergebnis Referenz 0.03 m/Pixel

Die Referenz für die Verifikation von Beispiel 3 mit dem gegebenen Objektmodell in A_0=0.03 m/Pixel (Abbildung 4.23 für das Primärobjekt Straße und Abbildung 4.24 für das lokale Kontextobjekt Fahrzeug) ist in Abbildung 4.46 dargestellt. Die Extraktion der linken Leitlinie als Objektteil der einfachen Straße ist durch das Fahrzeug über die Kontextrelation *Optional occlusion* gestört. Um die Stelle der Störung wird eine Kontextregion gebildet, in der die Extraktion des Fahrzeugs ausgeführt wird. Aus Testbild 5 konnte in der Kontextregion mit dem gegebenen Objektmodell ein Fahrzeug erfolgreich extrahiert werden. Damit wird die in der Straßenextraktion entstandene Lücke durch das Kontextobjekt gefüllt, so dass auch die Straßenextraktion im Testbild als erfolgreich bewertet wird.

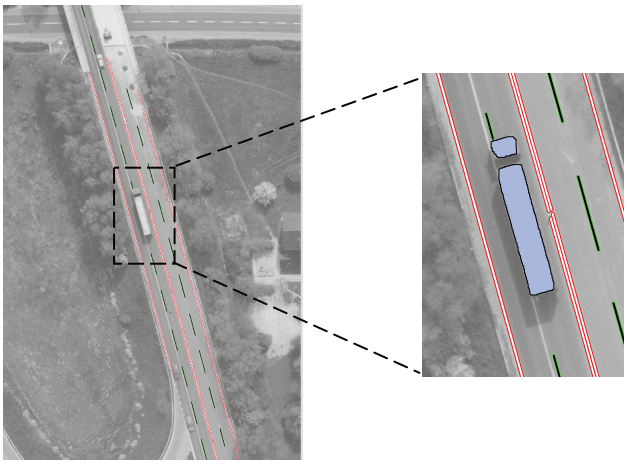


Abbildung 4.46: Extraktionsergebnisse für Bild 5 in 0.03 m/Pixel – Referenz (weiß/rot: Fahrbahnlinie, schwarz/grün: Leitlinie, blau/schwarz: Fahrzeug)

Extraktionsergebnisse Kontextobjekte A_{z1}=0.10 m/Pixel

Aus der Zielauflösung 0.10 m/Pixel ist die Extraktion der linken Leitlinie wie auch in der Referenz gestört. Die entstandene Kontextlücke kann durch die erfolgreiche Extraktion eines Fahrzeuges mit dem angepassten Objektmodell gefüllt werden. Alle Objektteile des Primärobjektes Straße werden damit als erkannt bewertet. Die Extraktionsergebnisse der einzelnen Objektteile des Primärobjektes und des Kontextobjektes aus dem Kontextbild mit den automatisch angepassten Operatoren zeigt Abbildung 4.47.

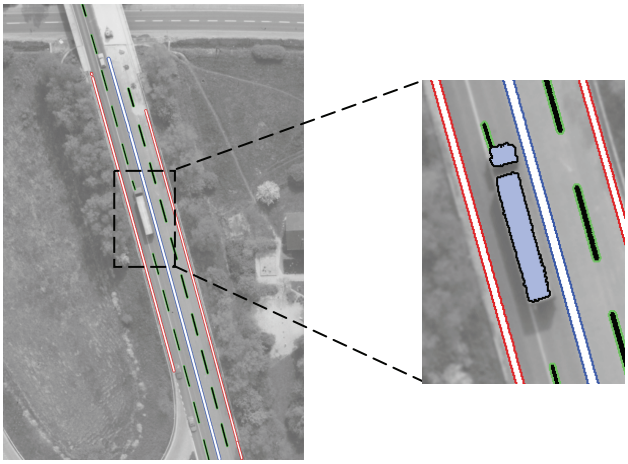
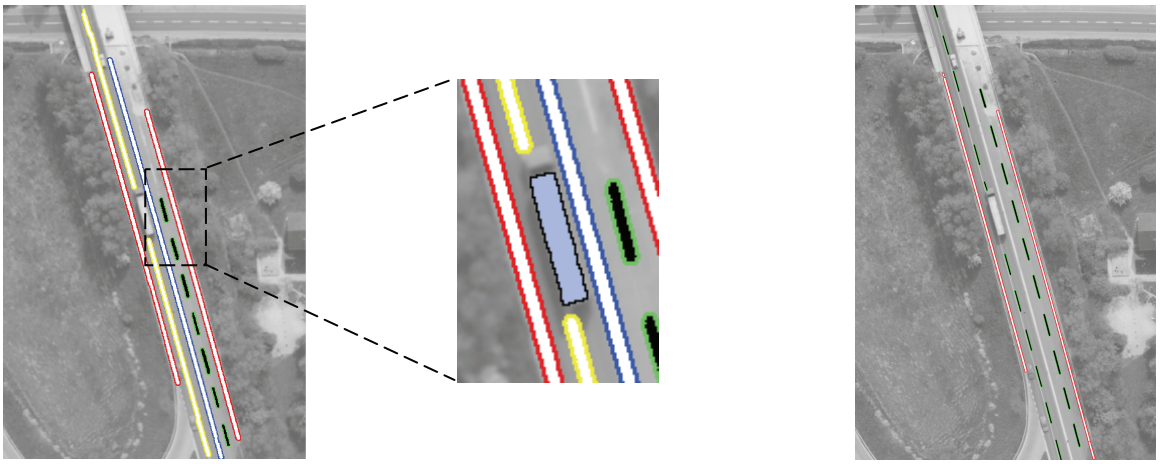


Abbildung 4.47: Extraktionsergebnisse für Bild 5 in 0.1 m/Pixel (weiß/rot: Fahrbahnlinie, schwarz/grün: Leitlinie, blau/schwarz: Fahrzeug)

Extraktionsergebnisse Kontextobjekte A_{z2}=0.20 m/Pixel



Modell 1 Modell 2

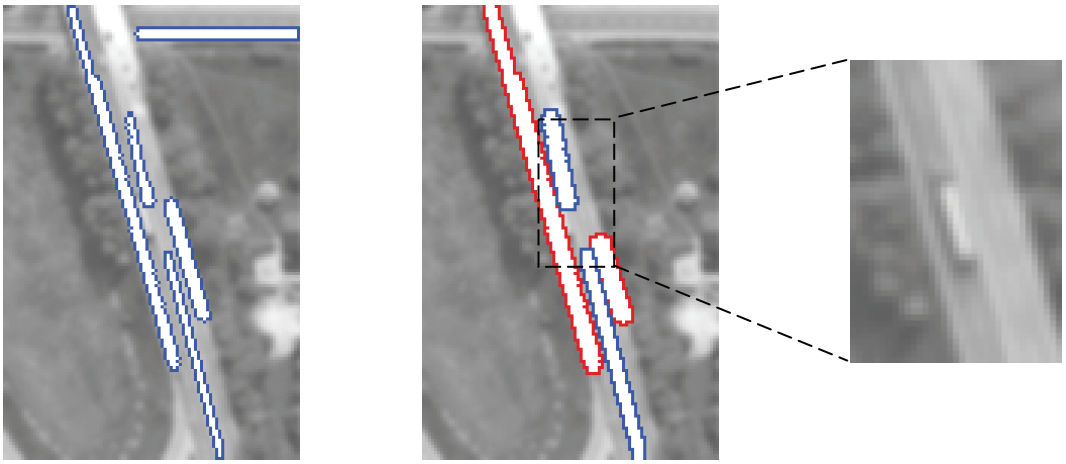
Abbildung 4.48: Extraktionsergebnisse für Bild 5 in 0.2 m/Pixel (weiß/rot: Fahrbahnlinie, schwarz/grün: Leitlinie, weiß/blau: Mittellinie, weiß/gelb: Baustellenmarkierung, hellblau/schwarz: Fahrzeug)

Aus dieser Bildauflösung ist die Extraktion der linken Leitlinie nicht mehr möglich, da sie mit der angrenzenden Baustellenmarkierung verschmolzen ist, die als nicht permanente Markierung im Modell nicht enthalten und somit auch in der Analyse des Skalenverhaltens nicht berücksichtigt ist. Die Extraktion dieser Baustellenmarkierung kann jedoch mit dem Operator für kontinuierliche Fahrbahnlinien durchgeführt werden. Das Ergebnis in Abbildung 4.48 zeigt, dass ihre Extraktion durch das Fahrzeug ebenfalls gestört ist. Die dabei entstandene Kontextlücke kann jedoch durch die Extraktion mit dem angepassten Operator für Fahrzeug nicht gefüllt werden, da die Front des Fahrzeugs nicht extrahiert werden konnte. Der vordere Teil der Front bildet bei der Segmentierung der Fahrzeugteile mit dem Flächenwachstumsverfahren zusammen mit der angrenzenden Baustellenmarkierung ein gemeinsames Segment. Bei der anschließenden Bewertung nach Form und Größe wird es nicht als Fahrzeugteil ausgewählt, da seine Eigenschaften nicht den vorgegebenen Attributen der Fahrzeugfront entsprechen. In diesem Beispiel tritt sehr deutlich der Einfluss der Umgebung der

Objektteile im Bild hervor. Einerseits ist die Extraktion der periodischen Leitlinie nicht möglich, da diese in der Zielauflösung 0.20 m/Pixel mit der nahen Baustellenmarkierung verschmolzen ist. Andererseits versagt auch die Extraktion eines Fahrzeugteils, da es im Bild durch das angewendete Segmentierungsverfahren nicht von der angrenzenden Markierung, die es verdeckt, zu unterscheiden ist.

Extraktionsergebnisse Kontextobjekte A_{z3}=1.00 m/Pixel

In der Zielauflösung von 1.00 m/Pixel ist keine Extraktion der linken Leitlinie oder Baustellenmarkierung möglich. Die Auslöschung der Leitlinie wurde für das Objektmodell der Straße prädiziert (siehe Abbildung 4.8). An der Stelle des Fahrzeugs ist daher keine Extraktion eines Objektteils der Straße gestört. Zur Verifikation des angepassten Fahrzeug-Modells wurde trotzdem eine Extraktion des Fahrzeugs mit dem angepassten Objektmodell vorgenommen. Auch vom Fahrzeug konnte allerdings kein Objektteil im Testbild 5 in der Zielauflösung 1.00 m extrahiert werden, da die Segmentierung keine rechteckähnlichen Ergebnisse im Bereich des Fahrzeugs lieferte. Die Extraktionsergebnisse aus dem Testbild 5 für die einzelnen Objektteile der Straße und des Fahrzeugs sind in Abbildung 4.49 dargestellt.



Modell 1 Modell 2

Abbildung 4.49: Extraktionsergebnisse für Bild 5 in 1.00 m/Pixel (weiß/rot: Fahrbahnlinie, weiß/blau: Mittellinie)

Qualitätsmaße

Während Vollständigkeit, Korrektheit und Qualität in der ersten Zielauflösung (0.10 m/Pixel) für eines der angepassten Objektmodelle 100 % beträgt, sinken die Werte der Qualitätsmaße für das Primärobjekt in den untersuchten niedrigeren Bildauflösungen auf bis zu 69 % (siehe Tabelle 4.3). Grund für das Absinken der Vollständigkeit des Primärobjektes ist vor allem das Versagen des Operators zur Extraktion der linken Leitlinie. Diese verschmilzt in der zweiten Zielauflösung (0.20 m/Pixel) im Testbild mit einer Baustellenmarkierung, die nicht im Modell enthalten ist. Diese Verschmelzung kann daher nicht prädiziert werden. Auch das lokale Kontextobjekt Fahrzeug kann in der zweiten Zielauflösung nicht vollständig extrahiert werden. Hier ist die nahe liegende Fahrbahnmarkierung, die vom Fahrzeug verdeckt wird, mit der Fahrzeug-Front verschmolzen. Dieses Skalenraumereignis kann mit dem gewählten separaten Ansatz nicht prädiziert werden, da dieser die Umgebung des lokalen Kontextobjektes nicht berücksichtigt und stattdessen einen homogenen Hintergrund voraussetzt. In der dritten Zielauflösung (1.00 m/Pixel) ist die Extraktion keines der Fahrzeugteile mit dem verwendeten Operator möglich. Aufgrund der starken Glättung der Kanten der Fahrzeugteile ist die Segmentierung mit den Flächenwachstumsverfahren nicht erfolgreich. Auch die Extraktion der gestrichelten Leitlinien des Primärobjektes Straße, die das Kontextobjekt Fahrzeug in der hohen Bildauflösung verdeckt, ist in dieser Zielauflösung nicht mehr möglich (siehe auch Beispiel 1). Allerdings ist eine Lücke der Extraktion der Mittellinie an der Stelle des Fahrzeugs zu beobachten, die durch eine erfolgreiche Extraktion des Fahrzeugs geschlossen werden könnte.

Tabelle 4.3: Qualitätsmaße der Extraktionsergebnisse der Objektteile für Beispiel 3

		0.10 m/Pixel		0.20 m/Pixel	1.00 m/Pixel	
		Modell 1	**Modell 2**		**Modell 1**	**Modell 2**
Bild 5 - Primärobjekt	Vollständigkeit	100 %	52 %	69 %	78 %	67 %
	Korrektheit	100 %	100 %	100 %	34 %	100 %
	Qualität	100 %	52 %	69 %	31 %	67 %
Bild 5 - Kontextobjekt	Vollständigkeit	100 %		90 %	0 %	
	Korrektheit	100 %		100 %	–	
	Qualität	100 %		90 %	–	
Bild 5 gesamt	Vollständigkeit	100 %	53 %	70 %	70 %	65 %
	Korrektheit	100 %	100 %	100 %	32 %	100 %
	Qualität	100 %	53 %	70 %	29 %	65 %

4.4 Fazit

In diesem Kapitel wurde eine Verifikation des entwickelten Verfahrens zu automatischen Anpassung von Bildanalyse-Objektmodellen durchgeführt. Anhand von drei Beispielen für Objektmodelle unterschiedlicher Typen wurden zuerst die jeweiligen Methoden zur automatischen Anpassung dieser Objektmodelle an drei verschiedene niedrigere Bildauflösungen demonstriert. Anschließend wurden diese angepassten Modelle einer Qualitätsbewertung zur Verifikation des, in dieser Arbeit entwickelten, Verfahrens unterzogen. Um eine umfassende Bewertung der Methoden vornehmen zu können, reichte der Umfang der Auflösungsänderung bei den Anpassungsbeispielen bis zum 150fachen der Ursprungsauflösung.

Mit den meisten der automatisch angepassten Objektmodelle konnte eine Objektextraktion aus den korrespondierenden Bilddaten erfolgreich durchgeführt werden. Die Korrektheit dieser Extraktion ist hoch, so lange genügend Evidenz durch viele Objektteile im Bild vorhanden ist, die für ein vollständiges Objekt vorhanden sein müssen. Bei sinkender Evidenz sinkt jedoch auch die Korrektheit der Extraktionsergebnisse. Die Evidenz nimmt normalerweise bei niedrigerer Bildauflösung im Zuge der Abstraktion ab, da aufgrund von Auslöschungen und Verschmelzungen immer weniger Objektteile extrahierbar sind. Die Vollständigkeit ist stark vom jeweiligen Bildmaterial abhängig. Durch das Vorliegen von Kontextobjekten und geringem Kontrast kann die Extraktion erheblich beeinträchtigt sein. Wie die Verifikation von Beispiel 1 zeigt (siehe Tabelle 4.1), kann vor allem Schattenwurf in niedrigerer Bildauflösung einen größeren Einfluss auf die Extraktionsfähigkeit des Objektes haben. Die Qualität des angepassten Objektmodells bzw. die Extraktion des Landschaftsobjektes mit diesem aus der Zielauflösung wird naturgemäß auch von der Güte des gegebenen Modells für die Ursprungsauflösung bestimmt. Wenn die Modellierung in der hohen Auflösung unzureichend oder nicht detailliert genug ist, wird auch in der Zielauflösung kein gutes Extraktionsergebnis erzielt werden.

Als besonders kritischer Faktor für den Extraktionserfolg stellte sich die Anpassung der Operatoren der einzelnen Objektteile an die niedrigere Bildauflösung heraus. Die Parameter eines Operators sind in der Regel auf die Darstellung des zu extrahierenden Objektteils in einer bestimmten Bildauflösung optimiert. Wenn für die Operatoren keine geeigneten Eingabeparameter in der automatischen Anpassung prädiziert werden können, schlägt die Extraktion bei einer großen Auflösungsänderung fehl. Je größer die Auflösungsänderung, desto eher versagt der Operator, dessen Parameter nicht geeignet angepasst wurden, da sich das Erscheinungsbild der Objekte in den Bilddaten mit gröber werdender Bildauflösung immer stärker ändert. Allerdings sind nicht alle Operatoren für eine Anpassung an eine niedrigere Bildauflösung geeignet. Stimmen ihre Eingabeparameter nicht mit den Attributen überein, die das Objektteil im Modell beschreiben, sind diese Operatoren mit dem entwickelten Verfahren nicht an eine niedrigere Bildauflösung anpassbar. Bei sehr großer Auflösungsänderung kann sich auch der Objekttyp ändern, so dass ein anderer Operatortyp zur Extraktion dieses Objektteils nötig ist.

Bei der Verifikation von Beispiel 2 zeigte sich, dass die Anpassung des Operators, der mit Template-Matching arbeitet, für die als flächenhafte Muster modellierten Richtungspfeile, über eine zu große Auflösungsänderung durchgeführt wurde. An dieser Stelle ist die Methode zur Prädiktion effektiver Auslöschungen von Mustern noch zu verbessern. Im derzeitigen Verfahren ist eine Prädiktion der effektiven Auslöschung für Objektteile, die mit Template-Matching extrahiert werden, nicht möglich, da die angepassten Musterbilder als Eingabeparameter für den Operator aus dem synthetischen Bild selbst entnommen werden. Ein zuverlässiger Test dieses Operators auf dem synthetischen Bild der Zielauflösung (wie für die anderen Operatoren) ist daher nicht möglich. Daher sollte ein anderes, besser geeignetes, Kriterium für diesen Operatortyp gefunden werden.

Ein Problem bei der Objektextraktion mit den automatisch angepassten Modellen aus realen Bilddaten ist der steigende Einfluss der Umgebung des modellierten Landschaftsobjektes mit gröber werdender Bildauflösung. Beim Übergang auf eine niedrigere Bildauflösung wird durch die Tiefpassfilterung auch die Umgebung eines Pixels in die Berechnung des resultierenden Grauwertes in der niedrigeren Bildauflösung herangezogen. Je gröber die angestrebte Zielauflösung ist, desto größer ist auch der für eine angemessene Glättung benötigte Tiefpassfilter. Ein größerer Filter bedeutet jedoch die Einbeziehung einer größeren Umgebung in das Ergebnis der Filterung. Der Einfluss der Umgebung auf das Skalenverhalten des modellierten Landschaftsobjektes steigt daher mit zunehmender Auflösungsänderung. In den gegebenen Modellen der hohen Bildauflösung ist die Umgebung jedoch häufig nicht modelliert. Aufgrund dieser Unkenntnis geht das vorgestellte Verfahren der automatischen Anpassung daher bei der Simulation der Objektszene von einer homogenen Umgebung aus. Bei einer homogenen Umgebung des Objektes bleibt der Einfluss auf das Aussehen des Objektes im niedriger aufgelösten Bild (vor allem der Grauwerte) gering. Je heterogener die Umgebung jedoch ist, desto schlechter sind die Resultate aus der Anpassung, wenn die Umgebung nicht modelliert und daher eine Prädiktion des gemeinsamen Skalenverhaltens des Objektes mit der Umgebung unmöglich ist. Der Einfluss der jeweiligen Umgebung eines Landschaftsobjektes in unterschiedlich homogenen Bildszenen wird besonders in der Verifikation von Beispiel 1 im Testbild 1–3 deutlich. Im Testbild 2 resultiert aufgrund der Häuser in der Umgebung für die zu extrahierende Straße ein höherer Grauwert, als aus der Analyse des Skalenverhaltens prädiziert worden ist. Ähnliches ist beim lokalen Kontextobjekt auf Testbild 5 zu beobachten, bei dem ein Teil des Kontextobjektes mit einer Fahrbahnmarkierung verschmilzt. Diese Verschmelzung wurde nicht prädiziert, da in der gewählten separaten Anpassung der Objektmodelle für das Primärobjekt und das lokale Kontextobjekt eine gemeinsame Analyse des Skalenverhaltens nicht erfolgt.

5 Schlussbetrachtungen und Ausblick

In dieser Arbeit wird erstmals ein Verfahren zur automatischen Anpassung von Objektmodellen zur wissensbasierten Bildanalyse, die als Konzeptnetze formuliert sind, an eine niedrigere Bildauflösung vorgestellt. Das Ergebnis sind ein oder mehrere für die niedrigere Auflösung geeignete Modelle. Somit kann die bisher für jede Bildauflösung notwendige manuelle Erstellung durch ein automatisches Verfahren ersetzt werden.

Objektmodelle können sich in ihrer Zusammensetzung hinsichtlich der beinhalteten Objekttypen unterscheiden und daher unterschiedliche Methoden zur automatischen Anpassung erfordern. Für die folgenden Arten von Objektmodellen wurden in dieser Arbeit Methoden entwickelt: Modelle mit parallelen linienartigen Objektteilen (Abschnitt 3.3.1), Modelle mit parallelen linienartigen Objektteilen variabler Anzahl (Abschnitt 3.3.2), Modelle mit flächenhaften beliebig orientierten Objektteilen (Abschnitt 3.3.3) und Modelle mit lokalen Kontextobjekten (Abschnitt 3.3.4). Die automatische Anpassung an eine niedrigere Bildauflösung ist allerdings nicht für beliebige Objektmodelle möglich. Je konkreter die Beschreibung des Objektes bzw. dessen Umgebung ist, desto eindeutiger lässt sich das Skalenverhalten des Objektes vorhersagen. Dagegen führt insbesondere eine Variabilität der Objektteile im Modell bzgl. relativer Position, Kontrast und Ausdehnung der Objektteile zu unterschiedlichem Skalenverhalten des modellierten Landschaftsobjektes. Dann ist keine eindeutige Prädiktion des Skalenverhaltens und somit keine Anpassung an ein einziges für die Zielauflösung geeignetes Modell möglich. Stattdessen resultieren, möglicherweise sehr viele, unterschiedliche Modelle für die Zielauflösung. Eine Extraktion des gesuchten Landschaftsobjektes mit vielen Modellen ist jedoch nicht praktikabel. Die begrenzte Variabilität der anpassbaren Modelle steht im Widerspruch zu einigen realen Modellen. Für eine erfolgreiche Extraktion eines Landschaftsobjektes aus realen Bilddaten müssen die Bildanalyse-Modelle sehr flexibel hinsichtlich des Kontrastes gestaltet sein, da die Landschaftsobjekte sehr variabel aufgrund unterschiedlicher Aufnahmeverhältnisse in den Bilddaten erscheinen können. Je flexibler die Modellierung jedoch ist, desto schwieriger ist die Prädiktion des Skalenverhaltens. Dieser Widerspruch bewirkt die wichtigste Einschränkung für den praktischen Einsatz des entwickelten Verfahrens.

Die automatische Anpassung wurde für drei Objektmodelle an einen großen Auflösungsbereich (bis zum 150fachen der ursprünglichen Bildauflösung) beispielhaft in Kapitel 4 demonstriert. Wie die anschließende Verifikation des entwickelten Verfahrens gezeigt hat, ist die Anpassung nicht für alle niedrigeren Bildauflösungen erfolgreich, sondern auf einen Auflösungsbereich begrenzt. Die automatisch angepassten Modelle erzielten für kleine Auflösungsänderungen (bis zum 10fachen der Ursprungsauflösung) sehr gute Ergebnisse. Danach verschlechtert sich die Qualität der Anpassung. Sehr große Auflösungsänderungen bedeuten in der Regel eine sehr starke Veränderung des Erscheinungsbildes des Landschaftsobjektes in den Bilddaten. Eine Ursache für die schlechter werdenden Ergebnisse der Anpassung bei zunehmender Auflösungsreduktion ist der steigende Einfluss der Umgebung der Bildszene. Da im Verfahren der automatischen Anpassung jedoch keine Information über die Umgebung verfügbar ist (so wie in den meisten Modellen für eine hohe Auflösung auch nicht), geht die Prädiktion des Skalenverhaltens des Objektes von einem homogenen Hintergrund aus. Wird diese Annahme in Bildszenen aus heterogenen Gebieten verletzt, werden die prädizierten Skalenraumereignisse der Objektteile und die Werte der Attribute nicht mit den tatsächlichen Gegebenheiten im Bild übereinstimmen.

Eine weitere Ursache für die Grenzen der Anpassung bei großer Auflösungsänderung sind Skalenraumereignisse, die mit zunehmender Auflösungsänderung häufiger auftreten. Evidenz aus der Extraktion von verschiedenen Objektteilen geht durch Auslöschungen und Verschmelzungen in gröber werdenden Bilddaten verloren. Dieser Verlust der Evidenz wird bei einer manuellen Erstellung von Modellen üblicherweise durch die Modellierung anderer Objekteigenschaften kompensiert, die jedoch erst in einer groben Auflösung nutzbar werden, wie beispielsweise die Netzwerkeigenschaft von Straßen. Für eine Berücksichtigung dieser Charakteristika bzw. deren Extraktionsmethoden auf der Grundlage der vorgestellten Methode müssen diese bereits im gegebenen Objektmodell der hohen Auflösung enthalten sein, denn sie können nicht im automatischen Verfahren erlernt werden. Weiterhin können durch Skalenraumereignisse neue Objekttypen entstehen, für die keine Operatoren im Modell der Ursprungsauflösung gegeben sind.

Auch der automatischen Anpassung der Operatorparameter für die Extraktion der jeweiligen Objektteile sind bei einer großen Auflösungsänderung Grenzen gesetzt. Wenn sich der Objekttyp in der groben Bildauflösung von dem in der hohen Bildauflösung unterscheidet, muss ein anderer geeigneter Operator die Extraktion des neuen Objekttyps über-

nehmen. Die Anpassungsfähigkeit der Operatoren ist ein kritischer Punkt für den Erfolg der automatischen Anpassung, wie am Beispiel für das Objektmodell mit flächenhaften Objekttypen in Kapitel 4 demonstriert. Für die automatische Anpassbarkeit der Operatoren muss deren Funktionsweise bekannt sein und im Anpassungsverfahren berücksichtigt werden. Dies für sämtliche Merkmalsextraktions-Operatoren durchzuführen, ist jedoch unpraktikabel. Daher sollten die Operatoren diejenigen Parameter als Eingangsgrößen nutzen, die im Verfahren der Anpassung als Attribute der Knoten für die Zielauflösung prädiziert werden. Alternativ könnte zur Verbesserung der automatischen Anpassung der Operatorparameter der jeweilige Operator mit den angepassten Attributwerten des korrespondierenden Objektteils iterativ getestet werden. Falls der Operator keinen Erfolg hat, werden die prädizierten Parameterwerte um kleine Schritte verändert und der Operator wird mit diesen erneut getestet. Diese Iteration wird wiederholt, bis die Extraktion erfolgreich ist, d.h. ein Objektteil extrahiert wird. Dabei liegt die Annahme zugrunde, dass die prädizierten Attributwerte relativ nah an den optimalen Parametern liegen. Dieses Vorgehen kann bei Operatoren, deren Parameter nicht exakt den Attributen entsprechen, zu besseren Ergebnissen führen.

Das Verfahren der automatischen Anpassung ist nicht auf Modelle zur Extraktion von Straßen beschränkt. Wenn die Beschreibung der Objekte in der Modellierung genügend genau ist, sind Informationen über die wichtigsten Parameter zur Prädiktion des Skalenverhaltens – Distanzen, Objektgröße, Kontrast – der Objektteile bekannt. Dann ermöglichen die entwickelten Methoden zur Analyse des Skalenverhaltens eine Anwendung auf andere Landschaftsobjekte, beispielsweise Gebäude, die über eine strukturelle Bildinterpretation extrahiert und konform zu den formulierten Randbedingungen des automatischen Anpassungsverfahrens modelliert werden können. Die automatische Anpassung von anderen Landschaftsobjekten wurde jedoch im Rahmen der vorliegenden Arbeit nicht untersucht.

Grundsätzlich ist das Verfahren der Anpassung nicht auf Semantische Netze beschränkt, sondern auch die Anpassung anderer Arten von Modellen möglich, solange diese eine ausreichend detaillierte Beschreibung der Objekte und ihrer Topologie gewährleisten. Semantische Netze dienen im vorgestellten Verfahren ausschließlich der Formulierung der Information über die Objekte, die für die Analyse des Skalenverhaltens des Objektes verwendet wird. Die Prädiktion der Objektdarstellung in den Bilddaten niedrigerer Auflösung – der Kern des Verfahrens – kann auch mit Informationen über das Objekt durchgeführt werden, die aus einer anderen Form der Wissensrepräsentation entnommen wurde. Die Zerlegung als erste Stufe des Verfahrens müsste in diesem Fall für eine automatische Verabeitung der Informationen auf die neue Repräsentationsform abgestimmt werden.

In der automatischen Bildinterpretation von Luft- und Satellitenbildern werden in jüngster Zeit deterministische Modellierungsformen, zu denen auch Semantische Netze gehören, immer häufiger durch eine statistische Modellierung ersetzt. Statistische Modelle können im Gegensatz zu deterministischen Modelltypen die Unsicherheit von Objekten und ihren Relationen über Wahrscheinlichkeiten berücksichtigen. Damit können die Modelle zur Objektextraktion flexibler werden und sind dadurch auf mehr unterschiedliche Szenen anwendbar – ein wichtiger Vorteil bei realen Bildszenen, die je nach Umgebung und Aufnahmeverhältnisse sehr unterschiedlich erscheinen können. Der Nachteil von statistischen Modellen liegt jedoch in der Schwierigkeit ihrer Erzeugung – die Wahrscheinlichkeitsdichten der einzelnen Objektattribute und Relationen müssen oftmals erst aufwändig ermittelt werden. Für den in dieser Arbeit vorgestellten Ansatz zur Anpassung von Bildanalyse-Objektmodellen könnten die Wahrscheinlichkeitsdichten der Objekte eine wertvollen Beitrag zu der korrekten Prädiktion des Skalenverhaltens von lokalen Kontextobjekten leisten, deren Position im Konzeptnetz des Primärobjektes unbestimmt ist. Die Unsicherheit über die Position der Kontextobjekte könnte mit Wahrscheinlichkeiten modelliert werden. Die Kontextobjekte könnten dann mit ihren Wahrscheinlichkeiten durch den Skalenraum propagiert und an die Zielauflösung angepasst werden. Dies bedarf jedoch weiterer Forschung, insbesondere bezüglich der Prädiktion der Skalenraumereignisse des Primär- mit den Kontextobjekten, die die Struktur der Objekte signifikant verändern.

Die automatische Anpassung eines gegebenen Modells, das für die Extraktion eines Landschaftsobjektes aus optischen Bilddaten erstellt wurde, an Bilddaten anderer Sensortypen, wie Radar oder Lidar, ist dagegen nur schwer möglich. Bei anderen Datentypen ist in der Regel die Annahme einer Gaußschen Glättung bei der Bildaufnahme aufgrund einer grundsätzlich anderen Aufnahmetechnik nicht erfüllt. Damit kann auch der lineare Skalenraum nicht zur Simulation der Bilddaten eingesetzt werden.

Über eine Erweiterung des entwickelten Verfahrens könnte die bestmögliche Bildauflösung für die Extraktion eines bestimmten Objektes gefunden werden. Dazu ist eine Analyse der Extrahierbarkeit im gesamten Skalenraum nötig statt in einer spezifischen zur Zielauflösung korrespondierenden Skalenebene, wie es im derzeitigen Verfahren durchgeführt wird. Für die Klassifikation von Fernerkundungsdaten existieren bereits ähnliche Ansätze. Beispielsweise wird die lokale Varianz der Grauwerte als Maß für die beste Skalenebene für eine Klassifikation ausgewählt [WOODCOCK & STRAHLER 1987, ATKINSON & APLIN 2004] oder die mittlere Entropie wird für die beste Separabilität von Klassen in Multispektraldaten minimiert [HAN et al. 2008]. Diese Ansätze nehmen jedoch weder Bezug zu Merkmalsoperatoren, noch sind sie objektbasiert. Sie beziehen sich stattdessen auf eine rein pixel-basierte Klassifikation.

Die Netzwerkeigenschaft als skaleninvariante Eigenschaft, die erst in gröberen Bildauflösungen für die Straßenextraktion wichtig wird, wurde in dieser Arbeit nur ansatzweise untersucht. In einem Beispiel wurde ein Modell für Kreuzungsarme gezeigt, deren detaillierte Modellierung die Voraussetzung für eine weitergehende Vernetzung der Straßen ist. Die Netzwerkeigenschaft leistet jedoch einen wichtigen Beitrag zur Straßenextraktion aus niedrigeren Bildauflösungen, wenn andere Evidenz zur Objektextraktion aus Bilddaten nicht mehr vorhanden ist. Es bleibt daher zu untersuchen, wie übergeordnete Charakteristika, die erst in der niedrigeren Auflösung relevant werden, in das Anpassungsverfahren eingebunden werden können. Dazu müssen allerdings a priori Informationen über diese Eigenschaften vorliegen, da diese nicht durch die Analyse des Skalenverhaltens in der jetzigen Form gewonnen werden können. In diesem Zusammenhang ist die Bestimmung derjenigen Bildauflösung wichtig, ab der eine Einbindung der zusätzlichen Eigenschaft sinnvoll bzw. notwendig für die Extraktion des Landschaftsobjektes ist.

Im derzeitigen Verfahren zur automatischen Anpassung wird die Extraktionsstrategie zum Modell als fest vorgegeben und unveränderlich in allen Auflösungen angenommen. Zusätzlich zur Anpassung des Objektmodells könnte jedoch auch die Extraktionsstrategie an die veränderten Gegebenheiten der niedrigeren Bildauflösung mit dem Modell gemeinsam angepasst werden. Der Vorteil einer gemeinsamen Anpassung ist die Abstimmung von Modell und Strategie, die ein besseres Extraktionsergebnis bewirken kann. Denkbar sind dafür verschiedene vorher definierte Module für den Extraktionsprozess, die dann je nach Bedarf in der niedrigeren Bildauflösung eingesetzt werden können. Alternativ könnte aus einem Sortiment einiger verschiedener Extraktionsstrategien die für eine bestimmte Zielauflösung optimale Strategie ausgewählt werden. Dagegen dürfte die Entwicklung einer komplett neuen Strategie in einem automatischen Anpassungsverfahren sehr schwierig sein.

Eine weitere Anwendungsmöglichkeit des Anpassungsverfahrens ist eine multiskalare multitemporale Bildanalyse. Oftmals sind Bilddaten eines bestimmten Beobachtungsgebietes zu unterschiedlichen Zeitpunkten nur von unterschiedlichen Sensoren verfügbar, die sich auch in ihrer Auflösung unterscheiden. Mit dem entwickelten Verfahren könnten die Modelle für die Objektextraktion aus den Bilddaten anderer Sensoren niedrigerer Auflösung automatisch abgeleitet werden. Allerdings ist eine Voraussetzung, dass die genutzten Bilddaten von optischen Sensoren stammen.

Diese Arbeit hat gezeigt, dass das Skalenverhalten von in Konzepten gegebenen Modellen zur Objektextraktion nicht immer eindeutig ist. Je flexibler die Modelle, desto weniger ist eine eindeutige Prädiktion des Skalenverhaltens möglich. Der Umgang mit dieser Mehrdeutigkeit des Skalenverhaltens erfordert weitere Untersuchungen. Forschungsbedarf besteht auch für eine Übertragung der entwickelten Methoden zur automatischen Anpassung auf Modelltypen für andere Landschaftsobjekte, die mit den bisherigen Methoden nicht angepasst werden können, jedoch von großem Interesse sind, z. B. Gebäude. Trotz der Komplexität der Prädiktion des Skalenverhaltens hat die vorliegende Arbeit Lösungen für einige Modelltypen präsentiert, die in begrenztem Rahmen Mehrdeutigkeit des Skalenverhaltens zulassen, und an realen Bildaten erfolgreich getestet. Die vorliegende Arbeit leistet damit einen wichtigen Beitrag zur Automatisierung der wissensbasierten Extraktion von Landschaftsobjekten aus optischen Fernerkundungsdaten. Bilddaten sind häufig ausschließlich in bestimmten Auflösungen verfügbar. Mit dem präsentierten Verfahren kann eine automatische Ableitung von Modellen zur Objektextraktion für eine verfügbare Bildauflösung aus einem bereits vorhandenen manuell erstellten Modell für eine höhere Bildauflösung durchgeführt werden. Damit kann der Aufwand zur Erstellung von geeigneten Modellen zur Objektextraktion erheblich reduziert werden.

Literaturverzeichnis

ADELSON, E. H., ANDERSON, C. H., BERGEN, J. R., BURT, P. J., OGDEN, J. M. (1984). Pyramid methods in image processing. *RCA Engineer* 29(6), S. 33–41.

ALTENKRÜGER, D. & BÜTTNER, W. (1992). *Wissensbasierte Systeme*. Vieweg-Verlag, Wiesbaden.

ARNOLD, V. (1992). *Catastrophe Theory*. Springer-Verlag, Berlin, 3. Auflage, 150 S.

ATKINSON, P. & APLIN, P. (2004). Spatial variation in land cover and choice of spatial resolution for remote sensing. *International Journal of Remote Sensing* 25(18), S. 3687-3702.

BABAUD, J., WITKIN, A., BAUDIN, M., DUDA, R. (1986). Uniqueness of the Gaussian Kernel for Scale-Space Filtering. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 8(1), S. 26–33.

BÄHR, H. & MÜLLER, M. (2004). Graphics and Language as Complementary Formal Representations for Geospatial Descriptions. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 36-4B, S. 216–221.

BAUMGARTNER, A., STEGER, C., MAYER, H., ECKSTEIN, W. (1997). Semantic objects and context for finding roads. In: D.M. McKeown, J.C. McGlone, O. Jamet (Hrsg.): *SPIE* 3072, S. 98–109.

BAUMGARTNER, A., STEGER, C., MAYER, H., ECKSTEIN, W., EBNER, H. (1999). Automatic road extraction based on multi-scale, grouping, and context. *Photogrammetric Engineering & Remote Sensing* 65(7), S. 777–785.

BAUMGARTNER, A. (2003). *Automatische Extraktion von Straßen aus digitalen Luftbildern*. Dissertation, Deutsche Geodätische Kommission C 564, München, 91 S.

BEIERLE, C. & KERN-ISBERNER, G. (2003). *Methoden wissensbasierter Systeme*. Vieweg-Verlag, Wiesbaden, 2. Auflage

BENZ, U., HOFMANN, P., WILLHAUCK, G., LINGENFELDER, I., HEYNEN, M. (2004). Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. *ISPRS Journal of Photogrammetry & Remote Sensing* 58, S. 239–258.

BLACK, M., SAPIRO, G., MARIMONT, D., HEEGER, D. (1998). Robust Anisotropic Diffusion. *IEEE Transactions on Image Processing* 7(3), S. 421–432.

BRESSON, X., VANDERGHEYNST, P., THIRAN, J.-P. (2006). Multiscale Active Contours. *International Journal of Computer Vision* 70(3), S. 197–211.

BÜCKNER, J., PAHL, M., STAHLHUT, O., LIEDTKE, C.-E. (2002). A Knowledge-Based System for Context Dependent Evaluation of Remote Sensing Data. In: 24. DAGM Symposium, Zürich, *Lecture Notes in Computer Sciences* 2449, Springer, Berlin, S. 58–65.

CANNY, J. (1986). A Computational Approach to Edge Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 8(6), S. 679–698.

CASTLEMAN, K. (1996). *Digital Image Processing.* Prentice Hall, New Jersey, 667 S.

CHEN, M. & YAN, P. (1989). A Multiscaling Approach Based on Morphological Filtering. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 11(7), S. 694–700.

DAMON, J. (1995). Local Morse Theory for Solutions to the Heat Equation and Gaussian Blurring. *Journal of Differential Equations* 115, S. 368–401.

DE GUNST, M. & VOSSELMAN, G. (1997). A Semantic Road Model for Aerial Image Interpretation. In: W.Förstner, L.Plümer (Hrsg.): *SMATI*, S. 107–122, Birkhäuser Verlag, Basel.

DEVOGELE, T., TREVISAN, J., RAYNAL, L. (1996). Building a Multiscale Database with Scale-Transition Relationships. In: *Proceedings of the 7th International Symposium on Spatial Data Handling*, 10 S.

ERIKSON, M. & OLOFSSON, K. (2005). Comparison of three individual tree crown detection methods. *Machine Vision and Applications* 16(4), S. 258–265.

FISCHLER, M., TENENBAUM, J., WOLF, H. (1981). Detection of roads and linear structures in low-resolution aerial imagery using a multisource knowledge integration technique. *Computer Graphics and Image Processing* 15(3), S. 201–223.

FLORACK, L., TER HAAR ROMENY, B., KOENDERINK, J., VIERGEVER, M. (1992). Scale and the differential structure of images. *Image and Vision Computing* 10(6), S. 376–388.

FLORACK, L., TER HAAR ROMENY, B., KOENDERINK, J., VIERGEVER, M. (1994). Linear scale-space. *Journal of Mathematical Imaging and Vision* 4(4), S. 325–351.

FLORACK, L. & KUIJPER, A. (2000). The topological structure of scale-space images. *Journal of Mathematical Imaging and Vision* 12(1), S. 65–79.

FÖRSTNER, W. (1993). A future of photogrammetric research. *NGT Geodesia* 93(8), S. 372–383.

FÖRSTNER, W. (1996). 10 Pros and Cons against performance characterization of vision algorithms. In: *European Conference on Computer Vision, Workshop on Performance Characteristics of Vision Algorithms*, S. 13–29.

FÖRSTNER, W. (2003). Notions of scale in geosciences. In: Neugebauer, H.J. & Simmer, C. (Hrsg.): Dynamics of Multiscale Earth Systems, *Lecture Notes in Earth Sciences* 97, Springer, Heidelberg, S. 17–39.

GAUCH, J. (1999). Image segmentation and analysis via multiscale gradient watershed hierarchies. *IEEE Transactions on Image Processing* 8(1), S. 69–79.

GERKE, M., BUTENUTH, M., HEIPKE, C., WILLRICH, F. (2004). Graph supported verification of road databases. *ISPRS Journal of Photogrammetry & Remote Sensing* 58(3/4), S. 152–165.

GRIFFIN, L. & LILLHOLM, M. (2006). Hypotheses for image features, icons and textons. *International Journal of Computer Vision* 70(3), S. 213–230.

GRIFFIN, L. & LILLHOLM, M. (2007). Maximum likelihood metameres for local 2nd order image structure of natural images. In: F. Sgallari, A. Murli, and N. Paragios (Hrsg.): SSVM 2007, *Lecture Notes in Computer Sciences* 4485, Springer, Berlin, S. 394–405.

HAKE, G., GRÜNREICH, D., MENG, L. (2002). *Kartographie*. Walter de Gruyter, Berlin, New York, 8. Auflage.

HAN, P., GONG, J.Y., LI, Z.L., CHENG, L. (2008). The study on the choice of optimal scale in image classification. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 37-B2, S. 385–389.

HAY, G., DUBE´, P., BOUCHARD, A., MARCEAU, D. (2002). A scale-space primer for exploring and quantifying complex landscapes. *Ecological Modelling* 153, S. 27–49.

HEIPKE, C., STEGER, C., MULTHAMMER, R. (1995). A hierarchical approach to automatic road extraction from aerial imagery. In: D. McKeown and I. Dowman (Hrsg.): Integrated Photogrammetric Techniques with Scene Analysis and Machine Vision II, *SPIE* 2486, S. 222–231.

HEIPKE, C., MAYER, H., WIEDEMANN, C., JAMET, O. (1998). External evaluation of automatically extracted road axes. *Photogrammetrie-Fernerkundung-Geoinformation* 1998(2), S. 81–94.

HELLER, J. & PAKZAD, K. (2005A). Adaption of object models for road extraction in images of different resolution. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 36-8/W27, Tempe, 6 S.

HELLER, J. & PAKZAD, K. (2005B). Scale-dependent adaptation of object models for road extraction. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 36-3/W24, Wien, S. 23–28.

HEUWOLD, J. (2006). Verification of a methodology for the automatic scale-dependent adaption of object models. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 36/3, Bonn, S. 173–178.

HEUWOLD, J. & PAKZAD, K. (2006). Automatic scale-dependent adaption of variable object models. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 36-2/W40, Hannover, S. 65–70.

HEUWOLD, J., PAKZAD, K., HEIPKE, C. (2007). Scale behaviour prediction of image analysis models for 2D landscape objects. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 36-3/W49A, München, S. 43–49.

HEUWOLD, J., PAKZAD, K., HEIPKE, C. (2008A). Automatic adaptation of image analysis models for 2D landscape objects to a coarser image resolution. *Photogrammetrie-Fernerkundung-Geoinformation* 2008(3), S. 197–206.

HEUWOLD, J., PAKZAD, K., HEIPKE, C. (2008B). Scale-dependent adaptation of image analysis object models incorporating local context objects. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 37-B3b, Peking, S. 435–440.

HINZ, S., BAUMGARTNER, A. (2000). Road extraction in urban areas supported by context objects. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 33-B3/1, S. 405–412.

HINZ, S., BAUMGARTNER, A., MAYER, H., WIEDEMANN, C., EBNER, H. (2001). Road extraction focussing on urban areas. In: E. Baltsavias, A. Grün, L. van Gool (Hrsg.): *Automatic Extraction of Man-Made Objects from Aerial and Space Images (III)*, Ascona, Schweiz, Balkema Publishers, Rotterdam, S. 255–265.

HINZ, S. & BAUMGARTNER, A. (2003). Automatic extraction of urban road networks from multi-view aerial imagery. *ISPRS Journal of Photogrammetry & Remote Sensing* 58, S. 83–98.

HINZ, S. (2003A). Integrating local and global features for vehicle detection in high resolution aerial imagery. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 34-3/W8, S. 119–124.

HINZ, S. (2003B). *Automatische Extraktion urbaner Straßennetze aus Luftbildern*. Dissertation, Deutsche Geodätische Kommission C 580, München, 130 S.

HINZ, S. (2004). Detection of Vehicles and Vehicle Queues in High Resolution Aerial Images. *Photogrammetrie-Fernerkundung-Geoinformation* 3/04, S. 201–213.

HINZ, S., LENHART, D., LEITLOFF, J. (2008). Traffic extraction and characterization from optical remote sensing data. *Photogrammetric Record* 23(124), S. 424 – 440.

JÄHNE, B. (1991). *Digitale Bildverarbeitung*. Springer-Verlag, 2.Auflage.

KIMIA, B., TANNENBAUM, A., ZUCKER, S. (1995). Shapes, Shocks, and Deformations I: The Components of Two-Dimensional Shape and the Reaction-Diffusion Space. *International Journal of Computer Vision* 15(3), S. 189–224.

KIMMEL, R., MALLADI, R., SOCHEN, N. (2000). Images as Embedded Maps and Minimal Surfaces: Movies, Color, Texture, and Volumetric Medical Images. *International Journal of Computer Vision* 39(2), S. 111–129.

KOENDERINK, J. J. (1984). The structure of images. *Biological Cybernetics* 50(5), S. 363–370.

KOJIMA, N., LABA, M., VELEZ LIENDO, X. M., BRADLEY, A. V. (2006). Causes of the apparent scale independence of fractal indices associated with forest fragmentation in Bolivia. *ISPRS Journal of Photogrammetry & Remote Sensing* 61, S. 84–94.

KOSSLYN, S.M. (1994). *Image and brain: the resolution of the imagery debate*. MIT Press, 516 S.

KREMS, J. (1989). *Expertensysteme im Einsatz – Erfahrungsberichte der 1. Generation*. Oldenbourg Verlag, München, Wien, 158 S.

KUIJPER, A. (2002). *The deep structure of Gaussian scale space images*. PhD thesis, Utrecht University

KUIJPER, A. (2004). Mutual information aspects of scale space images. *Pattern Recognition* 37, S. 2361–2373.

KUIJPER, A. & FLORACK, L. (1999). Calculations on critical points under Gaussian blurring. In: M. Nielsen, P. Johansen, O. Fogh Olsen, J. Weickert (Hrsg.): Scale-Space Theories in Computer Vision, *Lecture Notes in Computer Science* 1682, Springer-Verlag, Berlin, S. 318–329.

KUIJPER, A. & FLORACK, L. (2004). The relevance of non-generic events in scale space models. *International Journal of Computer Vision* 57(1), S. 67–84.

KUIJPER, A. & FLORACK, L. (2005). Using catastrophe theory to derive trees from images. *Journal of Mathematical Imaging and Vision* 23, S. 219–238.

KUIJPER, A., FLORACK, L., VIERGEVER, M. (2003). Scale space hierarchy. *Journal of Mathematical Imaging and Vision* 18, S. 169–189.

LAPTEV, I., MAYER, H., LINDEBERG, T., ECKSTEIN, W., STEGER, C., BAUMGARTNER, A. (2000). Automatic extraction of roads from aerial images based on scale space and snakes. *Machine Vision and Applications* 12(1), S. 21–31.

LIEDTKE, C.-E., BÜCKNER, J., GRAU, O., GROWE, S., TÖNJES, R. (1997). AIDA: A System for the Knowledge Based Interpretation of Remote Sensing Data. *3rd Int. Airborne Remote Sensing Conference & Exhibition* 2, S. 313–320.

LIEDTKE, C.-E., BÜCKNER, J., PAHL, M., STAHLHUT, O. (2001). Knowledge based system for the interpretation of complex scenes. In: E. Baltsavias, A. Grün, L. van Gool (Hrsg.): *Automatic extraction of man-made objects from aerial and space images (III)*, Ascona, Schweiz, Balkema Publishers, Rotterdam, S. 3–12.

LINDEBERG, T. (1990). Scale-space for discrete signals. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 12(3), S. 234–254.

LINDEBERG, T. (1992). Scale-Space Behaviour of Local Extrema and Blobs. *Journal of Mathematical Imaging and Vision* 1, S. 65–99.

LINDEBERG, T. (1993). Detecting salient blob-like image structures and their scales with a scale-space primal sketch: a method for focus-of-attention. *International Journal of Computer Vision* 11(3), S. 283–319.

LINDEBERG, T. (1994). *Scale-space theory in computer vision*. Kluwer Academic Publisher, The Netherlands, 423 S.

LINDEBERG, T. (1998A). Feature detection with automatic scale selection. *International Journal of Computer Vision* 30(2), S. 79–116.

LINDEBERG, T. (1998B). Edge and ridge detection with automatic scale selection. *International Journal of Computer Vision* 30(2), S. 117–154.

LEITLOFF, J., HINZ, S., STILLA, U. (2006). Automatic vehicle detection in satellite images. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 36-3, PCV06, Bonn, S. 221–227.

LOOG, M., DUISTERMAAT, J., FLORACK, L. (2001). On the behavior of spatial critical points under Gaussian blurring (a folklore theorem and scale-space constraints). In: M. Kerckhove (Hrsg.): *Lecture Notes in Computer Sciences*, S. 183–192.

LOWE, D. (2004). Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision* 20, S. 91–110.

LUCAS, C., MÜLLER, M., BÄHR, H. (2008). Integration of language in GIS: models in ownership cadastre and disaster management. *Photogrammetrie-Fernerkundung-Geoinformation* 2008(3), S. 217–225.

LUSTI, M. (1990). *Wissensbasierte Systeme*. Reihe Informatik, Band 77. BI-Wissenschafts-Verlag, Mannheim.

MALLADI, R., SETHIAN, J., VEMURI, B. (1995). Shape Modeling with Front Popagation: A Level Set Approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 17(2), S. 158–175.

MANDELBROT, B. (1977). *Fractals: form, chance and dimension.* Freeman, San Fransico, 365 S.

MANDELBROT, B. (1987). *Die fraktale Geometrie der Natur*. Birkhäuser Verlag, Basel, 491 S.

MARR, D. (1982). *Vision*. W.H. Freeman and Company, New York.

MATSUYAMA, T. (1987). Knowledge-based aerial image understanding systems and expert systems for image processing. *IEEE Transactions on Geoscience and Remote Sensing* 25(3), S. 305–316.

MATSUYAMA, T. & HWANG, V.S.-S. (1990). *SIGMA: A Knowledge-Based Aerial Image Understanding System*. Plenum Press, New York, 277 S.

MAYER, H. (1998A). *Automatische Objektextraktion aus digitalen Luftbildern.* Habilitation, Deutsche Geodätische Kommission C 494, München, 131 S.

MAYER, H. (1998B). Maßstabsräume: Theorie und Anwendung. *Photogrammetrie-Fernerkundung-Geoinformation* 1998(4), S. 197–208.

MAYER, H. & STEGER, C. (1998). Scale-space events and their link to abstraction for road extraction. *ISPRS Journal of Photogrammetry & Remote Sensing* 53, S. 62–75.

MAYER, H. (2008). Object extraction in photogrammetric computer vision. *ISPRS Journal of Photogrammetry & Remote Sensing* 63(2), S. 213–222.

MCGLONE, C. & SHUFELT, J. (1994). Projective and Object Space Geometry for Monocular Building Extraction. *Computer Vision and Pattern Recognition*, 21-23 June, S. 54–61.

MCKEOWN, D., WILSON, A., MCDERMOTT, J. (1985). Rule-based interpretation of aerial imagery. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 7(5), S. 570–585.

MENG, L. (2000). *ATKIS: Modell- und kartographische Generalisierung.* Vorstudien zum AdV-Forschungs- und Entwicklungsvorhaben. Lehrstuhl für Kartographie, Technische Universität München.

MÜLLER, S., WEIS, M., LIEDTKE, C.-E., PAHL, M. (2003). Automatic quality surveillance of GIS data with geoAIDA. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 36-3/W8, PIA03, München, S. 187–192.

NIEMANN, H., SAGERER, G., SCHRÖDER, S., KUMMERT, F. (1990). ERNEST: A Semantic Network System for Pattern Understanding. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 12(9), S. 883–905.

PAHL, M. (2003). Architektur eines wissensbasierten Systems zur Interpretation multisensorieller Fernerkundungsdaten. *Schriftenreihe des TNT der Universität Hannover* 3, ibidem-Verlag, Stuttgart, 145 S.

PAKZAD, K. (2001). Wissensbasierte Interpretation von Vegetationsflächen aus multitemporalen Fernerkundungsdaten. Dissertation, Deutsche Geodätische Kommission, C 543, München, 104 S.

PAKZAD, K. & HELLER, J. (2004). Automatic scale adaptation of semantic nets. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 35-B3, Istanbul, S. 325–330.

PERONA, P. & MALIK, J. (1990). Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 12(7), S. 629–639.

PESCHEL, G. (1992). Wissensbasierte Systeme in den Geowissenschaften. *Beiträge zur Mathematischen Geologie und Geoinformatik* 4, Verlag Sven von Loga, 147 S.

PLATEL, B., BALMACHNOVA, E., FLORACK, L., KANTERS, F., TER HAAR ROMENY, B. (2005). Using top-points as interest points for image matching. In: O.F. Olsen et al. (Hrsg.): *Lecture Notes in Computer Sceinces* 3753, S. 211-222.

QUILLIAN, M. (1968). Semantic memory. In: M. Minsky (Hrsg.): *Semantic information processing*. MIT Press, Cambridge, MA, S. 216–270.

RAVANBAKHSH, M. (2008). Road junction extraction from high resolution aerial imagery assisted by topographic database information. Dissertation, Deutsche Geodätische Kommission, C 621, München, 92 S.

RAVANBAKHSH, M., HEIPKE, C., PAKZAD, K. (2008A). Road junction extraction from high resolution aerial imagery. *The Photogrammetric Record* 23(124), S. 405-423.

RAVANBAKHSH, M., PAKZAD, K., HEIPKE, C. (2008B). Automatic extraction of traffic islands from aerial images. *Photogrammetrie-Fernerkundung-Geoinformation* 2008(5), S. 375-384.

REICHGELT, H. (1991). *Knowledge representation: an AI perspective*. Ablex Publishing Corporation, Norwood, New Jersey, 251 S.

REIMER, U. (1991). *Einführung in die Wissensrepräsentation.* Leitfäden der angewandten Informatik, B.G. Teubner, Stuttgart, 313 S.

RUSKONÉ, R., AIRAULT, S., JAMET, O. (1994). A road extraction system using the connectivity properties of the network. *Zeitschrift für Photogrammetrie und Fernerkundung* 1994(5), S. 174–180.

RUSKONÉ, R. (1996). *Road Network Automatic Extraction by Local Context Interpretation: Application to the Production of Cartographic Data*. These de Doctorat, Université de Marne-La-Vallee, Noisy-le-Grand, France.

SAPIRO, G. & TANNENBAUM, A. (1993). Affine Invariant Scale-Space. *International Journal of Computer Vision* 11(1), S. 25–44.

SCHIEWE, J. (2005). Status and future perspectives of the application potential of digital airbone sensor systems. *International Journal of Applied Earth Observation and Geoinformation* 6, S. 215–228.

SCHLOSSER, C., REITBERGER, J., HINZ, S. (2003). Automatic Car Detection in High Resolution Urban Scenes Based on an Adaptive 3D-Model. *URBAN 2003*, Berlin, S. 167–171.

SCHRAMM, M. (2005). *Untersuchungen zum Skalenverhalten von Bildanalyse-Operatoren zur automatischen Extraktion von Fahrbahnmarkierungen*. Diplomarbeit, Universität Hannover, 66 S.

SESTER, M. (2001). *Maßstabsabhängige Darstellungen in digitalen räumlichen Datenbeständen.* Habilitation, Deutsche Geodätische Kommission, C 544, München, 114 S.

SESTER, M., ANDERS, K.-H., WALTER, V. (1998). Linking Objects of Different Spatial Data Sets by Integration and Aggregation. *GeoInformatica* 2(4), S. 335–358.

SESTER, M., HAMPE, M., ANDERS, K.-H. (2003). Multiskalige Geodaten in einer Datenbank – Erzeugung und Nutzen. *Photogrammetrie-Fernerkundung-Geoinformation* 2003(6), S. 481–491.

SHANNON, C. E. (1949). Communication in the presence of noise. *Proceedings of the IRE* 37(1), S. 1–21. reprint in *Proceedings of the IEEE* 86(2), S. 447–457.

SOCHEN, N., KIMMEL, R., MALLADI, R. (1998). A general framework for low level vision. *IEEE Transactions on Image Processing* 7(3), S. 310–318.

SOWA, J. (1992). Semantic Networks. In: S. Shapiro (Hrsg.): *Encyclopedia of Artificial Intelligence* 2 (M-Z). John Wiley & Sons, New York, 2. Auflage, S. 1493–1511.

STEGER, C. (1998A). An unbiased detector of curvilinear structures. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 20(2), S. 113–125.

STEGER, C. (1998B). *An unbiased detector of curvilinear structures from 2D and 3D images*. Dissertation, Technische Universität München, Herbert Utz Verlag, München, 185 S.

STILLA, U., MICHAELSEN, E., SOERGEL, U., HINZ, S., ENDER, J. (2004). Airborne Monitoring of Vehicle Detection in Urban Areas. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 35-B3, S. 973–979

SUNDERMEYER, K. (1991). *Knowledge-Based Systems – Terminology and References*. BI Wissenschaftsverlag, Mannheim, 222 S.

TER HAAR ROMENY, B. M. (2002). *Front-End Vision and Multi-Scale Image Analysis: Multi-Scale Computer Vision Theory and Applications, written in Mathematica*. Kluwer Academic Publishers, 464 S.

TÖNJES, R., GROWE, S., BÜCKNER, J., LIEDTKE, C.-E. (1999). Knowledge-based interpretation of remote sensing images using semantic nets. *Photogrammetric Engineering & Remote Sensing* 65(7), S. 811–821.

TRINDER, J. & WANG, Y. (1998). Knowledge-based road extraction in aerial images. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 32-4, Stuttgart, 6 S.

VAN DEN BOOMGAARD, R. & SMEULDERS, A. (1994). The Morphological Structure of Images: The Differential Equations of Morphological Scale-Space. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 16(11), S. 1101–1113.

WEICKERT, J., ISHIKAWA, S., IMIYA, A. (1999). Linear scale-space has first been proposed in Japan. *Journal of Mathematical Imaging and Vision* 10(3), S. 237–252.

WIEDEMANN, C. (2002). *Extraktion von Straßennetzen aus optischen Satellitenbilddaten*. Dissertation, Deutsche Geodätische Kommission, C 551, München, 121 S.

WITKIN, A. (1983). Scale space filtering. *Proceedings of IJCAI*, S. 1019–1022.

WITKIN, A. P. (1986). Scale space filtering. In: A. Pentland (Hrsg.): *From Pixels to Predicates*, Ablex Publishing Corporation, New Jersey, S. 5–19.

WOODCOCK, C. & STRAHLER, A. (1987). The factor of scale in remote sensing. *Remote Sensing Environment* 21, S. 311–322.

ZHANG, C. (2004). Towards an operational system for automated updating of road databases by integration of imagery and geodata. *ISPRS Journal of Photogrammetry & Remote Sensing* 58(3-4), S. 166–186.

Dank

An dieser Stelle möchte ich allen Beteiligten danken, die zum Gelingen der Arbeit auf ganz unterschiedliche Art und Weise beigetragen haben:

Meinem Doktorvater Prof. Dr.-Ing. Christian Heipke danke ich für die hervorragende Betreuung der Arbeit, die stets äußerst konstruktiven kritischen Diskussionen, die sehr angenehme Zusammenarbeit und die immer mögliche Hilfe trotz Deines Zeitmangels. Vielen Dank auch für die Unterstützung in meiner Unterbrechungszeit, die für Dich selbstverständlich war.

Prof. Dr.-Ing. Monika Sester, Leiterin des Instituts für Kartographie und Geoinformatik der Leibniz Universität Hannover, möchte ich herzlich danken für die Übernahme des Referates, der Leitung des DFG-Bündels sowie die Anregungen zu stetigen offenen Diskussionen in der Bündelgruppe.

Prof. Dr.-Ing. Olaf Hellwich, Leiter der Fachrichtung Computer Vision am Institut für Informatik der Technischen Universität Berlin danke ich sehr für die Übernahme des Referates und die wertvollen Anmerkungen zur Verbesserung der Arbeit.

Dr.-Ing. Kian Pakzad danke ich für die Einführung in das Projekt und die Anleitung zur wissenschaftlichen Arbeit und zur Projektarbeit in der Anfangszeit des Projektes. Dass du jederzeit ansprechbar und offen gegenüber Fragen warst, hat mir sehr geholfen.

Danke an die Bildanalyse-Gruppe am IPI, die in veränderlicher Zusammensetzung während meiner Zeit am Institut durch offene, kritische Diskussion unserer Arbeit in regelmäßigen Treffen neue Anregungen zur Forschung gab und neben den internationalen Konferenzen für die nötige Qualitätskontrolle sorgte.

Dem gesamten IPI danke ich für die nicht selbstverständliche angenehme Arbeitsatmosphäre und die jederzeit vorhandene große gegenseitige Hilfsbereitschaft.

Mein Dank geht auch an die fleißigen Korrekturleser: Dr.-Ing. Franz Rottensteiner, der wertvolle Impulse zur Struktur und zum Aufbau der Arbeit gab, Dr.-Ing. Ralph Schmidt und Dr.-Ing. Stephan Gehrke, die sehr hilfreiche Anregungen zum Ausdruck und Format hatten, sowie an meine Schwiegereltern Inge und Hans-Peter Heuwold für die Hinweise zur Rechtschreibung, Grammatik und zum Ausdruck in einer sehr kurzen Zeit.

Der DFG möchte ich für die Finanzierung der Forschungsarbeit und die Ermöglichung der Unterbrechung der Projektzeit während meiner Elternzeit danken.

Durch das wunderbare Kümmern um und Beschäftigen von Paula sorgten meine Eltern dafür, dass wir ab und zu trotz des großen Zeitdrucks merklich entspannen konnten.

Danke an Niels für die unbeschreibliche Unterstützung bei der Fertigstellung der Dissertation. Durch die monatelange Übernahme der kompletten Familien-Alltags-Organisation konnte ich mich auf meine Arbeit konzentrieren. Danke auch für das ständige Aufmuntern, Bestärken, Nachfragen und Erinnern an wesentliche Organisationsaufgaben im Zusammenhang mit der Dissertation. Du hast mich oft erst aus dem Tunnel des Schreibens rausgeholt.

Meiner Tochter Paula danke ich dafür, dass sie mir täglich zeigt, warum unbändiger Forscherdrang die Grundvoraussetzung für Weiterentwicklung ist und wie diese uns verändert und voranbringt …

Lebenslauf

Persönliche Daten

Name	Janet Heuwold, geb. Heller
Geburtsdatum	7. April 1976
Geburtsort	Berlin
Familienstand	verheiratet, eine Tochter

Berufstätigkeit / Studienbegleitende Tätigkeiten

seit 04/2009	Referentin *Bundesministerium des Innern*
12/2003 – 04/2009	wissenschaftliche Mitarbeiterin *Institut für Photogrammetrie und GeoInformation, Leibniz Universität Hannover,*
	davon Mutterschutz und Elternzeit 04/2007 – 12/2007
10/2003 – 12/2003	Mitarbeiterin *Informus GmbH*
07/2003 – 10/2003	Praktikum *Deutsche Bahn AG* – DB Projektbau GmbH
11/2002 – 04/2003	wissenschaftliche Mitarbeiterin *Fraunhofer IPK*
04/2000 – 10/2002	Tutorin *Institut für Geodäsie und Geoinformationstechnik, Technische Universität Berlin*
05/2001 – 04/2002	Diplomandin *Institut für Weltraumsensorik und Planetenerkundung, Deutsches Zentrum für Luft- und Raumfahrt e.V.*

Ausbildung

02/1999 – 02/2000	Geomatics *University of Melbourne*, Australien
10/1995 – 08/2002	Vermessungswesen (Diplom) *Technische Universität Berlin*
09/1990 – 07/1995	Friedrich-List-Gymnasium (Abitur), Berlin-Pankow
09/1982 – 08/1990	Polytechnische Oberschule, Berlin-Pankow