

Dissertation:

Integrierte Verarbeitung von Satellitenbild-
und gescannter Karteninformation

Zur Erlangung des akademischen Grades eines

DOKTOR-INGENIEURS

von der Fakultät für

Bauingenieur- und Vermessungswesen
der Universität Fridericiana zu Karlsruhe (TH)

genehmigte

DISSERTATION

von

Ing. M.Sc. Jorge Silva Centeno

aus La Paz, Bolivien

Tag der mündlichen Prüfung: 16.02.2000

Hauptreferent: Prof. Dr.-Ing. Hans-Peter Bähr

Korreferent: Prof. Dr.-Ing. Dieter Morgenstern

Karlsruhe, 16. Februar 2000

Vorwort

Im Rahmen der Förderung der internationalen Zusammenarbeit in Hochschulen und Wissenschaft, erhielt ich ein vom DAAD (Deutscher Akademischer Austauschdienst) zugesprochenes Stipendium zur Promotion am Institut für Photogrammetrie und Fernerkundung der Universität Karlsruhe. Betreuer dieser Arbeit war Prof. Dr.-Ing. habil. Hans-Peter Bähr, Leiter des Institutes. Es handelte sich um eine Forschungsarbeit aus dem Bereich des Einsatzes von Fernerkundung und raumbezogenen Daten unter Berücksichtigung der Lage in Südamerika.

Während dieser Zeit hat mich mein Doktorvater Prof. Dr.-Ing. habil. Hans-Peter Bähr gefördert und unterstützt. Für die Übernahme dieser Arbeit und die freundliche Beziehung die er ermöglichte, bedanke ich mich bei ihm. Ebenso möchte ich mich bedanken bei allen ehemaligen und jetzigen Kollegen des Institutes, die mir durch ihre Freundschaft und Verständnis geholfen haben. Aber vor allem bin ich für die Unterstützung meiner Frau dankbar, denn sie war immer bereit, mit Geduld und Liebe, mich durch alle Situationen zu begleiten.

Abstract

The present work deals with the integration of different data sources for land-use classification. Satellite imagery and scanned topographic maps are used to obtain an updated description of the Earth's surface at regions where data availability is precarious, as is often the case in developing countries.

Scanned maps are analysed and information is derived from the raster representation of this. The first step is to separate colour layers from a RGB image of the map. This task is performed using a fuzzy logic inference schema, based on rules that enable a segmentation algorithm to adapt to the special properties of the image to be segmented. Information about the objects represented in the map is later derived from the separated colour layers. Linear and area features are identified and binary layers for each land-use class are produced.

A rule-based system is used as basis for the integration of the information from the map and the satellite imagery. Therefore, the cartographic information from the binary layers, derived from the scanned map, and the spectral information, obtained from the result of a Maximum Likelihood classification, are represented as fuzzy sets and stored as new grey level layers. These are combined according to user-specified rules in order to generate a classified image.

Since the classification is based on fuzzy-set theory, the definition of the necessary membership functions is a key point in the classification. A first test was performed using functions that were defined by an human operator. These were proposed based on an interactive trial-and-error analysis of the effect of each function on the result. This task is defficient, since a great amount of functions and parameters are involved. Hence, an alternative to estimate the functions was studied. The solution is based on the principles of genetic algorithms. An evolutive optimisation procedure is designed to search for the optimal solution according to rules that resemble the natural evolution. The result of the optimisation process produced a set of membership functions that were later used to obtain a thematic image, which summarises the information of both data sources and has an hybrid character.

The test area is located in south Brazil, within the Rio dos Sinos basin, near the city of Porto Alegre. Landsat imagery and a Spot image of the same area were used to derive information. Selected areas were classified and used as ground truth for the calibration process using genetic algorithms.

Zusammenfassung

Die vorliegende Arbeit behandelt die Problematik der Kombination unterschiedlicher Datenquellen zur Gewinnung einer Beschreibung der Erdoberfläche. Den Schwerpunkt bildet dabei die gemeinsame Verarbeitung von gescannten Karten und Satellitenbildern. Da immer die Möglichkeit besteht, daß die zur Verfügung stehenden Karten veraltet sind und eventuell nicht mehr der Wirklichkeit entsprechen, bedarf die Verknüpfung dieser Informationen mit aus Satellitenbildern gewonnenen Daten einer sorgfältigen Analyse.

Das Grundproblem in diesem Zusammenhang besteht darin, basierend auf Informationen aus unterschiedlichen Quellen (Satellitenbildern und gescannten Karten), jedes Bildelement einer Bedeutung zuzuweisen. Da bei jeder Art von Datengewinnung andere Aspekte der Realität registriert werden, ist es nötig, Hinweise auf die Zugehörigkeit der einzelnen Pixel zu den Klassen aus den unterschiedlichen Datenquellen abzuleiten und sie später miteinander zu kombinieren, so daß am Ende eine Beziehung zwischen den Pixeln und den Landnutzungsklassen besteht.

Die Datenbasis für die digitale Bildverarbeitung besteht aus Satellitenbildern und Karten; das sind Daten, die auch relativ leicht in Entwicklungsländern erhältlich sind.

Basierend auf dem gescannten Bild der Karte wird ein binäres Bild für jede vorhandene Farbe erzeugt. Dafür wird ein fuzzy-logic basiertes Optimierungsverfahren angewandt. Die vollständige Erkennung aller Kartenobjekte wird nicht untersucht, sondern man beschränkt sich auf linien- und flächenhafte Elemente, die auch in den Satellitenbildern vorkommen. Die Klassengrenzen werden ermittelt und als Fuzzy-Mengen dargestellt, so daß ihr Wachstum oder ihre Verkleinerung modelliert werden kann. Aus den Satellitenbildern werden, mit Hilfe des Maximum-Likelihood-Verfahrens, neue Hinweise über mögliche Klassen abgeleitet.

Unschärfe Mengen werden benutzt, um die Informationen darzustellen und sie miteinander zu kombinieren. Unschärfe Logik eignet sich gut für diesen Zweck, denn damit ist man in der Lage, Unsicherheit über die berechnete Zugehörigkeitsfunktion ausdrücken und zu verarbeiten. Das Fuzzy-Klassifikationsschema bietet ein offenes Modell, das an die Eigenschaften der Szene angepaßt werden kann, denn es basiert auf Regeln und unscharfen Funktionen. Zur Schätzung der notwendigen unscharfen Funktionen wird ein Verfahren vorgeschlagen, welches auf dem Prinzip der genetischen Algorithmen basiert.

Inhaltsverzeichnis

1	Einleitung	9
1.1	Motivation und Problemstellung	10
1.2	Zielsetzung	12
1.3	Gliederung der Arbeit	12
2	Verwendete Datensätze	13
2.1	Satellitenbilder	13
2.2	Karten	14
2.3	Entzerrung	15
3	Theoretische Grundlagen	17
3.1	Modelle für das Bildverstehen: Szene vs. Bild	17
3.2	Methoden zur gemeinsamen Verarbeitung unterschiedlicher Daten	18
3.2.1	Möglichkeiten zur Integration von GIS und Satellitenbildern	20
3.2.2	Statistischer Ansatz	22
3.2.3	Dempster-Shafer-Kombinationsregeln	24
3.2.4	Fuzzy-Logik	24
4	Ableitung von Information aus Karten	28
4.1	Farbsegmentierung	28
4.1.1	Wahl des Farbraumes	28
4.1.2	Farbtrennung	32
4.1.3	Histogrammbasierte Farbsegmentierung	34
4.1.4	Optimierung der Farbtrennung	35
4.1.5	Bewertung	39
4.1.6	Ergebnisse	41
4.2	Erkennung von Landnutzungsgruppen	44
4.2.1	Erfassung von Linien	44
4.2.2	Erfassung von Flächen	44
4.2.3	Erfassung von Mustern und parallelen Linien	46
4.2.4	Das Wasserscheide-Verfahren	46
4.2.5	Erfassung des Straßennetzes mit dem Wasserscheide-Verfahren	49
4.2.6	Erkennung von Zwischenräumen mit dem Wasserscheide-Verfahren	50
4.2.7	Die Delaunay-Triangulation	52
4.2.8	Erkennung der Zwischenräume mit der Delaunay-Triangulation	53
4.2.9	Erkennung von Flächen mit Mustern	54

5	Ableitung von Information aus Satellitenbildern	56
5.1	Ableitung spektraler Informationen	56
5.2	Ableitung von Information aus der Nachbarschaft	57
6	Integrierte Verarbeitung der Information	60
6.1	Einbeziehung der Unsicherheit der Daten	60
6.1.1	Berücksichtigung der Lageunsicherheit	61
6.1.2	Berücksichtigung der Unsicherheit der Semantik	62
6.2	Inferenzschema	63
6.2.1	Regelbasiertes System	64
6.2.2	Linguistische Variable und Terme	65
6.2.3	Regeln	66
6.2.4	Fuzzy-Funktionen	66
7	Automatische Schätzung der Fuzzy-Funktionen	68
7.1	Genetische Algorithmen	68
7.2	Bildung der genetischen Kette	71
7.3	Operatoren	72
7.3.1	Initialisierung	72
7.3.2	Auswahlprozeß	72
7.3.3	Paarung	72
7.3.4	Mutation	72
7.3.5	Ersetzung	72
7.3.6	Anpassungskriterium	73
7.3.7	Homogenitätsmaß	74
7.3.8	Ende des Suchprozesses	74
8	Experimentelle Ergebnisse	75
8.1	Beschreibung des Testgebietes anhand einer Karte im Maßstab 1:50.000	75
8.1.1	Kartenhinweise	76
8.2	Beschreibung des Testgebietes anhand der Satellitenbilder	79
8.3	Regeln	81
8.4	Fuzzy-Funktionen	83
8.4.1	Landnutzungsklassen	83
8.4.2	Schätzung der Fuzzy-Funktionen für die linguistischen Terme	83
8.4.3	Anwendung der genetischen Algorithmen zur Bestimmung der Fuzzy-Funktionen	86

9 Diskussion	93
9.1 Kartenerkennung	93
9.2 Inferenzschema	94
9.3 Modellierung der unscharfen Mengen	95
9.4 Klassifikation	95
9.5 Schlußbetrachtungen und Ausblick	96
Literaturverzeichnis	97

1 Einleitung

Das schnelle Bevölkerungswachstum in manchen Regionen der Erde verändert den Lebensraum mit hoher Geschwindigkeit, zum Beispiel am Rande und in der Nähe großer Städte in Entwicklungsländern. Die damit verbundene Intensivierung der Landwirtschaft bewirkt Veränderungen der Landschaft und beeinflusst die Art der Flächennutzung. Neue Industriegebiete entstehen am Rande großer Städte, und die notwendige Infrastruktur wird ausgebaut. Diese Veränderungen geschehen oft ohne Planung. Die Konsequenz daraus ist eine ungeordnete Nutzung der vorhandenen Flächen. Eine Grundlage zur Verbesserung der Lebensbedingungen in Ländern der sogenannten Dritten Welt ist die Kenntnis der jeweiligen Gegebenheiten, zum Beispiel des Wasserhaushalts, der Siedlungsstrukturen, der Bodennutzung, der Bewaldung, und der Infrastruktur (Ajayi [1993]).

Der Stand der Kartenbedeckung in den Maßstabsbereichen 2 (1:40.000 bis 1:75.000) und 3 (1:100.000 bis 1:126.726) wurde von Brandenberger & Ghosh [1990] dokumentiert. Gemäß dieser Zahlen beträgt die Bedeckung in den Südamerikanischen Ländern 30 Prozent im zweiten Maßstabsbereich und 53 Prozent im dritten (Tab. 1). Die Kartenbedeckung in dieser Region ist niedrig, und im Gegensatz zu industrialisierten Ländern liegen die Daten in Form analoger Karten vor. Durch moderne technische Entwicklungen können manche Lücken in der Datengewinnung, zur Aktualisierung oder Erweiterung der Datenbasis, gefüllt werden. Fernerkundung, zusammen mit den Verfahren der computergestützten Kartographie und Bildverarbeitung, eröffnet neue Wege in der Bereitstellung von Grundlageninformationen für eine sinnvolle Bewirtschaftung und Planung (Haefner [1997]), die besonders interessant in Entwicklungsregionen sind (Ehlers [1995b]).

Tabelle 1: Prozentsatz Kartenbedeckung in Südamerika, nach Brandenberger & Ghosh [1990].

Region	Maßstabsbereich 2 1:40.000 bis 1:75.000 ($\sim 1 : 50.000$)	Maßstabsbereich 3 1:100.000 bis 1:126.726 ($\sim 1 : 100.000$)
Südamerika	30	53
Argentinien	17	50
Brasilien	14	68
Chile	98	0
Ecuador	56	34
Peru	14	46

Die Auswertung von Satellitenbildern basiert meistens auf den pixelbezogenen spektralen Eigenschaften der Bildelemente. Üblicherweise wird bei der Klassifizierung von Satellitenbildern angenommen, daß die spektralen Eigenschaften ausreichen um die Klassen bzw. Objekte im Bild voneinander zu trennen. Die Erkennung von Objekten in Bildern ergibt sich in der Praxis nicht nur aufgrund ihrer spektralen Eigenschaften, sondern auch aufgrund anderer Merkmale, wie Größe, Form oder Nachbarschaftsbeziehungen (Segl [1996], Quint [1997], Brügelmann [1998]). Die Anzahl und Form der Objekte in einem Bild ist vom Aufnahmesystem und den Aufnahmebedingungen abhängig. Solche Informationen werden in der Regel nicht direkt bei der multispektralen Analyse berücksichtigt, sondern später durch die Interaktion eines Menschen in den Prozeß eingeführt. Diese Tatsache ist in der Bildanalyse bekannt, und die Einbeziehung neuer Merkmale in den Analyseprozeß wird momentan untersucht (Bähr [1998]).

Die vielen verschiedenen Methoden zur Erfassung der Erdoberfläche ermöglichen ihre Beobachtung aus unterschiedlichsten Perspektiven, wobei jeweils bestimmte Merkmale in den Vordergrund treten. Die vorliegende Arbeit bezieht sich auf die Problematik der Kombination unterschiedlicher Datenquellen zur Gewinnung einer Beschreibung der Erdoberfläche zu einem bestimmten Zeitpunkt. Den Schwerpunkt bildet dabei die gemeinsame Verarbeitung von gescannten Karten und Satellitenbildern.

Bereits zur Verfügung stehende Karten im Maßstab 1:50.000 bilden dabei die Grundlage. Da immer die Möglichkeit besteht, daß diese Daten veraltet sind und eventuell nicht mehr der Wirklichkeit entsprechen, bedarf die Verknüpfung dieser Informationen mit aus Satellitenbildern gewonnenen Daten einer sorgfältigen Analyse.

Das gewählte Untersuchungsgebiet liegt in Südamerika, genauer in Brasilien (Abb. 1). Da von diesem Gebiet nur sehr wenige aktuelle Karten existieren, ist die Notwendigkeit zur Benutzung von Satellitenbildern noch sehr viel größer. Zwei Datenquellen werden untersucht. Zum einen Satellitenbilder (Landsat und Spot) und zum anderen Karten des IBGE (Instituto Brasileiro de Geografia e Estatística).



Abbildung 1: Brasilien (<http://edcwww.cr.usgs.gov>).

1.1 Motivation und Problemstellung

Aufgrund der rasanten technologischen Entwicklung der letzten Jahre, erlebten die Methoden der Bildverarbeitung tiefgreifende Veränderungen, die sich auf die Art der Problemlösung in der Fernerkundung auswirkten. Ehlers [1995b] sieht die rasante Hardware- und Software-Entwicklung im GIS-Bereich (Geographische-Informationssysteme), die wachsende Verfügbarkeit an hochauflösenden Sensordaten und die Fortschritte in der Automation auf dem Feld der Bildanalyse als die bedeutendsten Ursachen an. Vor zehn Jahren bestand das Hauptproblem in der Fernerkundung darin, geeignete Anwendungsmöglichkeiten für die neuen Daten zu finden, die mit vorhandenen technischen Mitteln zu bewerkstelligen waren. Den Wissenschaftlern der Gegenwart stellt sich das Problem der anwachsenden Datenmengen, ihrer Kombination und der Auswahl des besten Datentyps zur Lösung eines bestimmten Problems. Dafür werden neue Werkzeuge wie z.B. GIS-Systeme (raumbezogene Informationssysteme) entwickelt, die Werkzeuge zur Speicherung und Verarbeitung unterschiedlicher Datentypen wie z.B. Satellitenbildern, sozio-ökonomischer Daten, Karten und anderen Meßdaten bieten (Ehlers [1995a]).

Der Begriff „Fernerkundung“ umfaßt die Methoden und Technologien die benutzt werden, um Daten über die Erdoberfläche durch Messung und Interpretation der von ihr reflektierten oder emittierten elektromagnetischen Energie zu erfassen (DIN 18716/3 in Bähr & Vögtle [1998]). Obwohl es möglich ist, die Sensoren auf verschiedenen Plattformen zu installieren (z.B. auf Flugzeugen), ist die Aufnahme von Satelliten aus die am häufigsten unter Fernerkundung verstandene Einsatzart. Durch die systematische Messung der Energie innerhalb eines Spektralbereiches werden Bilder in verschiedenen

spektralen Kanälen aufgenommen. Heutzutage steht eine große Vielfalt an Fernerkundungssystemen zur Verfügung. Es werden z.B. Bilder mit Hilfe optischer Sensoren und RADAR aufgenommen.

Bähr [1984] erwähnt, daß die „anfängliche Euphorie“ der Automatisierung im Bereich der Bildklassifikation schnell sank, da die Klassifikationsergebnisse für die Nutzer „vielfach enttäuschend blieben“. Die Annahme, daß die Fernerkundung allein alle Probleme lösen kann, ist falsch, denn zur Auswertung dieser Daten sind zusätzliche Informationen und Wissen über die Natur der Objekte und ihrer Umwelt notwendig.

Die Fernerkundung wird oft als ein optimales Werkzeug zur Datenerfassung in Entwicklungsländern angesehen. Satellitenbilder liefern eine aktuelle, aber begrenzte Beschreibung der Oberfläche unseres Planeten. Die Daten der Fernerkundung spiegeln die Variation der spektralen Eigenschaften der Erdoberfläche wider. Der Nutzer dieser Daten ist nicht an den Daten selbst interessiert, sondern verlangt vielmehr Informationen zu einem bestimmten Thema, wie zum Beispiel dem Zustand der Vegetation oder der Fruchtbarkeit eines Bodens. Der Übergang „von Daten zu Information“ erfolgt durch die Interpretation, basierend auf Kenntnissen und Erfahrungen eines menschlichen Experten.

Raumbezogene Informationen (in Form von Karten oder GIS-Datenbanken) bilden heutzutage die Grundlage für Raumplanung und Umweltschutz. Diese Daten können die Bildinterpretation unterstützen, denn sie beschreiben andere Merkmale derselben Oberfläche. Der Ansatz, Methoden der künstlichen Intelligenz zur Analyse von Satellitenbildern und zur Integration dieser mit raumbezogenen Daten einzusetzen, wird momentan untersucht. Bähr [1998] stellt verschiedene Methoden zur Repräsentation von Wissen zur Unterstützung einer automatischen Bildauswertung vor, die eine menschliche Interaktion teilweise ersetzen sollen. Wegen der großen Vielfalt an Arten von digitalen Bildern (Satellitenbilder, Radarbilder, gescannte Karten) wurde eine allgemein einsetzbare Methode zur Auswertung bildhafter Informationen noch nicht gefunden, aber Bähr [1998] zeigt, daß die Einbeziehung von externem Wissen dazu beitragen kann, einige Fragen bei der Suche nach einer automatischen Bildanalyse zu lösen. Eine dieser Fragen bezieht sich auf die Kombination von Satellitenbildern mit anderen raumbezogenen Daten.

Faktoren wie Geologie, Bodentyp, Topographie, Bevölkerung und Infrastruktur beeinflussen das Erscheinungsbild der Erdoberfläche (Bronsveld *et al.* [1994]), deshalb basiert die Bildinterpretation im Grunde genommen auf der Detektion dieser einzelnen Wechselbeziehungen. Durch die Kombination aus Daten und Kenntnissen über die Wechselbeziehungen zwischen einzelnen Faktoren in einem Gebiet (Wissen) werden Informationen abgeleitet, welche die Interpretation eines Bildes unterstützen. Hierbei werden visuelle Analyse oder computergestützte Methoden benutzt. Die multispektralen Daten werden in Form von Karten (Rasterbilder oder Vektordaten) oder in statistischen Tabellen abgelegt, in denen zum Beispiel die Lage und Verteilung der Objekte in der Szene beschrieben werden. Die Interpretation eines Bildes ist auch vom Ziel abhängig, denn für landwirtschaftliche Planung werden andere Informationen als für ökologische Forschungen benötigt.

Da man oft nicht ausreichende theoretische Kenntnisse der Beziehungen zwischen den Informationen und den Daten besitzt, sind nach Shea [1991] flexible Modelle und anpassungsfähige, intelligente Algorithmen nötig, um das Wissen formal auszudrücken und auszunutzen. Der Umfang des enthaltenen Wissens spielt bei der Modellierung eine entscheidende Rolle. Der Aufbau und die Verwaltung der geo- und kartographischen Wissensbasis bildet die Grundlage der Interpretation der Daten (Shea [1991]), und der Zweck der Analyse bestimmt ihre Struktur.

In Entwicklungsländern ist die Verfügbarkeit einer aktuellen geographischen Datenbasis nicht immer gegeben, und oft sind veraltete Karten die einzigen Informationsquellen. Trotzdem muß man nicht auf die Integration dieser Daten in die Bildanalyse verzichten, da nicht unbedingt alle Informationen in einer Karte veraltet sind (manche Objekte bleiben unverändert). Es werden deshalb Methoden gesucht die es erlauben Daten auszuwählen, ihre Aktualität und Bedeutung zu bewerten und sie dann im Bildanalyseprozeß gewichtet zu integrieren.

Gesucht wird ein Verfahren zur Klassifizierung der Erdoberfläche ausgehend von Satellitenbildern und

von bereits zur Verfügung stehenden Daten. Das Grundproblem in diesem Zusammenhang besteht darin, basierend auf Informationen aus unterschiedlichen Quellen, jedes Bildelement (oder Cluster von Elementen) einer Klasse (oder Bedeutung) zuzuweisen. Da bei jeder Art von Datengewinnung andere Aspekte der Realität registriert werden ist es nötig, Hinweise auf Zugehörigkeit zu der Klassen aus den unterschiedlichen Datenquellen abzuleiten und sie später zu kombinieren, so daß am Ende eine Beziehung zwischen den Pixeln und den Landnutzungsklassen besteht.

1.2 Zielsetzung

Ziel dieser Forschungsarbeit ist die Beantwortung der Frage, wie Satellitenbilder und Karten, zwei Datentypen die üblicherweise in Entwicklungsländern vorhanden sind, zusammen genutzt werden können um Informationen über die Landnutzung zu gewinnen. Da die Karten in Entwicklungsländern meist in Form gedruckter Blätter zu finden sind, werden auch Methoden untersucht die helfen können, die notwendige Information aus diesen Blättern abzuleiten. Eine vollständige Erkennung des Karteninhaltes wird nicht untersucht, denn nicht alle graphischen Elemente der Karte sind relevant für die Bildanalyse, sondern es wird nur nach Hinweisen auf Zugehörigkeit zu der Landnutzungsklassen gesucht, die zur Unterstützung der Auswertung von Satellitenbildern dienen können.

1.3 Gliederung der Arbeit

Die Arbeit ist folgendermaßen gegliedert: In Kapitel 3 sind die Grundlagen der in der Literatur bekannten Verfahren zu finden, die eine gemeinsame Verarbeitung von Satellitenbildern und Karten betreffen. Kapitel 4 hat die Bearbeitung der gescannten farbigen Karten auf der Signalebene zum Thema. In diesem Kapitel werden Verfahren zur Trennung der einzelnen Farbschichten, so wie dem Erhalt der räumlichen Verteilung unterschiedlicher Landnutzungsklassen beschrieben. Kapitel 5 hat die Bearbeitung der Satellitenbildern zum Thema. Die Methoden zur gemeinsamen Analyse von Karten und Satellitenbildern werden in Kapitel 6 beschrieben. Diese Methoden basieren auf unscharfer Logik. Eine Option zur Schätzung der unscharfen Funktionen, die im Fuzzy-Inferenzschema teilnehmen, wird in Kapitel 7 eingeführt. Im darauffolgenden Kapitel werden die entsprechenden Experimente und Ergebnisse mit Daten aus Brasilien beschrieben. Die Arbeit wird mit einer Diskussion der Ergebnisse, der Schlußbetrachtung und einem Ausblick abgeschlossen.

2 Verwendete Datensätze

Die Analyse des Zusammenhanges zwischen Mensch und Umwelt ist eine sehr komplizierte Aufgabe, die eine große Menge an Informationen verlangt. Diese Information ist in Industrieländern leichter zu erhalten, weil in diesen Ländern die systematische Datensammlung als ernste Aufgabe angesehen wird. Dagegen stehen die notwendigen Daten, wie zum Beispiel konventionelle Datenerhebungen und Karten, in Entwicklungsländern oft nicht zur Verfügung. Sind Datenerhebungen vorhanden, dann sind sie oft veraltet und lange nicht mehr aktualisiert. Neue Informationsquellen werden deswegen untersucht und benutzt. Dazu zählen Satellitenbilder, die systematisch aufgenommen und gespeichert werden.

Unter diesen Bedingungen sind für die Erfassung der Erdoberfläche vor allem Satellitenbilder, Karten und GIS-Daten von Bedeutung, denn sie beschreiben flächenhafte Veränderungen. Für die Arbeit wurden Daten aus einem Gebiet in Süd-Brasilien, in der Nähe der Stadt Porto Alegre, benutzt. Es handelt sich um den unteren Teil Einzugsgebietes des Rio dos Sinos zwischen Porto Alegre und Novo Hamburgo in Bundesland Rio Grande do Sul (Abb. 2). In diesem Gebiet bestehen Nutzungskonflikte zwischen Landwirtschaft, Industrie und Siedlungen, die zu Veränderungen in der Bodennutzung führen. Bei dieser Arbeit stehen zwei Satellitenbilder zur Verfügung, eine multispektrale Aufnahme des Landsat-TM und eine panchromatische Spot-Szene. Daneben wird kartographische Information aus einer Karte (im Maßstab 1:50.000) benutzt.

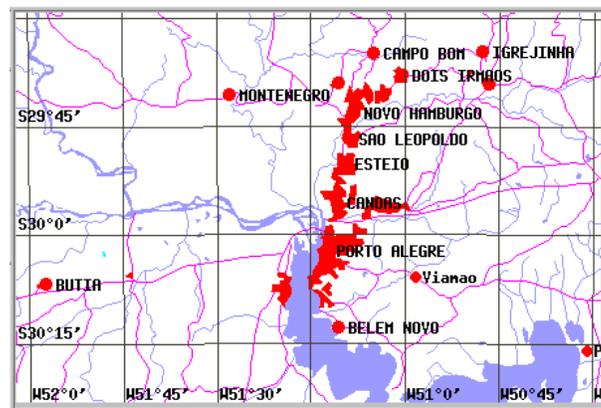


Abbildung 2: Unterer Teil des Rio-dos-Sinos-Einzugsgebietes in Brasilien (<http://edcwww.cr.usgs.gov>).

2.1 Satellitenbilder

Satellitenbilder ermöglichen eine flächenhafte Erfassung der Erdoberfläche, sind aber auf pixelbezogene spektrale Eigenschaften beschränkt. Sie bestehen aus einer Menge von Bildelementen, die Daten über die reflektierte oder emittierte Strahlung der Erdoberfläche speichern. Diese Messungen werden in engen Bereichen des elektromagnetischen Spektrums gemacht.

Nach Konecny [1990] lassen sich drei Erderkundungssystemgruppen, nach ihren Anwendungsmöglichkeiten, unterscheiden: Die erste Gruppe bilden Satellitensysteme für meteorologische und ozeanographische Zwecke. Sie haben zwar nur eine niedrige räumliche Auflösung, bieten dafür aber die Möglichkeit, die Bilder in kürzeren Zeitintervallen zu übermitteln. Die zweite Gruppe bilden Erderkundungssatelliten, deren räumliche Auflösung in der Größenordnung von 20 m liegt. Diese bieten den Vorteil einer höheren radiometrischen Auflösung, haben dafür allerdings einen längeren Wiederholungszyklus. In der dritten Gruppe werden Sensorensysteme zusammengefaßt, die eine höhere räumliche Auflösung (von etwa 10m) bieten. Diese Systeme haben eine geringe spektrale und zeitliche Auflösung und werden bevorzugt zur Herstellung topographischer Karten verwendet.

In Abb. 3 werden räumliche und spektrale Auflösung sowie die Wiederholungsrate der heute verfügbaren Typen von Fernerkundungssatelliten verglichen. NOAA und Meteosat sind Satellitensysteme für meteorologische und ozeanographische Zwecke, während Luftbilder sich am besten für die Herstellung topographischer Karten eignen.

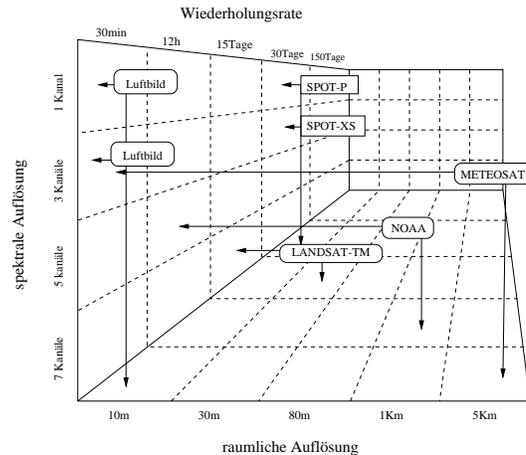


Abbildung 3: Vergleich der Wiederholungsrate und räumlichen und spektralen Auflösung von Fernerkundungssatelliten (aus Chuvieco [1990]).

Die Systeme Landsat und Spot gehören zur zweiten Gruppe. Obwohl das Landsat-System nicht die beste räumliche Auflösung bietet, ist es wegen seinem historischen Archiv, der guten Verfügbarkeit und guten spektralen Auflösung das verbreitetste System zur Beobachtung der Erdoberfläche. Dieses System ist ein Abtastersystem, das eigentlich für die thematische Bestandsaufnahme konzipiert wurde. Aufgrund der spektralen Unterschiede in den verschiedenen Kanälen (7) lassen sich die Elemente auf der Erdoberfläche, z.B. Vegetation, Gewässer oder Siedlungen, relativ einfach identifizieren. Das Spot-System hat weniger spektrale Kanäle als Landsat, aber sein größter Vorteil liegt in seiner besseren räumlichen Auflösung, besonders im panchromatischen Kanal, der erlaubt kleinere Objekte oder feinere Texturen in den Bildern zu erkennen (Chevres *et al.* [1981]).

Die Datenbasis für die multispektrale Bildverarbeitung in dieser Arbeit enthält Satellitenbilder der Serie Landsat-TM (WRS221.81B-04, BSQ12345-7) vom 28.08.89 und Daten des panchromatischen Sensors von Spot (KJ 710-410/0 PAN vom 16.09.95). Diese Bilder wurden mit Hilfe von Passpunkten, die in den Bildern deutlich zu erkennen sind, und deren Koordinaten aus einer Karte 1:50.000 abgegriffen wurden, entzerrt. Die Korrektur der geometrischen Verzerrungen geht der Ableitung kartographisch relevanter Größen aus den Bildern voraus. Durch diese Korrektur werden die verschiedenen Bilder in eine vorgegebene Kartenprojektion überführt.

2.2 Karten

Räumliche Information in Form von Karten (analog oder digital) wird üblicherweise für Landnutzungsmanagement gebraucht. Die Nutzung von Karten zur Bildanalyse erfordert, daß der Rechner die kartographische Darstellung automatisch interpretieren kann. Hierbei treten folgende Probleme auf:

1. Die Darstellung der Objekte in der Karte entspricht nicht genau der Lage und Geometrie in der realen Welt, da die Abbildung der Objekte der realen Welt in die Karte eine Generalisierung erfordert. Man kann die Generalisierung in zwei Kategorien teilen: (a) die statistische Manipulierung der Daten, die die Vereinfachung und Klassifizierung der Elemente umfaßt, und (b) die Verwendung von Symbolen („Versinnbildlichung“) (Morrison [1980]). Nachteile der Generalisierung sind, daß die Geometrie der Karte leicht verändert wird, und daß die Elemente, die man auf

der Karte erkennt, nicht unbedingt in ihrer wirklichen Form dargestellt werden und deswegen nicht der Form, die man in einer senkrechten Aufnahme sieht, entsprechen.

2. Die auf einer Karte gezeichneten Grenzen zwischen Objekten werden vom menschlichen Auswerter aufgrund seiner Erfahrung zur Generierung einer deutlichen und verständlichen Repräsentation der Welt gebildet. Oft widerspricht diese Darstellung der Wirklichkeit, denn diese Grenzen sind nicht immer scharf. Die Veränderung der Lagegenauigkeit zugunsten einer besseren Lesbarkeit (Hake & Grünreich [1994]) spielt bei der Kombination von Daten auch eine entscheidende Rolle.
3. Objekte werden in einer Karte nominal beschrieben. Ihre Eigenschaften werden nicht explizit genannt, sondern mit Hilfe von Farben und Symbolen. Es wird angenommen, daß die Objekte spektral homogen sind. Diese Annahme widerspricht der Natur, ist aber eine Abstraktion, die für die Vereinfachung der Lesbarkeit einer Karte wichtig ist (Hake & Grünreich [1994]). Räumliche Variationen der Objekte auf der Erdoberfläche, zum Beispiel die Dichte der Vegetation oder Farbe des Bodens, werden in der Karte nicht explizit dargestellt, haben aber großen Einfluß auf die spektralen Daten der Satellitenbilder. Von einem statistischen Standpunkt aus gesehen, hat diese nominale Darstellung auch den Nachteil, daß die Klassen nicht normalverteilt sind. Deswegen ist die Anwendung mancher statistischen Ansätze nicht möglich.

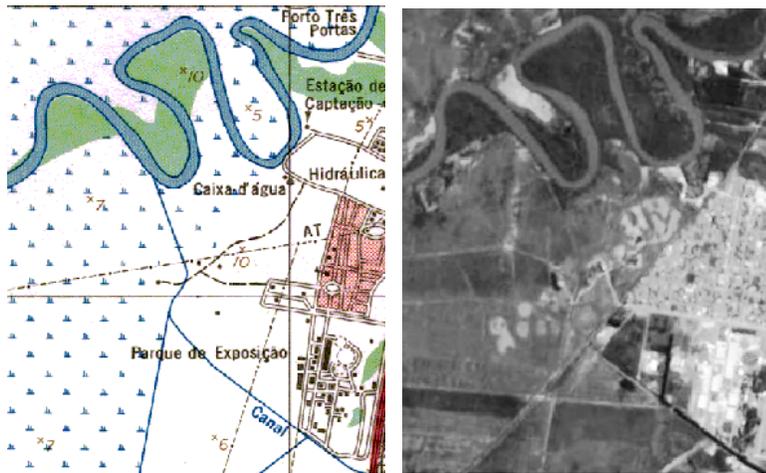
Die topographische Karten, die in dieser Arbeit benutzt werden, liegen im Maßstab 1:50.000 vor. Sie werden gemäß den Normen vom IBGE (Instituto Brasileiro de Geografia e Estatística) aus Luftbildern hergestellt. Jede Karte deckt ein Gebiet von 24,1x27,7 km ab. Für ihre Herstellung im Jahr 1978 wurden hauptsächlich Luftbilder photogrammetrisch und stereoskopisch ausgewertet. Die Befliegung zur Aufnahme der Bilder fand im Jahre 1975 statt. Sie wurde durch terrestrische Zusatzmessungen aus dem Jahr 1976 ergänzt. Ihr Inhalt beschreibt die Situation der Region 1975, mit Höhenlinien (20 m Äquidistanz) und Information über die Landnutzung: Gewässer, Straßennetz, Siedlungsgebiete und Wälder. Die Karten enthalten auch andere Informationen in Form von Texten und Ziffern, sowie Bahnlinien und andere weitere geographische Informationen. Ein Gitter von 2 km, ermöglicht das Ablesen der geographischen Position der Punkte im UTM-System.

Diese Karten sind teilweise veraltet, besonders im Bereich zwischen Porto Alegre und São Leopoldo, in Rio Grande do Sul, Brasilien, wo die Landschaft in den letzten Jahren stark verändert wurde. Ein Ausschnitt der Folie *São Leopoldo* wird in Abb. 53 gezeigt. Abb. 4 zeigt einen Vergleich zwischen einer kleinen Region dieser Karte und dem panchromatischen Spot-Bild von 1995. Auf diesen Abbildungen kann man sehen, daß die Vegetationsflächen (dunklere Flächen im Satellitenbild) große Veränderungen erfuhr und daß die Stadt am rechten Rand der Karte (Abb. 4a) gewachsen ist. Ähnliche Situationen sind vor allem in der Nähe der bereits existierenden Städte auf der ganzen Karte zu finden.

2.3 Entzerrung

Für die Kombination der Informationen auf Rasterbasis ist die Entzerrung der Bilder und Karten auf eine gemeinsame geometrische Basis nötig. Ein Rasterbild mit 10 m Pixelgröße wurde als Basis gewählt. Diese Größe entspricht der Pixelgröße eines Spot-Bildes. Eine erste geometrische Korrektur wurde mit Hilfe von Passpunkten auf dem Spot-Bild durchgeführt. Danach wurde das Landsat Bild auf dieses neue Bild entzerrt. Dafür wurden Punkte benutzt, die auf dem Spot-Bild und auf dem Landsat-Bild deutlich zu erkennen sind.

Die Karten wurden zuerst mit einer Pixelgröße von 85 μm abgetastet. Diese Größe entspricht 2,9 m auf der Erdoberfläche und ist im Vergleich zur Auflösung der Satellitenbilder sehr hoch. Die Kartenerkennung wurde mit dieser Pixelgröße durchgeführt. Die Größe des Pixels ist eine entscheidende Frage bei der Kartenerkennung, denn sie bestimmt die Qualität der Ergebnisse. Kleinere Pixel erlauben mehr Details und bessere Trennung der graphischen Elemente, aber erzeugen auch eine große Menge



(a) Karte

(b) SPOT-Szene

Abbildung 4: Unterschiedliche Datenquellen: (a) Analoge Karte und (b) panchromatisches SPOT-Bild.

an Daten. Nach der Erkennung der graphischen Elemente der Karte wurden die Ergebnisse auf das Spot-Bild entzerrt.

3 Theoretische Grundlagen

3.1 Modelle für das Bildverstehen: Szene vs. Bild

Daten, auch Bilder der Fernerkundung, vermitteln nur eine Projektion der „wirklichen Welt“, die bestimmte Merkmale der „Realität“ (die Szene) beschreibt. Satellitenbilder und gescannte Karten sind beide ikonische Darstellungen derselben Realität (Quint [1997]). Ihre Interpretation bezieht sich deshalb nur auf bestimmte Eigenschaften der realen Szene und kann nicht als vollständige Beschreibung betrachtet werden. Zum Beispiel werden räumliche Beziehungen, die in der Wirklichkeit dreidimensional sind, in den Bildern in nur zwei Dimensionen dargestellt. Deswegen können Beziehungen im Bild nicht direkt in Beziehungen der Szene übersetzt werden. Zwei *benachbarte* Objekte in einem Bild liegen in der Wirklichkeit nicht unbedingt *nebeneinander*.

Die Interpretation eines Bildes (das Auffinden seiner Semantik) ergibt sich durch ein Modell (Abb. 5). In diesem Modell werden die Beziehungen zwischen den Objekten und deren Geometrie sowie anderen Merkmalen, die für die Interpretation von Bedeutung sind, abgelegt. Ein allgemeines Modell wird in der Praxis nicht angewandt, sondern ein Modell, das der besonderen „Situation“ im Augenblick der Erfassung des Bildes, oder der Daten für die Kartierung, entspricht. Diese „Situation“ besteht in der Fernerkundung zum Beispiel in den atmosphärischen Bedingungen, dem Abstand zwischen Sensor und Objekt oder dem Aufnahmewinkel. Die Berücksichtigung der besonderen Bedingungen ist für die Interpretation der spektralen Daten besonders wichtig. Ein passendes Modell ist aber in der Regel dann schwierig zu finden, wenn Daten aus unterschiedlichen Quellen vorliegen. Eine Karte beschreibt die Oberfläche der Erde im allgemeinen nicht zum Zeitpunkt der Aufnahme der Satellitenbilder.

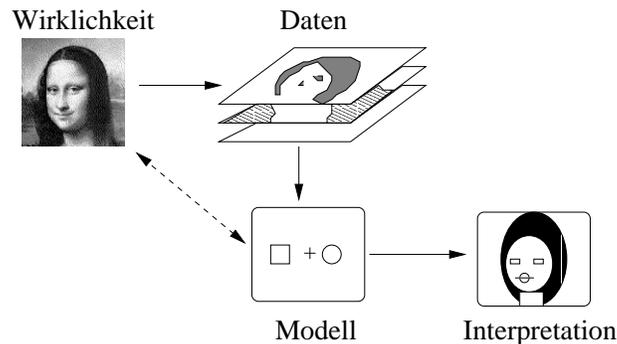


Abbildung 5: Interpretation eines Bildes.

Das Modell zur Analyse einer Szene umfasst drei Wissensebenen (Kanade [1980]). Sie werden *semantische Ebene*, *physikalische Ebene* und *Signalebene* (oder *ikonische Ebene*) genannt. Die oberste Ebene, die *semantische Ebene*, bezieht sich auf die semantische Interpretation der Szene im Bild. Sie umfasst die Beschreibung der Objekte aufgrund der Geometrie und der Beziehungen zwischen ihren Komponenten in der Szene. In der semantischen Ebene wird ein Bild als Ensemble von realen *Objekten* behandelt. Diese Ebene wird von Förstner [1992] als „Objektebene“ bezeichnet. Beispielsweise würde man eine Pyramide durch ihre Dimension, die Anzahl von Seiten und Winkeln und andere Eigenschaften (Farbe, Textur) beschreiben. Die physikalische Ebene, von Förstner [1992] auch Merkmalebene genannt, besteht aus einer strukturellen Beschreibung des Bildes, die normalerweise nach der Phase der Mustererkennung aufgebaut wird.

Im Gegensatz dazu umfasst die Signalebene die erfassten Daten und die Beziehungen zwischen diesen Beobachtungen, z.B. Grauwerte der Pixel und Nachbarschaftsbeziehungen wie die Werte der Korrelation zwischen Pixeln im Bild. Die Signalebene ist vom Datenaufnahmesystem oder der Art der Darstellung der Daten abhängig. Zum Beispiel wird eine Linie in einem Rasterbild als eine dünne

Fläche dargestellt, denn die minimale Einheit eines Rasterbildes ist ein rechteckiges Pixel. Die ursprüngliche gerasterte Struktur, in der die Daten erfaßt werden, kennzeichnet die Signalebene eines digitalen Bildes (Förstner [1992]).

Auf der *physikalischen Ebene* versucht man die Beziehungen zwischen der *semantischen Ebene* und den Bilddaten festzustellen. Im Fall des Satellitenbildes spielen die Eigenschaften der Objekte, die Aufnahmewinkel und die Beleuchtungsbedingungen eine große Rolle. Bei einer Pyramide würde man z.B. erwarten, daß diese als eine Menge von benachbarten Flächen, die verschiedene Intensitätswerte haben, abgebildet wird. Deshalb beginnt die Analyse auf der *Signalebene* mit der Suche nach benachbarten Pixeln gleicher Intensität.

Wald [1999], bei der Diskussion über eine Zusammenführung der Daten, bezeichnet drei andere Ebenen. Er spricht von Messungen, Merkmalen und Entscheidungsebenen (auf Englisch: *measure*, *attribute* und *decision*). Im Bezug zur Modellierung einer Szene in der Bildauswertung spricht Quint [1997] von drei Ebenen: einer *konzeptionellen Ebene*, einer *Repräsentationsebene* und einer *Implementierungsebene*, wobei die Repräsentationsebene und die Implementierungsebene zusammenfallen können.

Auf dem Gebiet der Datenintegration existieren unterschiedliche Terminologien. In dieser Arbeit wird die von Förstner [1992] vorgeschlagene benutzt, vor allem deshalb, weil die drei Kernbegriffe (Signal, Merkmal und Objekt) leichter auf die Bildanalyse übertragen werden können.

Die Benutzung unterschiedlicher Datenquellen zur Erfassung der Erdoberfläche verlangt die Anpassung des Modells an diese Daten auf der Signalebene und der Merkmalsebene. Die Merkmale, die zur Erkennung von Objekten in einem Bild oder einer Karte untersucht werden, unterscheiden sich je nach der Beschaffenheit der erfaßten Daten. Ein Sumpfgebiet wird in der Regel in einer Karte durch ein bestimmtes Muster repräsentiert, das in der Wirklichkeit nicht existiert und deswegen nicht im Bild vorkommt (Abb. 4). Dementsprechend ändert sich die Verarbeitung der Daten, die Segmentierung oder das Klassifikationsverfahren auf der Signalebene, je nach zugrunde liegendem Datenträger. Dadurch verändert sich auch die Analyse in der Merkmalsebene, da sie als Verknüpfung zwischen der Datenebene und der semantischen Ebene steht, also direkt von der Signalebene abhängt.

3.2 Methoden zur gemeinsamen Verarbeitung unterschiedlicher Daten

Satellitenbilder und Karten haben sich als Basis für die Analyse der Erdoberfläche als geeignet erwiesen (Hutchinson [1982]). Obwohl sie am Anfang getrennt behandelt wurden, stellte sich mit der Zeit immer deutlicher heraus, daß die Benutzung beider Datenquellen zusammen (Synergie-Effekt) zu einer Verbesserung der Analyse beitragen kann (Kunz *et al.* [1998]). Das Ziel der Kombination dieser Daten kann sehr unterschiedlich sein. Die Liste umfaßt die Bildanalyse, Change-Detection, automatische Passpunktfindung, Trainingsgebietsbestimmung oder Darstellung einer Szene (z.B. 3D-Darstellung). Das Thema der Landnutzungsklassifizierung ist ebenfalls ein Beispiel hierfür und wird in den folgenden Kapiteln behandelt. Beispiele für die Benutzung unterschiedlicher Karten gemeinsam mit Satellitenbildern sind in der Literatur beschrieben. Oft handelt es sich um Höhendaten, Neigung und Exposition (Hutchinson [1982]; Binaghi *et al.* [1997]; Peddle [1995]), Landnutzungsklassen (Bronsveld *et al.* [1994]; Solberg *et al.* [1996]; Moon [1990]), geophysische und geologische Daten (Moon [1990]), Bodentypskarten oder andere Geodaten (Burger [1992]; Bronsveld *et al.* [1994]).

Beispiele der Einbeziehung von raumbezogenen Daten bei größeren Maßstäben in der Bildanalyse können in Quint [1997] und Silva Centeno *et al.* [1999] gefunden werden. In beiden Fällen handelt es sich um die Erkennung von Gebäuden in Bildern (Luftaufnahmen oder Laserscanner-Daten) mit höherer räumlicher Auflösung.

Die Form, in der die raumbezogenen Daten in der Bildanalyse einbezogen werden, unterscheidet sich von Autor zu Autor. Die Unterschiede in den Begriffen *data fusion*, *concatenation* und *integration* und ihre Bedeutung bei der Bildanalyse werden von Baltasvias & Hahn [1999] und Wald [1999] diskutiert.

Oft sind diese Unterschiede nicht sehr groß, und man kann diese Begriffe als Synonyme betrachten (Baltsavias & Hahn [1999]). Der Begriff *Integration* bedeutet nach Baltsavias & Hahn [1999]:

„*Integration bedeutet, daß unterschiedliche Komponenten zusammenwirken und zu einem besseren Ergebnis führen, besser als das Ergebnis ohne Integration.*“

Dabei unterscheiden diese Autoren *Verkettung (concatenation)* und *Integration*, da die *Verkettung* die einfache Sammlung aller Daten in einem einzigen Merkmalsvektor darstellt, während die *Integration* nicht auf die Datenebene beschränkt ist. Drei Anwendungsarten der Integration von Satellitenbildern mit raumbezogenen Daten lassen sich gemäß Baltsavias & Hahn [1999] trennen: (a) Erhalt von a priori Information durch raumbezogene Daten. Dabei kann es sich um die automatische Bestimmung von Trainingsgebieten für eine Klassifizierung handeln. (b) Die Benutzung der raumbezogenen Daten in einer Vor- oder Nachbearbeitungsphase der Satellitenbildanalyse. Einfache Ansätze wie die „*Schichtung*“ der Fernerkundungsdaten anhand raumbezogener Daten (Hutchinson [1982]) sind leicht anwendbar und vereinfachen das Problem der Klassifizierung, indem sie nur die Klassen, deren Trennung aufgrund der spektralen Merkmale nicht möglich ist, mit Hilfe der Information einer Karte unterscheiden. (c) Die dritte Möglichkeit ist die totale Überlagerung dieser Daten im Analyseprozeß (z.B. Quint [1997]).

Die in der Literatur beschriebenen Ansätze für eine gleichzeitige Verarbeitung der gesamten Information aller Quellen sind unterschiedlich, obwohl eine Tendenz in Richtung der Evidenztheorie zu erkennen ist (Moon [1990]). Die Repräsentation der Ungewißheit in der Evidenztheorie wird als ihr größter Vorteil bezeichnet (Lee *et al.* [1987]). Dieser Vorteil ist besonders wichtig, wenn Teile der Information nicht zur Verfügung stehen, z.B. in Regionen der Karte, in denen keine bedeutende Information dargestellt wurde, oder in Regionen, wo Satellitenbilder nicht genügend Hinweise liefern. Die Integration von Hinweisen anstatt der Daten selbst, wurde von Baltsavias & Hahn [1999] als ein Thema, das immer mehr die Aufmerksamkeit der Wissenschaftler erweckt, bezeichnet.

Die Umwandlung einer gescannten Karte in eine Informationsebene im Analyseprozeß verlangt umfangreiche menschliche Interaktion. Das gilt auch für die Kombination dieser Information mit den Bildern. Oft werden Regeln aufgrund der menschlichen Kenntnisse und Erfahrung vorgeschlagen (Moon [1990]; Bronsveld *et al.* [1994]). Die entscheidenden Zugehörigkeitsfunktionen (Fuzzy-Logik, Evidenztheorie), die die graphische Information einer Karte mit den Landnutzungsklassen verbinden, entsprechen Kenntnissen, die nicht explizit in Karten dargestellt werden, sondern von Besonderheiten wie der Landwirtschaft des Gebietes oder der Phänologie der vorhandenen Vegetation abhängen.

Die Aufgabe Satellitenbilder und den Inhalt einer Karte zu kombinieren wird von mehreren Faktoren eingeschränkt. Gemäß Baltsavias & Hahn [1999] können folgende Probleme zitiert werden:

- Der Unterschied zwischen der Information über die Landnutzung eines GIS oder Karte (eng: *land use*) und das Aussehen der selben Oberfläche in einem Bild (Landbedeckung eng: *land cover*),
- Mängel bei der Qualitätskontrolle der Daten,
- unterschiedliche geometrische Genauigkeit,
- Unsicherheit über die Klassifizierung,
- Generalisierung,
- unterschiedliche Darstellung und Abstraktionsniveau (Auflösung und Maßstab),
- Anwendung unterschiedlicher Modelle bei der Analyse einzelner Datensätze und
- niedrige Homogenität der Daten.

Eine weitere Frage, die in diesem Fall eine entscheidende Rolle spielt, ist, wie groß die menschliche Teilnahme im Prozeß sein soll. Ein menschlicher Bearbeiter hat den Vorteil, daß seine Entscheidung auf den Daten, seiner eigenen Erfahrung, Wissen und Gefühl basiert, während bei einer rechnerischen Analyse die Daten und bestimmte Regeln überwiegen. In den letzten Jahren wurde versucht, immer mehr Entscheidungen dem Rechner zu überlassen, so daß der Prozeß automatisch oder teilautomatisch stattfindet (Bähr [1998]).

3.2.1 Möglichkeiten zur Integration von GIS und Satellitenbildern

Raumbezogene Daten können die Analyse von Satellitenbildern an drei Stellen des Prozesses unterstützen (Hutchinson [1982]). Sie können vor der Analyse zur Aufteilung der spektralen Daten (*stratification*), zusammen mit den spektralen Daten im Klassifikationsschritt (*classification operations*) oder als Hilfsmittel zur Optimierung der Ergebnisse (*postclassification sorting*) benutzt werden.

„*Stratification*“ ist die Aufteilung der multispektralen Daten vor der Klassifizierung in einfachere Datenmengen (Abb. 6). Jede neu erzeugte Schicht, beispielsweise eine kleinere Fläche im Bild, kann danach getrennt verarbeitet werden. Das Ziel der Schichtung ist es, die Homogenität der Eingabedaten zu erhöhen und dadurch die Qualität der Klassifizierung zu verbessern. Man erhofft sich dadurch, daß unterschiedliche Klassen, deren Instanzen im Bild sehr ähnliche spektrale Eigenschaften haben, wodurch diese Klassen schwer trennbar sind und es zu Fehlklassifikation kommen kann, in den neuen Schichten leichter zu trennen sind. Als Basis für diese Trennung benötigt man eine Beschreibung der Szene, in der die Klassen sinnvoll und leicht zu trennen sind. Eine solche Beschreibung kann aus einer Karte gewonnen werden.

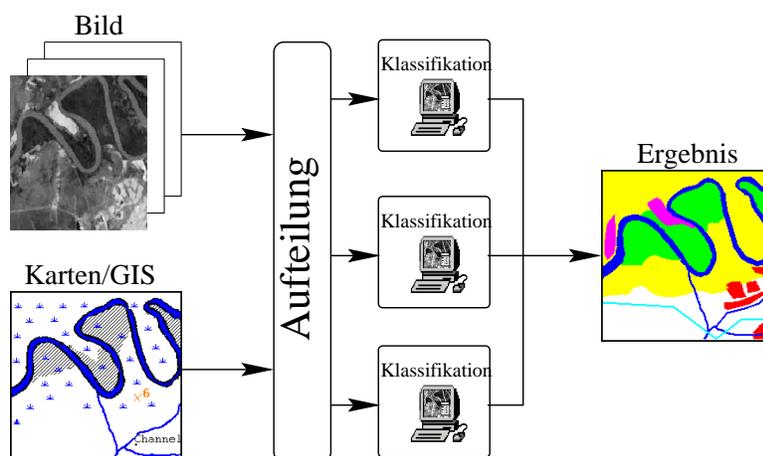


Abbildung 6: Prinzip der Schichtung der Daten.

Diese Operation ist sehr einfach, aber ihre Anwendung erfordert gute Kenntnis der Daten, da eine falsche Wahl der Kriterien unerwünschte Auswirkungen auf die Struktur der Daten haben kann. Die Grenzen der Aufteilung werden sehr genau definiert. Deswegen hat die Teilung der raumbezogenen Daten sehr großen Einfluß auf die Ergebnisse. Die spektralen Grenzen der Klassen werden mit Hilfe von Regionen im Bild, deren wahre Klassifikation bekannt ist (Trainingsprobe) genau definiert. Die Ergebnisse hängen stark von der gewählten Trainingsprobe und der Erfahrung des Anwenders ab.

Die Möglichkeit, raumbezogene Daten und Satellitenbilder gleichzeitig zu analysieren, bietet sich geradezu an (Abb. 7). Da alle Daten zusammen verarbeitet werden, wird die Anzahl der Merkmale erhöht und ihre Struktur komplexer. Das entspricht den Konzept der „*Verkettung*“ (Wald [1999]). Die Verarbeitung dieser neuen Datenmenge ist, wie erwartet, mühsamer und komplizierter als die Aufteilung der Fernerkundungsdaten vor der Klassifizierung (*stratification*). Manche Daten, z.B. binäre Daten aus einer gerasterten Karte, sind nicht normalverteilt, und die Anwendung statistischer Standardverfahren ist nicht ohne weiteres möglich. Der Hauptvorteil der gleichzeitigen Benutzung aller

Daten liegt darin, daß die Reihenfolge, in der die Daten benutzt werden, nicht im voraus festgelegt werden muß.

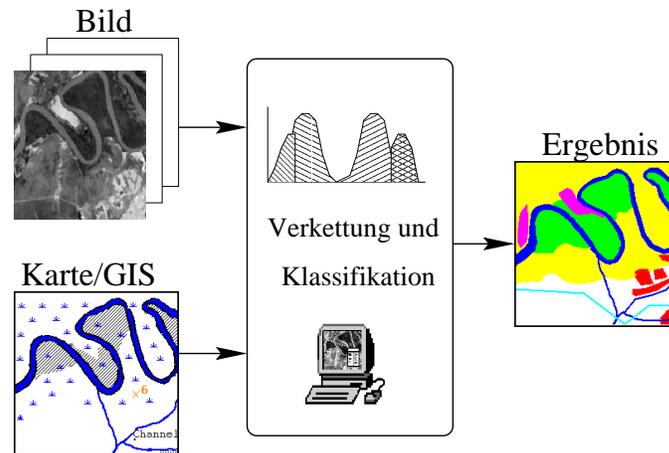


Abbildung 7: Gleichzeitige Analyse der raumbezogenen und spektralen Daten.

Die letzte Möglichkeit ist die Nachbearbeitung der Ergebnisse einer spektralen Klassifikation mit Hilfe des Karteninhaltes. In diesem Fall werden die Ergebnisse einer rein spektralen Analyse mit Informationen aus der Karte verfeinert (Abb. 8). Das Grundprinzip ist ähnlich wie bei der Vorverarbeitung der Daten (*stratification*), weswegen dies oft „*post-stratification*“ genannt wird. Die Klassen, die geringe spektrale Unterschiede haben und deswegen im Ergebnisbild zusammengefaßt werden, werden mit Hilfe des Karteninhaltes einfacher getrennt und ein neues Ergebnisbild, mit mehr Klassen, erzeugt.

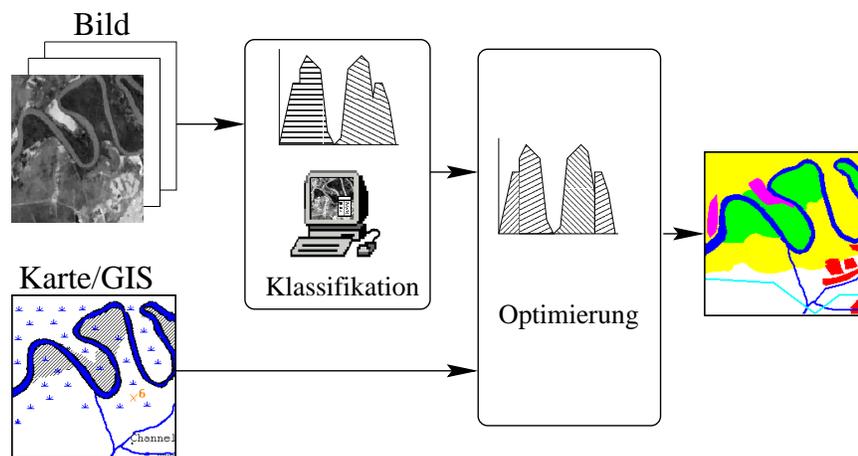


Abbildung 8: Benutzung der raumbezogenen Daten nach der Spektralanalyse der Bilder.

Die zweite Möglichkeit, die *Verkettung* der Daten in einem einzigen Vektor, ist eine komplexe Aufgabe, denn die Daten besitzen eigene Strukturen, die einen einheitlichen Ansatz verhindern. Faktoren wie die unterschiedliche geometrische Genauigkeit, Klassifikationsfehler, Generalisierung, Auflösung, Maßstab und Erfassungsmethoden müssen berücksichtigt werden. Die Ansätze, die diese Aufgabe erfüllen sollen, müssen flexibel sein. Drei mögliche Ansätze werden auf den folgenden Seiten behandelt: ein statistischer Ansatz, ein auf den Dempster-Shafer-Kombinationsregeln basierender Ansatz und ein auf unscharfer Logik (Fuzzy-Logic) aufbauender Ansatz. Sie werden in der Praxis oft benutzt, obwohl es auch andere Möglichkeiten gibt, die hier nicht behandelt werden, wie Blackboard Systeme, Semantische Netze und Frames. In (Baltasvias & Hahn [1999]) werden Beispiele zu diesen letztgenannten Ansätzen gegeben.

3.2.2 Statistischer Ansatz

Die Klassifizierung eines Pixels in eine der möglichen Klassen kann mit Hilfe statistischer Ansätze durchgeführt werden. Die Pixel $\mathbf{x} \in X$ werden nach bestimmten Kriterien in Klassen zusammengefaßt. Einer der bekanntesten statistischen Ansätze ist das Maximum-Likelihood-Verfahren. Dabei wird angenommen, daß die digitalen Werte normalverteilte Zufallsvariablen sind. Jedes Pixel wird durch einen Merkmalsvektor \mathbf{x} beschrieben. Im Fall eines Satellitenbildes enthält der Vektor \mathbf{x} die digitalen Werte der verschiedenen spektralen Kanäle im n -dimensionalen Merkmalsraum. Die n -dimensionale Wahrscheinlichkeitsdichte einer Klasse ($w_j \in W$, $W = w_1, w_2, \dots, w_m$) ist unbekannt und wird mit Hilfe von Trainingsgebieten geschätzt (Mittelwertvektor \mathbf{m}_j und Kovarianzmatrix \mathbf{C}_j):

$$p(\mathbf{x}|w_j) = (2\pi)^{-n/2} |\mathbf{C}_j^{-1}|^{-1/2} \exp \left\{ -\frac{1}{2} (\mathbf{x} - \mathbf{m}_j)^\top \mathbf{C}_j^{-1} (\mathbf{x} - \mathbf{m}_j) \right\} \quad (1)$$

Ein Pixel wird einer Klasse zugewiesen ($\mathbf{x} \in w_j$), wenn:

$$p(w_j|\mathbf{x}) > p(w_k|\mathbf{x}) \quad \forall j \neq k \quad . \quad (2)$$

Die gesuchte, bedingte *a posteriori* Wahrscheinlichkeit $p(w_j|\mathbf{x})$ kann, nach dem Bayes-Theorem, aus der Dichtefunktion $p(\mathbf{x}|w_j)$ und der *a priori*-Wahrscheinlichkeit $p(w_j)$ berechnet werden:

$$p(w_j|\mathbf{x}) = \frac{p(\mathbf{x}|w_j)p(w_j)}{\sum_k^m p(\mathbf{x}|w_k)p(w_k)} \quad , \text{ mit } k = 1 \dots m. \quad (3)$$

Der normierende Faktor $\sum_k^m p(\mathbf{x}|w_k)p(w_k)$ ist beim direkten Vergleich der Wahrscheinlichkeitsfunktionen identisch (Gl.1), deswegen ist er für die Entscheidung über die richtige Klassifizierung des Pixels (Gl.2) nicht relevant. Dem Pixel wird die Klasse w_j zugewiesen, wenn:

$$p(\mathbf{x}|w_j)p(w_j) > p(\mathbf{x}|w_k)p(w_k) \quad \forall j \neq k \quad . \quad (4)$$

Die Entscheidung über die Klassenzugehörigkeit basiert auf der verfügbaren Information. In der Fernerkundung wird oft angenommen, daß eine Unterscheidung aufgrund der Reflektions- und Austrahlungseigenschaften der Objekte ausreicht um Landnutzungsklassen zu trennen. In diesem Fall enthält \mathbf{x} die spektralen Messungen in verschiedenen Kanälen. Diese implizite Voraussetzung stimmt leider oft nicht. Es gibt Situationen, in denen die Landnutzungsklassen ähnliche spektrale Eigenschaften haben. Die Einbeziehung (Verkettung, Wald [1999]) neuer Merkmale aus anderen Informationsquellen (z.B. Karten) könnte die Trennung verbessern. Die Verkettung ist aber aufgrund der unterschiedlichen Natur der Daten und der Erhöhung der Rechenzeiten nicht besonders attraktiv. Eine weniger aufwendige Alternative ist es, zuerst die Beziehung zwischen dem Merkmalsvektor \mathbf{x} und den Klassen D_S in einem bestimmten Datensatz (Quelle S) anzunähern, diese Beziehung als eine Funktion darzustellen und später die Funktionen zu kombinieren. Ein statistischer Ansatz, der diese Integration erlaubt, wird im folgenden nach Lee *et al.* [1987] beschrieben.

Nicht alle Klassen w_j können in jedem Datensatz S erkannt und von den anderen getrennt werden. Manche Klassen, die auf einer Karte eingezeichnet sind, sind oft in einem Bild nicht erkennbar. Die Anzahl der Klassen d_{S_i} , die in einem Datensatz erkennbar sind, ist von den Merkmalen, die man aus dem Datensatz ableiten kann, abhängig. Die Information der Quelle S wird mittels dem spezifischen Merkmalsvektor \mathbf{x}_s dargestellt.

Die Funktion, welche die Beziehung zwischen den Pixeln und der i -ten Klasse d_{si} beschreibt, wird $f(d_{si}, \mathbf{x})$ genannt. Ein Beispiel ist in Gl. (3) angegeben. Für jeden Datensatz (Karte oder Bild) kann man ähnliche Funktionen schätzen. Ihre Kombination (F) ermittelt eine Beschreibung der Beziehungen zwischen den Klassen w_j und \mathbf{x} . Die Form der Funktion F_j ist von der Informationsquelle abhängig.

$$F_j = F_j\{f(w_j|d_{si}), \alpha\} \quad (5)$$

mit $i = 1, \dots, m_s$ Klassen und $s = 1, \dots, N$ Quellen. Der Faktor α beschreibt die Zuverlässigkeit der Quelle, die Funktion F_j den Grad der Beziehung zwischen den Pixeln und den Klassen. Die Klasse mit dem maximalen Wert wird dem Pixel zugewiesen. Diese globale Funktion kann gemäß der Bayes-Theorie berechnet werden (Lee *et al.* [1987]):

$$F_j(\mathbf{x}) = p(w_j|\mathbf{x}) = p(w_j|\mathbf{x}_1, \dots, \mathbf{x}_N) \quad (6)$$

Wenn man bedingte Unabhängigkeit zwischen Klassen annimmt gilt:

$$p(w_j|\mathbf{x}_1, \dots, \mathbf{x}_N) = p(\mathbf{x}_1|w_j) \cdots p(\mathbf{x}_N|w_j) \quad (7)$$

und

$$F_j(\mathbf{x}) = \frac{p(\mathbf{x}_1, \dots, \mathbf{x}_N|w_j)p(w_j)}{p(\mathbf{x}_1, \dots, \mathbf{x}_N)} \quad (8)$$

$$F_j(\mathbf{x}) = \frac{p(\mathbf{x}_1|w_j) \cdots p(\mathbf{x}_N|w_j)p(w_j)}{p(\mathbf{x}_1, \dots, \mathbf{x}_N)} \quad (9)$$

$$F_j(\mathbf{x}) = H \frac{p(w_j|\mathbf{x}_1) \cdots p(w_j|\mathbf{x}_N)p(w_j)}{p(w_j)^N}, \quad (10)$$

wobei

$$H = \frac{p(\mathbf{x}_1) \cdots p(\mathbf{x}_N)}{p(\mathbf{x}_1, \dots, \mathbf{x}_N)} \quad (11)$$

konstant für alle j ist, deswegen:

$$F_j(\mathbf{x}) = p(w_j|\mathbf{x}_1) \cdots p(w_j|\mathbf{x}_N)p(w_j)^{1-N}. \quad (12)$$

$F_j(\mathbf{x})$ wird auf Basis der einzelnen spezifischen, quellenbezogenen Funktionen berechnet:

$$p(w_j|\mathbf{x}_s) = \sum_{i=1}^N \frac{p(w_j, d_{si}, \mathbf{x}_s)}{p(\mathbf{x}_s)} \quad (13)$$

$$p(w_j|\mathbf{x}_s) = \sum_{i=1}^N \frac{p(\mathbf{x}|w_j, d_{si})p(d_{si}|w_j)p(w_j)}{p(\mathbf{x}_s)} \quad (14)$$

$$p(w_j|\mathbf{x}_s) = \sum_{i=1}^N p(w_j|d_{si}, \mathbf{x}_s)p(d_{si}, \mathbf{x}_s) \quad (15)$$

3.2.3 Dempster-Shafer-Kombinationsregeln

Statistische Ansätze, wie das Maximum-Likelihood-Verfahren, sind nicht zur Datenanalyse geeignet, wenn (Peddle [1995]):

- (a) der Merkmalsraum zu groß ist,
- (b) die Daten nicht statistisch modelliert werden können,
- (c) Daten aus unterschiedlichen Datenquellen, mit entsprechenden Unterschieden, Fehlern und Unsicherheiten vorhanden sind und
- (d) es grundsätzliche Unterschiede in der Datenerfassung gibt (nominale Daten vs. ordinale Daten).

Hier kann aber die Dempster-Shafer-Evidenztheorie angewandt werden (Shafer [1976]). Laut der Evidenztheorie kann die Beziehung zwischen jedem Pixel und einem Ereignis des Ereignisraumes $A \in \Theta$ durch einen positiven Verdacht, ausgedrückt in dem sog. *Belief-Maß* ($Bel(A)$), beschrieben werden, dessen Wert im Intervall $0 \leq Bel(A) \leq 1$ liegt. Der positive Verdacht wird als untere Grenze des Vertrauensmaßes betrachtet. Die obere Grenze ist durch die Plausibilität $Pl(A)$ gegeben. Die Summe $Bel(A) + Pl(A)$ ist nicht gleich eins und läßt dadurch Platz für die Modellierung einer gewissen Unsicherheit.

$$Bel(A) = \sum_{A_i \subseteq A} m(A_i) \geq Bel(A) \quad (16)$$

$$Pl(A) = 1 - Bel(\neg A) \quad (17)$$

Die Kombination zweier Hinweise $F1$ und $F2$ unterschiedlicher Datenquellen erfolgt nach den Kombinationsregeln von Dempster, die kommutativ und assoziativ sind. Aus zwei Maßfunktionen $m_1(F1)$ und $m_2(F2)$ wird ein Verdacht über die Zugehörigkeit eines Pixels zu einer Klasse $m_1 \oplus m_2$ berechnet.

$$m_1 \oplus m_2 = \frac{1}{1-K} \sum_{F1 \cap F2} m_1(F1)m_2(F2) \quad (18)$$

$$K = \sum_{F1 \cap F2=0} m_1(F1)m_2(F2) \quad (19)$$

Der Term $(1-K)$ hat die Aufgabe, die Werte zu normieren. Diese Gleichung erlaubt die Kombination der Hinweise aus einer Karte und aus einem Bild, ungeachtet der statistischen Verteilung der Daten. Die Berechnung des positiven Verdachtes eines Datensatzes $m_1(F1)$ erfolgt unabhängig vom zweiten Datensatz $m_2(F2)$.

3.2.4 Fuzzy-Logik

Die Zugehörigkeit eines Elementes zu einer Menge wird in der herkömmlichen Mathematik binär dargestellt. Zwei Zugehörigkeitsgrade sind möglich. Entweder gehört das Element zu einer Klasse (1) oder nicht (0). Die Theorie der Fuzzy-Mengen (unscharfe Mengen) erlaubt, den Grad der Zugehörigkeit als eine reelle Funktion mit beliebig vielen Werten zwischen 0 und 1 darzustellen.

$$A = \{x \in X \mid 0 \leq \mu_A(x) \leq 1\} \quad (20)$$

Die Funktion $\mu_A(x)$ wird Zugehörigkeitsfunktion (*membership function*) genannt. $\mu_A(x) = 0$ entspricht der Nichtzugehörigkeit, während $\mu_A(x) = 1$ für totale Zugehörigkeit steht. Zwischenwerte sind

möglich. Diese Darstellung eignet sich gut zur Beschreibung der Beziehung zwischen den Grauwerten eines Pixels und den Landnutzungsklassen. Die Form dieser Funktionen ist beliebig. Dreiecksförmige Funktionen werden oft benutzt, da sie einfach zu behandeln sind.

Für die Bearbeitung von unscharfen Mengen stehen eine Reihe von Operatoren zur Verfügung. Die Grundoperatoren sind Komplement, Vereinigung (ODER) und Durchschnitt (UND).

$$\mu_{\neg A}(x) = 1 - \mu_A(x) \quad (21)$$

$$\mu_A(x) \vee \mu_B(x) = \max\{\mu_A(x), \mu_B(x)\} \quad (22)$$

$$\mu_A(x) \wedge \mu_B(x) = \min\{\mu_A(x), \mu_B(x)\} \quad (23)$$

Der Inferenzprozeß benutzt eine Reihe von Zugehörigkeitsfunktionen und Regeln, um eine Schlußfolgerung aus den Eingabedaten X zu ziehen (Abb. 9) (Stanley & Zhu [1995]). Die Zugehörigkeitsfunktionen für die Eingabedaten und die Schlußfolgerungen werden zuerst im Schema definiert. Die Regeln sind in der Form WENN-DANN gegeben:

$$\begin{array}{llllll} \text{Regel}_1: & \text{Wenn } x_1 = A_{11} & \text{und } x_2 = A_{12} & \dots & \text{und } x_n = A_{1n} & \text{dann } y = B_1 \\ \text{Regel}_2: & \text{Wenn } x_1 = A_{21} & \text{und } x_2 = A_{22} & \dots & \text{und } x_n = A_{2n} & \text{dann } y = B_2 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \text{Regel}_m & \text{Wenn } x_1 = A_{m1} & \text{und } x_2 = A_{m2} & \dots & \text{und } x_n = A_{mn} & \text{dann } y = B_m \end{array}$$

x_j sind die Eingabedaten, und y ist das Schlußfolgerungsergebnis. A_{ij} und B_i sind ihre linguistischen Terme, oder die Adjektive, welche die Variablen beschreiben. Das Übereinstimmungsmaß der Regeln wird mittels der Fuzzy-Operatoren min und max (Gl. 21) bestimmt.

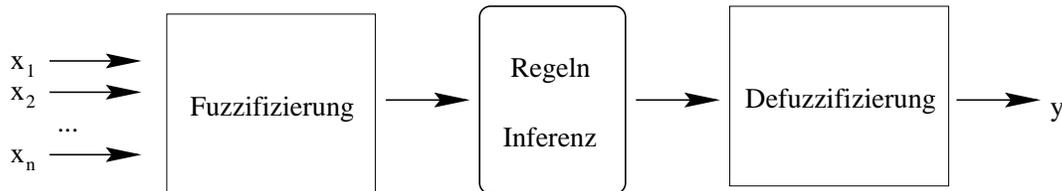


Abbildung 9: Fuzzy-Logik-regelbasiertes System.

Als Beispiel nehmen wir die Funktionen aus Abb. 10. Sie beschreiben:

- x_1 = *Grauwert im Bild* (Abb. 10a) und
- x_2 = *Farbe auf der Karte* (Abb. 10b),
- y = *Klasse des Pixels* (Abb. 10c).

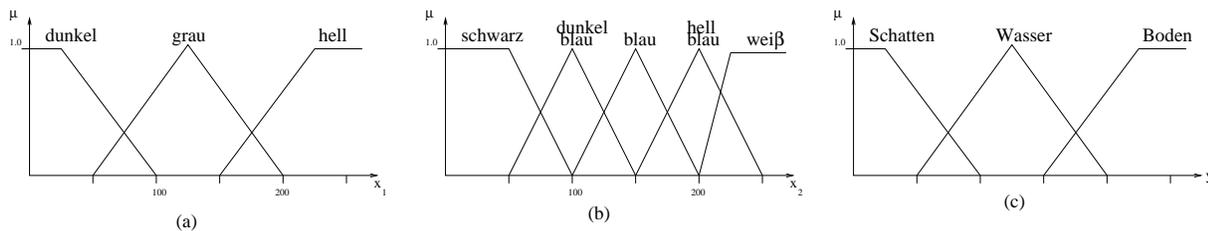


Abbildung 10: Fuzzy-Zugehörigkeitsfunktionen.

Die Regeln sind:

- R_1 : Wenn $x_1 = \textit{dunkel}$ und $x_2 = \textit{blau}$, dann $y = \textit{Wasser}$
- R_2 : Wenn $x_1 = \textit{grau}$ und $x_2 = \textit{hellblau}$, dann $y = \textit{Boden}$
- R_3 : Wenn $x_1 = \textit{hell}$ und $x_2 = \textit{weiß}$, dann $y = \textit{Boden}$
- R_4 : Wenn $x_1 = \textit{grau}$ und $x_2 = \textit{blau}$, dann $y = \textit{Wasser}$

Eine Schlußfolgerung für $X = (118, 165)$ wird gesucht.

Erst werden die entsprechenden Werte der Zugehörigkeitsfunktionen berechnet (*Fuzzifizierung*). Die *Fuzzifizierung* der Werte $X = (118, 165)$ wird in Abb. 11 dargestellt.

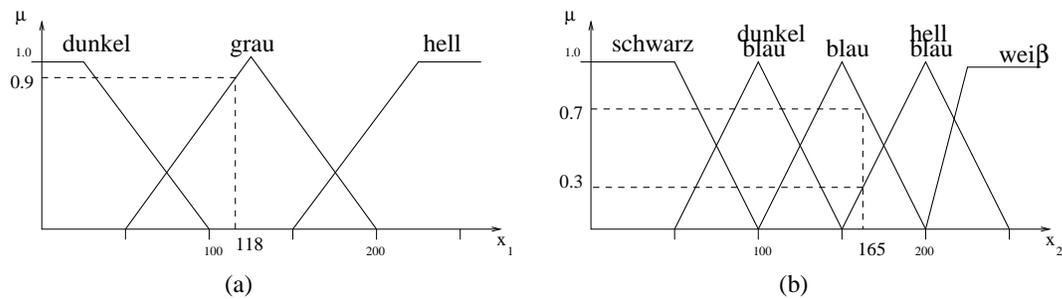


Abbildung 11: Zugehörigkeitsgrad für $X = (118, 165)$.

Danach werden die aktiven Regeln bestimmt. Aktiv sind diejenigen Regeln, bei denen das Übereinstimmungsmaß größer als Null ist.

R_1 : nicht aktiv

$R_2 : H2 = \wedge(\mu_{x_1=grau}, \mu_{x_2=hellblau}) = \wedge(0.9, 0.3) = 0.3$

R_3 : nicht aktiv

$R_4 : H4 = \wedge(\mu_{x_1=grau}, \mu_{x_2=blau}) = \wedge(0.9, 0.7) = 0.7$

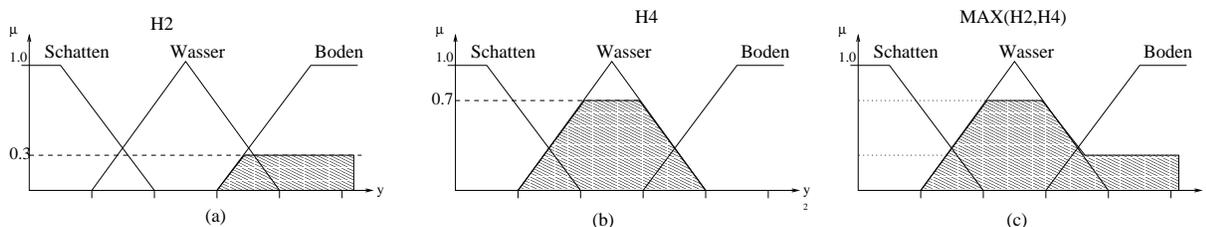


Abbildung 12: Übertragung der Übereinstimmungsmaß.

Bei unserem Beispiel sind die Regeln 1 und 3 nicht aktiv. Dagegen sind die Regeln 2 und 4 aktiv und ein Übereinstimmungsmaß kann berechnet werden, das beschreibt, wie gut die Ausgangslage mit der jeweiligen Regel übereinstimmt.

Das Übereinstimmungsmaß der Schlußfolgerung wird nun auf das Schlußfolgerungsergebnis übertragen. Da zur Fuzzy-Inferenz der unscharfe min-Operator benutzt wird, entspricht das einem Schneiden der unscharfen Ergebnismengen y in der Höhe der Übereinstimmungsmaße der Schlußfolgerung (Abb. 12a und b). Die resultierenden Ausgangs-Fuzzy-Mengen werden durch den max-Operator (ein logisches ODER) verknüpft und eine neue Fuzzy-Menge wird erzeugt (Abb. 12c).

Das Produkt der Fuzzy-Operationen ist eine neue Fuzzy-Menge (Abb. 12c). Diese neue Menge ist wieder unscharf und spiegelt die Sicherheit über das Ergebnis oder Schlußfolgerung mit Werten zwischen Null und Eins.

Falls ein numerisches Ergebnis benötigt wird, wäre die Defuzzifizierung der Ergebnis-Menge der letzte Schritt. Dafür gibt es verschiedene Methoden (Abb. 13). Zwei der allgemeineren Techniken sind die Methoden des *Maximalwertes* und des *Schwerpunktes* (Lin [1994]). Beim ersten wird ein Wert aus dem Maximum der Fuzzy-Ausgangs-Menge der Regel mit dem höchsten Erfüllungsgrad abgeleitet. Es wird angenommen, daß der höchste Wert der unscharfen Funktion dem Ergebnis am besten entspricht. Oft gibt es kein einzelnes Maximum, sondern einen Maximumbereich. In diesem Fall wird ein Wert in diesem Bereich ausgewählt. Ein Beispiel wird in Abb. 13 gezeigt. Die Schwerpunktmethode berechnet den Wert aus dem Flächenschwerpunkt der resultierenden Fuzzy-Menge (schraffierte Fläche in Abb. 12c).

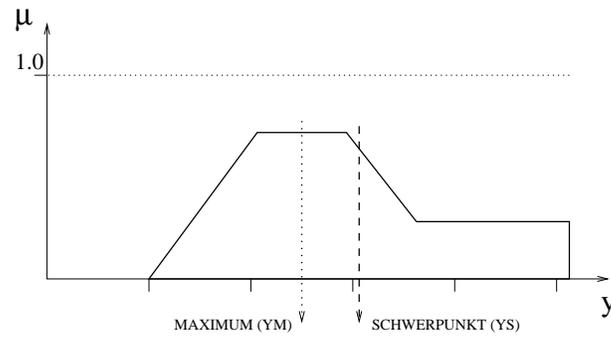


Abbildung 13: Ermittlung eines scharfen Wertes aus der unscharfen Ergebnismenge.

Die Schwerpunktmethode ist das gebräuchlichste Verfahren. Es berechnet einen unscharfen Wert aus dem Schwerpunkt der gesamten Menge, die Y-Koordinate des Schwerpunktes in Abb. 13, ist gegeben durch:

$$YS = \frac{\int_0^{\infty} y \cdot \mu(y) dy}{\int_0^{\infty} \mu(y) dy} \quad (24)$$

Wenn die unscharfen Mengen aus Dreiecken und Trapezen gebildet werden, ist die Berechnung des Schwerpunktes leicht, und die Integrationen können durch einfache Summen ersetzt werden. Der Schwerpunkt und der Y-Wert des Maximums sind in der Regel nicht identisch.

4 Ableitung von Information aus Karten

Eine digitale Rasterkarte entsteht aus einer analogen Karte, indem man die analoge Karte mit einem lichtempfindlichen Sensor abtastet (Vögtle [1998]). Dabei wird die Vorlage beleuchtet und die Intensität der reflektierten Strahlung mit Hilfe eines CCD-Chips gemessen. Die Karte wird so Punkt für Punkt, zeilenweise abgetastet. Als Ergebnis enthält man eine Matrix, in der Information über die Intensität der reflektierten Strahlung für jeden Punkt der Karte gespeichert ist. Durch die Verwendung von Filtern kann die reflektierte Strahlung in ihre Farbanteile zerlegt werden. Oft werden die Filter so gewählt, daß Informationen über die drei komplementären Farben rot, grün und blau abgeleitet werden können. Damit steht eine farbige Rasterkarte zur digitalen Weiterverarbeitung zur Verfügung. Durch die Mustererkennung wird anschließend versucht, den graphischen Strukturen automatisch eine Bedeutung zuzuweisen (Hake & Grünreich [1994]) und die Informationen über die Szene zu rekonstruieren.

Die automatische Erkennung des Karteninhaltes ist immer noch ein nicht vollständig gelöstes Problem (Klauer [1993], Doermann [1998], Brügelmann [1998], Yang [1989]). Von einer allgemeinen Lösung ist die Forschung noch weit entfernt. Bisher sind die entwickelten Lösungen noch stark von der Natur und dem Zweck des graphischen Dokuments abhängig. Es haben sich nicht nur unterschiedliche Zeichenvorschriften für verschiedene Karten des gleichen Inhalts entwickelt, sondern selbst die Inhalte sind je nach Zweck der Karte unterschiedlich: eine farbige topographische Karte enthält zum Beispiel andere Elemente als ein Bauplan. Je nach Karte ergeben sich also unterschiedliche Objektmerkmale, und es werden unterschiedliche Verarbeitungsschritte nötig.

Nicht immer ist bei Bildanalyseaufgaben eine vollständige Erkennung aller Kartenelemente gefordert. Bei dieser Arbeit stellt sich die Aufgabe, die Karten zu benutzen um Hinweise über die Landnutzungsklassen zu erhalten. Diese Hinweise aus den Karten sollen mit der Information aus dem Bildern der Satellitensensoren verknüpft werden. Eine derartige Vorgehensweise nennen Baltsavias & Hahn [1999] „Integration von Zwischenergebnissen“ (*partial results*).

4.1 Farbsegmentierung

Der erste Schritt bei der Extraktion graphischer Information aus farbigen Karten ist meist die Farbsegmentierung. Durch die Struktur der Bildelemente der Farbsegmentierung werden alle folgenden Verarbeitungsschritte entscheidend beeinflusst. Das heißt, daß gerade bei der Farbsegmentierung das Verfahren sorgfältig ausgewählt werden muß.

Die drei Farbkomponenten Rot, Grün und Blau (RGB) spannen einen Farbraum auf, in den jedes Bildelement abgebildet wird. Je nachdem welche Aspekte der Farbe man betonen will, kann man sich verschiedene Transformationen vorstellen, die einen Farbraum in viele andere Farbräume abbilden. Je nach Wahl des Farbraumes unterscheidet sich die Häufigkeitsverteilung der abgebildeten Pixel. Das heißt, daß zwei Faktoren die Einteilung des Farbraumes bestimmen: Die Wahl des Farbraumes bestimmt die Häufigkeitsverteilung und das ausgewählte Segmentierungsverfahren bestimmt das Kriterium, mit dem die Regionen unterteilt werden. In den folgenden Abschnitten wird das Verfahren, das für die Trennung der Farbschichten einer Karte vom IBGE (*Instituto Brasileiro de Geografia e Estatística*) benutzt wurde, vorgestellt. In Abb. 14 wird ein Ausschnitt aus der Karte *São Leopoldo* im Maßstab 1:50.000 gezeigt.

4.1.1 Wahl des Farbraumes

Die Trennung der einzelnen Farbschichten einer Karte basiert auf der Erkennung von Gruppen im Farbraum. Das Produkt der Abtastung wird üblicherweise mittels drei Farbkomponenten im RGB-System (Rot, Grün und Blau) gespeichert. Da nicht alle Farben durch die lineare Kombination der

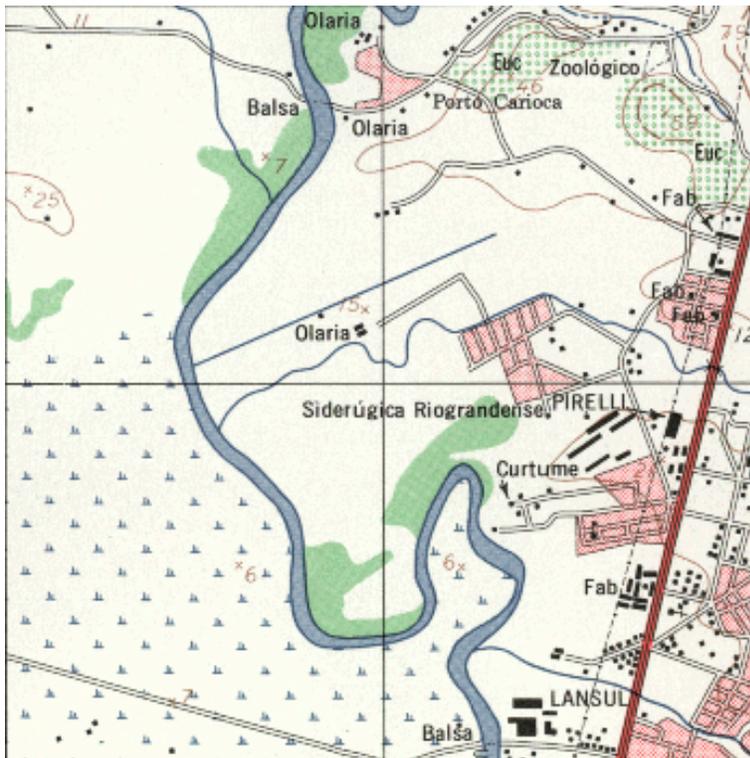


Abbildung 14: Ausschnitt einer abgetasteten Karte vom IBGE im Maßstab 1:50.000. Blatt São Leopoldo.

RGB-Komponenten erzeugt werden können (Hunt [1992]), sind andere Systeme geeigneter um das menschlichen Farbempfinden im Rechner zu simulieren. Ein schnell zu berechnendes Farbsystem, das dem psychologischen Farbempfinden des Menschen entspricht hätte den Vorteil, daß man Regeln im Segmentierungsprozeß finden könnte, die die menschlichen Interpretation nachahmen. Lauterbach [1993] erläutert die Vor- und Nachteile unterschiedlicher Farbsysteme für die Farbtrennung. Zu Farbsystemen, die dem menschlichen Farbempfinden nicht entsprechen, gehören RGB, IHS, CIE-XYZ, YIQ (Basis für das NTSC-Farbsystem) und das von Ohta *et al.* [1980] aus den Hauptkomponenten der Farbvektoren entwickelte I1,I2,I3-System. Die Beschreibungen einiger möglicher Systeme sowie die entsprechenden Transformationen in einander findet man in Hunt [1992], Lauterbach [1993] und Ohta *et al.* [1980].

Nach dem RGB- ist das IHS-System das vielleicht bekannteste Farbsystem. Die drei Komponenten dieses Systems sind Intensität (I), Farbfrequenz (H, englisch „hue“) und Sättigung (S). Sie haben physikalische Bedeutung und können einfacher von einem Menschen interpretiert werden. Konzepte wie „Intensität“ sind leichter als die Summe der RGB-Komponenten zu verstehen. Diese Komponenten sind nicht normiert, und da die Farbfrequenz und die Sättigung als Winkel gemessen werden, entstehen an manchen Stellen ihres 3D-Raumes Diskontinuitäten.

Ein System, das dem menschlichen Empfinden voll entspricht wurde bis jetzt noch nicht entwickelt, aber als gute Annäherungen gelten CIE-L*u*v (1976), CIE-L*a*b (1976) und CIE-UVW (1964). Farbunterschiede werden in Form von euklidischen Abständen zwischen Punkte des Farbraumes dargestellt. Dieser Abstand ist leicht zu berechnen. Abstände in diesen drei Systemen entsprechen besser den vom Menschen empfundenen Farbunterschieden und eignen sich besser zur digitalen Verarbeitung als zum Beispiel das RGB- oder IHS-System (Lauterbach [1993]). Zwei Farben, die nahe beieinander in einen dieser Farbräume liegen, werden von einem Menschen als sehr ähnlich empfunden. Das ist aber zum Beispiel beim IHS-System nicht der Fall, in dem die Farbfrequenz als Winkel angegeben wird.

In der Literatur werden Farbsysteme oft im Zusammenhang mit Ansätzen zur Bildsegmentierung beschrieben (Ohta *et al.* [1980]; Celenk [1990]). Je nach Anforderungen der speziellen Aufgabe muß

nach einem geeigneten System gesucht werden. Lauterbach [1993] empfiehlt das CIE-L*u*v-Farbsystem zur Segmentierung von farbigen Karten.

Die Segmentierung sollte möglichst schnell arbeiten. Durch die Vereinfachung eines dreidimensionalen Farbsystems zu einem System niedrigerer Dimension erreicht man eine Beschleunigung der Gruppierung (Lauterbach [1993]). Bei der Vereinfachung des Farbraumes muß man aufpassen, daß die Segmentierung nicht nur schnell arbeiten, sondern daß auch möglichst wenig Information verloren gehen soll.

Bei der Untersuchung unterschiedlicher Farbsysteme zur Segmentierung von Bildern wurde von Ohta *et al.* [1980] gezeigt, daß die drei orthogonalen Farbkomponenten (I_1, I_2, I_3) die aus der Hauptkomponententransformation abgeleitet werden, die Farbinformation für die Aufgabe der Segmentierung geeignet darstellen.

$$I_1 = (R + G + B)/3 \quad (25)$$

$$I_2 = (R - B)/2 \quad (26)$$

$$I_3 = (2G - R - B)/4 \quad (27)$$

Die Hauptkomponententransformation besteht aus einer Verschiebung und Drehung der Achsen des Farbsystems (Abb. 15), so daß die neuen Achsen parallel zur Richtung der größten Variationsrichtung der Daten im alten Farbraum liegen. Die Parameter der Rotation und Verschiebung werden aus den Eigenvektoren und Eigenwerten der Varianz-Kovarianz-Matrix der Farben berechnet. Eine ausführliche Beschreibung wird in Johnson & Wichern [1998] gegeben.

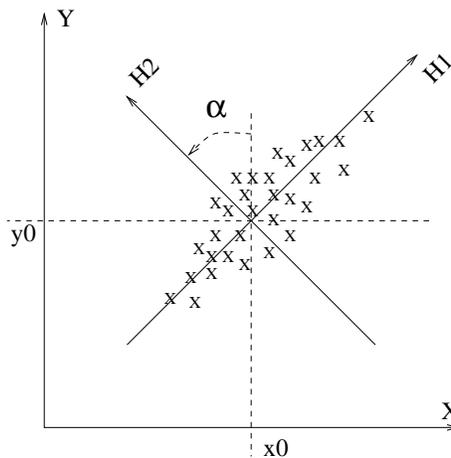


Abbildung 15: Hauptkomponententransformation.

Abb. 15 zeigt ein Beispiel der Hauptkomponententransformation im zweidimensionalen Raum. Die Punkte in diesem Beispiel bilden eine Wolke mit sehr hoher Korrelation zwischen den Komponenten X und Y . Die Korrelation kann nach Gl. 28 berechnet werden.

$$r = \sum_{i=1}^n \frac{(x_i - m_x)(x_i - m_y)}{(s_x s_y)} \quad (28)$$

mit

s_x = Standardabweichung der Komponente X

s_y = Standardabweichung der Komponente Y .

Die Koordinatendifferenzen der Punkte bezüglich X - und Y -Achse fallen geringer aus als die Koordinatendifferenzen der Punkte auf der $H1$ -Achse .

Betrachtet man eine Achse allein, dann werden die Punkte am deutlichsten auf der Achse $H1$ getrennt. Das neue System ($H1$ und $H2$) ist orthogonal und entsteht durch die Verschiebung (x_0, y_0) und Drehung (α) des ursprünglichen Systems:

$$\begin{vmatrix} H_1 \\ H_2 \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} * \begin{vmatrix} x_1 \\ y_2 \end{vmatrix} + \begin{vmatrix} b_1 \\ b_2 \end{vmatrix}.$$

Die Transformationsparameter werden aus der Varianz-Kovarianzmatrix C berechnet. Die Skalare λ , die die Gleichung 29 erfüllen, werden Eigenwerte der Matrix C genannt (Johnson & Wichern [1998]), wobei I die Einheitsmatrix ist:

$$|C - \lambda I| = 0. \quad (29)$$

Die Eigenvektoren sind die Vektoren $(\mathbf{v}, \mathbf{v} \neq 0)$, die die Gleichung 30 erfüllen:

$$C\mathbf{v} = \lambda\mathbf{v} \quad (30)$$

Jede Lösung dieses Systems ergibt einen Eigenvektor \mathbf{v}_i , der zum entsprechenden Eigenwert λ_i gehört. Diese Vektoren bilden ein neues, orthogonales, System, das parallel zur größten Variation der Daten liegt. Die Richtung der größten Variation wird von dem Vektor angegeben, der zu dem größten normierten Eigenwert gehört.

Basierend auf den Eigenschaften der Hauptkomponenten entwickelte Ohta *et al.* [1980] ein System für die Farbsegmentierung. In diesem Experiment wurden acht unterschiedliche Farbaufnahmen analysiert. Dabei wurde festgestellt, daß die Eigenvektoren der Bilder ziemlich ähnlich aussahen. Die erste Hauptkomponentenachse zeigte bei allen eine Tendenz in Richtung $(1/3, 1/3, 1/3)^\top$. Die Zweite war von den Relationen $(1/2, 0, -1/2)^\top$ oder $(-1/2, 0, 1/2)^\top$ geprägt, und die dritte zeigte Werte um $(-1/4, 1/2, -1/4)^\top$.

Die Möglichkeit, eine günstige Kombination von Farbraum und Algorithmus zu benutzen, wird für die Farbsegmentierung der IBGE-Karte untersucht. Dafür wird ein Farbsystem gesucht, in dem die Aufteilung des Farbraumes parallel zu seinen drei Achsen möglich ist. Diese Aufteilung basiert auf der Analyse der Histogramme der Farbkomponenten. Dies erlaubt die Anwendung von Schwellwertverfahren, mit denen Histogramme in einzelnen Bereiche zerlegt werden können. Dafür ist es aber besonders günstig, wenn die Farbcluster parallel zu den Farbraumachsen liegen.

Tabelle 2 hilft bei der Analyse der Korrelationen zwischen den Farbkomponenten in einer IBGE-Karte. Diese Werte wurden aus der in Abb. 14 gezeigten abgetasteten Karte berechnet. Die RGB-Daten wurden dafür in das *IHS*, *Lab*, *Luv*, *YIQ* und *Ohta* System transformiert. Die Parameter der Transformationen wurden aus der Literatur (z.B. Ohta *et al.* [1980]) genommen.

Die Korrelation zwischen den einzelnen Farbkomponenten ist am geringsten bei den *Luv* und *Ohta* Farbsystemen. Das Farbsystem von *Ohta* zeigt die kleinsten Korrelationen. Das ist auch zu erwarten, da dieses System aus den Hauptkomponenten des *RGB*-Farbraumes abgeleitet wird. Gegenüber dem *Luv*-System hat das *Ohta*-System den Vorteil, daß es einfacher zu berechnen ist.

Die Analyse der Hauptkomponenten der in der IBGE-Karte enthaltenen Farben zeigt ähnliche Ergebnisse wie die von Ohta *et al.* [1980]. Tabelle 3 zeigt die Eigenvektoren des in Abb. 14 gezeigten Bildes. Der erste Vektor entspricht den von Ohta *et al.* [1980] erwähnten Ergebnissen und ist nichts weiter als die Intensität im *IHS*-System. Die zweite und dritte Komponente könnte durch die Vektoren $(-1/2, 1/4, 1/4)'$ und $(0, 1/2, 1/2)'$ angenähert werden. Da diese beiden Vektoren senkrecht zur ersten

Tabelle 2: Korrelationsmatrix der Farbkomponenten in der IBGE-Karte.

Farbraum	Kovarianz-Matrix		
RGB	1.000000	0.935979	0.937669
	0.935979	1.000000	0.978121
	0.937669	0.978121	1.000000
IHS	1.000000	0.268679	-0.591064
	0.268679	1.000000	-0.218811
	-0.591064	-0.218811	1.000000
YIQ	1.000000	-0.381630	0.598551
	-0.381630	1.000000	-0.274456
	0.598551	-0.274456	1.000000
Lab	1.000000	0.506998	-0.412796
	0.506998	1.000000	-0.067603
	-0.412796	-0.067603	1.000000
Luv	1.000000	-0.233736	-0.258219
	-0.233736	1.000000	0.362677
	-0.258219	0.362677	1.000000
Ohta	1.000000	0.139828	-0.033255
	0.139828	1.000000	-0.129086
	-0.033255	-0.129086	1.000000

Tabelle 3: Eigenwerte und Eigenvektoren der Farben der Karte vom IBGE, bzgl. dem Ohta System.

Eigenwerte	7585.17	204.30	56.09
Normierte Eigenwerte	0.796	0.115	0.089
Eigenvektor 1	0.336	0.340	0.325
Eigenvektor 2	-0.497	0.269	0.234
Eigenvektor 3	0.015	-0.473	0.511

Hauptkomponente liegen, beschreiben sie dieselbe Ebene wie die von Ohta *et al.* [1980] beschriebenen Vektoren, sie sind gegenüber denen nur etwas gedreht. Wegen des geringeren Rechenaufwands und der kleineren Korrelationswerte im *Ohta*-System wurde dieses für die Trennung der Farben gewählt.

Die Größe der normierten Eigenwerte ist proportional zum Informationsinhalt jeder Komponente. Deswegen wäre es sinnvoll, bei der Gruppierung der Farben im *Ohta*-System der Reihe nach die Farben auf die Achsen zu projizieren und zu trennen.

4.1.2 Farbtrennung

Die Qualität der Farbtrennung hängt von der Implementierung der Verfahren zur Detektion spektraler Homogenität und der Gruppierung der Pixel in sinnvolle Regionen ab. Ansätze, die diesen Zweck erfüllen, basieren auf Kantenextraktion, Flächenwachstum, Histogrammanalyse oder Clustering. Zwei Arten von Segmentierungsverfahren werden in der Literatur beschrieben (Levine [1985]): Teilsegmentierung (*partial segmentation*) und vollständige Segmentierung (*complete segmentation*). Die erste teilt das Bild in homogene Regionen ohne Berücksichtigung der Beziehung zwischen den Pixel und den Objekten in der Szene (oder graphischen Elementen der Karte) und kann für mehrere Zwecke benutzt werden. Sie findet nur in der *Signalebene* statt. Die bekanntesten low-level-Segmentierungsverfahren gehören zu dieser Gruppe. Kanade [1980] unterscheidet zwischen rein spektralen (*histogram analy-*

sis oder *clustering*) und auf spektralen und räumlichen Merkmalen basierenden Verfahren (*split and merge*, Flächenwachstum, *recursive splitting*). Die vollständige Segmentierung, von Ton *et al.* [1991] auch *knowledge guided segmentation* genannt, teilt das Bild in Regionen, deren Grenzen den Grenzen der Objekte in der realen Welt genau entsprechen (Levine [1985]). Wissen auf der physikalischen und semantischen Ebene ist dafür nötig.

Um die Farben im digitalen Bild der Karte zu trennen wurde ein Verfahren angewandt, das basierend auf Eigenschaften des Bildes, wie spektrale Homogenität der Farben, räumlicher Zusammenhang von Regionen mit derselben Farbe und Kontrast zwischen benachbarten Regionen, eine Aufteilung des Farbraumes optimiert. Das Verfahren ist vom Typ „*Split-and-Merge*“ und kann mit Hilfe von Abb. 16 verdeutlicht werden.

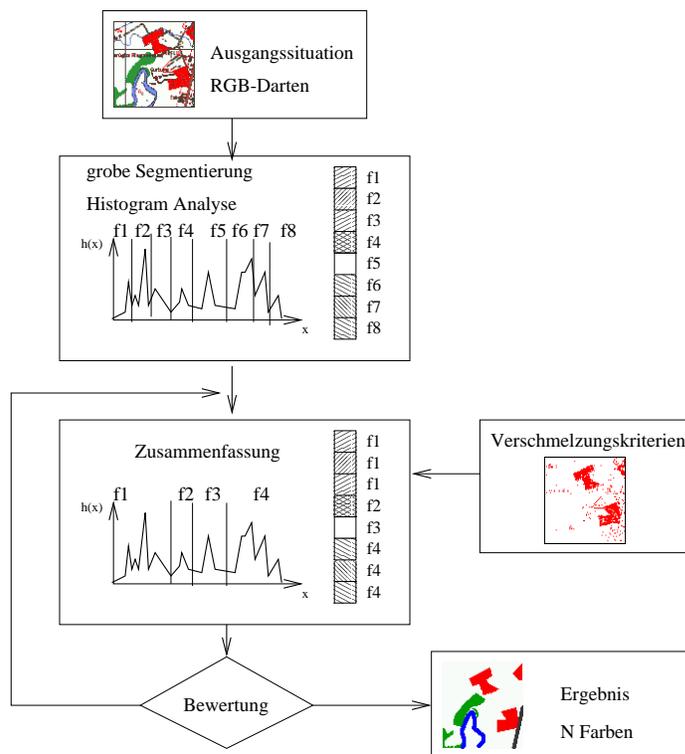


Abbildung 16: Schritte des Farbtrennungsverfahrens.

Die Eingangswerte sind die drei Farbmatrizen des farbigen Bildes. Jedes Pixel wird durch drei Farbkomponenten (im I_1, I_2, I_3 -Farbsystem) beschrieben. Das heißt, sein Merkmalsvektor (I_1, I_2, I_3) enthält nur die Farbinformation. Diese Daten werden in einem ersten Schritt in viele, mehr als gewünschte, Farbgruppen aufgeteilt. Am Ende dieser Trennung entstehen viele homogene Gruppen. Diese Gruppen entsprechen nicht unbedingt den gewünschten Farbgruppen, und ihre Verschmelzung in sinnvolle Gruppen ist nötig. Es werden Gruppen gesucht, die gleichartige Objekte in der Karte repräsentieren. Sie sollen spektral und räumlich homogen sein und sich von anderen benachbarten gut unterscheiden (guter Kontrast).

Die Reduktion der Zahl der Gruppen wird als ein Optimierungsproblem betrachtet, bei dem die Farbgruppen iterativ kombiniert werden. Wenn eine Gruppe nicht bestimmten Merkmalen entspricht, dann wird sie nach vorgegebenen Regeln mit anderen Gruppen verschmolzen. Die Merkmale der Gruppen werden aus der unscharfen Darstellung ihrer spektralen und räumlichen Eigenschaften abgeleitet.

Zum Beispiel soll vermieden werden, daß die Komponenten einer Farbgruppe über das ganze Bild verstreut sind. Dazu wird eine unscharfe Beschreibung eingeführt, die den Grad der Granularität der Flächen beschreiben kann. Sind die Regionen stark zergliedert und verstreut, dann soll die Farbgruppe die Beschreibung „schwach verbunden“ erhalten. Es sollen allerdings nur stark verbundene Regionen

erhalten werden. Deshalb wird in einem nachfolgenden Schritt diese „schwach verbundene“ Farbgruppe mit der am nächsten verwandten Farbgruppe verschmolzen werden.

4.1.3 Histogrammbasierte Farbsegmentierung

Ein globaler histogrammbasierter Ansatz wird angewandt, um den Farbraum in homogene Regionen zu teilen. Das Grundprinzip dieses Verfahrens ist die Suche nach relativen Minima (Täler) und Maxima (Spitzen) im Histogramm. Da dieser Ansatz nur von den Farbwerten abhängig ist, wird er als „global“ bezeichnet. Verschiedene Ansätze zur Erkennung von Gruppen in Histogrammen werden in der Literatur beschrieben, z.B. in Fu & Mui [1981], Kanade [1980], Sahoo *et al.* [1988]. Der mathematische Ansatz hierbei lautet:

$$g(x) = c \quad \text{if } T_c \leq f(x) \leq T_{c+1} \quad (31)$$

$$c = 0, 1, 2, 3, \dots, m \quad (32)$$

In diesem Zusammenhang beschreibt x die relative Position des Pixels im Bild, $f(x)$ steht für den Grauwert oder die Farbe dieses Pixels, g steht für das Segmentierungsergebnis und T für den Schwellwert.

Rosenfeld & de la Torre [1988] schlagen die Analyse der Histogrammvariation als Basis für die Schätzung der Gruppen vor. Die positiv- und negativkonkaven Bereiche erhält man mit Hilfe der konvexen Hülle (dem flächenmäßig kleinsten Vieleck, welches das Histogramm einschließt). Eine Verbesserung dieses Algorithmus, bei dem statt der Histogrammwerte selbst ihre Logarithmen $l(i)$ benutzt werden, wird von Whatmough [1991] beschrieben. In diesem Fall wird die obere exponentielle Hülle berechnet.

$$l(i) = \log(h(i)) \quad \text{wenn } i > 0 \quad (33)$$

$$l(i) = 0 \quad \text{wenn } i = 0 \quad (34)$$

Die obere konvexe Hülle $q(i)$ von $l(i)$ wird als die minimale obere Funktion, die immer größer als $l(i)$ ist, definiert. Die Täler des Histogramms werden aus dem Unterschied zwischen der logarithmischen Funktion und der dazugehörige Hülle ermittelt $r(i) = q(x) - l(x)$. Die Grauwerte, an denen sich Gruppengrenzen befinden können, werden durch die lokalen Maxima von $r(i)$ angegeben (s. Abb. 17).

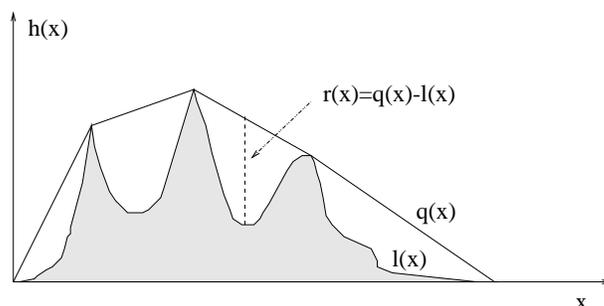


Abbildung 17: Schätzung der Farbgruppengrenzen aus der exponentiellen Hülle des Histogramms.

Sezan [1990] beschreibt ein anderes Verfahren zur Bestimmung der Grenzen und Maxima von Gruppen im Histogramm. Dieses Verfahren basiert auf der Analyse der Variation der Histogrammwerte. Es ist rechnerisch einfach und hat den Vorteil, daß man nur einen Parameter braucht, um die Empfindlichkeit des Verfahrens zu regeln. Ein „Bergerkennungssignal“ (peak-detection signal) wird aus der Differenz

zwischen zwei Funktionen des Histogramms berechnet. Die erste Funktion wird aus der Ansammlung der Histogrammwerte gebildet. Die zweite Funktion ist das Ergebnis der Filterung der Ansammlung des Histogramms mit einem low-pass-Filter (Sezan [1990]):

$$rd(i) = ch(i) - mh(i) \quad (35)$$

wobei ch für die Summe der Histogrammwerte $h(i)$ steht:

$$ch(i) = \sum_{j=0}^i h(j) \quad (36)$$

$mh(i)$ ist das Ergebnis der Glätung von ch mit einem low-pass Filter.

Folgende Merkmale kennzeichnen die Position der „Berge“ im Histogramm:

1. Negative Nulldurchgänge: das sind Übergänge von positiven zu negativen Werten, die den Anfang der Berge kennzeichnen.
2. Der nächste Übergang von negativen zu positiven Werten, der eine Näherung für den Grauwert der „relativen Maximum“ darstellt.
3. Das lokale Maximum von $rd(i)$ zwischen zwei Nulldurchgängen entspricht dem Ende eines lokalen Maximums.

Die Größe der Nachbarschaft, die in der Filterung benutzt wird, bestimmt die Empfindlichkeit des Verfahrens. Je kleiner die Nachbarschaft, desto mehr kleinere Berge können gefunden werden, und umso feiner kann man das Histogramm aufteilen.

4.1.4 Optimierung der Farbtrennung

Zeichenvorschriften für Karten sehen meist nur wenige Farben vor, aber durch die Variation bei der Herstellung, durch den Alterungsprozeß und durch die Ungenauigkeiten bei der Wandlung von „analog“ nach „digital“ entstehen mehrere Farbtöne. An den Übergängen entstehen durch die räumliche Diskretisierung Mischpixel, und durch die Überlagerung unterschiedlicher Farben werden neue Farbtöne erzeugt. Diese neuen Farbtöne können sogar häufiger als die ursprüngliche Farbe vorkommen, die Abwesenheit der ursprünglichen Farben vortäuschen und das Ergebnis der Segmentierung mit einem herkömmlichen Verfahren stark beeinträchtigen.

Der histogrammbasierte Farbsegmentierungsalgorithmus (Abs. 4.1.3) ist eine herkömmliche Lösung. Die in der Gl. 31 beschriebenen Schwellwerte sind nur von der Häufigkeit der Grauwerte im Bild abhängig. Andere Merkmale, zum Beispiel die Geometrie der farbigen Flächen, werden dabei nicht berücksichtigt, obwohl sie zur Verfeinerung der Segmentierung beitragen könnten. Das Interesse bei der Farbsegmentierung ist jedoch, nicht nur homogene Farbgruppen zu erkennen, sondern daß die Pixel eines Objektes im Bild gruppiert werden. Deswegen ist es sinnvoll, nach einer globalen Farbsegmentierung die Farbgruppen mit Hilfe anderer Merkmale (räumlicher und spektraler Merkmale) zu analysieren und zu entscheiden, ob sie zu Objekten in der Szene passen.

Im folgenden wird gezeigt, wie dieses Segmentierungsverfahren optimiert werden kann:

1. Das Bild wird mit einem globalen Verfahren segmentiert. In dieser Anfangsphase werden mehr Farbgruppen generiert, als es ursprünglich Farben in der Karte gab. Diese erste Lösung wird als Ausgangssituation verwendet.

2. Die räumliche Verteilung und die spektralen Eigenschaften der Pixel jeder Farbgruppe werden untersucht. Für jede Gruppe werden folgende Eigenschaften numerisch dargestellt:

- Spektrale Homogenität,
- Gestalt (räumlicher Zusammenhang der Pixel einer Gruppe),
- Kontrast zwischen einer Gruppe und ihre Nachbarn.

Diese drei Eigenschaften sind unscharfe Begriffe, deswegen werden sie als unscharfe Mengen (mit Hilfe der Fuzzy-Logik) behandelt. Sie beschreiben das Aussehen der farbigen Flächen in der Karte und ermöglichen eine Bewertung der Qualität des Ergebnisses.

3. Je nach Wert dieser Qualitätsbeschreibung werden gewisse Gruppen beseitigt, das heißt einer anderen, stärkeren Gruppe zugewiesen. Die Entscheidung über die Erhaltung oder Beseitigung einer Gruppe wird aufgrund der Qualitätsparameter aller Gruppen getroffen.

4. Es wird überprüft, ob die von den drei Merkmalen beschriebene Qualität der neuen Gruppen ausreichend ist (größer als ein Schwellwert). Wenn eine Gruppe nicht ausreichend spektral homogen ist, ihre Pixel auf dem Bild zu stark gestreut sind oder der Kontrast dieser Gruppe mit den benachbarten Regionen zu niedrig ist, dann wird der Prozeß ab dem zweiten Schritt wiederholt. Der Prozeß wird beendet, wenn alle Gruppen die gewünschten Eigenschaften erreichen oder keine weitere Verbesserung möglich ist.

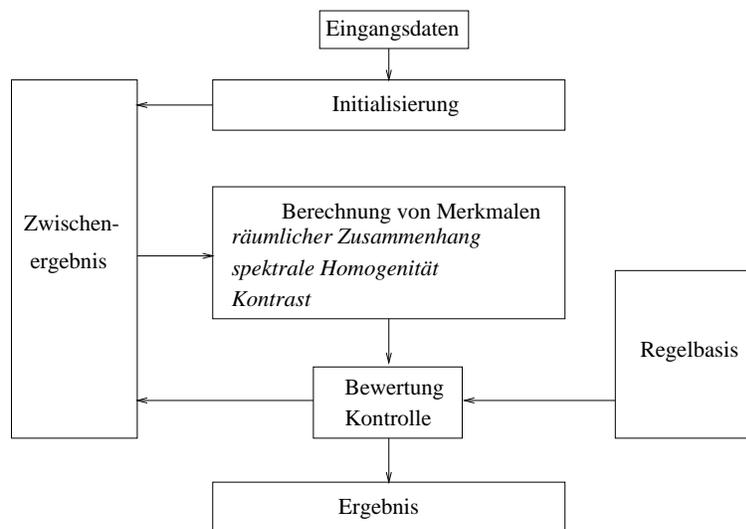


Abbildung 18: Optimierungsprozeß.

Ein Kontroll-Modul steuert die Entwicklung des Prozesses. Die notwendigen Bewertungskriterien werden durch Regeln explizit definiert. Die Regeln sind vom Typ:

„Wenn die Gruppe spektral homogen ist *und* der Kontrast zwischen ihren Pixeln und der Nachbarschaft groß ist *und* die Pixel der Gruppe kompakte Flächen bilden, *dann* stellt diese Gruppe mindestens ein graphisches Kartenelement dar und sollte beibehalten werden. *Sonst* kann diese Gruppe gelöscht werden.“

Das hier benutzte System ist modular aufgebaut (Abb. 18).

Ein Modul, das Bildverarbeitungsmethoden umfaßt, ermöglicht die quantitative Beschreibung des aktuellen Zustandes. Die Fortschritte und Ergebnisse des Prozesses werden in einem Zwischenspeicher gespeichert und als Ausgangssituation des nächsten Schrittes benutzt.

Die Regelbasis ist vom analysierten Bild und den low-level Bildverarbeitungsmethoden unabhängig. Das hat den Vorteil, daß ihre Regeln je nach der Art des Bildes oder dem Zweck der Segmentierung modifiziert werden können, aber während des Prozesses konstant bleiben. Diese Regeln werden vom Nutzer explizit definiert.

Die Schätzung der Homogenität und des Kontrastes in einem digitalen Bild erfolgt durch die Untersuchung der Variation seiner Grauwerte und ist im wesentlichen bekannt (Haralick & Shapiro [1997]). Das Problem besteht darin, wie man aus diesen Messungen eine Entscheidung über die nächsten Schritte des Segmentierungsprozesses treffen kann und die bislang erreichte Qualität des Ergebnisses bewerten kann. Diese Entscheidung soll ähnlichen Prinzipien und Kriterien entsprechen, mit denen ein menschlicher Operator das Bild bewerten würde.

Die Qualität des Zwischenergebnisses und das Ziel der Segmentierung können gut mit Hilfe von unscharfen Termen beschrieben werden, denn auch die einbezogenen Beurteilungskriterien sind meist unscharf. Die spektrale Homogenität wird nicht quantitativ, sondern in einer qualitativen Form beschrieben, zum Beispiel „*die gefärbten Flächen haben eine homogene Farbe*“. Die hierfür „ausreichende“ Homogenität ist aber nicht explizit definiert. Nicht alle Farben einer Fläche müssen gleich sein, um diese Region als homogen zu bezeichnen. Es entsteht ein Spielraum für eine gewisse Unsicherheit und Toleranz.

Die Merkmale, die für die Beschreibung des aktuellen Zustands der Farbgruppen benutzt werden, werden im folgenden diskutiert.

Spektrale Homogenität - die Homogenität einer Menge ist umgekehrt proportional zur Streuung und kann aus der Varianz geschätzt werden.

Wenn A die Anzahl der Pixel in der k -ten Gruppe ist ($x \in G_k$), dann können der Mittelwert und die Varianz dieser Farbgruppe in der j -ten Farbkomponente aus den Grauwerten ihrer Pixel gemäß Gl. 37 berechnet werden.

$$m_j(G_k) = \sum_{x_i \in G_k} f(x_{ij})/A \quad (37)$$

$$s_j^2(G_k) = \sum_{x_i \in G_k} (f(x_{ij}) - m_j(G_k))^2/A \quad (38)$$

A = Anzahl der Pixel der k -ten Gruppe.

Eine Schätzung der Streuung St in der Menge kann aus der Varianz der drei Farbkomponenten abgeleitet werden, zum Beispiel als die mittlere Varianz s_m^2 . Die Homogenität der Gruppen U_k wird dann als Funktion der Streuung geschätzt.

$$s_{mk}^2 = \sum_{j=0}^3 s_j^2(x)/3 \quad (39)$$

$$St_k = s_{mk}^2/s_{max}^2 \quad (40)$$

$$s_{max}^2 = \max(s_{mk}^2) \quad (41)$$

$$U_k = 1 - \sqrt{St_k} \quad (42)$$

$$0 \leq U_k \leq 1 \quad (43)$$

Abb. 19 zeigt den Unterschied der spektralen Verteilung zweier Farbgruppen. Der erste Farbgruppe besitzt eine geringere Streuung und Varianz. Die spektrale Homogenität der ersten Gruppe ist größer als die spektrale Homogenität der zweiten.

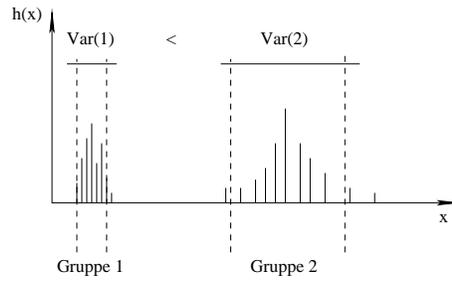


Abbildung 19: spektrale Homogenität.

Räumlicher Zusammenhang - Der räumliche Zusammenhang der Pixel einer Gruppe ist proportional zur relativen Häufigkeit, mit der benachbarte Pixel zur selben Gruppe gehören (Abb. 20). Die Schätzung der relativen Häufigkeit erfolgt mit Hilfe eines 3x3 Fensters, das über das ganze Bild geschoben wird. Für jede mögliche Position des Fensters im Bild werden die Nachbarn des zentralen Pixels analysiert.

Wenn i das zentrale Pixel des Fensters und j eines seiner Nachbarn ist, dann wird die relative Häufigkeit, mit der dieselbe Farbgruppe im benachbarten Pixel auftritt durch Gl. 44 gegeben.

$$H(k, k) = H(i \in k \cap j \in k) = \frac{\sum N(i \in k, j \in k)}{A_k} \quad (44)$$

i beschreibt das zentrale Pixel des 3x3 Fensters und j einen seiner 8 Nachbarn.

A_k steht für die gesamte Fläche der Gruppe k .

$H(k, k)$ wächst mit der Zahl der homogenen Flächen im Bild. Wenn die Gruppen im Bild zerstreut sind, dann wird $H(k, k)$ kleiner. Deswegen kann $H(k, k)$ als Parameter zur Beschreibung des räumlichen Zusammenhanges R der Gruppe k benutzt werden.

$$R(k) = H(k, k) = \frac{\sum N(i \in k, j \in k)}{A_k} \quad (45)$$

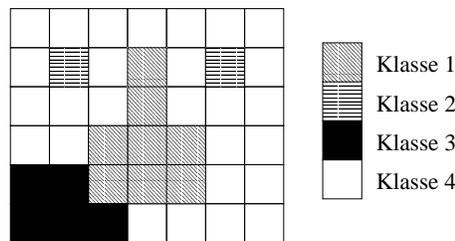


Abbildung 20: Räumlicher Zusammenhang.

In Abb. 20 ist der räumliche Zusammenhang der Pixel der Gruppe 1 größer als der räumliche Zusammenhang der Gruppe 2, denn Pixel der Gruppe 2 kommen nur isoliert vor.

Kontrast - Der Kontrast einer Gruppe wird aus dem spektralen Unterschied zwischen ihrem Mittelwert und dem Mittelwert anderer Gruppen in der Nachbarschaft in allen Farbkomponenten berechnet.

Der Kontrast zweier Gruppen ($k \neq l$) wird nach Gl. 46 berechnet.

$$K(k, l) = \sum_{j=0}^3 \frac{m_j(G_k) - m_j(G_l)}{m_j(G_k) + m_j(G_l)} \quad (46)$$

Mit: $m_j(G_k)$ = Mittelwert der Gruppe G in Farbkomponente j .

Der globale Kontrast der Gruppe k wird als die gewichteten Summe der Kontraste dieser Gruppe und ihrer Nachbarn berechnet. Die Gewichtsfunktion $L(k, l)$ wird aus der Häufigkeit, mit der die beiden Gruppen in benachbarten Pixeln vorkommen, abgeleitet.

$$K(k) = \sum_{l=0}^{NC} K(k, l) * L(k, l) \quad (47)$$

mit:

$K(k)$ = Kontrast der Gruppe k

$L(k, l)$ = relative Länge der Grenze zwischen Gruppe k und l ,

NC = Anzahl der Gruppen $l \neq k$.

Die Ränder der Gruppen werden untersucht, um das Gewicht jeder Kombination zu berechnen.

$$L(k, l) = \frac{\text{Länge}(k, l)}{\text{Länge}(k)} \quad (48)$$

Mit:

$\text{Länge}(k, l)$ = Länge der Grenze zwischen Gruppen k und l

$\text{Länge}(k)$ = Länge der gesamten Konturen der Gruppe k .

Dieser Parameter kann mit Hilfe der Abb. 20 verdeutlicht werden. Die Pixel der Gruppe 1 haben als Nachbarn nur Pixel der Gruppen 3 und 4. Deswegen wird der Wert des Kontrastes für Gruppe 1 nur aus dem Kontrast zwischen 1 und 3 ($K(1, 3)$) und 1 und 4 ($K(1, 4)$) berechnet. Da die Grenzen zwischen Gruppe 1 und 4, im Bezug zur Gruppe 1, länger sind, wird das Gewicht der Gruppe 4 dementsprechend größer.

4.1.5 Bewertung

Die genaue Bestimmung der Werte der obengenannten Parameter ist möglich, aber genaue numerische Werte verlieren in einer qualitativen Analyse an Bedeutung, da die Phänomene und Konzepte, die sie beschreiben, üblicherweise nicht genau definiert sind. Eine Kombination dieser Werte mittels einer mathematischen Formel ist möglich aber schwierig zu formulieren, denn Begriffe wie „Homogenität“ oder „Qualität“, die in der Analyse von Bildern benutzt werden, sind unscharf. Deswegen ist es konsequent sie als unscharfe (*fuzzy*) Definitionen zu beschreiben. Das hat noch den Vorteil, daß die Kombination der Hinweise mittels Fuzzy-Logik eine Bewertung der Qualität in linguistischer Form ermöglicht. Deswegen hat die Formulierung dieser Bewertung in Fuzzy-Termen Vorteile gegenüber einer expliziten Gleichung, die die Parameter mathematisch kombiniert und deren Schwellwerte nicht genau bekannt sind.

Die „Fuzzifizierung“ der Parameterwerte ist mit Hilfe von Zugehörigkeitsfunktionen möglich. Sie wandeln die numerischen Werte einer Aussage in eine Fuzzy-Menge um. Beispiele dieser Funktionen werden

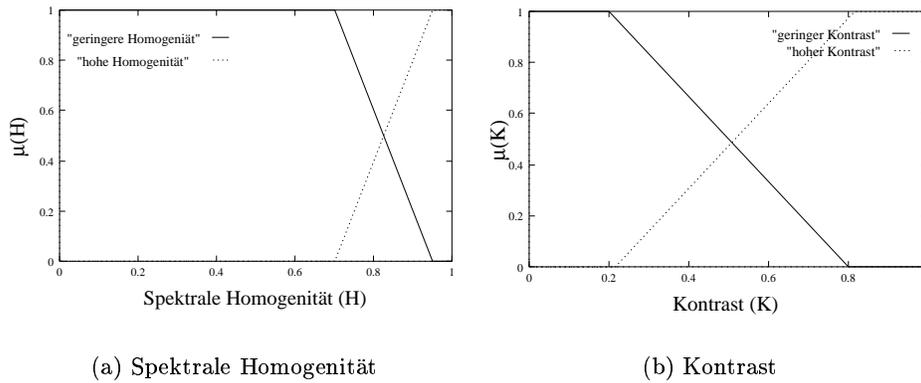


Abbildung 21: Zugehörigkeitsfunktionen: spektrale Homogenität und Kontrast.

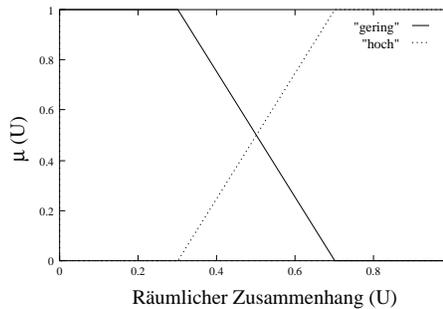


Abbildung 22: Zugehörigkeitsfunktion: räumlicher Zusammenhang.

in Abb. 21 gezeigt. Auf der x -Achse werden die Werte der Parameter und auf der y -Achse die Werte der Fuzzy-Menge dargestellt.

Die Variation der Fuzzy-Werte als Funktion der Parameter stellt den Einfluß der Parameter, ihre Bedeutung, auf die Qualität der Gruppe dar. Die Anpassung dieser Funktionen verbessert die Segmentierung. Da die Homogenität in abgetasteten Bildern hoch sein sollte, sind die Funktionen steil zu wählen (Abb. 21a). Extreme Werte (0.0 oder 1.0) entsprechen Sicherheit über die Homogenität der Gruppe. Entweder ist die Gruppe sehr homogen (Werte um 1.0) oder gar nicht homogen (Werte um 0.0). Zwischenwerte sollen die Unsicherheit einer Aussage über die Homogenität modellieren.

Am Anfang der Analyse müssen auch die Regeln explizit definiert, und in der Regelbasis gespeichert werden. Linguistische Aussagen sind in diesem Fall spezifisch dargestellt. Sie stellen die Bedingungen dar, die die resultierende Entscheidung konditionieren. Ein Beispiel dieser Regeln wird im Tab. 4 gezeigt.

Tabelle 4: Regeln für die Optimierung der Segmentierung: R = räumlicher Zusammenhang; K = Kontrast; U = spektrale Homogenität.

IF	K ist Niedrig	AND	R ist Niedrig	AND	U ist Niedrig	THEN	Gruppe Löschen
IF	K ist Niedrig	AND	R ist Niedrig	AND	U ist Hoch	THEN	Gruppe Löschen
IF	K ist Niedrig	AND	R ist Hoch	AND	U ist Hoch	THEN	Gruppe Behalten
IF	K ist Niedrig	AND	R ist Hoch	AND	U ist Niedrig	THEN	Gruppe Löschen
IF	K ist Hoch	AND	R ist Niedrig	AND	U ist Niedrig	THEN	Gruppe Löschen
IF	K ist Hoch	AND	R ist Niedrig	AND	U ist Hoch	THEN	Gruppe Behalten
IF	K ist Hoch	AND	R ist Hoch	AND	U ist Niedrig	THEN	Gruppe Behalten
IF	K ist Hoch	AND	R ist Hoch	AND	U ist Hoch	THEN	Gruppe Behalten

Der Entscheidungsprozeß fängt mit der Bestimmung der Fuzzy-Menge μ_K , μ_H und μ_U der Homogenität, Kontrast und Gestalt für alle Gruppen G_k an.

Eine unscharfe Entscheidungsmenge wird aus diesen Mengen für jede Gruppe abgeleitet, und zwar durch ihre unscharf-logische Kombination, nach dem Prinzip:

$$E_{kt} = \mu_H(G_k) \cap \mu_K(G_k) \cap \mu_U(G_k) > E_{min} \quad (49)$$

Die Aussagen „Die Gruppe C_k wird gelöscht“ und „Die Gruppe C_k bleibt erhalten“ werden damit formuliert. Der Wert E_{kt} beschreibt die Sicherheit der Aussage.

In jeder Iteration werden unscharfe Mengen für die beiden Aussagen abgeleitet. Die schwächste Gruppe wird identifiziert. Die Aussage „Die Gruppe C_k bleibt erhalten“ wird dann für diese Gruppe bewertet. Falls die Zuverlässigkeit dieser Aussage nicht hoch genug ist, wird die Gruppe gelöscht und die Analyse mit der neuen Menge wiederholt. Sonst wird der Prozeß beendet und ein Endergebnis abgeleitet.

4.1.6 Ergebnisse

Abb. 23 zeigt einen Ausschnitt aus dem Ergebnisbild. Zuerst wurde das Bild mit dem histogrammbasierten Verfahren nach spektralen Gesichtspunkten segmentiert und die resultierenden Farbgruppen mit Hilfe des Optimierungsverfahrens zusammengefaßt. Man erkennt in diesem Bild, daß die Farben richtig gruppiert wurden und daß die festgelegten Anforderungen erfüllt worden sind. Die Gruppen zeigen, außer ihrer Homogenität, auch guten räumlichen Zusammenhang und guten Kontrast.

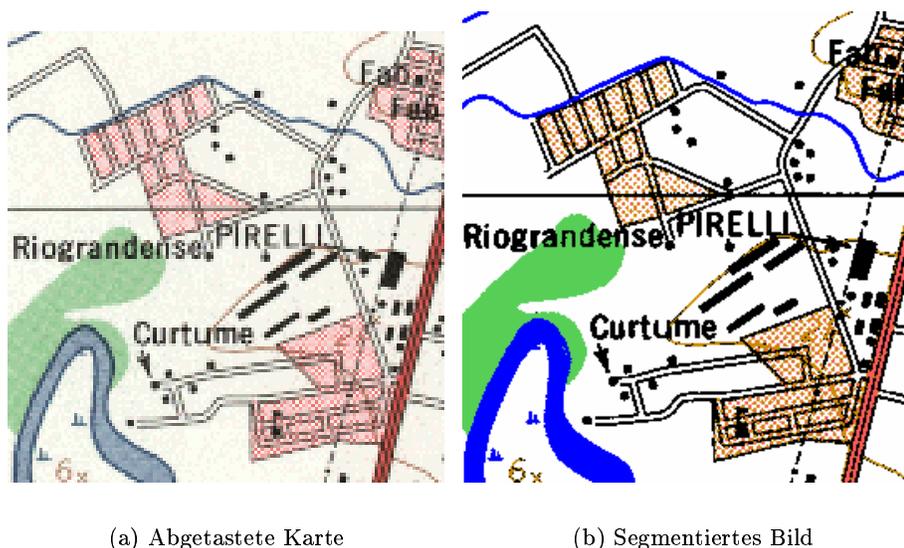


Abbildung 23: Ergebnis der Farbtrennung.

Abb. 24 zeigt die Entwicklung des Homogenitätsmaßes für jede Iteration. Das Homogenitätsmaß der Gruppen insgesamt sinkt nach jeder Iteration. Wenn verschiedene Farbgruppen zusammengefaßt werden, wird der Variationsbereich der neuen Gruppe dadurch erweitert und die Varianz erhöht.

Der Gesamtwert der Homogenität bleibt hoch, vor allem weil es sich um eine gerasterte Karte handelt, deren Farbgruppen kleine Varianzen haben. Die Veränderung nach jeder Iteration ist gering, denn nur spektral ähnliche Farbgruppen werden gemischt.

Wenn eine Farbgruppe den Qualitätsanspruch (Gl. 49) nicht erfüllt, dann wird sie in der spektral ähnlichsten Gruppe eingeschlossen. Voraussetzung dafür ist, daß die Pixel beider Gruppen häufig

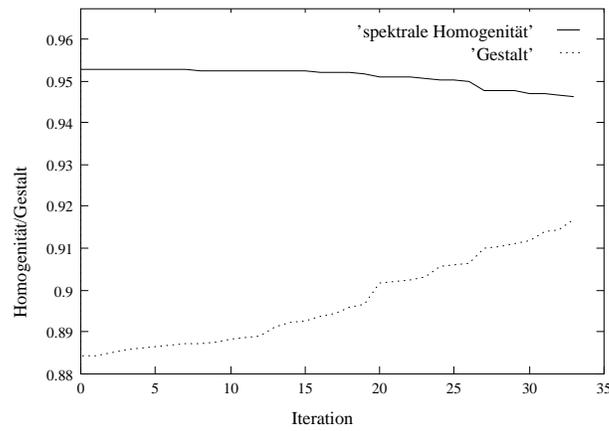


Abbildung 24: Evolution der spektralen Homogenität und Zusammenhang.

benachbart im Bild vorkommen. Dies läßt sich mit einem graphischen Beispiel zeigen. Abb. 25 zeigt drei Gruppen. Wenn die Gruppe 2 gelöscht werden soll und der Kontrast zwischen dieser Gruppe und den anderen (1 und 3) gleich wäre,

$$K(1, 2) = K(2, 3) \quad (50)$$

dann wäre es sinnvoller, die Pixel der Gruppe 2 der Gruppe 3 zuzuweisen, denn die Grenze (Gl. 48) zwischen diesen beiden Gruppen ist länger (Farben 1 und 2 kommen öfter zusammen vor).

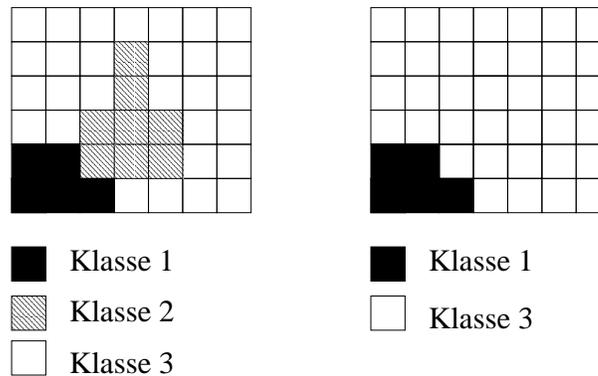


Abbildung 25: Beispiel der Verschmelzung einer schwachen Farbgruppe (Gruppe 2).

Da die Verschmelzung benachbarter Gruppen, die spektral ähnlich sind, zur Bildung größerer Flächen führt, verbessert sich auch der räumliche Zusammenhang der Gruppen, wie man in Abb. 24 sehen kann.

Der Kontrast wächst nach jeder Iteration (Abb. 26), denn durch die Mischung spektral ähnlicher Gruppen werden Kanten mit schwachem Kontrast beseitigt.

Die Bewertungsfunktion dieser drei Parameter spiegelt die Qualität des Bildes nach jeder Iteration wider. Die Abb. 27 zeigt einen Vergleich der Qualität, beschrieben nach Gl. 49, zwischen den ursprünglichen Farben (insgesamt 46 Farben) und den zusammengefaßten Farben (13 Farben). Die Farben, die ursprünglich eine niedrige Qualität zeigten, wurden nicht alle beseitigt, denn durch ihre Kombination wurden neue Gruppen erzeugt, die besseren Gestalt und Kontrast besitzen. Das kann anhand der Gruppen 5 oder 32 verdeutlicht werden. Dies ist eine Folge der räumlich-spektralen Bewertung, welche diesen Prozeß steuert.

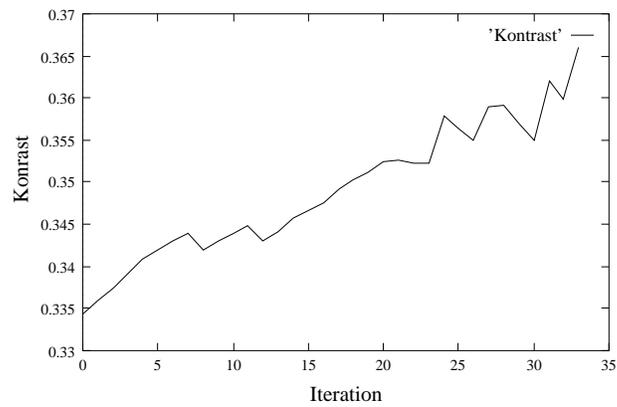


Abbildung 26: Evolution des Kontrastes.

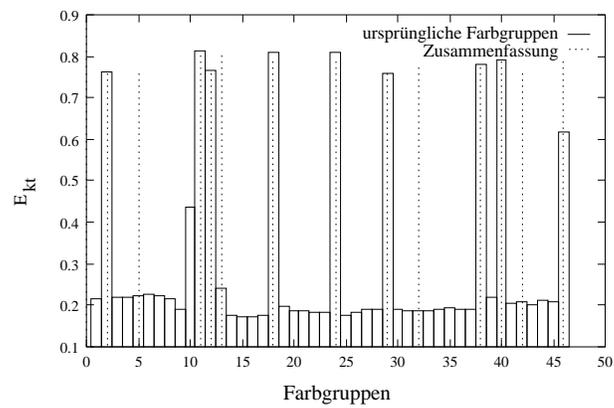


Abbildung 27: Zusammenfassung der Farbgruppen.

4.2 Erkennung von Landnutzungsgruppen

Grundelemente der kartographischen Gestaltungsmittel sind Punkt, Linie und Fläche, sowie zusammengesetzte Zeichen, Signaturen, Diagramme, Halbtöne und Schriftzüge (Hake & Grünreich [1994]).

Die in einem Satellitenbild sichtbaren Landnutzungsgruppen werden meist als Linien und Flächen in einer Karte dargestellt. Zu den linienhaften kartographischen Elementen in einer IBGE Karte (1:50.000) gehören Objekte wie Flüsse und Straßen. Kleine Flüsse werden als durchgehende/massive oder strichlierte blaue Linie dargestellt. Eine Straße ist als eine einzige Linie oder durch zwei parallele Linien in der Karte dargestellt. Andere Klassen wie Siedlungsgebiete, breitere Wasserflächen (größer als 4 mm im jeweiligen Maßstab, IBGE [1993]), Sumpfgebiete, Wälder oder Obstplantagen, werden als Flächen dargestellt. Dabei kennt man zwei Arten der Darstellung: geschlossene Flächen, wie zum Beispiel Wälder, und flächenhafte Signaturen, wie zum Beispiel Siedlungsgebiete.

4.2.1 Erfassung von Linien

Unter dem Begriff „linienhafte Elemente“ versteht man „alle nicht unterbrochenen Striche, die eine Lage angeben“ (Hake & Grünreich [1994]). Ihre graphische Variation ist hinsichtlich der Farbe oder der Strichbreite möglich. Linien werden benutzt, um Lageangaben über die Abgrenzung von flächenhaften Objekten (Umfang) anzugeben oder bei sehr schmalen Objekten zur Darstellung des Objekts selbst. In diesem Fall wird das Objekt durch eine Mittellinie dargestellt. Ein Spezialfall ist die Verwendung von zwei parallelen Linien, z.B. zur Repräsentation einer Straße; dabei ist der leere Zwischenraum ein wichtiger Bestandteil ihrer Darstellung.

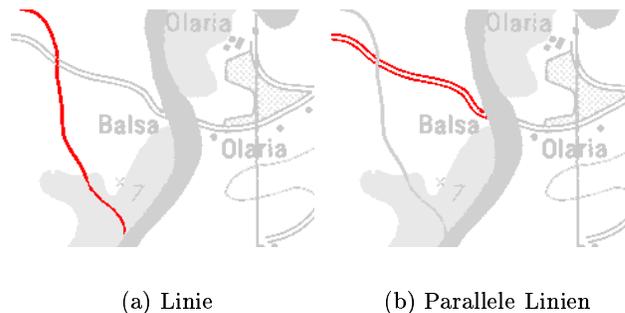


Abbildung 28: Zwei mögliche Darstellungen linienhafter kartographischer Elemente.

Die Erfassung linienhafter Elemente aus binären Bildern ist im wesentlichen ein Problem der Vektorsierung von Linien und wird in Bässmann & Besslich [1989] sehr detailliert beschrieben. Ein Großteil der bekannten Ansätze fängt mit der Merkmalsextraktion in einer low-level Ebene (*Signalebene*) an und versucht dann später Linien anzunähern. Zur Schätzung der Mittelachse der Linien bieten sich zwei Möglichkeiten an. Die erste ist die iterative Verdünnung der erkannten Regionen bis zu dem Punkt, an dem die Flächen zu einer Kette von einzelnen Pixeln reduziert werden. Diese Punkte werden zu Linien verkettet (Linienverfolgung). Der alternative Ansatz berechnet die Mittelachse aus der Konturen der Regionen. Benachbarte Liniensegmente werden später zu größeren Linien aufgrund ihrer Abstände und Richtungen zusammengefaßt (Bässmann & Besslich [1989]).

4.2.2 Erfassung von Flächen

Als flächenhafte Elemente versteht man nicht nur die Regionen einer Karte, die in Form einer kompakten Menge gleichfarbiger Pixel vorliegen, sondern auch die Regionen, die durch Muster und Schraffuren

dargestellt werden (Abb. 29).

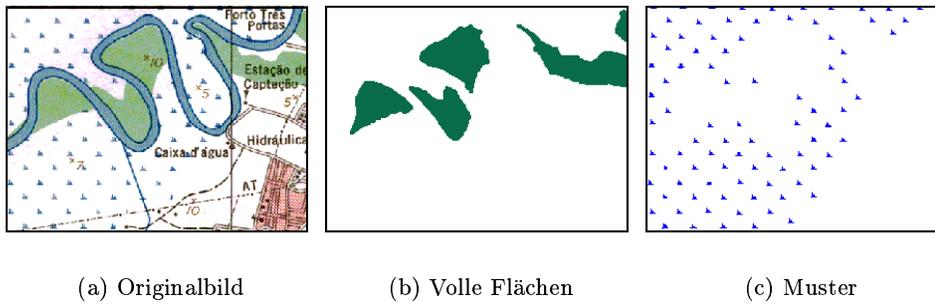


Abbildung 29: Zwei mögliche flächenhafte Regionen.

Zur Erfassung gleichfarbiger Regionen werden, ähnlich wie bei der oben genannten Erfassung von Linien, Methoden der Vektorisierung angewandt. In diesem Fall werden die Konturen der Regionen untersucht. Wichtig ist hier die Entscheidung, ob eine Region als Linie oder Fläche erfasst wird. Lauterbach [1993] benutzt das von Wahl [1984] vorgeschlagene Kriterium zur Unterscheidung von linien- und flächenhafter kartographischer Elemente. Dieses Verfahren basiert auf der Bereichstiefe T der Bildpunkte i , die mit Hilfe der euklidischen Distanz zum Hintergrund (DH) berechnet wird. Anhand der Form des Histogramms von DH ($h(DH)$) läßt sich eine Entscheidung über die Art der Region treffen (Abb. 30).

$$DH_i = \sum_{i=1}^n d(T_i - i) \quad (51)$$

Hier gilt:

$d(k) = 1$, wenn $k = 0$,
sonst $d(k) = 0$.

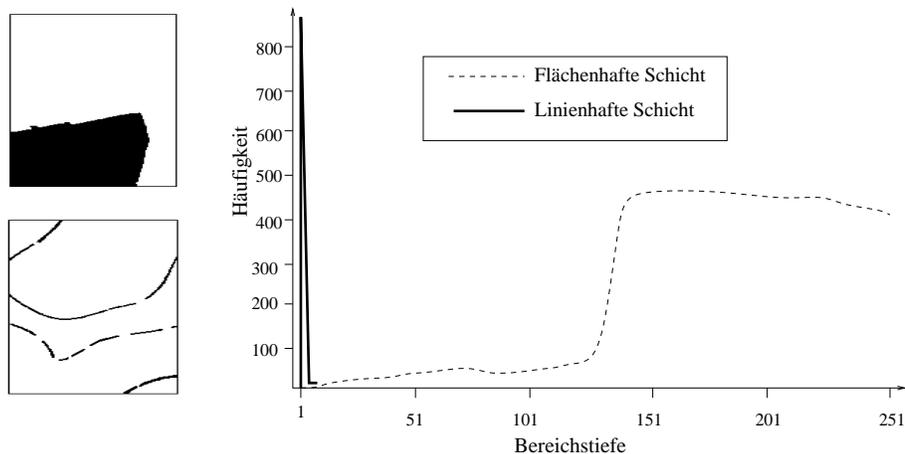


Abbildung 30: Unterscheidung zwischen Flächen- und linienhaften Regionen, Lauterbach [1993].

4.2.3 Erfassung von Mustern und parallelen Linien

Die Erfassung von Mustern und Linien ist schwieriger, da die Flächen nicht mehr allein aufgrund einer Farbe, die die gesamte Region ausfüllt, erkannt werden. Zwischen den Symbolen und Linien entstehen leere Räume. Diese leeren Räume zwischen parallelen Linien, sowie in flächenhaften Signaturen, können als Hinweis für ihre Erfassung benutzt werden, denn sie sind ein wichtiger Bestandteil dieser graphischen Elemente. Beispiele werden in Abb. 31 gezeigt. Die grauen Flächen sind Zwischenräume und liegen zwischen den Pixeln des Vordergrundes (schwarz).

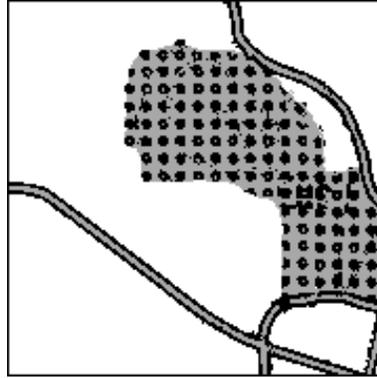


Abbildung 31: Zwischenräume in einem binären Bild.

Da die farbige Schicht ein binäres Bild ist, haben ihre Pixel nur zwei mögliche Bedeutungen, sie stellen entweder Symbole, Linien, Flächen oder Muster (Vordergrund) dar, oder sie sind Teil des Hintergrunds. Die Zwischenräume gehören zum Hintergrund, sie sind von Vordergrundpixeln umgeben, wobei die Vordergrundpixel sie nicht unbedingt vollständig einrahmen müssen. Drei getrennte Vordergrundpixel können einen Zwischenraum im zweidimensionalen Raum definieren. Wenn drei Punkte q_a , q_b und q_c nicht auf einer Geraden liegen, dann wird als gemeinsame innere Region das Dreieck definiert, dessen Scheitelpunkte durch diese drei Punkte repräsentiert werden (Abb. 32).

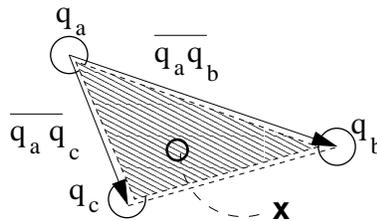


Abbildung 32: Raum zwischen drei Vordergrundpixeln q_a , q_b und q_c .

Die Erkennung der Zwischenräume ist ein Problem der Interpolation, denn diese Regionen werden nicht explizit im Bild dargestellt, sondern aus den Grenzen der Vordergrundpixel in der Umgebung abgeleitet. Hier werden zwei Methoden diskutiert. Die erste basiert auf einem Verfahren, welches benutzt werden kann, um Wasserscheiden in einem digitalen Geländemodell zu suchen. Dieses Verfahren ist rasterbasiert. Im Gegensatz dazu benutzt das zweite Verfahren die vektorielle Darstellung der Vordergrundpixel um die Zwischenräume zu erkennen. Es basiert auf der Darstellung der Information mit Hilfe einer Delaunay-Triangulation. Diese beide Methoden werden zunächst erklärt.

4.2.4 Das Wasserscheide-Verfahren

Das Wasserscheide-Verfahren wurde ursprünglich zur Erkennung von Einzugsgebieten in Digitalen GeländeModellen (DGM) entwickelt. Das Prinzip dieses Verfahrens wird in Mark [1983] und Jenson &

Domingue [1988] bei der Anwendung für hydrologische Zwecke beschrieben. Das Ziel ist es, den Abfluß des Wassers auf der von einem digitalen Geländemodell dargestellten Oberfläche zu modellieren.

Die Grundidee dabei ist, die möglichen Wege für den Oberflächenabfluß auf dem DGM aus den lokalen Gradienten zu schätzen. Da es sich beim DGM meist um ein quadratisches Gitter handelt, gibt es nur acht mögliche Flußrichtungen für jede Pixelposition. Es wird angenommen, daß das Wasser in die Richtung des größten Gradienten fließt.

Zuerst wird die Flußrichtung für jedes Pixel im DGM berechnet. Dafür wird in einem 3x3 Fenster der Gradient in den 8 möglichen Richtungen geschätzt. Der größte Gradient bestimmt die Flußrichtung an der jeweiligen Stelle, und diese Richtung wird in einer neuen Matrix gespeichert.

Der lokale Gradient bestimmt die Flußrichtung eines Pixels, und die Verbindungslinien dieser Flußrichtungen bestimmen die Wege für den Oberflächenabfluß auf dem DGM. Auf einer idealen Oberfläche, würden alle Verbindungslinien am Rande des DGM enden, bzw. der Abfluß endet außerhalb der Matrix.

Wenn Vertiefungen im DGM vorliegen, kann dort der Abfluß nicht modelliert werden. Deswegen müssen diese Senken aufgefüllt werden. Das geschieht, indem man sie identifiziert und ihre Einzugsgebiete bis zu der kleinstmöglichen Höhe, die einen Abfluß aus der Senke garantiert, auffüllt.

Das Wasserscheide-Verfahren wurde auch in der Bildverarbeitung angewendet (Vincent & Soille [1991]). Das DGM wird in diesem Fall durch die Variation der Grauwerte im Bild ersetzt. Dabei wird ein künstliches digitales Geländemodell (KDGM) erzeugt. Die Erkennung der Vertiefungen in diesem KDGM entspricht dann die Segmentierung von Regionen.

Eine schwache Variation der Grauwerte, die der Höhenvariation eines Geländemodells ähnelt, gibt es in einem binären Bild nicht. Da bei binären Bildern nur zwei Werte erlaubt sind, würde das Verfahren kein brauchbares Ergebnis liefern. Es ist deswegen nötig, zuerst ein künstliches DGM mit einer stetigen Variation aus dem binären Bild abzuleiten. Zu diesem Zweck kann man eine Funktion, die umgekehrt proportional zum Abstand jedes Pixels zu dem nächstliegenden Vordergrundpixel ist, benutzen.

Wenn x ein Pixel der Klasse „Hintergrund“ und y das nächstliegende Vordergrundpixel zu diesem Pixel ist, dann ist $d(x, y)$ der kürzeste Abstand zwischen x und y . Eine Funktion, die kleiner wird, je größer der Abstand jedes Pixels zu dem nächstliegenden Vordergrundpixel wird, erzeugt ein künstliches DGM:

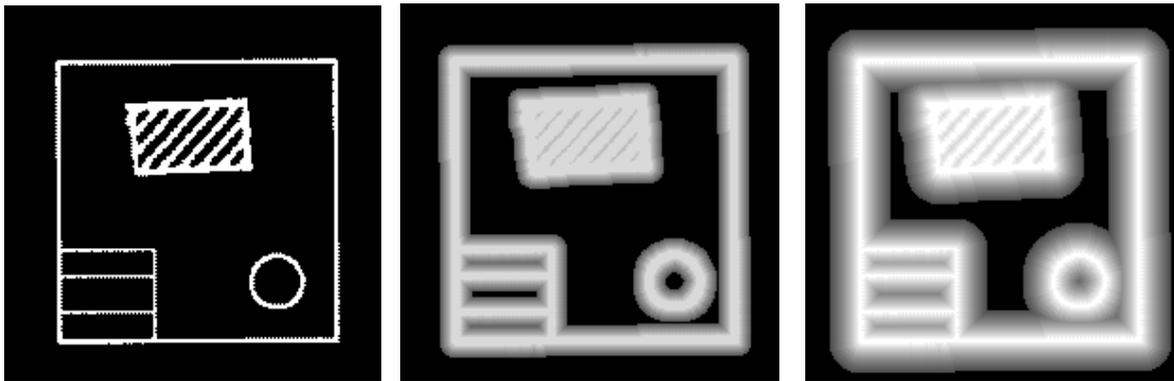
$$f(x) = 255 - d(x, y) \quad \text{wenn } d(x, y) \leq d_m \quad (52)$$

$$f(x) = 0 \quad \text{wenn } d(x, y) > d_m \quad (53)$$

$$h(x) = \max(f(x)) \quad (54)$$

Abb. 33 zeigt zwei KDGM, die aus einem einfachen binären Bild (a) erzeugt werden können. Das erste Modell (b) wird mit einem maximalen Abstand von 8 Pixeln erzeugt, das zweite Modell (c) mit einem maximalen Abstand von 16 Pixeln.

Der Schwellwert d_m , mit dem die Grauwertvariation erzeugt wird (Gl. 54), bestimmt die Länge der Talwege im künstlichen Geländemodell. Wenn die Vordergrundpixel weit voneinander entfernt liegen, dann enden die Talwege an einem Hintergrundpixel. Wenn der Abstand zwischen den Vordergrundpixeln kleiner als $2 * d_m$ ist, dann werden „geschlossene Täler“ bzw. Vertiefungen erzeugt, zum Beispiel zwischen den parallelen Linien in Abb. 33c. Die Erkennung dieser Vertiefungen entspricht der Erkennung von Zwischenräumen im binären Bild.



(a) Binäres Bild

(b) Künstliches Geländemodell mit maximalem Abstand von 8 Pixeln

(c) Künstliches Geländemodell mit maximalem Abstand von 16 Pixeln

Abbildung 33: Erzeugung eines KDGM aus einem binären Bild.

Abb. 34 illustriert das Verfahren zur Erkennung der inneren Regionen durch einen einfachen Fall. Hier wird nur eine Zeile eines binären Bildes bearbeitet. Abb. 34a zeigt die ursprüngliche binäre Zeile. Der Hintergrund wird mit Null und der Vordergrund mit dem maximalen Wert dargestellt. Aus dieser Zeile können neue künstliche Geländemodelle (Zeilen) mit unterschiedlichem Radius um die Vordergrundpixel erzeugt werden (Abb. 34b und d). Diese neuen Werte sind alle kleiner als die Werte der Vordergrundpixel. Der Unterschied wächst mit dem Abstand zwischen den Pixeln und den Vordergrundpixeln. Abb. 34b zeigt ein KDGM mit maximalem Abstand von 3 Pixeln, und Abb. 34d zeigt ein KDGM mit maximalem Abstand von 4 Pixeln. Je nach dem Wert des maximalen Abstands werden „Täler“ zwischen den Vordergrundpixeln gebildet. Die *Talwege*, die in einem Hintergrundpixel enden, kennzeichnen offene Strecken die nicht berücksichtigt werden (gekennzeichnet mit gestrichelten Linien). Benachbarte „Vertiefungen“ können zu größeren Regionen zusammengefaßt werden (Abb. 34.c und e).

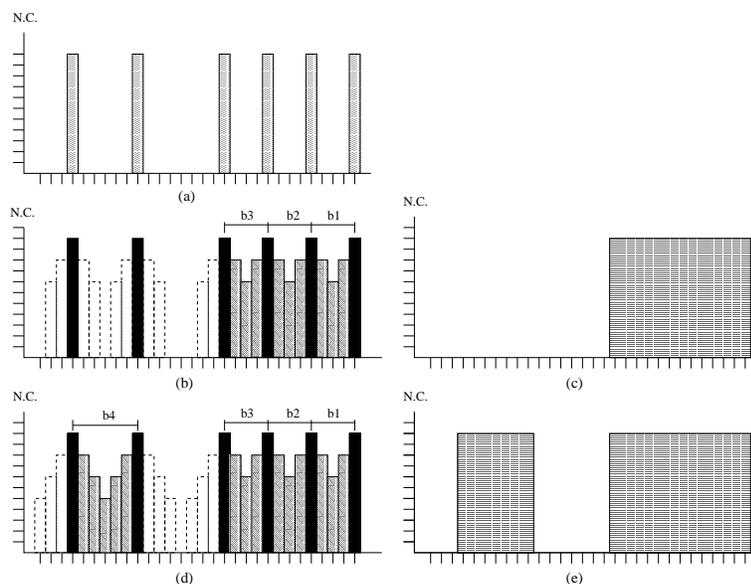


Abbildung 34: Erkennung von inneren Räumen mit dem Wasserscheide-Verfahren. Die Abbildung zeigt die Zeile eines binären Bildes. (a) Binäres Bild. (b) Künstliches DGM mit $d=3$ Pixeln. (c) Vertiefungen für $d=3$. (d) Künstliches DGM mit $d=4$ Pixeln. (e) Vertiefungen für $d=4$.

Ein Beispiel der Anwendung dieses Prinzips auf einen zweidimensionalen Datensatz wird in Abb. 35 gezeigt. Je nach der Größe von d_m (Gl. 54) werden verschiedene Zwischenräume erkannt. Die Wahl von d_m steuert die Fähigkeit des Verfahrens, Zwischenflächen zu erkennen. Dieser Wert muß sorgfältig gewählt werden, um bestimmte Flächen, z.B. die Räume zwischen den parallelen Linien, zu erfassen.

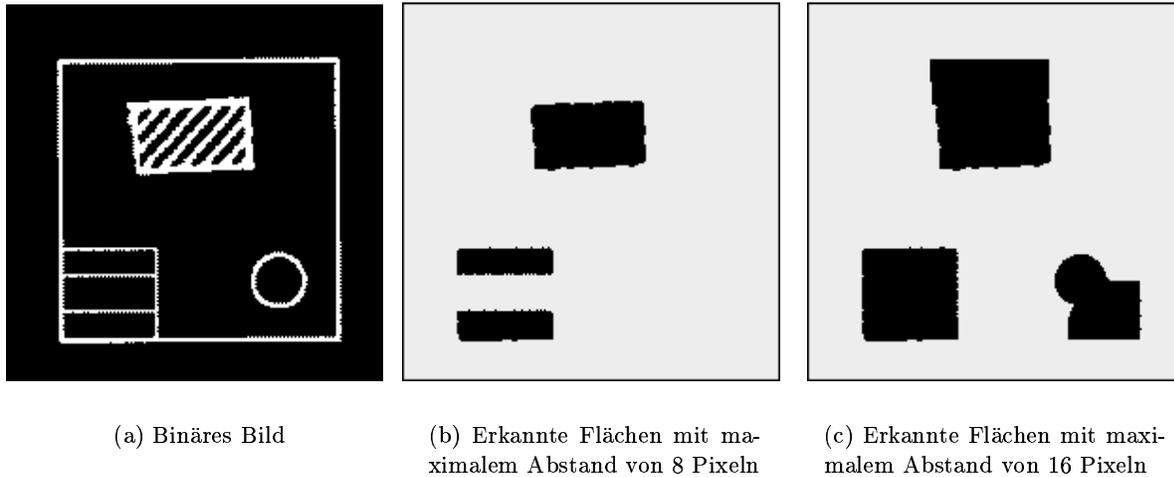


Abbildung 35: Erkannte Zwischenräume mit Hilfe des Wasserscheide-Verfahrens.

4.2.5 Erfassung des Straßennetzes mit dem Wasserscheide-Verfahren

Das Straßennetz in einer Karte vom IBGE mit dem Maßstab 1:50.000 wird als Netz von parallelen Linien dargestellt (Abb. 28).

Parallele Linien bleiben nach der Analog-Digital-Wandlung nicht notwendigerweise erhalten. Stattdessen kann man behaupten, daß ein kleiner maximaler Abstand zwei parallele Linien in einem gerasterten Bild trennt. Da der Raum zwischen den Linien sehr eng ist, reicht ein kleiner Abstand aus, um sie mit Hilfe des Wasserscheide-Verfahrens zu erfassen.

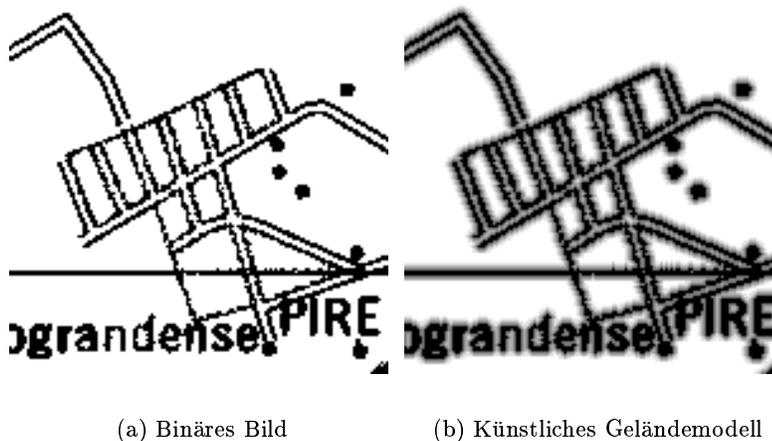


Abbildung 36: Erzeugung eines künstlichen Höhenmodells im Zwischenraum paralleler Linien.

Abb. 36b zeigt das DGM, das aus dem binären Bild (Abb. 36a) mit einem maximalen Abstand von drei Pixeln erzeugt wurde. Höhere Werte wurden dunkel dargestellt und eine weiße Farbe steht für Null. Die Vertiefungen befinden sich nicht nur zwischen parallelen Linien, sondern auch in den Lücken und am Rande der Buchstaben. Parallele Linien, die weiter als drei Pixel voneinander entfernt liegen, bilden kein geschlossenes Einzugsgebiet (z.B. die Siedlungsblöcke).

Abb. 37a zeigt die geschlossenen Einzugsgebiete. Die engen Räume innerhalb der Schrift und zwischen anderen Linien wurden mit den Straßen zusammengefaßt. Um diese Elemente zu beseitigen und nur die Straßen zu behalten, wurden die erfaßten Flächen nach der Größe sortiert. Das Ergebnis nähert das Straßennetz sehr gut an. Die Abb. 37b stellt die erkannte Straßen als schwarze Linien dar. Dieses Netz zeigt jedoch Lücken, die eine vollständige Darstellung verhindern.

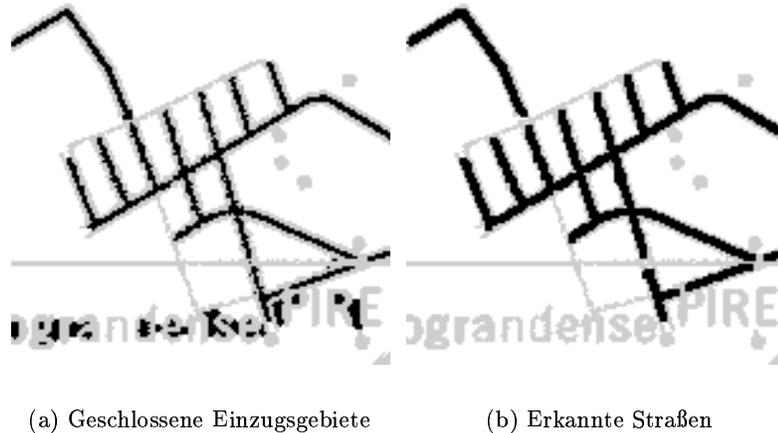


Abbildung 37: Erkennung der Zwischenräume bei parallelen Linien.

Diese Lücken können geschlossen werden, wenn man eine Analyse der Nachbarschaft durchführt. Wenn ein Ende einer Linie frei ist, wird der Raum um den Punkt nach anderen freien Enden untersucht. Falls ein entsprechender Punkt innerhalb eines bestimmten Radius gefunden wird, werden beide mit einem neuen Liniensegment verbunden (Bässmann & Besslich [1989]). Die genaue Verbindung der Linien ist aber für die Ziele dieser Arbeit nicht wichtig, denn die Informationen werden als Fuzzy-Mengen bearbeitet. In Abs. 6.1.1 wird auf dieses Thema genauer eingegangen.

4.2.6 Erkennung von Zwischenräumen mit dem Wasserscheide-Verfahren

Ein maximaler Radius kennzeichnet die Räume zwischen Mustern (Abb. 38). Zwischenräume innerhalb dieses Radius können mit dem Wasserscheide-Verfahren identifiziert werden, wenn man aus dem binären Bild ein künstliches DGM erzeugt und die Vertiefungen in dieser Oberfläche erkennt. Ein künstliches DGM einer gemusterten Region wird in Abb. 38b dargestellt. Die Erkennung der Vertiefungen in diesem DGM entspricht der Erkennung innerer Regionen im binären Bild.

Abb. 39 zeigt ein Beispiel der Anwendung dieses Verfahrens zur Erkennung von Flächen mit Mustern. Der Abstand zwischen den Symbolen dieser Regionen ist insgesamt ähnlich. Die dunkleren Vielecke zeigen die Konturen, die aus diesen Flächen gebildet werden. Das Verfahren ist in der Lage, die Zwischenräume innerhalb der gemusterten Regionen richtig zu erfassen.

Die Leistungsfähigkeit des Verfahrens an den Rändern der Regionen ist nicht von der Vollständigkeit der Symbole abhängig, denn sie erfolgt nur auf Grund der Analyse der Zwischenräume. Trotzdem

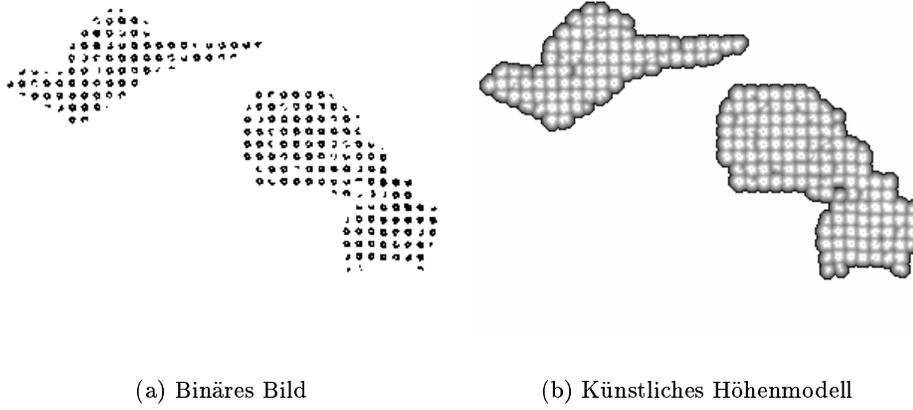


Abbildung 38: Generierung eines künstlichen Höhenmodells aus dem binären Bild.

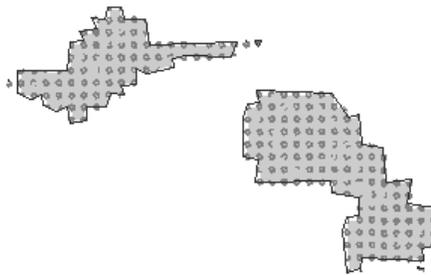


Abbildung 39: Ergebnis der Segmentierung von Muster nach dem Wasserscheide-Verfahren.

können Probleme an diesen Stellen auftreten. Wenn die Muster kollinear angeordnet sind, dann werden sie nicht erfaßt, denn sie definieren keinen zweidimensionalen Zwischenraum. Das geschieht vor allem in Bereichen wo die Flächen Spitzen bilden. Ein entsprechendes Beispiel ist in Abb. 39 dargestellt.

4.2.7 Die Delaunay-Triangulation

Eine mögliche Lösung zur Segmentierung der Karte aufgrund des Abstands zwischen gleichfarbigen Pixeln ist die Anwendung von zweidimensionalen Netzen wie dem Delaunay-Netz, das aus den Punkten eines Musters gebildet wird. Eine Beschreibung der relativen Position und des Abstands zweier nah beieinander liegenden Pixel kann aus diesem Netz abgeleitet werden.

Zweidimensionale Netze, wie die Voronoi- und Delaunay-Netze, haben große Bedeutung für die Analyse der räumlichen Beziehung zwischen benachbarten graphischen Elementen. Eine Delaunay-Triangulation besteht aus Dreiecken, die benachbarte Punkte im Bild verbinden, wobei diese Punkte auf dem Rand eines leeren Kreises liegen (Okabe *et al.* [1992]). In einem klassifizierten oder segmentierten Bild wird die räumliche Beziehung zwischen Pixeln, die zur selben Klasse gehören, untersucht.

Aus einer Menge P von n Punkten im Bereich Ω^2 des zweidimensionalen Raumes eines Bildes (R^2) werden jeweils zwei Punkte (p_i und p_j) durch Segmente der Delaunay-Triangulation verbunden (eine Kante wird gebildet) wenn ein dritter Punkt ($p_k \in \Omega^2$) existiert, der nicht kollinear zu diesen Punkten liegt, und kein anderer Punkt innerhalb des Kreises, der diese drei Punkte verbindet, existiert (Abb. 40).

$$\exists \Delta(i, j, k) \Leftrightarrow \exists c \mid c \in \Omega^2 \wedge r = d(c, p_i) = d(c, p_j) = d(c, p_k) \wedge r < d(c, p_x) \forall x \neq i, j, k \quad (55)$$

mit:

r = Radius des Kreises,

c = Zentrum des Kreises.

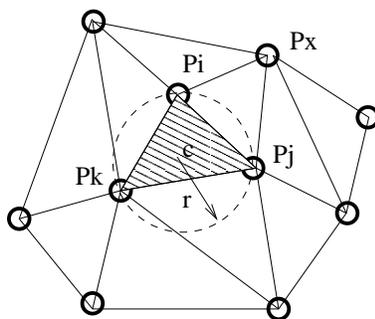


Abbildung 40: Delaunay Dreieck.

Weindorf [1994] beschreibt die Benutzung der Delaunay-Triangulation zur Zusammenfassung homogener Regionen in Satellitenbildern. Der Abstand zwischen den Pixeln wurde mit Hilfe des Umfangs der Triangulationsdreiecke ausgedrückt. Die Grundidee dieses Ansatzes ist, größere Dreiecke, die aus voneinander entfernten Pixeln bestehen, zu beseitigen und die übrigen Dreiecke zu größeren Flächen zu gruppieren.

Die Wahl des maximalen akzeptierbaren Umfangs, der die Trennung von kleinen und großen Dreiecken steuert, ist entscheidend. Ein kleiner Wert schließt nur nah beieinander liegende Pixel ein. Je größer dieser Wert ist, um so größere Dreiecke werden in die Flächen eingefügt. Dadurch kann es passieren, daß auch Pixel, die relativ weit voneinander entfernt liegen, als eine Fläche zusammengefaßt werden. Ein Nebenprodukt dieses Ansatzes ist die Bildung kleinerer Flächen, die aus isolierten kleinen Dreiecken entstehen. Sie werden mit Hilfe eines Schwellwertes für die minimal erlaubte Fläche beseitigt (Vögtle & Schilling [1995]).

4.2.8 Erkennung der Zwischenräume mit der Delaunay-Triangulation

Flächen mit Mustern bestehen aus zwei Komponenten: den Vordergrundpixeln, die Symbole darstellen, und den leeren Räumen (Hintergrund) zwischen diesen Symbolen (Abb. 38).

Die Triangulation verbindet die Mittelpunkte der Vordergrundpixel. Der Umfang der Dreiecke liefert Hinweise über den Abstand, in dem die Symbole voneinander liegen und kann zur Extraktion der Regionen benutzt werden. Dreiecke, die Pixel desselben graphischen Elements (ein Symbol, zum Beispiel) verbinden, haben oft einen minimalen Umfang. Andererseits werden Dreiecke, die Pixel von zwei oder drei verschiedenen Symbolen verbinden, durch einen größeren Umfang gekennzeichnet. Dieser Unterschied wird in Abb. 41 veranschaulicht. Das linke Bild zeigt das binäre Bild einer Region mit Mustern. Die Triangulation innerhalb der markierten Region in Abb. 41a wird in Abb. 41b dargestellt. Sehr kleine Dreiecke verbinden Pixel desselben graphischen Elements, mittelgroße Dreiecke verbinden naheliegende Elemente und die großen Dreiecke (grau) entstehen durch die Verbindung von weit entfernten Pixeln.

Die Dreiecke können nach ihrem Umfang sortiert werden. In Abb. 41b wurden die Dreiecke, deren Umfang größer als ein Schwellwert (10 Einheiten) ist, grau dargestellt. Die übrigen Pixel gehören zu graphischen Elementen der gemusterten Region. So werden die Pixel dieser gemusterten Region von den Pixeln außerhalb der Region getrennt (Vögtle & Schilling [1995]).

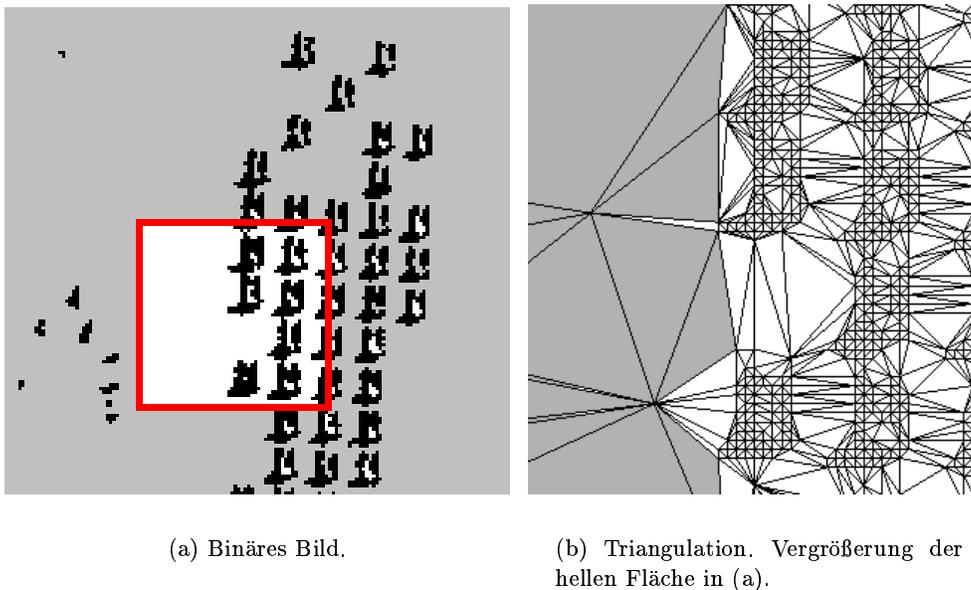


Abbildung 41: Beispiel der Segmentierung mit der Triangulation.

Die Erkennung von Flächen mit unterschiedlicher Dichte erfolgt durch die Variation des Schwellwertes für den Umfang. Zuerst werden die Regionen mit höherer Dichte erkannt. Sie werden später nach der Größe sortiert, und die größten werden gespeichert. Danach wird der Schwellwert vergrößert, um Regionen mit kleinerer Dichte zu extrahieren. Kleinere Flächen werden wieder beseitigt und die größeren Regionen gespeichert. Auf diese Weise werden die Flächen nach ihrer Dichte getrennt.

Um die gefüllten Flächen zu erkennen reicht ein minimaler Schwellwert, weil alle ihre Pixel aneinander grenzen. Diese Regionen werden dann aus dem Bild gelöscht, und nur die Regionen mit Mustern und das Rauschen bleiben übrig. Ein größerer Schwellwert erlaubt, die Regionen mit Mustern von dem Rauschen zu trennen.

Ein Beispiel der Trennung von Regionen nach ihrer Dichte wird in der Abb. 42 dargestellt.

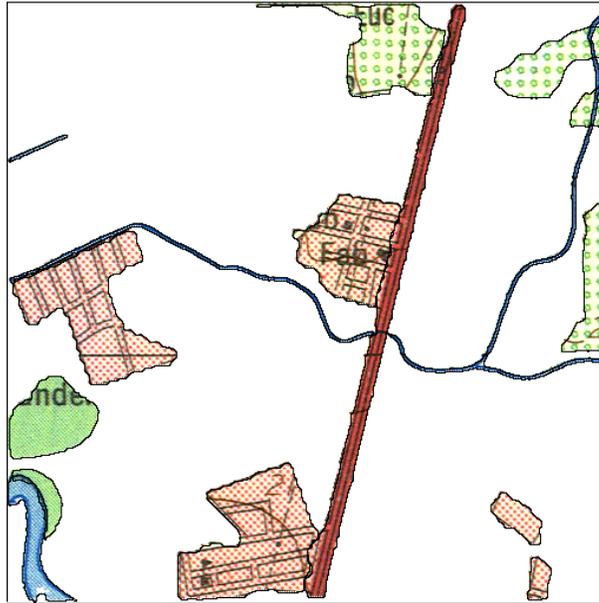


Abbildung 42: Ergebnis der Segmentierung von Farbschichten nach dem Umfang der Dreiecke in der Delaunay-Triangulation.

4.2.9 Erkennung von Flächen mit Mustern

Beide Ansätze, die Triangulation und das Wasserscheide-Verfahren haben Vor- und Nachteile.

Die Triangulation ist eine rechenintensive Arbeit. Das wäre im Prinzip kein Nachteil, wenn man genügend Rechnerleistung zur Verfügung hätte. Besonders bei der Verarbeitung einer gescannten Karte, mit einer großen Anzahl von Pixeln, wird die Anzahl von Punkten, Kanten und Dreiecken sehr groß. Der wichtigste Nachteil wäre, daß das Auswahlkriterium, das benutzt wird um die Dreiecke zu sortieren, sehr empfindlich auf Rauschen in der Nähe der Grenzen der Region reagiert. Ein Kriterium, welches nur auf der Länge des Umfangs der Dreiecke basiert, ist nicht in der Lage Rauschen zu erkennen. Ein Beispiel wurde in Abb. 41b gezeigt.

Das Wasserscheide-Verfahren ist ein rasterbasiertes Verfahren. Seine Prinzipien sind einfach und deswegen kann es leicht implementiert werden. Der größte Nachteil ist die Tatsache, daß es scheitert wenn die Zwischenräume kein geschlossenes Einzugsgebiet (Senken im künstlichen DGM) bilden.

Das Wasserscheide-Verfahren wurde hauptsächlich benutzt, um Flächen mit Mustern zu erkennen, weil diese Regionen selten problematische Spitzen haben. Nur in Fällen, in denen das Wasserscheide-Verfahren scheiterte, wurde die Triangulation benutzt. Der Vorteil bei dieser Kombination war, daß nach der Erkennung eines großen Teils des Bildes mit dem Wasserscheide-Verfahren, die Anzahl der Punkte sehr stark reduziert wurde.

Abb. 43a zeigt einen Teil eines Ergebnisses des Wasserscheide-Verfahrens. Die Konturen folgen den Grenzen der geschlossenen Einzugsgebiete und ihr Verlauf ist unregelmäßig. Die Fläche der Region wird auf diese Weise unterschätzt. Diese erste Näherung wird durch die Berechnung der Hülle dieser Vielecke deutlich verbessert (Abb. 43b).

Abb. 44 zeigt das Prinzip der Berechnung der Hülle der Konturen. Es wird versucht, kleine Abweichungen zu detektieren ($d1$ in Abb. 44a) und sie durch die Hülle der Figur zu ersetzen. Größere Krümmungen werden nicht beseitigt, denn sie gehören zur Form der Figur. Der Prozeß fängt mit der Analyse am ersten Punkt der Vielecke an, zum Beispiel p_0 in Abb. 44b, und sucht nach hohlen Konturabschnitten. Die Richtung der Geraden, die diesen Punkt mit dem folgenden verbindet, wird analysiert. Diese Richtung wird als der Winkel zwischen der Geraden $(p, p + 1)$ und $(p, p + j)$, mit $j > 1$, gemessen. An einer hohlen Strecke wächst dieser Winkel mit der Entfernung zum Referenz-

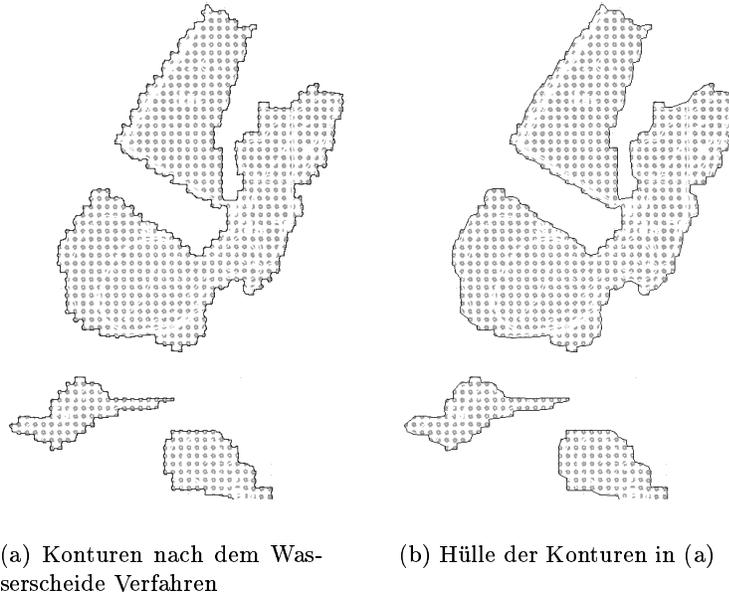


Abbildung 43: Konturen.

punkt, z.B. $p + 2$ und $p + 3$ in Abb. 44b. Der Punkt, der unter dem maximalen Winkel erscheint, markiert den gesuchten Verbindungspunkt. Dieser Konturabschnitt kann dann durch die neue Gerade ersetzt werden, wenn der maximale Abstand zwischen der Geraden und den ursprünglichen Punkten ($d1$ und $d2$ in Abb. 44a) kleiner als ein Schwellwert ist. Nach der Ersetzung der Strecke wird dieselbe Analyse ab dem letzten Punkt wiederholt ($p + 3$ im Beispiel).

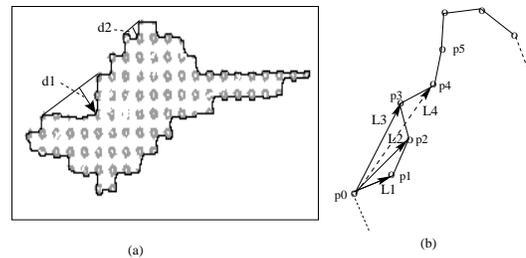


Abbildung 44: Berechnung der Hülle.

Das Ergebnis ist eine einfache Kurve, glatter als die ursprüngliche. Es besteht immer die Gefahr, daß mit den kleinen Krümmungen ein Teil der Information verloren geht. Ein Vorteil dieser Konturapproximation ist, daß die Anzahl von Punkten stark reduziert wird.

5 Ableitung von Information aus Satellitenbildern

Die Aufgabe der Bildanalyse ist es, Objekte der realen Welt in einer ikonischen Darstellung zu erkennen. Bei der Analyse von Satellitenbildern versucht man Objekte der Erdoberfläche zu erkennen. Objekte zu erkennen heißt, sie von anderen Objekten zu unterscheiden. In der Bildanalyse nutzt man zur Unterscheidung Merkmale der Objekte wie Farbe, Form oder Textur (Lillesand & Kiefer [1994]). Daneben lassen sich die Objekte auch durch ihre unterschiedlichen Relationen, die sie zu anderen Objekte aufbauen, unterscheiden.

Satellitenbilder haben die Eigenschaft, daß sie eine geringere räumliche Auflösung aufweisen. Im Gegensatz dazu liefern sie aber oft eine höhere spektrale Auflösung.

Zwischen geometrischer Auflösung und erkennbaren Merkmale in einem Bild gibt es einen Zusammenhang: Die geometrischen Strukturen (z.B. Form und Textur) gehen bei geringer Auflösung immer mehr verloren. Die Objekte lassen sich aber auch bei geringer Auflösung immer noch an den spektralen Merkmalen erkennen. Diese Eigenschaft nutzt man bei den Satellitenbildern aus, um auch bei geringer Auflösung noch Objekte zu erkennen.

Es haben sich viele Klassifikationsverfahren entwickelt, die aufgrund der spektralen Signatur eines Pixels auf eine bestimmte Klassenzugehörigkeit schließen. Dabei kann nicht mehr das einzelne Objekt in seiner ganzen Form und Anordnung beschrieben werden, sondern nur die wahrscheinlichste Klasse des Objektes, das das Pixel beschreibt.

Derartige rasterbasierte Klassifikatoren analysieren nur die Daten eines isolierten Pixels und vernachlässigen die Information der Nachbarschaft. Räumliche Beziehungen, die oft mehr Information über die Klassenzugehörigkeit eines Pixels liefern könnten, werden nicht berücksichtigt.

Zu den Ansätzen, die die Pixelnachbarschaft untersuchen, sind auch die texturbasierten Verfahren (Pratt [1974]) zu rechnen. Mit der Berücksichtigung räumlicher Beziehungen werden neue Informationen über die Objekte in die Analyse einbezogen, welche die spektralen Informationen ergänzen und die Erkennung der Klassen erleichtern.

Die Nutzung von räumlichen Beziehungen zwischen Objekten zur Analyse von Satellitenbildern ist bisher noch wenig untersucht. Es existieren lediglich einfache Ansätze, die die unmittelbare Pixelnachbarschaft bei der Klassifikation berücksichtigen (z.B. Haralick & Shapiro [1997]).

In dieser Arbeit werden spektrale Merkmale und räumliche Beziehungen zwischen Objekten bei der Analyse von Satellitenbildern benutzt um Information über die Szene abzuleiten.

5.1 Ableitung spektraler Informationen

Bei multispektralen Fernerkundungsdaten erfolgt die Klassifikation zum Beispiel im sogenannten Spektralraum. Jedes Pixel eines Multispektralbildes mit N Kanälen wird als N -dimensionaler Vektor von Grauwerten repräsentiert. Den so definierten Vektorraum bezeichnet man als Spektralraum. Diejenigen Pixel, die Materialien mit ähnlichen spektralen Eigenschaften abbilden, liegen im Spektralraum dicht beieinander. Pixel einer Klasse bilden also Häufungen im Spektralraum. Bei der Klassifikation wird versucht diese Häufungen durch Parameter zu beschreiben. Zum Beispiel läßt sich die spektrale Signatur einer Klasse durch ihren Mittelwertvektor im Spektralraum beschreiben.

Zur Zuordnung der Pixel zu einer Klasse werden Entscheidungen gebildet, die den Spektralraum in fest abgegrenzte Gebiete unterteilen. Zum Beispiel entscheidet man sich bei einem Pixel für diejenige Klasse, die im Spektralraum die geringste Distanz (oder größte Wahrscheinlichkeit) zum Mittelvektor der Klasse aufweist.

Aus derartigen Entscheidungsregeln folgt, daß ein Pixel immer nur zu einer einzigen Klasse gehören kann. Der Zugehörigkeitsgrad wird oft nur binär dargestellt. Die Information über die Distanz (oder bedingte Wahrscheinlichkeit) wird meist nicht weiterverwendet.

Sobald aber der Vektor des Pixels im Spektralraum nicht mit dem Klassenmittelwertvektor zusammenfällt, besteht ein gewisser Grad an Unsicherheit.

In dieser Arbeit wird die Wahrscheinlichkeit über die Zugehörigkeit zu einer Klasse mit Hilfe der Methoden der unscharfen Logik modelliert.

In der Fuzzy-Logik wird der Zugehörigkeitsgrad als eine unscharfe Menge dargestellt. Das Problem liegt darin, die Funktion des Zugehörigkeitsgrades abzuleiten. Wang [1990] schlägt einen Vergleich des Mahalanobis-Abstands zwischen dem Pixel und den Klassen vor, als Hinweis für den Grad mit dem das Pixel zu jeder Klasse gehört. Unter der Annahme, daß die Merkmalsvektoren von Pixeln in homogenen Bereichen näher am Mittelwert als die Vektoren von gemischten Pixeln liegen, wäre es möglich, unterschiedliche Mischungsverhältnisse zu schätzen. Weiterhin schlägt er vor, daß der Wert des Abstands zur Schätzung der Zuverlässigkeit der unscharfen Klassifikation benutzt werden kann.

Die Modellierung einer mehrfachen Zugehörigkeit ist mit der Fuzzy-Logik möglich. Die in der bedingten Wahrscheinlichkeit enthaltenen Informationen bleiben damit erhalten. Da für jede Klasse eine Zugehörigkeitsfunktion definiert wird, entsteht eine unscharfe Teilung des Merkmalsraums, in der jeder Punkt zu mehreren Klassen gehören kann. Diese Funktion kann, nach Wang [1990], aus der Gaußschen Dichtefunktion abgeleitet werden:

$$f_j(\mathbf{x}) = \frac{p^*(w_j|\mathbf{x})}{\sum_{k=1}^m p^*(w_k|\mathbf{x})} \quad (56)$$

in Gl. 56, steht $p^*(w_j|\mathbf{x})$ für die Dichtefunktion der Klasse w_j , unter Berücksichtigung des Fuzzy-Mittelwertvektors (\mathbf{mf}_j) und der Fuzzy-Kovarianzmatrix (CF_j), die eine Erweiterung von unscharfem Mittelwert und Kovarianzmatrix sind.

$$\mathbf{mf}_j = \sum_{i=1}^n \frac{\mu_j(\mathbf{x}_i)\mathbf{x}_i}{n} \quad (57)$$

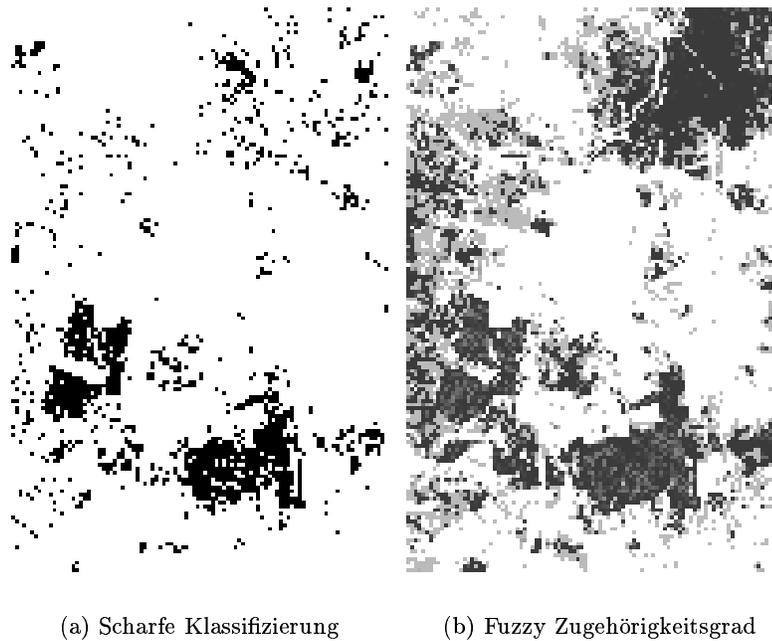
$$CF_j = \sum_{i=1}^n \frac{\mu_j(\mathbf{x}_i)(\mathbf{x}_i - \mathbf{mf}_j)^2(\mathbf{x}_i - \mathbf{mf}_j)}{n} \quad (58)$$

$\mu_j(\mathbf{x}_i)$ steht für die Zugehörigkeit des Vektors \mathbf{x}_i zur Klasse w_j . Die mögliche Werte für μ_j liegen zwischen Null und Eins. Wenn Trainingsgebiete zur Verfügung stehen, können die unscharfen Parameter für jede Klasse bestimmt und die Zugehörigkeitsfunktion geschätzt werden.

Mit diesem Ansatz wird mit der Maximum-Likelihood-Klassifikation ein Zugehörigkeitsgrad für jedes Pixel zu jeder Klasse berechnet. Ein Beispiel wird in Abb. 45 gezeigt. Das Bild zeigt die berechneten Werte für die Klasse „Wald“. Höhere Werte des Zugehörigkeitsgrades werden als dunkle Punkte dargestellt, während die hellen Töne niedrigerer Zugehörigkeit entsprechen.

5.2 Ableitung von Information aus der Nachbarschaft

Wenn die Objekte kleiner als die Bildauflösung sind, dann wird die reflektierte oder ausgestrahlte Energie zusammen mit der Energie ihrer Umgebung in demselben Pixel gespeichert. Falls sich mehrere kleine Objekte innerhalb einer Region wiederholen, wird diese Region von der Variation der Grauwerte gekennzeichnet, ihrer Textur. Diese Eigenschaft wird in der Bildverarbeitung oft benutzt, um spektral ähnliche Flächen zu unterscheiden. Das Ziel solcher Verfahren ist es, die Variation der Grauwerte innerhalb der Umgebung jedes Pixels quantitativ zu bestimmen und anhand dieser Information das Bild in homogene Regionen, gemäß der Textur, zu teilen.



(a) Scharfe Klassifizierung

(b) Fuzzy Zugehörigkeitsgrad

Abbildung 45: Zugehörigkeitsgrad nach der Klassifizierung der Klasse „Wald“: (a) Klassifiziertes Bild, (b) Zugehörigkeitsgrad. Grauwerte stehen für Werte zwischen 0 (weiß) und 1 (schwarz=volle Zugehörigkeit).

Nach Pratt [1974] lassen sich zwei große Texturgruppen erkennen: Künstliche und natürliche Texturen. Künstliche Texturen werden durch die systematische Wiederholung eines graphischen Elementes erzeugt. Diese Elemente können Linien, Punkte oder Symbole sein. Natürliche Texturen sind Produkte der Variation der Grauwerte in Bildern natürlicher Szenen und kommen in Regionen vor, in denen die einzelnen Objekte kleiner als die Auflösung sind. Ein Beispiel sind Siedlungsgebiete in Landsat-Bildern mit 30 m Auflösung.

Die Berechnung von Texturparametern aus der Nachbarschaft des Pixels wird verwendet, um die lokale Variation der Grauwerte zu beschreiben und Flächen mit ähnlicher Textur zu erkennen. Diese Parameter haben die Aufgabe, die Homogenität (Variation) der Grauwerte innerhalb einer Region darzustellen. Die Schätzung der Textur aus der Pixelumgebung maskiert die Anwesenheit der einzelnen Objekte (z.B. Häuser oder Straßen in Siedlungsbereichen bei einer Auflösung von 30 m) und ermöglicht eine kompakte Beschreibung der gesamten Region, aber mit einem signifikanten Verlust an geometrischer Auflösung im Bild. Diese Generalisierung ist für die kompakte Abbildung geschlossener Flächen mit unregelmäßiger Textur notwendig (Hagg [1998]).

Die Texturparameter von Haralick (Haralick & Shapiro [1997]) werden aus der Variation der Grauwerte in der Pixelumgebung berechnet. Diese Parameter werden aus einer *Co-occurrence-Matrix* berechnet, in der die Häufigkeit, mit der zwei Grauwerte in einer festgelegten relativen Anordnung im Bild vorkommen, gespeichert wird. Für die Berechnung der *Co-occurrence-Matrix* $P[i, j]$ wird zuerst ein Verschiebungsvektor $v = (dx, dy)$ definiert. Der Vektor v stellt die relative Lage zweier Pixel dar. Die Häufigkeit, mit der zwei Grauwerte (i, j) in dieser Anordnung im Bild vorkommen, wird danach gezählt (Jain *et al.* [1995]). Aus dieser Matrix können unterschiedliche Parameter abgeleitet werden, z.B. der *Kontrast*, die *Korrelation* oder die *lokale Homogenität*. Eine ausführliche Darstellung findet man in Haralick & Shapiro [1997].

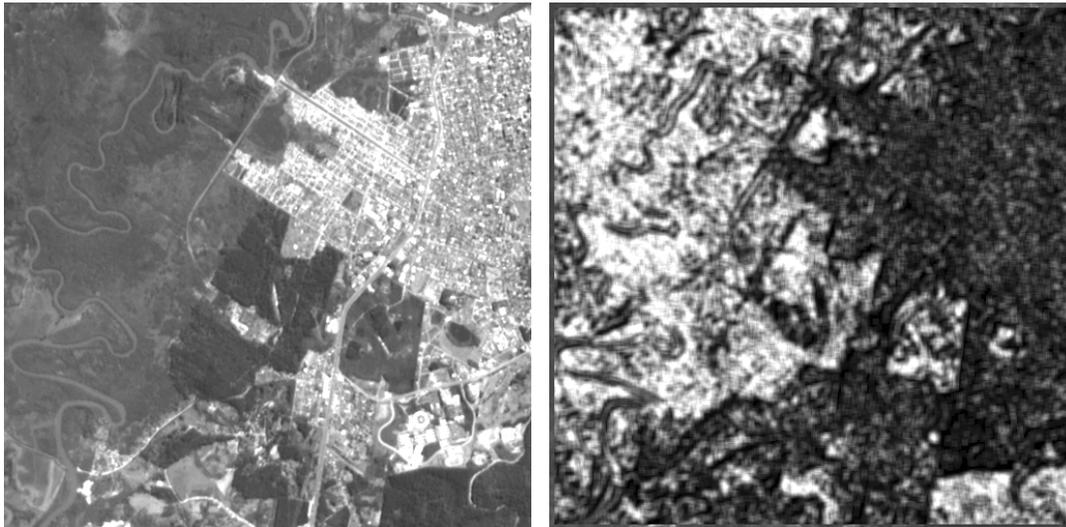
$$\textit{Entropie} = \sum_i \sum_j P^2[i, j] \quad (59)$$

$$\textit{Energie} = \sum_i \sum_j (i - j)^2 P[i, j] \quad (60)$$

$$\textit{Kontrast} = \sum_i \sum_j P[i, j] \log P[i, j] \quad (61)$$

$$\textit{Homogenität} = \sum_i \sum_j \frac{P[i, j]}{1 + |i - j|} \quad (62)$$

Die Textur tritt in Satellitenbildern um so mehr hervor, je höher die Auflösung ist. Deswegen wurde die Co-occurrence-Matrix aus dem panchromatischen Spot-Bild berechnet. Als Maß der Textur wurde die lokale Homogenität ausgewählt. Dieser Parameter läßt sich leicht und intuitiv interpretieren, da er einem üblichen linguistischen Begriff entspricht. Die lokale Homogenität wurde auf Werte zwischen 0 und 255 normiert, damit sie als Pixel eines Grauwertbildes gespeichert werden konnte. In diesem Bild erscheinen sehr homogene Regionen als dunkle Flächen und Regionen mit höherer Variation hell. Andere Grauwerte beschreiben Regionen, über die man nicht sicher sagen kann, ob sie homogen sind, oder nicht. Abb. 46 zeigt ein Beispiel der Berechnung der lokalen Homogenität. Man kann erkennen, daß die Siedlungsbereiche eine sehr niedrige lokale Homogenität besitzen. Dagegen zeigen die landwirtschaftlichen Flächen eine hohe Homogenität. Das Ergebnis stimmt mit den Schlüssen eines menschlichen Beobachters überein.



(a) panchromatisches Spot-Bild

(b) Lokale Homogenität

Abbildung 46: Berechnung der lokalen Homogenität.

6 Integrierte Verarbeitung der Information

Abgetastete Karten und Satellitenbilder sind Datenmengen unterschiedlicher Natur und ikonische Darstellungen derselben Szene. Obwohl sie zwar beide im Rasterformat vorliegen, ist ihr Informationsgehalt ein anderer: die abgetastete Karte ist das Ergebnis generalisierter menschlicher Objektbildung, während das Satellitenbild eine ungeneralisierte Abbildung der Realität darstellt. Bei der Aktualisierung des Kartenwerks durch die Satellitenbilder soll nun nicht nur die Information des Satellitenbildes genutzt werden, sondern auch die Information aus den bestehenden Karten. Bei einer derartigen integrierten Nutzung von verschiedenen Datenquellen spricht man von Datenfusion. Förstner [1992] und Kanade [1980] teilen die Datenfusion in der Bildverarbeitung in drei Verarbeitungsebenen ein: in die Signalebene, die Merkmalsebene und die Objektebene. Es besteht immer noch die Möglichkeit, Verfahren zu benutzen, die diese Ebenen miteinander kombinieren (die sogenannten hybriden Verfahren).

Datenfusion auf der niedrigsten Ebene, der Signalebene, ist die einfachste Option. Da die Daten auf der Signalebene im Rasterformat vorliegen, sind üblicherweise die Ergebnisse auch Rasterbilder. Bei der Datenfusion auf der Merkmalsebene werden nicht mehr nur einzelne Pixel als Einheit der Fusionsentscheidung angesehen, sondern komplexere Gebilde, wie Graphen, Listen oder relationale Beziehungen werden herangezogen (Ehlers [1995b]), um aus ihrer gemeinsamen Auswertung neue Informationen zu gewinnen. Bei der Datenfusion auf der Objektebene ist die Einheit, auf der die Fusionsentscheidung stattfindet, bereits ein Objekt oder eine Szene.

Nach Ehlers [1995b] werden in der Bildverarbeitung Verfahren zur Datenkombination auf der Signalebene bevorzugt, besonders weil sie einfacher und intuitiv sind und eine gute Leistungsfähigkeit bewiesen haben. Weil sie auf einfachen Ansätzen basieren, kann die Kombination der Daten auf der Signalebene nur zur Lösung einfacher Probleme benutzt werden.

6.1 Einbeziehung der Unsicherheit der Daten

Grundsätzlich werden in dieser Arbeit zwei Sorten von Daten benutzt: die multispektralen Messungen der Satellitenbilder und die aus einer Karte extrahierten kartographischen Objekte. Beide Datensätze haben Fehler, welche aus der Datenerfassung oder den Vorverarbeitungsschritten stammen und berücksichtigt werden müssen. Ehlers & Shi [1995] und Glemser & Fritsch [1998] nennen zwei Fehlerquellen bei der Integration von Satellitenbildern und raumbezogenen Daten: geometrische Ungenauigkeit und semantische Ungenauigkeit.

Geometrische Fehler bei abgetasteten Karten entstehen zum Beispiel bei der Verdrängung durch die Generalisierung, Zeichenungenauigkeiten bei der Herstellung oder Verzerrung der Papierkarte durch Feuchtigkeit. Die Fehlerquellen für geometrische Ungenauigkeiten in Satellitenbildern sind z.B. Abbildungsfehler und Georeferenzierungsfehler.

Bei der automatischen Extraktion von Objekten können Fehler bei der Definition der Objektgrenzen auftreten, die vom jeweiligen Extraktionsverfahren abhängen. Bei der automatischen Interpretation des Bildinhaltes entstehen Fehler in der semantischen Beschreibung, z.B. bei der Klassifikation des Bildes.

Wenn ein Bild mit der konventionellen scharfen Logik klassifiziert wird, kommen zwei Fehlersorten vor (Zhu [1997]): Wenn man sich für die „wahrscheinlichste“ Klasse (die Klasse, deren Mittelwert am nächsten zum untersuchten Pixel im Spektralraum liegt) entscheidet und annimmt, daß der Merkmalsvektor des Pixels und der Mittelwertvektor der Klasse identisch sind, dann entsteht ein Fehler zweiter Art (*commission error*). Die Tatsache, daß der Zugehörigkeitsgrad des Elements zu den anderen Klassen ignoriert wird, entspricht einem Fehler erster Art (*omission error*).

6.1.1 Berücksichtigung der Lageunsicherheit

Ein signifikanter Anteil des Fehlers bei der Erkennung kartographischer Elemente bezieht sich auf die Lage der Pixel. Lagefehler entstehen auch bei der Analog-Digital- und Raster-Vektor-Wandlung. Zur Behandlung vom Lagefehler eines Punktes kann das Fehlermodell der Vermessungskunde benutzt werden (Scheuring [1995]).

Die Lageinformation eines Punktes kann in zwei Komponenten geteilt werden. Die erste Komponente ist der genaue Wert der Punktkoordinaten. Die zweite Komponente bezieht sich auf eine zufällige Variation um den echten Wert. Die Variation der Lage eines Punktes im zweidimensionalen Raum wird als eine normalverteilte Zufallsvariable modelliert. Der Mittelwert dieser Variation ist Null, und die Varianz mit Hilfe der Varianz-Kovarianz-Matrix dargestellt.

$$p_i = \begin{bmatrix} x_i \\ y_i \end{bmatrix} \sim N_i \left(\begin{bmatrix} \mu_i \\ \nu_i \end{bmatrix}, \begin{bmatrix} \sigma_{xx} & \sigma_{xy} \\ \sigma_{yx} & \sigma_{yy} \end{bmatrix} \right) \quad (63)$$

Das Ergebnis der Kartenerkennung erklärt man im Raster- oder Vektorformat. Die Kartenelemente werden als Flächen und Linien dargestellt, seltener in Form von Punkten. Im Vektorformat werden Linien und Flächen mit Hilfe von Segmenten dargestellt. Zwei Ansätze werden zur Behandlung von Fehlern in Segmenten vorgestellt: das Epsilon Band Modell von Perkal [1956] und der stochastische Ansatz von Ehlers & Shi [1995]. Der erste Ansatz ist relativ einfach und basiert auf der Bestimmung einer Pufferzone mit konstanter Breite um das Segment. Es wird angenommen, daß die Punkte des Segmentes innerhalb dieser Region liegen. Der Ansatz von Ehlers & Shi [1995] bestimmt auch eine Region um das Liniensegment, aber deren Breite ist nicht konstant, sondern wird als Funktion der Lageunsicherheit der beiden Extreme dargestellt. Für ein Segment, definiert durch zwei Punkte mit Koordinaten p_1 und p_2 (Gl. 63) gilt:

$$L = (1 - r)p_1 + rp_2 \quad \text{mit } 0 \leq r \leq 1 \quad (64)$$

Nach Ehlers & Shi [1995] ist es möglich, andere Punkte in diesem Segment (p_i) als lineare Kombination der beiden Punkte p_1 und p_2 zu definieren. Dadurch kann die Unsicherheit der Lage als eine normalverteilte Variable modelliert werden (Gl. 65, Gl. 66).

$$p_i \sim (1 - r) \left(N \left(\begin{bmatrix} \mu_1 \\ \nu_1 \end{bmatrix}, \begin{bmatrix} \sigma_{xx} & \sigma_{xy} \\ \sigma_{yx} & \sigma_{yy} \end{bmatrix} \right) \right) + r \left(N \left(\begin{bmatrix} \mu_2 \\ \nu_2 \end{bmatrix}, \begin{bmatrix} \sigma_{xx} & \sigma_{xy} \\ \sigma_{yx} & \sigma_{yy} \end{bmatrix} \right) \right) \quad (65)$$

$$p_i \sim N \left(\begin{bmatrix} (1 - r)\mu_1 + r\mu_2 \\ (1 - r)\nu_1 + r\nu_2 \end{bmatrix}, ((1 - r)^2 + r^2) \begin{bmatrix} \sigma_{xx} & \sigma_{xy} \\ \sigma_{yx} & \sigma_{yy} \end{bmatrix} \right) \quad (66)$$

Mit dieser Funktion wird eine Vertrauensregion um das Segment beschrieben. Jedem Punkt der Region wird ein Vertrauenswert zugeordnet. Größere Vertrauenswerte werden in der Nähe des Segmentes erwartet. Die Werte sinken proportional zum Abstand zwischen dem Punkt und dem Segment. Das gilt auch für Flächen, die durch ihren Umfang repräsentiert werden. Innerhalb dieser Fläche gilt das maximale Vertrauensmaß.

Die Vertrauensregion ist in der Nähe der Endpunkte p_1 ($r = 0$) und p_2 ($r = 1$) breiter und in der Mitte, für $r = 0.5$, schmaler. Sie hat unscharfe (Fuzzy-) Konturen, denn es gibt keinen festen Übergang zwischen totalem Vertrauen und keinem Vertrauen.

Gemusterte Regionen sind ein Sonderfall, denn sie werden nicht als eine einzige Fläche dargestellt, sondern durch die Wiederholung von Symbolen, die aus Linien bestehen. Obwohl sie nicht explizit gezeichnet werden, sind die Zwischenräume dieser Region mit Sicherheit Bestandteil der Fläche. Deswegen wird den Zwischenräumen der maximale Vertrauenswert zugewiesen, d.h. gemusterte Regionen werden als geschlossene Flächen betrachtet. Die Variation des Vertrauens in ihren inneren Zwischenräumen wird nicht berücksichtigt. Da es schwierig ist, den genauen Konturverlauf aus dem Muster abzuleiten (Silva Centeno [1998]), wird die Unsicherheit an den Grenzen der Symbole außerhalb der Fläche mit Hilfe von Gl. 65 dargestellt (Abb. 47).

Eine vereinfachte Darstellung der Vertrauensregion wird von Ehlers & Shi [1995] vorgeschlagen. Dabei genügt eine dreieckige Verteilung anstatt der Normalverteilung in Gl. 66, um die Lageunsicherheit zu modellieren.

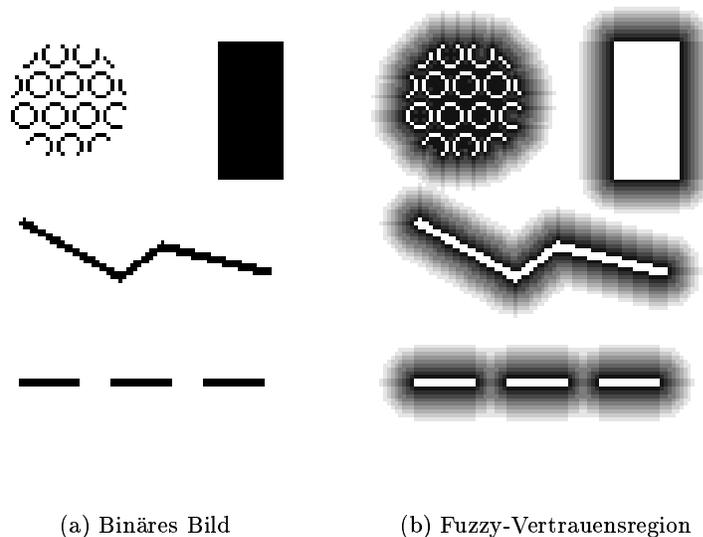


Abbildung 47: Fuzzy Vertrauensregion unterschiedlicher Elemente. Die ursprünglichen Pixel wurden in (b) weiß gekennzeichnet. Graue Töne stehen für Werte zwischen 0 (weiß) und 1 (schwarz).

6.1.2 Berücksichtigung der Unsicherheit der Semantik

Semantische Fehler beziehen sich auf die falsche Klassifikation eines Pixels und treten auf wenn die Daten und/oder die Segmentierungs- oder Klassifikationsverfahren nicht leistungsfähig genug sind, um die wahren Attribute des Pixels zu schätzen. Mögliche Gründe dafür sind:

- Die Daten stellen nur eine Abstraktion der Szene dar, eine ikonische Darstellung die manchmal nicht ausreicht, um den wahren Zustand der Szene darzustellen. Die spektralen Daten der Fernerkundung beziehen sich auf die reflektierte oder emittierte elektromagnetische Energie und reichen nicht aus, um manche Elemente zu trennen. Zum Beispiel ist es schwierig, Wasser und Schatten im Bereich des nahen Infrarots nur durch die Analyse der spektralen Daten zu unterscheiden.
- Die Daten werden abgetastet und quantisiert und als natürliche Zahlen im Rechner gespeichert. Das führt dazu, daß viele Objekte mit nur wenigen Pixeln dargestellt oder sogar zu einem einzigen Pixel zusammengefaßt werden und damit nicht mehr erkennbar sind. Aus der Abtastung

entstehen Mischpixel, welche die Informationen mehrerer Objekten tragen. Das Signal, das gemessen wird, enthält auch Störungen (Rauschen), welche die wahre Messung überlagern.

- Es werden viele Vereinfachungen in die Datenverarbeitung eingeführt. Die Binarisierung der abgetasteten Karte ist das beste Beispiel, weil sie eine sehr starke Vereinfachung ist. Dabei wird eine sehr scharfe Entscheidung getroffen, die den gesamten Prozeß der Interpretation beeinflussen kann.
- Die Klassifikation oder Segmentierung des Bildes basiert auf Annahmen über die Natur der Daten, welche nicht immer zutreffen.
- Nicht alle Klassen werden bei der Klassifikation berücksichtigt, sondern oft nur die Klassen, an denen man Interesse hat. Andere Klassen werden deswegen zu einer der berücksichtigten Klassen falsch zugeordnet.
- Die Klassifikationsansätze funktionieren nach Prinzipien der klassischen Mengentheorie, in der ein Element nur zu einer Klasse gehören kann, und zwar mit voller Zugehörigkeit. In natürlichen Szenen wird selten ein Pixel von einer einzigen Klasse bedeckt.
- Die wahren Parameter der benutzten Klassen sind unbekannt. Anstatt der wahren Parameter werden nur Schätzungen, die aus Trainingsgebieten des Datensatzes abgeleitet werden, verwendet (Swain & Davis [1978]).
- Der Inhalt einer Karte oder eines Bildes kann veraltet sein oder, im Fall einer Karte, Generalisierungsfehler enthalten.

Obwohl ein signifikanter Anteil der Fehler in der Kartenerkennung sich auf die Geometrie bezieht, treten oft semantische Fehler auf, vor allem wenn die Erkennung automatisch oder teilautomatisch erfolgt. Diese Fehler sind auch bei der Klassifikation von Satellitenbildern bekannt. Die traditionellen Methoden sind leider nicht in der Lage, volle Sicherheit über die Klassifikation zu garantieren. Die Interpretation dieser Daten muß diese Fehler deswegen berücksichtigen, um deren Einfluß in einer integrierten Analyse zu verringern.

6.2 Inferenzschema

Bei der Klassifikation wird eine Entscheidung über die möglichste Klasse für jeden Pixel getroffen. Die Sicherheit oder Unsicherheit dieser Entscheidung hängt von der Merkmale des Pixels ab. Unsicherheit läßt sich verkleinern wenn man nicht auf eine Datenquelle zurückgreift, sondern das Urteil aufgrund mehrerer Datenquellen fällt.

Manche Fusionstheorien gehen davon aus, daß das Fusionsergebnis umso besser ist, je früher die Datenquellen gemeinsam ausgewertet werden. In der Bildanalyse trifft dies nur bedingt zu. Bis zu einem gewissen Punkt ist es sinnvoll die Bilder getrennt voneinander auszuwerten, da die Auswerteverfahren als voneinander unabhängig zu betrachten sind. Diese Auswerteverfahren sind an die jeweiligen Eigenschaften des Datensatzes angepaßt.

Die Datenfusion setzt einen Rahmen voraus, in dem sich die vorhandene Information darstellen läßt, und in dem sich die notwendigen Schlüsse ziehen lassen. Einen derartigen Rahmen bildet die Theorie der unscharfen Logik. Der Zugehörigkeitsgrad eines Pixels zu einer Klasse, basierend auf Informationen aus einer bestimmten Quelle, wird gemäß der in den vorigen Abschnitten beschriebenen Methoden als eine reelle Zahl dargestellt.

Hinweise über die richtige Klassifikation eines Pixels können aus einem Datensatz abgeleitet werden. Die Stärke dieses Hinweises wird mit Hilfe einer unscharfen Funktion dargestellt. Da nicht alle Klassen in jedem Datensatz erkennbar sind, können sich diese Fuzzy-Funktionen widersprechen. Ein Pixel kann

mit zwei oder mehreren unterschiedlichen Klassen verbunden sein, da keine boolesche Logik benutzt wird.

Die Ergebnisse der multispektralen Klassifikation und der Kartenerkennung werden in diesem Sinne nicht als die absolute Wahrheit betrachtet, sondern nur als Hinweise auf die wahrscheinlichste Klasse des Pixels. Ihre Kombination ermittelt eine bessere Beschreibung der Beziehung zwischen den Pixeln und den Klassen. Die Form der Zugehörigkeitsfunktion (μ_j), die diese Beziehung darstellt, ist von der Natur der Daten abhängig.

Die Beseitigung der Konflikte und die Ableitung einer einheitlichen Interpretation wird mit Hilfe eines Fuzzy Inferenzschemas möglich. Dieses Inferenzschema basiert auf expliziten Regeln mit welchen versucht wird, die Erfahrung eines menschlichen Benutzers zu simulieren. Für die Kombination der Hinweise stehen die Operatoren der unscharfen Mengen zur Verfügung (Gl.21).

6.2.1 Regelbasiertes System

Der Inferenzprozeß basiert auf einer Reihe von Zugehörigkeitsfunktionen und Regeln zur Schlußfolgerung (Abschnitt 3.2.4). Diese Regeln werden explizit definiert. Dafür werden bekannte Fakten und die menschliche Erfahrung berücksichtigt. Die Regeln sind in der Form WENN-DANN gegeben und werden in einer Regelbank gespeichert. Die explizite Deklaration der Regeln erfolgt mit Hilfe von linguistischen Variablen und Termen. Jede Regel entspricht einer linguistischen Aussage.

Beispiele für linguistische Variable sind:

*Nach dem Satellitenbild, handelt es sich bei der Klasse um „Wald“ oder
Nach der Karte, handelt es sich bei der Klasse um „Wald“.*

Diesen Variablen werden Adjektive (linguistische Terme) zugeordnet. Sie beschreiben den möglichen Zustand der linguistischen Variablen. Beispiele dafür sind

*handelt es sich bestimmt,
handelt es sich vielleicht oder
handelt es sich eher nicht.*

Die Schlußfolgerung (das Ergebnis) wird ebenfalls sprachlich ausgedrückt. Eine einzige linguistische Variable mit mehreren linguistischen Termen kann das Ergebnis einer Klassifikation darstellen. Diese Variable wird sprachlich ausgedrückt:

das Pixel gehört zur Klasse

und ihre linguistischen Terme haben die Form:

*Wald,
Siedlungsbereich oder
Wasser,*

also die Landnutzungsklassen.

Durch die Kombination der Variablen und Terme ist es möglich, linguistische Sätze zu bilden, die eine Regel definieren. Zum Beispiel:

WENN

Nach dem Satellitenbild, es sich bestimmt bei der Klasse um Wald handelt

UND

nach der Karte, es sich bestimmt bei der Klasse um Wald handelt,

DANN

gehört das Pixel bestimmt zur Klasse Wald.

Diese sprachlichen Aussagen werden in den Regeln verbunden und mittels unscharfer Funktionen dargestellt. Für die Aussage:

Nach dem Satellitenbild, handelt es sich bestimmt bei der Klasse um Wald

wird eine Funktion benutzt:

$$\mu_j(\mathbf{x}) = \mu_{Wald,bestimmt}(\mathbf{x}, Sat.Bild) = \mu_{w_j,g}(\mathbf{x}, S_A) \quad (67)$$

\mathbf{x} sind die Eingabeinformationen aus dem Datensatz (S_A), während w_j eine der Klassen aus der Menge w ist und g für eine Adjektiv in Beziehung zur linguistischen Variable steht. Aus der Kombination mehrerer linguistischer Aussagen wird eine Ausgangsgröße (y) berechnet (Abschnitt 3.2.4). Die endgültige Klassifikation basiert auf den Hinweisen aller Datensätze und den Regeln.

6.2.2 Linguistische Variable und Terme

Voraussetzung für den Erfolg des regelbasierten Systems ist es, die Beziehung zwischen Fakten und Schlußfolgerungen, die sprachlich von einem Menschen formuliert werden, durch Regeln und algorithmischen Berechnungsverfahren darzustellen. Die Fuzzy-Linguistik bietet dafür eine passende Lösung. Eine sprachliche Problemstellung wird mit Hilfe von fuzzy-linguistischen Variablen und Termen ausgedrückt.

Grundsätzlich wurden drei Variablensorten definiert:

1. Variable, die die spektrale Information des Satellitenbildes darstellen,
2. eine Variable, welche die Textur des Spot Bildes darstellt und
3. Variablen, welche die Informationen aus der Karte darstellen.

Aus der Klassifikation des Landsat-Bildes wurde eine Menge linguistischer Variablen abgeleitet, welche die möglichen Zustände der Punkte auf der Erdoberfläche beschreiben. In einem Satellitenbild lassen sich mehr Objekte als in einer Karte erkennen (z.B. unterschiedliche Vegetationsarten). Deswegen wurden ähnliche Objekte in umfassenden Klassen zusammengefaßt. Damit war es möglich, die Information aus dem Bild mit der Information aus der Karte zu vergleichen.

Passende linguistische Variable wurden für die Klassen „Wasser“, „Wald“, „Siedlungsbereich“, „Reisanbau“ und „Landwirtschaft“ definiert. Dazu wurden jeweils zwei linguistische Terme benutzt, „niedrig“ und „hoch“, zur Beschreibung der Zustände dieser Variablen. Auf diese Weise wurde die sprachliche Formulierung von Aussagen wie:

„die Wahrscheinlichkeit, nach dem Satellitenbild, daß es sich bei der Klasse um Wald handelt ist hoch“
im Inferenzschema möglich.

Eine einzige Variable wurde zur Darstellung der Textur des Spot Bildes benutzt. Die Variation der „lokalen Homogenität“ im Spot-Bild wurde als Basis für diese Variable genommen. Ihre linguistischen Terme sind „klein“, „mittel“ und „groß“.

Die aus der Karte abgeleiteten Informationen wurden, gemäß der in Abschnitt 6.1.2 erläuterten Methode, als unscharfe Mengen dargestellt. Die berücksichtigten linguistischen Variablen sind deswegen „Wasser“, „Wald“, „Sumpfgebiet“, „Siedlungsbereiche“ und „Straßenetz“. Für diese Variablen wurden entweder zwei oder drei linguistische Terme vorgeschlagen: „klein“, „groß“ und, wenn nötig, „mittel“.

Tabelle 5 zeigt eine Zusammenfassung aller Variablen. Die ersten 6 Variablen beschreiben die spektralen Eigenschaften des Bildes, die nächsten 5 die Informationen aus der Karte und die letzte die Textur (räumliche Information) des Bildes. Das Ergebnis wird durch eine einzige Variable mit 5 linguistischen Termen dargestellt.

Tabelle 5: Inferenzschema: Linguistische Variablen und ihre Adjektive.

Variable	Nr.	Adj. 1	Adj. 2	Adj. 3	Adj. 4	Adj. 5
Klasse	5	Wasser	Landwirtschaftl. Flächen	Siedlungsflächen	Wald	unbekannt
Sat. Wasser	2	niedrig	hoch			
Sat. Wald	2	niedrig	hoch			
Sat. Siedlungsgebiet	2	niedrig	hoch			
Sat. Sumpfgebiet	2	niedrig	hoch			
Sat. Landwirtschaft	2	niedrig	hoch			
Sat. Reisanbau	2	niedrig	hoch			
Karte Sumpfgebiet	2	niedrig	hoch			
Karte Wasser	2	niedrig	hoch			
Karte Wald	2	niedrig	hoch			
Karte Siedlungsgebiet	3	niedrig	hoch	sehr hoch		
Karte Straßennetz	3	niedrig	hoch	sehr hoch		
Lokale Homogenität	3	niedrig	mittel	hoch		

6.2.3 Regeln

Das System basiert auf einer Menge von konstanten Regeln, die das Ziel der Analyse definieren. Sie verbinden die Eingangsdaten (Informationen aus den unterschiedlichen Datensätzen) mit den gewünschten Ergebnissen (Landnutzungsklassen). Die Regeln können dem gewünschten Klassifikationsziel angepaßt werden. Diese Anpassungsmöglichkeit verleiht dem System eine große Flexibilität.

Obwohl es möglich ist, alle Fakten mit einer Schlußfolgerung durch die Regeln zu verbinden, ist die Berücksichtigung aller Kombinationen nicht nötig. Nicht alle Kombinationen sind für eine bestimmte Aufgabe sinnvoll. Dadurch, daß die Regeln von einem Menschen ausgewählt werden, müssen nicht alle Kombinationen getestet werden.

Die Bestimmung der Regeln ist nur mit einer guten lokalen Kenntnis der Landschaft und wirtschaftlichen Aktivität der Region möglich. Wenn keine ausreichenden lokalen Kenntnisse zur Verfügung stehen, können die Regeln nicht sehr spezifisch formuliert werden, sondern eine größere Anzahl von Regeln muß überprüft werden. Die Integration des Wissens eines menschlichen Analysten ist in diesem Schritt entscheidend.

Wichtige Merkmale der Regelbasis sind die Anzahl von Regeln, die Anzahl von linguistischen Variablen und die Anzahl von linguistischen Termen für jede Variable.

Die Informationen aller Quellen wurden als geokodierte Ebenen gespeichert. In dieser Datenbank wird die Information jedes Pixels als eine digitale ganze Zahl zwischen 0 und 200 dargestellt. Da der Variationsbereich der Pixelwerte recht hoch ist, kann auf eine Fließkommadarstellung verzichtet werden.

6.2.4 Fuzzy-Funktionen

Der Erfüllungsgrad der Komponenten einer Regel wird von einer Menge unscharfer Funktionen dargestellt. Zum Beispiel, wird eine unscharfe Funktion benutzt, um die Zuverlässigkeit der Aussage „Die Wahrscheinlichkeit, daß das Pixel zur Waldklasse gehört ist hoch“ auszudrücken. Die Zugehörigkeitsfunktionen für die Eingabedaten und die Ausgangsgröße werden dann in einem ersten Schritt definiert.

Die Form der Funktionen bestimmt die Operation des Regelsystems. Die Analyse wird erleichtert,

wenn einfache Funktionen benutzt werden, denn dadurch wird ihre Speicherung und die Berechnung der Werte leichter.

Die unscharfen Funktionen der linguistischen Terme müssen die echte Beziehung zwischen den zur Verfügung stehenden Informationen und der Realität annähern. Sie können auf der Basis von Erfahrung, und durch mehrere Versuche von einem Benutzer geschätzt werden.

Die manuelle Bestimmung der Funktionen ist nicht trivial, denn es handelt sich dabei um eine große Anzahl von Funktionen, die zusammenwirken. Die Veränderung einer Funktion wirkt auf mehrere Regeln und muß deshalb in engem Zusammenhang mit den anderen mitwirkenden Regeln durchgeführt werden.

Die Kenntnis der Fakten spiegelt sich durch die Form der Funktionen wieder. Eine sehr stark aufsteigende Kurve gibt eine große Sicherheit wider, denn bereits kleine Merkmalsänderungen verändern die Zugehörigkeitsfunktion beträchtlich. Andererseits stellen Funktionen mit einer flachen Steigung dieses Übergangs eine Situation dar, in der die Definition des Zugehörigkeitsgrades nicht deutlich ist. Im extremen Fall würde eine waagerechte Gerade die totale Ungewißheit darstellen, denn es kann keine Entscheidung getroffen werden. Die Funktionen in Abb. 48 zeigen zwei verschiedene Situationen. Die Funktion $f2$ stellt eine fast binäre Funktion dar. Es gibt weniger Platz für Zwischenwerte. Der Übergang von Eins zu Null bei der anderen Funktion ($f1$) ist sanfter in Vergleich zu $f2$ und Zwischenwerte kommen öfter vor.

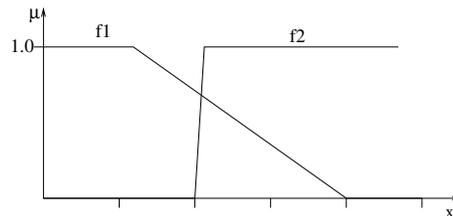


Abbildung 48: Beispiel Fuzzy-Funktionen.

Die Aufgabe des Benutzers ist es, den Verlauf der Funktionen zu bestimmen, so daß sie zu einer guten Klassifikation führen. Ein Versuch wurde in den Tabellen 8, 9 und 10 dokumentiert.

Als zweite Option zur Bestimmung der unscharfen Funktionen der Adjektive bietet sich eine rechnerische Annäherung mit Hilfe eines numerischen Verfahrens an. Im folgenden Abschnitt wird eine Optimierungsmethode zur Berechnung der Parameter dieser Funktionen beschrieben.

7 Automatische Schätzung der Fuzzy-Funktionen

Der im Abschnitt 6.2 beschriebene Ansatz benutzt eine große Menge von Funktionen zur Beschreibung der linguistischen Terme. Diese Funktionen können von einem menschlichen Benutzer explizit definiert oder mit Hilfe eines Rechners geschätzt werden. Da die Funktionen keine feste mathematische Beziehung beschreiben, ist ihre direkte Berechnung schwierig und eine numerische Lösung ist günstiger.

Genetische Algorithmen (Goldberg [1989]) gehören zu den stochastischen Optimierungsverfahren, die auf den Prinzipien der natürlichen Evolution basieren. Es handelt sich dabei um Suchalgorithmen, die einen mehrdimensionalen Variablenraum nach einer optimalen Lösung durchsuchen. Diese Verfahren eignen sich gut zur Schätzung unbekannter Parameter einer oder mehrerer Funktionen. Die Suchstrategie ähnelt den natürlichen Auswahlprozessen und basiert auf Prinzipien der Evolution. Deswegen werden genetische Algorithmen zu den naturanalogen Verfahren gerechnet.

In den folgenden Abschnitten wird ein Ansatz beschrieben, der auf genetischen Algorithmen basiert und die Schätzung der unscharfen Funktionen des Regelsystems unterstützt. Das Ziel des Optimierungsverfahrens ist in diesem Fall, die Qualität der Klassifizierung zu erhöhen.

7.1 Genetische Algorithmen

Bei den genetischen Algorithmen werden neue, verbesserte Lösungen aus den besten Schätzungen eines vorherigen Schrittes abgeleitet. Diese Methode benutzt nicht nur eine einzige optimale Lösung, sondern mehrere gute Lösungen, die nach jeder Iteration bewertet, aussortiert und kombiniert werden.

Der Prozeß fängt mit der zufälligen Auswahl einer große Menge möglicher Lösungen an. Die Qualität jeder Lösung wird bewertet. Die Lösungen mit niedriger Qualität, die zum Problem nicht passen, werden beseitigt. Die besten Lösungen werden dann kombiniert, um neue Lösungen zu erzeugen. Dieser Prozeß wird wiederholt, bis ein optimaler Punkt erreicht ist.

Als Vorbild wird eine Menge von *Individuen* genommen, die in einer natürlichen Umwelt zu überleben versuchen. Die Anpassung jedes *Individuums* an die Umwelt wird durch eine *Tauglichkeitsfunktion* (engl. *fitness*) gemessen.

Im Fall der Klassifikation wird diese Funktion aus dem direkten Vergleich des Ergebnisbildes der Klassifikation mit einem idealen Bild, einem Sollbild das die richtige Klassifikation jedes Pixels enthält, gewonnen. Ziel dieses Prozesses ist es, den Unterschied zwischen Ergebnis und Sollbild zu minimieren.

Folgende Bestandteile sind für die Suche notwendig:

- Die *Tauglichkeitsfunktion* muß explizit definiert werden.
- Die Anzahl von Parameter und ihr Variationsbereich muß bekannt sein, damit der Suchraum und die Art der Kodierung der *genetischen Kette* bestimmt werden können.
- Die Anzahl (Länge der *Kette*) und die Form (Zahl der Bits) der *Chromosomen* wird vorge-schrieben. Diese Kette enthält die Parameter der Funktion und damit eine mögliche Lösung des Problems.

Der Prozeß besteht aus den folgenden Schritten (Pham & Liu [1995]):

1. Initialisierung:

In einem ersten Schritt wird eine Menge von *Individuen* erzeugt (die erste *Bevölkerung*). Die Größe der *Bevölkerung* wird vorher festgelegt. Mehrere Lösungen werden zufällig ausgewählt, und

jede Lösung als eine binäre Kette kodiert. So entsteht die erste *Generation*, mit einer vielfältigen *Bevölkerung*. Die Anpassung jedes Elementes wird danach überprüft, so daß jedes Element durch ihre *genetische Kette* und einen *Tauglichkeitswert* gekennzeichnet wird.

2. Binäre Kodierung:

In diesem Schritt werden die numerischen Werte der Parameter einer Lösung in eine binäre Kette umgewandelt. Da eine binäre Einheit (Bit) nur zwei mögliche Werte speichern kann, ist zur Darstellung größerer Werte eine Kombination mehrerer binärer Einheiten erforderlich. Die Länge dieser Elementarkette hängt von der Darstellung des Wertes ab. Zum Beispiel, zur Darstellung positiver natürlicher Zahlen zwischen 0 und 255 sind 8 Bits nötig. Aus jedem Wert entsteht ein String, der Bestandteil einer längeren Bit-Kette (die *genetische Kette*) ist. Jedes Bit dieser Kette wird als Chromosom interpretiert. Nach der Umwandlung kann ein *Individuum* durch seine genetische Kette beschrieben werden. Eine kleine aber ausdrucksstarke Kette wird empfohlen.

Abb. 49 stellt ein Beispiel der binären Kodierung dar. Eine trapezförmige Funktion wird von vier Punkten definiert: (p_1/Y_2) , (p_2/Y_1) , (p_3/Y_1) und (p_4/Y_2) . Es wird angenommen, daß die Y -Koordinaten Konstanten sind, und deswegen reichen die vier X -Koordinaten aus, um ein Trapez zu beschreiben. Jede Zahl wird binär dargestellt, und diese Bitsequenzen werden danach verkettet. Als Ergebnis entsteht eine längere binäre Kette.

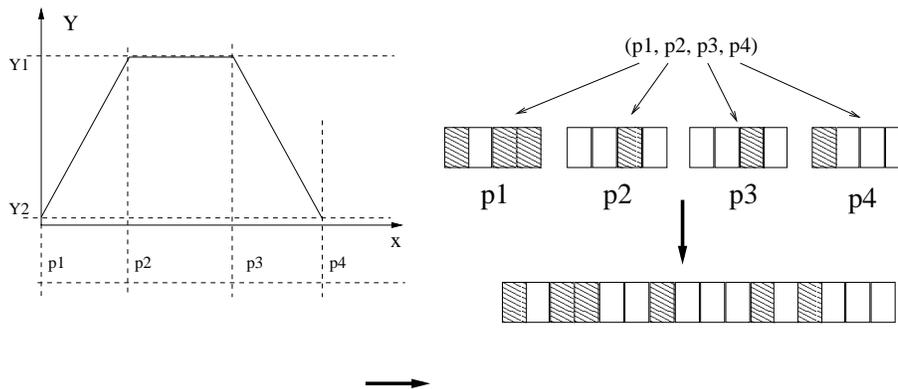


Abbildung 49: Binäre Kodierung einer Funktion.

Die ersten Schätzungen bilden die erste *Generation* und auch die erste *Elternmenge*. Diese *Individuen* werden in den folgenden Schritten durch die *genetischen Operationen* optimiert.

3. Überprüfung der *Bevölkerung*:

Die *genetischen Ketten* (bzw. die möglichen Lösungen) werden nach jeder Iteration analysiert. Es wird dabei überprüft, ob die *Bevölkerung* in Richtung eines *Individuums* (der optimalen Lösung) konvergiert. Wenn die Homogenität der *Bevölkerung* einen Schwellwert überschreitet, wird angenommen, daß die optimale Lösung gefunden wurde, und der Prozeß wird angehalten. Die Kette mit dem besten Anpassungswert stellt die optimale Lösung dar.

4. Auswahl von *Eltern*:

Der *Tauglichkeitsgrad* der *Individuen* wird durch die Dekodierung der *genetischen Kette* berechnet. *Individuen* mit höherem *Tauglichkeitsgrad* werden ausgewählt, um aus diesen *Individuen* (*Eltern*) neue verbesserte *Individuen* zu erzeugen. Der Anpassungsgrad der *Individuen* wird berücksichtigt (*Auswahlprozeß*) um zu sichern, daß nur gut geeignete *Individuen* in der *Paarung* benutzt werden.

Es gibt unterschiedliche Strategien zur Auswahl der *Eltern*. Zum Beispiel:

Turnier: Ein Paar *Individuen* wird zufällig ausgewählt und der schwächste, derjenige mit niedrigstem *Tauglichkeitsgrad*, wird beseitigt. Der stärkste bleibt erhalten.

Auswahl nach einem *Tauglichkeitskriterium*: Ein *Individuum* wird zufällig ausgewählt, und seiner *Tauglichkeitsgrad* mit einem Schwellwert verglichen. Falls sein *Tauglichkeitsgrad* über dem Schwellwert liegt, besteht er die Prüfung und überlebt.

Zufällige Auswahl: Ein Paar von *Individuen* wird zufällig ausgewählt und im *Reproduktionsprozeß* benutzt. Diese Methode garantiert keine Konvergenz.

5. Paarung:

Ein neues *Individuum*, oder besser gesagt eine neue binäre Kette, wird aus zwei *Eltern* erzeugt. Beliebige neue binäre Ketten werden aus der Kombination von *Elternpaaren* erzeugt. Zur Kombination der Chromosomen gibt es mehrere Möglichkeiten. Eine einfache Möglichkeit besteht darin, die Bits der *Eltern* zufällig kombiniert an das Kind weiterzugeben.

Eine andere Variante ist es, Teile der Kette aus den alten *Individuen* zu kopieren. Dafür wird ein Übergangspunkt bestimmt. Die Elemente der Kette, die vor diesem Übergangspunkt liegen, werden aus einer der *Eltern* kopiert. Die übrigen Bits werden aus dem Zweiten kopiert (Abb. 50). Durch die Dekodierung seiner Kette wird der *Tauglichkeitsgrad* dieses neuen *Individuums* berechnet.

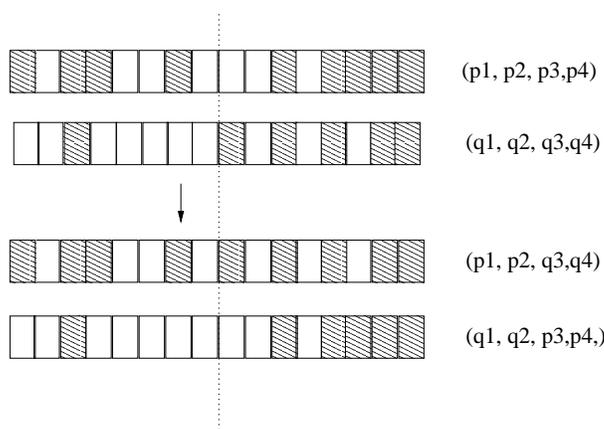


Abbildung 50: Paarung zweier binären Ketten.

6. Mutation:

Die Mutation ist eine Operation, die ein beliebiges Bit eines *Individuums* verändert. Diese Veränderung soll zufällig und nur in wenigen Fällen erfolgen. Sie hat die Aufgabe, neues *Genmaterial* in den Prozeß einzuführen, und ist für die Unterbrechung der Suchversuche, die in einem relativen Minimum stehen geblieben sind, nötig.

Abb. 51 zeigt zum Beispiel eine *Mutation* einer kleineren Kette, mit 16 *Chromosomen*. In dieser Kette wird ein Zahl (p_1 , p_2 , p_3 und p_4) mit 4 Bits gespeichert. Die *Mutation* erfolgt zufällig. In diesem Beispiel wird das neunte Bit umgekehrt. Dadurch entsteht ein neuer Wert (P_5) für die dritte Zahl, das heißt, daß sich für die neue Kette die Zahlen p_1 , p_2 , P_5 und p_4 ergeben.

Die *Mutation* ermöglicht die Untersuchung neuer Lösungen, die nicht unbedingt von der ersten Schätzung abhängig sind. Als Optionen stehen die Veränderung eines beliebig ausgewählten Bits, die Inversion eines bestimmten Teils einer Kette oder die Inversion zweier zufällig ausgewählte *Chromosomen* zur Verfügung.

Eine weitere Operation ist der *Elitismus*, d.h. die Erhaltung einer bestimmten Anzahl von *Eltern* in der neuen *Generation*.

7. Ersetzung:

Neue *Individuen* können später die alten *Eltern* ersetzen, so daß eine neue *Generation* entsteht.

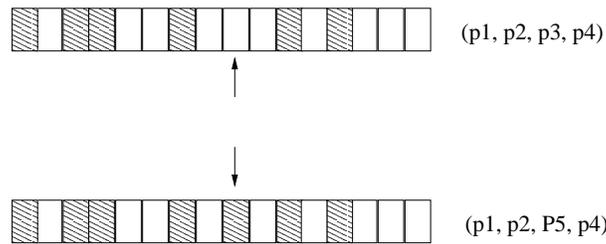


Abbildung 51: Mutation einer binären Kette.

Die Schritte (3) bis (6) werden systematisch wiederholt, bis eine optimale Lösung erreicht ist. Der Prozeß könnte im Prinzip beliebig oft wiederholt werden. In der Praxis werden Konvergenzkriterien eingeführt, um den Zeitpunkt zu dem der Prozeß angehalten werden soll zu bestimmen, zum Beispiel, wenn der Anpassungswert eines bestimmten Teils der *Bevölkerung* einen Schwellwert überschreitet, wenn die *Individuen* alle ähnlich aussehen, oder nach einer festen Anzahl von Iterationen.

Die Homogenität der *Bevölkerung* kann mittels eines Abstands gemessen werden, wie zum Beispiel die Summe der Abstände der *Individuen* vom Mittelwert der *Bevölkerung*, oder die Summe der Abstände zwischen allen *Individuen* (Gl. 69).

$$hom = \sum_{i=1}^N \sum_{j=1}^{NCR} (Chromo(i, j) - media(j)) \quad (68)$$

$$hom = \sum_{i=1}^N \sum_{j=1}^i \sum_{k=1}^{NCR} (Chromo(i, k) - Chromo(j, k)) \quad (69)$$

mit:

NCR= Anzahl von Chromosomen

N= Größe der *Bevölkerung*.

7.2 Bildung der genetischen Kette

In Kap. 6 wurde beschrieben, wie aus einer Karte oder aus dem Satellitenbild Hinweise für die Klassifizierung der Pixel abgeleitet werden können. Diese Hinweise werden als natürliche Zahlen zwischen 0 und 200 kodiert und gespeichert. Dieses Format bestimmt den Variationsbereich der unscharfen Funktionen. Auf der waagerechten Achse sind Werte zwischen 0 und 200 erlaubt, und auf der senkrechten Achse Werte zwischen 0 und 1. Die Frage ist: Welche Kurve soll benutzt werden, um die unscharfen Funktionen darzustellen?

Eine einfache Kurve hat den Vorteil, daß sie die Interpolation vereinfacht. Die einfachste Form wäre ein Dreieck, dessen Darstellung nur 5 Punkte verlangt. Dreieckige Funktionen haben den Nachteil, daß ihre Spitzen eng sind und nur eine begrenzte Menge von Funktionen darstellen.

Ein Trapez wäre eine bessere Option. Sechs Punkte sind für die Darstellung eines Trapezes nötig, wie es in Abb. 52 dargestellt wird. Zur Vereinfachung der Darstellung der Funktionen wurde angenommen, daß die Y-Werte (μ) dieser sechs Punkte konstant sind, und daß der erste und der letzte Punkt fest sind. Die Form des Trapezes wird von den Abständen zwischen den ersten fünf Punkten auf der X-Achse definiert, also reichen vier Werte aus um den Trapezverlauf zu beschreiben (Abb. 52).

Unter diesen Umständen kann eine trapezförmige Funktion mit nur vier ganzen Zahlen dargestellt werden, und zwar mit den Abständen zwischen den ersten fünf Punkten auf der X-Achse. Die binäre Kodierung dieser Werte ist leicht, da diese Zahlen kleiner als 256 sind und als eine Kette von acht Bit

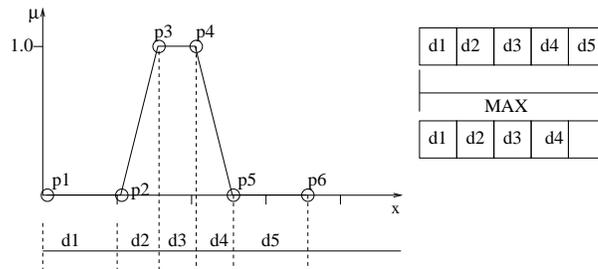


Abbildung 52: Allgemeine Form einer unscharfe Menge.

in einem Rechner gespeichert werden können. Für eine kleine Kette mit vier Zahlen werden 32 Bits gebraucht.

7.3 Operatoren

7.3.1 Initialisierung

Vor dem Start des Optimierungsprozesses wird eine große Menge von Lösungen zufällig generiert und kodiert. Gemäß dem Prinzip der genetischen Algorithmen werden diese *Individuen* systematisch sortiert und ausgewählt, so das eine Evolution in Richtung einer optimalen Lösung stattfindet.

7.3.2 Auswahlprozeß

Der Auswahlprozeß der implementierten Version der genetischen Algorithmen vergleicht jedes Ergebnis mit dem angestrebten Ziel. 70 Prozent der Lösungen werden beseitigt. Dafür werden die *Individuen* nach ihrer *Tauglichkeit* sortiert und die besten für die Paarung ausgewählt.

7.3.3 Paarung

Zur Paarung wird die Methode des *Übergangspunktes* benutzt. Als Übergangspunkt wurde in dieser Arbeit die Mitte der Kette gewählt.

7.3.4 Mutation

Der Anzahl der Mutanten ist konstant in einem 1:10 Verhältnis zur gesamten *Bevölkerung*. Alle *Individuen*, mit Ausnahme des besten Individuums, sind Kandidaten zur *Mutation*. Fünf *Individuen* werden ausgewählt und eines ihrer Bits ein- oder ausgeschaltet, so daß eine neue Lösung entsteht. Die *Mutation* erfolgt ganz zufällig.

Zur Verbesserung des Evolutionsprozesses wurde verhindert, daß das beste *Individuum* mutiert, da es die bisher beste Lösung enthält. Damit kann der maximale *Tauglichkeitsgrad* nach jeder Iteration nur wachsen und bleibt unverändert, wenn keine bessere Lösung in der Iteration erzeugt wird.

7.3.5 Ersetzung

70 Prozent der *Eltern* werden von den neuen *Individuen* ersetzt und eine neue Iteration fängt an.

7.3.6 Anpassungskriterium

Die Grundidee dieser Methode ist der Auswahl der *Individuen*, die am besten den Umweltbedingungen angepaßt sind. Diese Anpassung muß numerisch gemessen werden.

Das Ziel des Prozesses ist, eine gute Klassifikation der Pixel zu erreichen, weshalb das Maß der *Tauglichkeit* auf der Klassifikationsgenauigkeit basieren soll. Da nicht für alle Pixel die genaue Sollklassifikation bekannt ist, wird in der Praxis das Genauigkeitsmaß aus Testregionen abgeleitet. Für diese Regionen wird die richtige Klassifikation mit anderen Methoden bestimmt und als richtig angenommen.

Aus dem direkten Vergleich zwischen dem Ergebnis der Klassifikation der Trainingsgebiete und dem Sollbild kann man die Anzahl richtig klassifizierter Pixel ermitteln. Ein Maß für die Genauigkeit (*Qua*) wird aus der Hauptdiagonalen der Konfusionsmatrix abgeleitet.

$$Qua = \sum_{i=0}^n Tr(K)/Size \quad (70)$$

mit:

Size= Größe der Prüfungsmenge,

Tr(K)= Spur der Konfusionsmatrix K und

n= Anzahl der Klassen

Dieses Maß ist global und hängt nur von der Anzahl richtig klassifizierter Pixel ab. Es berücksichtigt nicht die Größe der Testdaten der einzelnen Klassen. Um eine ausgeglichene Schätzung der Qualität zu erhalten, muß die Größe der Testdaten ungefähr gleich sein, sonst würden die Funktionen mit mehr Testpixeln bevorzugt. Eine ausgeglichene Messung erhält man mit Hilfe der mittleren Genauigkeit der Klasse.

$$Fit = \sum_{i=0}^n N(i)/Size(i) \quad (71)$$

mit:

Fit=Qualitätsmaß (*Tauglichkeit eines Individuums*)

Size(i)= Anzahl der Testpixel der Klasse i,

N(i)= Anzahl der richtig klassifizierten Pixel der Klasse i und

n= Anzahl der Klassen

Die Zahl der Testdaten muß relativ groß sein um eine ausreichende Beschreibung der Klassen zu erreichen. Alle Funktionen müssen über den gesamten Variationsbereich geschätzt werden. Es reicht nicht den genauen Verlauf um dem Mittelwert zu schätzen, denn die Funktionen werden über den gesamten Variationsbereich überprüft.

Sehr homogene Trainingsgebiete eignen sich nicht für die Bestimmung eines Anpassungskriteriums, weil sie nur die Schätzung eines bestimmten Bereichs der Funktion ermitteln. Informationen an den Übergängen zwischen Klassen oder Informationen über Regionen, in denen manche Klassen nicht vorkommen, werden gebraucht um die Funktionen zu schätzen. Die Prüfungsgebiete müssen gut und homogen im Bild verteilt sein um die ganze räumliche Variation der Information darzustellen.

Die Zahl der Testdaten verlängert die Zeit die gebraucht wird, um das Anpassungsmaß einer *Generation* zu berechnen, denn für jedes *Individuum* (eine Menge von Funktionen), wird eine Klassifizierung des Trainingsgebietes durchgeführt. Da eine *Generation* aus vielen *Individuen* besteht, braucht das Verfahren eine lange Zeit für jede Iteration.

7.3.7 Homogenitätsmaß

Nach jeder Iteration wird die Ähnlichkeit der gesamten *Bevölkerung* berechnet. Wenn alle *Individuen* einander sehr ähnlich sind, dann wird vermutet, daß die *Bevölkerung* optimiert wurde und der Prozeß angehalten werden kann. Die Homogenität der *Bevölkerung* wird aus dem Abstand zwischen den Gen-Vektoren berechnet.

Der Abstand zwischen zwei Ketten, $CHROMO_1$ und $CHROMO_2$, wird gemäß Gl.72 berechnet:

$$d_{1,2} = \sum_{i=0}^n |CHROMO_1(i) - CHROMO_2(i)| \quad (72)$$

mit:

$d_{1,2}$ = Abstand zwischen zwei Ketten 1 und 2

$CHROMO_1(i)$ = Wert des i-ten Elements der Kette $CHROMO_1$

n = Länge der Kette

Das Homogenitätsmaß wird aus dem globalen Abstand dg berechnet.

$$dg = \sum_{k=0}^M \sum_{j=0}^k \sum_{i=0}^n |CHROMO_j(i) - CHROMO_k(i)| \quad (73)$$

mit:

$CHROMO_j(i)$ = Wert des i-ten Elements der Kette des j-ten *Individuums*

n = Länge der Kette

M = Größe der *Bevölkerung*.

Die Tendenz ist, daß das Homogenitätsmaß nach jeder Iteration kleiner wird, denn die Information der *Eltern* wird kombiniert. Wenn, aufgrund einer *Mutation*, eine sehr große Entwicklung stattfindet, dann kann der Wert der Homogenität wachsen, wird aber in den folgenden Iterationen wieder sinken. Wenn dieser Wert sich an Null annähert, dann wird angenommen, daß die optimale Lösung erreicht ist. Die Lösung kann aus der Kette des Individuums mit dem besten *Tauglichkeitsgrad* dekodiert werden.

Individuen mit sehr feinen genetischen Unterschieden können dieselbe Qualität haben, obwohl sie nicht unbedingt sehr gleich aussehen, vor allem wenn die Lösung sich an einem relativen Minimum innerhalb einer flachen Region befindet. Daraus folgt, daß manche kleine Abweichungen der Fuzzy-Funktionen keine Auswirkung auf das Ergebnis haben. Dieses Problem hängt mit der Definition des Zieles und der Rechenvorschrift für das *Tauglichkeitsmaß* zusammen. Vor allem bei der Klassifizierung kann es geschehen, daß es nicht nur eine einzige Lösung gibt, sondern daß mehrere ähnliche *genetischen Ketten* zum besten Ergebnis führen.

7.3.8 Ende des Suchprozesses

Die Iteration wird wiederholt, bis die Ketten sehr ähnlich sind, oder bis eine bestimmte Anzahl von Iterationen erreicht wird.

Der letzte Schritt ist die Dekodierung der Kette in mehrere unscharfe Funktionen, gemäß der in Abschnitt 7.2 beschriebenen Methode. Diese Funktionen können dann zur Klassifikation eines Bildes benutzt werden.

8 Experimentelle Ergebnisse

Abb. 53 zeigt einen Ausschnitt einer IBGE-Karte (Instituto Brasileiro de Geografia e Estatística) des Rio-dos-Sinos-Einzugsgebietes in Südbrasilien. Im Bild kann ein Teil der Stadt Sapucaia do Sul, die BR-110 Landstraße und der Rio Dos Sinos deutlich erkannt werden. Bei diesem Testgebiet handelt es sich um eine stark besiedelte Region, mit einem höheren Bevölkerungswachstum in den letzten Jahren. Der Ausschnitt hat eine Ausdehnung von 512×512 Pixel. Die Pixelgröße bei der Abtastung beträgt $85\mu\text{m}$.

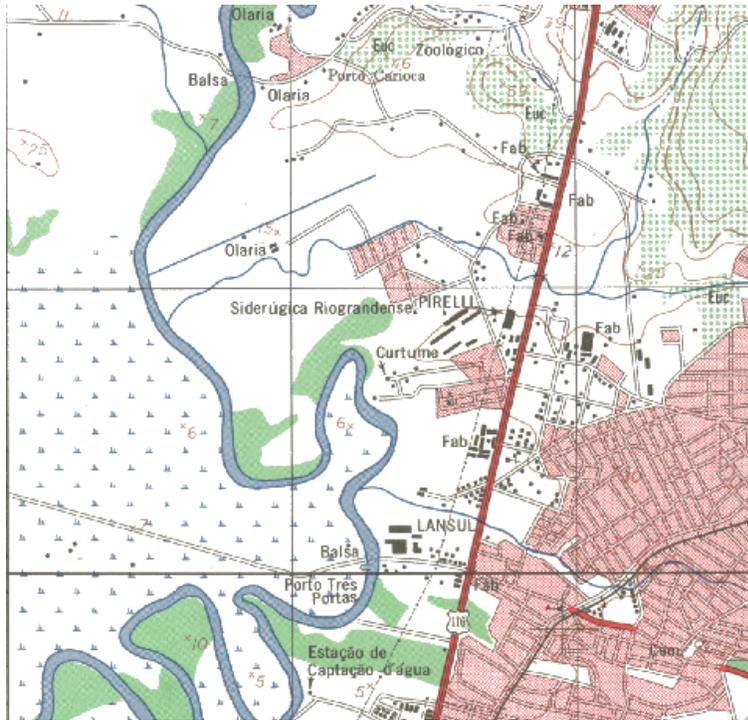


Abbildung 53: Ausschnitt aus der topographischen Karte von IBGE, Maßstab 1:50.000.

8.1 Beschreibung des Testgebietes anhand einer Karte im Maßstab 1:50.000

Um die Ergebnisse des hier vorgestellten Verfahren zu diskutieren, muß eine Referenz am Boden vorliegen. Als Referenz wird die in Abb. 53 gezeigte Karte benutzt. In dieser Abbildung sind verschiedene Landnutzungsklassen als Flächen oder Linien dargestellt.

Die blaue Farbe kennzeichnet die Wasserflächen, wie Flüsse und Dämme, aber auch feuchten Flächen, wie das Sumpfgebiet um den Rio Dos Sinos. Dämme sind flächenhafte Elemente und werden als hellblaue Flächen dargestellt. Flüsse kommen in zwei unterschiedlichen Formen vor. Die Grenzen größerer Flüsse werden mit dunkelblauer Farbe genau gezeichnet. Der eigentliche Fluß wird hellblau dargestellt. Die Breite dieser Flüsse ist größer als der minimalen Abstand, der in einer 1:50.000 Karte gezeichnet werden kann ($0,3\text{mm}$ auf der Karte). Die Form und Größe dieser Flächen unterscheidet sie von anderen blauen Flächen, wie zum Beispiel von Dämmen. Andererseits werden kleine Flüsse und Bäche als dünne durchgezogenen blaue Linien gezeichnet. Der oberen Teil eines Flusses wird auch als gestrichelte blaue Linie dargestellt, um zu zeigen, daß das Wasser auf dieser Strecke nicht das ganzen Jahr fließt.

Die Landschaft um den *Rio dos Sinos* wurde vom Fluß überformt: Es handelt sich um ein flaches Tal. Die Ablagerung von Sedimenten im Unterlauf des Flusses hat Einfluß auf die Entwicklung des Gebietes. Die fruchtbaren Böden begünstigen die Landwirtschaft. Diese Flächen werden oft, vor allem während

der Regenzeit, vom Rio dos Sinos überflutet. Sie sind als „Sumpfbereich“ auf der Karte gekennzeichnet. Weiterhin wird auf diesen Flächen auch oft Reis angebaut.

Auf der Karte erscheint das Sumpfbereich als eine gemusterte Fläche (vgl. linke unteren linken Ecke der Abb. 53). Die genauen Konturen dieser Flächen werden nicht explizit gezeichnet, sondern ergeben sich aus den Grenzen anderer Objekte bzw. der Verteilung der sonstigen Muster.

Das Straßennetz bildet Teil der linienhaften Bildelemente. Rote Linien, dick oder gestrichelt, zwischen dünnen schwarzen Linien, kennzeichnen die Landstraßen. In der Mitte des Bildes kann man eine Landstraße in Süd-Nord-Richtung sehen. Die Breite dieser Linie auf der Karte entspricht nicht der wahren Breite der Straße, sondern ist das Ergebnis einer Generalisierung.

Die anderen Straßen, die den größten Teil des Straßennetzes bilden, werden als parallele schwarze Linien gezeichnet. Sie umfassen oft die Siedlungsbereiche. Kleinere Wege erscheinen als einfache schwarze durchgezogene Linien auf der Karte.

Besiedelte Flächen werden mit hellrotem Muster dargestellt. Diese Flächen liegen immer am Straßennetz und werden von diesem Netz um geschlossen.

Die Vegetation ist grün gekennzeichnet. Zwei Klassen werden unterschieden: Ursprüngliche Vegetation (grüne Flächen) und bewaldete Flächen, die als Flächen mit grünen Mustern (Kreise) vorkommen.

Andere kartographische Elemente mischen sich mit den obengenannten, wie zum Beispiel die Höhenlinien (Braun-Orange), Häusern und andere wichtigen Gebäuden. Desweiteren finden sich Schrift und Einzelsymbole in unterschiedlichen Farben. Ein schwarzes Koordinatengitter soll dem Betrachter bei der Orientierung helfen.

Man kann die Flächen einer abgetasteten Karte in zwei Gruppen teilen: Regionen, die mit der Zeit wenige thematische Veränderungen erfahren, wie ein Stadtzentrum, und dynamische Regionen, deren Landnutzung, aufgrund der Entwicklung des Gebietes, verändert wird. In einer Siedlungsfläche kann man beide Gruppen finden. Zur ersten Gruppe gehören die Pixel des alten Stadtzentrums und die älteren Stadtteile. Da die Stadt im Laufe der Zeit gewachsen ist, finden am Rande der Stadt Veränderungen statt.

8.1.1 Kartenhinweise

Das Bild der Karte wurde gemäß der im Kapitel 4 beschriebenen Methoden segmentiert. Folgende fünf Klassen wurden berücksichtigt: „Wasser“, „Sumpfbereich“, „Straßennetz“, „Siedlungsgebiete“ und „Vegetation“. Andere kartographische Elemente wurden nicht berücksichtigt, obwohl sie auch wertvolle Informationen enthalten können: zum Beispiel Gebäude, die als Siedlungsfläche klassifiziert oder als potentielle Kerne für die Entstehung neuer Siedlungsbereiche betrachtet werden könnten. Abb. 54 und 55 zeigen Ausschnitte der binären Bilder jeder Klasse. Auf diesen Bildern haben die Informationen eine binäre Bedeutung: Entweder gehört das Pixel zur Klasse oder nicht. Durch diese Darstellung können Widersprüche entstehen. Ein Vergleich zwischen den Flächen am Ufer des Rio dos Sinos auf der Abb. 55a und 54 zeigt, daß ein Pixel als mehreren Klassen zugehörig klassifiziert werden kann. Zum Beispiel, wird die grüne Fläche am unteren linken Ecke des Bildes sowohl von der Klasse „Sumpfbereich“ umfaßt als auch von der Klasse „Vegetation“. Das liegt daran, daß die Erkennung jeder Objektklasse unabhängig von der anderen erfolgt und die Konflikte nicht gelöst werden.

Die Widersprüche der binären Bilder sind Produkte der einfachen Modellierung der Objekte (Hier wurden einfache Verfahren zur Erkennung der Objekte angewendet). Mehr als eine Bedeutung für ein Pixel zu erlauben wäre trotzdem in manchen Fälle nicht ganz falsch, denn die Bodenbedeckung ist nicht homogen. Dagegen sind manche Kombinationen nicht logisch, wie zum Beispiel eine landwirtschaftliche Fläche in der Stadt.

Die Konturen der Flächen mit Mustern sind schwierig aus einem einzigen binären Bild zu bestimmen. Ein Grund für die schwache Abgrenzung ist, daß die graphischen Elemente getrennt erkannt werden.



Abbildung 54: Binäres Ergebnisbild der Kartenerkennung: grüne Flächen.

Obwohl die anderen Objekte Informationen über die Grenzen der Flächen mit Mustern enthalten, werden sie in den einzelnen Schritten des Erkennungsverfahrens nicht berücksichtigt. Trotzdem wurden die Grenzen dieser Klassen gut geschätzt. Als Beispiel können die Flächen des Sumpfgebietes der Abb. 55a und die bewaldeten grünen Flächen am oberen rechten Eck der Abb. 54 genannt werden. Der kritische Fall ist das Sumpfgebiet, denn seine Muster sind sehr weit voneinander entfernt, und seine Grenzen werden oft interpoliert. Andererseits liegen die Muster der bewaldeten Flächen sehr dicht beieinander, weshalb die Konturen genauer definiert werden können.

Die Position der Straßen wurde mittels einer rasterbasierten Skelettierung der inneren Räume zwischen parallelen Linien bestimmt. Diese Methode wird in der digitalen Bildverarbeitung oft benutzt (Bässmann & Besslich [1989]). Aber, wie schon in Brügelmann [1998] erwähnt wird, liegt ein großer Nachteil der Skelettierung in der Kartenerkennung darin, daß die Mittelachse (Skelett) nur eine Approximation der abgetasteten Linien darstellt. Durch die Analog-Digitalewandlung entstehen Fehler und Lücken. Die spätere Approximation durch einfache Kurven wird weiterhin durch die Verschmelzung mehrerer Linienstücke und durch die Reduktion der Punktedichte der ursprünglichen Linien verfälscht. Gerade die Linien des Straßennetzes sind für solche Fehler anfällig. Die Extraktionsergebnisse werden nicht auf derartige Fehler untersucht. Deshalb hat das Ergebnis nicht die semantische Bedeutung „Straße“ oder „Straßennetz“, sondern es handelt sich nur um Linien in der Karte, die das Straßennetz darstellen könnten. Besser gesagt, es wurde nicht überprüft, ob diese Linien sinnvolle Straßen bilden. Jedoch reicht das Ergebnis aus, Informationen über die mögliche Lage mancher Objekte abzuleiten. Zum Beispiel kann man Hinweise über Siedlungsbereiche durch die hohe Konzentration von Straßen in manchen Regionen der Karte ableiten.

Ein Bild des gesamten Testgebietes ist sehr groß aufgrund der hohen räumlichen Auflösung. Zur Verarbeitung des Rasterbildes der Karte wurde es in zwei Bilder geteilt, ein oberes und ein unteres Bild. Diese Teilung hatte Folgen, vor allem bei der Erkennung der Straßen. Die doppelte Skelettierung einiger Straßen verursachte doppelte Linien im Überlappungsbereich zwischen beiden Bildern. Diese Fehler konnten mit Hilfe der mathematischen Morphologie verbessert werden. Das Straßennetz wurde zuerst ausgedehnt und danach wieder erodiert, um die redundanten Linien mit den nächstliegenden Linien zu verschmelzen. Diese nachträgliche Bereinigung der Linien erfolgt unkontrolliert und kann negative Wirkungen auf die genaue Definition der Linien haben (Brügelmann [1998]).

Aus der strengen Darstellung der binären Bilder wurden neue Bilder mit einem sanfteren Übergang zwischen Vordergrund und Hintergrund berechnet, gemäß den im Kapitel 6 beschriebenen Methoden. Ausschnitte dieser Bilder werden in Abb. 56 und 57 gezeigt. Der Übergang stellt die Unsicherheit über die genaue Position der Grenze dar.



(a) Sumpfgebiet

(b) Wasser



(c) Straßennetz

(d) Siedlungsbereiche

Abbildung 55: Binäre Ergebnisbilder der Kartenerkennung: (a) Sumpfgebiet, (b) Wasser, (c) Straßennetz und (d) Siedlungsbereiche.

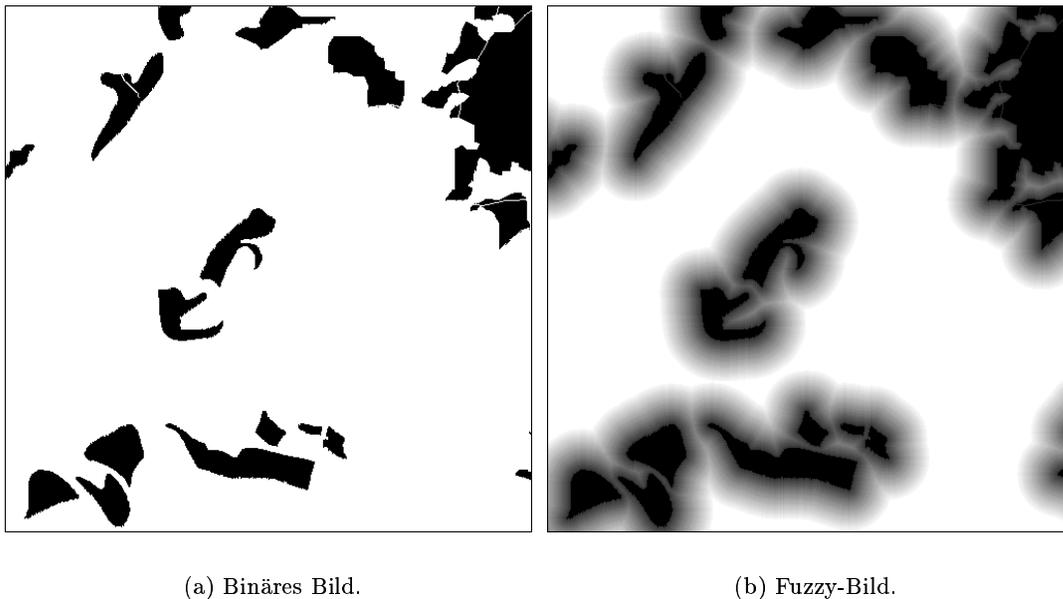


Abbildung 56: Ergebnis der Fuzzifizierung eines binären Bildes: (a) ursprüngliches Bild: grüne Flächen (b) Fuzzy-Bild.

Der Vergleich zwischen Abb. 55c und Abb. 57c zeigt, daß die genaue Bestimmung der Straßen innerhalb der Siedlungsbereiche nicht von großer Bedeutung ist, da durch die hohe Linienkonzentration diese Regionen von Information überflutet werden. Der geringe Abstand zwischen den Linien erzeugt höhere Werte in den Zwischenräumen, so daß der genaue Verlauf der Linien keine große Veränderung bedeuten würde. Das ist nicht der Fall außerhalb der besiedelten Flächen, in denen falsche Informationen nicht von anderen Linien verbessert werden können.

Abb. 55b zeigt, daß die Grenzen des Rio dos Sinos fehlerhaft erkannt wurden. Die Verschmelzung der Muster und dem Fluß verursacht kleine Spitzen an den Rändern des Flusses. Wenn die genaue Erkennung des Flusses von Interesse wäre, sollten diese Ungenauigkeiten bereinigt werden. Jedoch spielen sie bei der Umwandlung des binären Bildes in ein unscharfes Bild eine sekundäre Rolle (Abb. 57b).

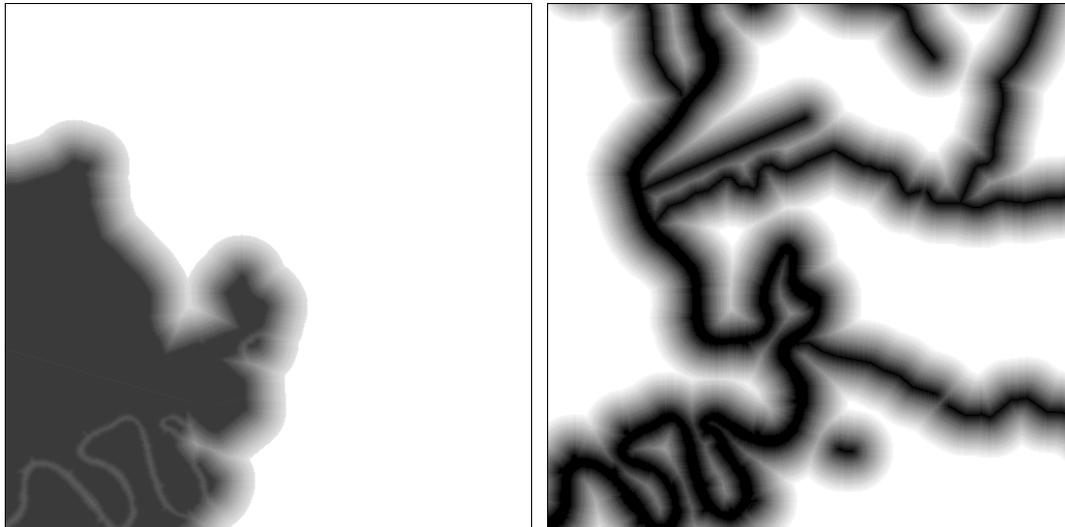
Die unscharfe Darstellung der Ergebnisse der Kartenerkennung wurde im Rasterformat gespeichert.

Es gibt Regionen, in denen keine Information zur Verfügung steht. Diese Pixel gehören entweder zum Hintergrund oder zu anderen nicht berücksichtigten kartographischen Elementen. Die Interpretation dieser Pixel aus der Information der Karte ist unmöglich, weil keine Hinweise zur Verfügung stehen, die eine Entscheidung unterstützen könnten. Dies kommt sehr häufig vor, da eine Karte eine vereinfachte (ikonische) Darstellung der Szene ist, in der viele kleine Einzelheiten vernachlässigt werden, um eine bessere Lesbarkeit des graphischen Dokuments zu erreichen (Generalisierung). Es ist deswegen sinnvoll, diese Lücken mit „Null“ (keinem Hinweis für eine bestimmte Klasse) in allen Ebenen darzustellen.

8.2 Beschreibung des Testgebietes anhand der Satellitenbilder

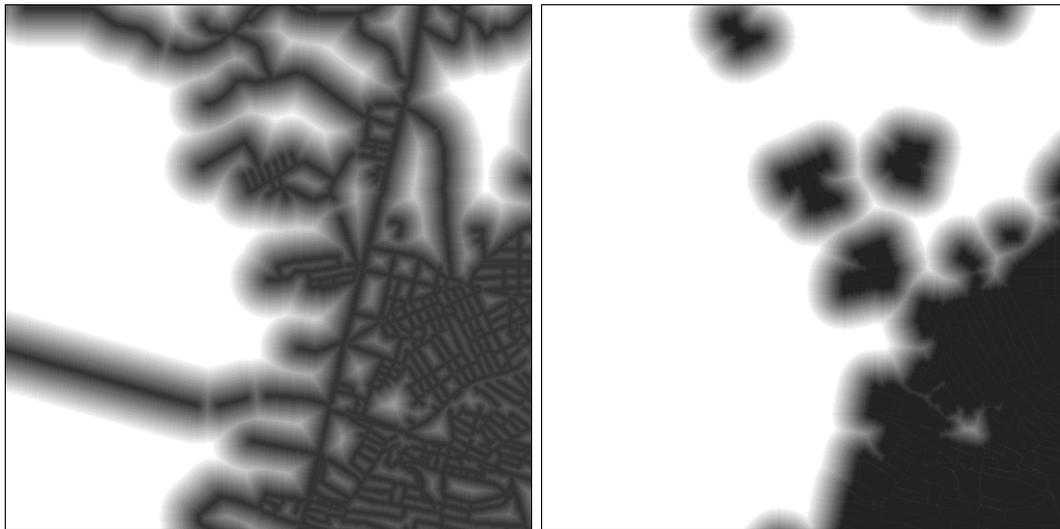
Die multispektrale Klassifikation eines Landsat-TM Bildes des Gebietes wurde gemäß des Prinzips des Maximum-Likelihood-Verfahrens durchgeführt. Dazu wurden drei Kanäle dieses Systems benutzt: TM5, TM4 (Infrarot) und TM3 (Sichtbar, rot). Trainingsgebiete mit typischen spektralen Eigenschaften der Landnutzungsklassen wurden ausgewählt, und aus diesen Pixeln wurden die notwendigen statistischen Parameter (Mittelwertvektor und Kovarianzmatrix) für jede Klasse berechnet.

Das Gebiet ist durch eine Vielfalt von Landnutzungen geprägt. Am Rande einer Autobahn lassen sich deutlich Siedlungsbereiche erkennen. Diese Flächen sind in der Regel nicht homogen, sondern es wird



(a) Sumpfgebiet

(b) Wasser



(c) Straßennetz

(d) Siedlungsbereiche

Abbildung 57: Ergebnis der Fuzzifizierung der binären Bilder: (a) Sumpfgebiet, (b) Wasser, (c) Straßennetz und (d) Siedlungsbereiche.

eine größere Variation der Grauwerte durch die Mischung der Objekte (Mischpixel) erzeugt. In einer Stadt sind sowohl Gebäude-, Vegetationsflächen als auch unversiegelte Flächen vertreten. Aufgrund der kleinen Objektdimensionen sind diese Elemente in einem einzigen Pixel nicht zu trennen und werden deshalb mit einem einzigen Grauwert dargestellt. Trotzdem lassen sich Siedlungsflächen mit gemeinsamen Merkmalen, wie zum Beispiel stark versiegelte Flächen und Flächen niedriger Dichte, unterscheiden. Die Kleingliederung, welche insbesondere kleinere brasilianischen Städten auszeichnet, spiegelt sich durch unversiegelte Flächen an Stadträndern und in neuen Stadtteilen wider. Allein aufgrund der spektralen Signatur sind diese Flächen schwer von anderen Flächen, außerhalb der Siedlungsbereiche, zu trennen.

In der Nähe des Rio dos Sinos wird eine intensive landwirtschaftliche Bodennutzung betrieben. Die Region eignet sich besonders für Reisanbau. Die landwirtschaftlichen Flächen lassen sich daher in zwei große Klassen einteilen: Reisanbau (bewässert) und andere Kulturen.

Tabelle 6 zeigt eine Zusammenfassung der Klassen, die aus dem Landsat-TM Satellitenbild ermittelt wurden. Abb. 58 zeigt das Ergebnis der Maximum-Likelihood-Klassifikation.

Tabelle 6: Klassen der spektralen Klassifikation.

Klasse	Anzahl der Unterklassen
Wiese	5
Siedlungsbereiche	8
Vegetation	3
Mandwirtschaft	10
Wasser	4
Boden	16
Sumpfgebiet	1
Gras	2
Landwirtschaftliche Flächen	10

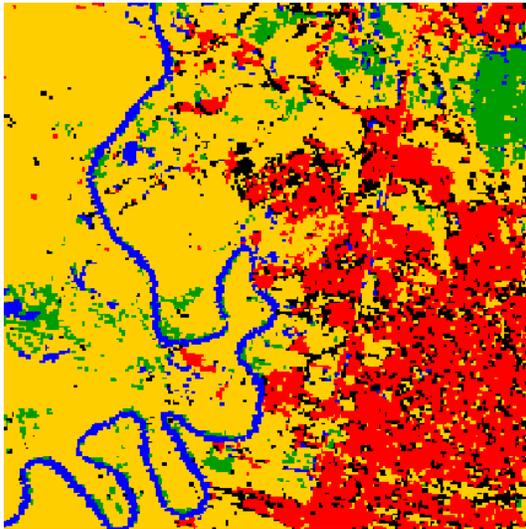
Die aus der Maximum-Likelihood-Klassifikation abgeleitete Wahrscheinlichkeitsfunktionen wurden zusammengefaßt, und die Klassen „Wasser“, „Wald“, „Siedlungsbereiche“, „Reis“ und „Landwirtschaft“ wurden in getrennten Ebenen gespeichert. Da das Ergebnis der multispektralen Klassifikation eindeutig ist wurden sie nicht direkt benutzt. Statt dessen wurden unscharfe Menge aus der Dichtefunktionen abgeleitet. Diese Werte wurden mit Grauwerten zwischen 0 und 200 kodiert und in der Datenbasis gespeichert.

Die Variation der lokalen Homogenität wurde aus einem Spot-Bild desselben Gebietes errechnet (Abb. 8.2).

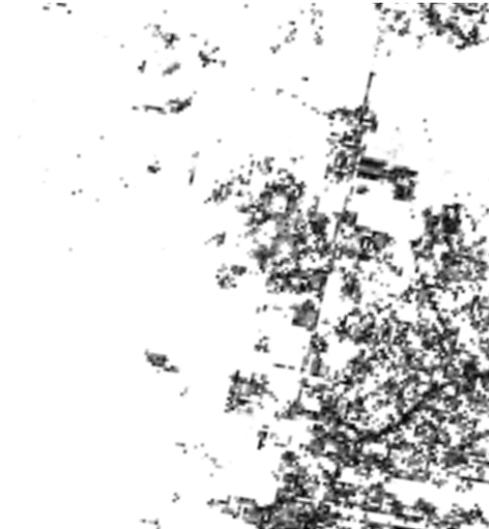
8.3 Regeln

Anhand der Beschreibung der spektralen Eigenschaften der Szene, ihrer Textur und der Information aus der Karte wurden im folgenden Schritt passende Regeln definiert. Diese Regeln stellen eine Beziehung zwischen den Hinweisen und dem gewünschten Ergebnis dar. Weiterhin verbinden diese Regeln Fakten, die in den unterschiedlichen Ebenen dargestellt werden, mit Schlußfolgerungen über die Klassifikation des Pixels.

Das Wissen über die Szene wurde benutzt, um 11 Regeln für das Testgebiet vorzuschlagen. Manche kombinieren Informationen aus dem Satellitenbild und der Karte, andere benutzen nur eine einzige Quelle. Wenn die Hinweise der Karte ausreichen, zum Beispiel bei der Identifizierung des Straßennetzes, wurden sehr einfache Regeln formuliert.



(a) Maximum Likelihood. Rot: Siedlungsgebiete, grün: Wald, gelb: Landwirtschaft, blau: Wasser, schwarz: nicht klassifiziert.



(b) Wahrscheinlichkeitsfunktion für die Klasse Siedlungsgebiete. Helle Pixel entsprechen höhere Funktionswerten.

Abbildung 58: Beispiel der Maximum Likelihood Wahrscheinlichkeitsfunktionen: (a) Ergebnis der Klassifikation, (b) Wahrscheinlichkeitsfunktion für die Klasse Siedlungsgebiete.

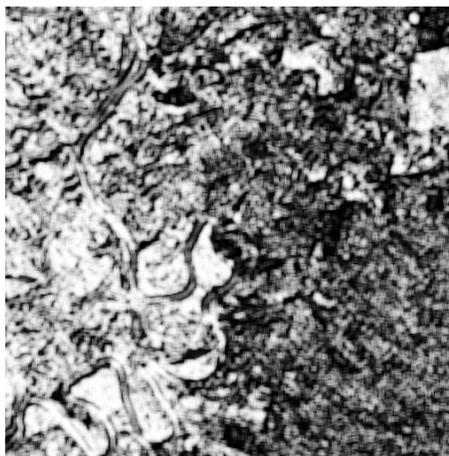


Abbildung 59: Spektrale Homogenität im Spot-Bild. Helle Pixel stellen höhere Werte dar.

Die Tabelle 7 gibt ein Beispiel der Regeln für das Gebiet um die Stadt Sapucaia do Sul.

Tabelle 7: Regeln für die Klassifikation des Gebiets um die Stadt Sapucaia do Sul. Hinweise für: SU: Siedlungsbereiche im Satellitenbild; SA: Reis im Satellitenbild; SW: Wald im Satellitenbild; SB: Sumpfbereich im Satellitenbild; SL: Landwirtschaft im Satellitenbild; SR: Wasser im Satellitenbild; MU: Siedlungsbereiche auf der Karte; MW: Wald auf der Karte; MB: Sumpfbereich auf der Karte; MR: Wasser auf der Karte; MS: Straßennetz auf der Karte; HO: Lokale Homogenität.

Nr.	IF	AND	AND	AND	THEN
1	SU hoch	MU hoch	HO niedrig		Siedlungsbereich
2	SU hoch	MU hoch	HO mittel		Siedlungsbereich
3	MU sehr groß				Siedlungsbereich
4	SU hoch	HO niedrig			Siedlungsbereich
5	SW hoch	MW hoch	HO hoch		Wald
6	SW hoch	MW hoch	HO mittel		Wald
7	SR hoch	MR hoch			Wasser
8	SL hoch	MB hoch			Landwirtschaft
9	SA hoch	MB hoch			Landwirtschaft
10	SL hoch	SW niedrig	MU niedrig		Landwirtschaft
11	SA hoch	SW niedrig	MU niedrig	SU niedrig	Landwirtschaft

8.4 Fuzzy-Funktionen

In den folgenden Abschnitten wird die Definition der Zugehörigkeitsfunktionen für die linguistischen Terme der Eingangs- und Ausgangsvariablen beschrieben.

8.4.1 Landnutzungsklassen

Es wurden dreieckige Funktionen zur Darstellung der linguistischen Terme der Schlußfolgerung benutzt (Abb. 60). Alle Funktionen haben die gleiche Form und Fläche, so daß ein Gleichgewicht zwischen den Klassen entsteht.

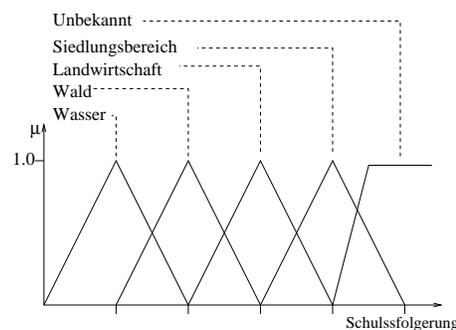


Abbildung 60: Linguistische Terme für die Landnutzungsklassen als unscharfe Ausgangsmengen.

8.4.2 Schätzung der Fuzzy-Funktionen für die linguistischen Terme

Der Versuch, die unscharfen Funktionen der Eingangsgrößen empirisch zu ermitteln, ist eine schwere Aufgabe, da es sich um eine große Anzahl zusammenwirkender Funktionen handelt. Die Veränderung einer Funktion wirkt auf mehrere Regeln und muß deswegen in engem Zusammenhang mit den anderen mitwirkenden Regeln durchgeführt werden.

Ein Mensch kann nicht gleichzeitig alle Daten der Bilder wahrnehmen, und die Analyse einzelner Punkte ist nicht repräsentativ für die Gesamtheit der Szene. Als erste Näherung werden die unscharfe Mengen aus einer beschränkten menschlichen Analyse abgeleitet: Es werden die Erfahrungen eines menschlichen Operators mit offensichtlichen Eigenschaften der Bilder als erste Näherung verwendet. Als Zugehörigkeitsfunktionen werden einfache trapezförmige Funktionen gewählt. Der Vorteil des Trapezes gegenüber dem Dreieck liegt darin, daß es nicht nur einen Punkt, sondern für ein ganzes Intervall eine maximale Bewertung zuläßt. Die Übergänge zwischen „Null“ und „Eins“ können sehr steil oder sanft sein, das hängt von den Koordinaten der Ecken des Trapezes ab. Ein Versuch ist in den Tabellen 8, 9 und 10 dokumentiert.

Die Tabelle 8 zeigt die Funktionen der Variable *lokaler Homogenität*. Der Variationsbereich dieser Variablen verteilt sich sehr homogen von Null bis 200. Um diese Variation besser darzustellen, wurden drei Adjektive benutzt: „niedrig“, „mittel“ und „hoch“. Das erste Adjektiv („niedrig“) kennzeichnet Regionen mit homogenen Grauwerten, zum Beispiel Ackerflächen. Für die Definition der Funktion („hoch“) wurde die Variation der Grauwerte innerhalb der Siedlungsbereiche berücksichtigt. Zwischenwerte, deren Bedeutung nicht unbedingt mit einer bestimmten Klasse verbunden ist, werden mit einer dritten Variablen („mittel“) dargestellt.

Tabelle 8: Fuzzy-Funktionen lokaler Homogenität. HO: Lokale Homogenität.

Funktion	p1	p2	p3	p4	p5	p6
HO hoch	0. 0.	50. 0.	90. 1.	200. 1.	201. 0.	201. 0.
HO mittel	0. 0.	1. 0.	20. 1.	50. 1.	100. 0.	201. 0.
HO niedrig	0. 0.	1. 1.	50. 1.	60. 0.	200. 0.	201. 0.

Die spektralen Informationen des Satellitenbildes wurden mit 200 Zwischenwerten gespeichert. Die Verteilung dieser Werte ist nicht homogen, sondern extreme Werte, entweder niedrig oder sehr hoch, kommen öfter vor. Deswegen werden nur zwei Terme für diese linguistischen Variablen verwendet. Der Übergang zwischen „hoch“ und „niedrig“ ist sehr steil (Tab. 9) und nähert sich damit der booleschen Logik. Die Schwierigkeit besteht darin, einen harmonischen und richtigen Verlauf aller Funktionen gleichzeitig zu bestimmen.

Tabelle 9: Fuzzy-Funktionen spektraler Hinweise. SU: Siedlungsbereich im Satellitenbild; SA: Reis im Satellitenbild; SW: Wald im Satellitenbild; SB: Sumpfbereich im Satellitenbild; SL: Landwirtschaft im Satellitenbild; SR: Wasser im Satellitenbild.

Funktion	p1	p2	p3	p4	p5
SL hoch	0. 0.	3. 0.	5. 1.	200. 1.	201. 0.
SL niedrig	0. 0.	1. 1.	8. 1.	10. 0.	201. 0.
SR hoch	0. 0.	13. 0.	15. 1.	200. 1.	201. 0.
SR niedrig	0. 0.	1. 1.	13. 1.	15. 0.	201. 0.
SA hoch	0. 0.	3. 0.	5. 1.	200. 1.	201. 0.
SA niedrig	0. 0.	1. 1.	8. 1.	10. 0.	201. 0.
SB hoch	0. 0.	196. 0.	198. 1.	200. 1.	201. 0.
SB niedrig	0. 0.	1. 1.	196. 1.	198. 0.	201. 0.
SU hoch	0. 0.	15. 0.	20. 1.	200. 1.	201. 0.
SU niedrig	0. 0.	1. 1.	40. 1.	50. 0.	201. 0.
SW hoch	0. 0.	8. 0.	10. 1.	200. 1.	201. 0.
SW niedrig	0. 0.	1. 1.	8. 1.	10. 0.	201. 0.

Für die Informationseinheiten aus der Karte kommen vor allem extreme Werte vor. Das kann man auf die Berechnung der Unsicherheitswolken um die erkannten Elemente zurückführen. Es wird angenom-

men, daß die Lagegenauigkeit der Karte gut ist. In diesem Fall reicht es aus schmale Regionen um die Kartenelemente vorzugeben. Die Funktionen der linguistischen Terme der Karte werden in Tabelle 10 gezeigt.

Tabelle 10: Fuzzy-Funktionen der Informationen aus der Karte. MU: Siedlungsbereich auf der Karte; MW: Wald auf der Karte; MB: Sumpfbereich auf der Karte; MR: Wasser auf der Karte; MS: Straßennetz auf der Karte.

Funktion	p1	p2	p3	p4	p5
MB hoch	0. 0.	190. 0.	191. 1.	200. 1.	201. 0.
MB niedrig	0. 0.	1. 1.	13. 1.	15. 0.	201. 0.
MS sehr hoch	0. 0.	199. 0.	192. 1.	200. 1.	201. 0.
MS hoch	0. 0.	13. 0.	180. 1.	200. 1.	201. 0.
MS niedrig	0. 0.	1. 1.	13. 1.	15. 0.	201. 0.
MU sehr hoch	0. 0.	198. 0.	199. 1.	200. 1.	201. 0.
MU hoch	0. 0.	190. 0.	195. 1.	198. 1.	201. 0.
MU niedrig	0. 0.	1. 1.	190. 1.	195. 0.	201. 0.
MW hoch	0. 0.	8. 0.	198. 1.	200. 1.	201. 0.
MW niedrig	0. 0.	1. 1.	198. 1.	200. 0.	201. 0.

Die Abb. 61 stellt eine Klassifikation vor, die man mit empirischen Fuzzy-Funktionen erhält. Das Ergebnis zeigt viele Lücken in Regionen, in denen es weniger Hinweise für eine Klassifikation gab. Diese Lücken treten eher an landwirtschaftlichen Flächen auf, für die keine Informationen aus der Karte abgeleitet werden können. Sie könnten mit der Information aus dem multispektralen Bild ergänzt werden, wenn man für diese Pixel das Ergebnis der Maximum-Likelihood-Klassifikation als die einzig gültige Information annimmt (Abb. 62).

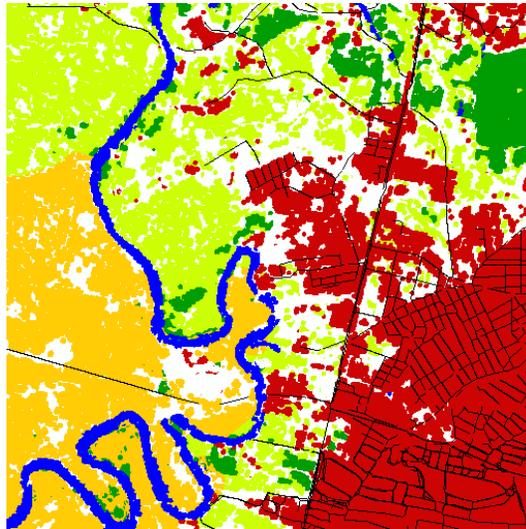


Abbildung 61: Ergebnis der Klassifikation. rot: Siedlungsgebiete, grün: Wald, gelb: Landwirtschaft (Reisanbau), hellgrün: Landwirtschaft, blau: Wasser, schwarz: Straßen, weiß: nicht klassifizierte Pixel.

Die besiedelten Flächen erscheinen im Ergebnisbild sehr homogen. Die Pixel der Stadt *Sapucaia do Sul* (unten links im Bild) sind nicht zerstreut, sondern bilden eine kompakte Fläche. Diese Eigenschaft folgt aus der abgeleiteten Karteninformation. Aus dem Vergleich zwischen der ursprünglichen Karte und dem Ergebnis (Abb. 62) wird deutlich, daß das Wachstum der Stadt im neuen Ergebnis erkannt würde.

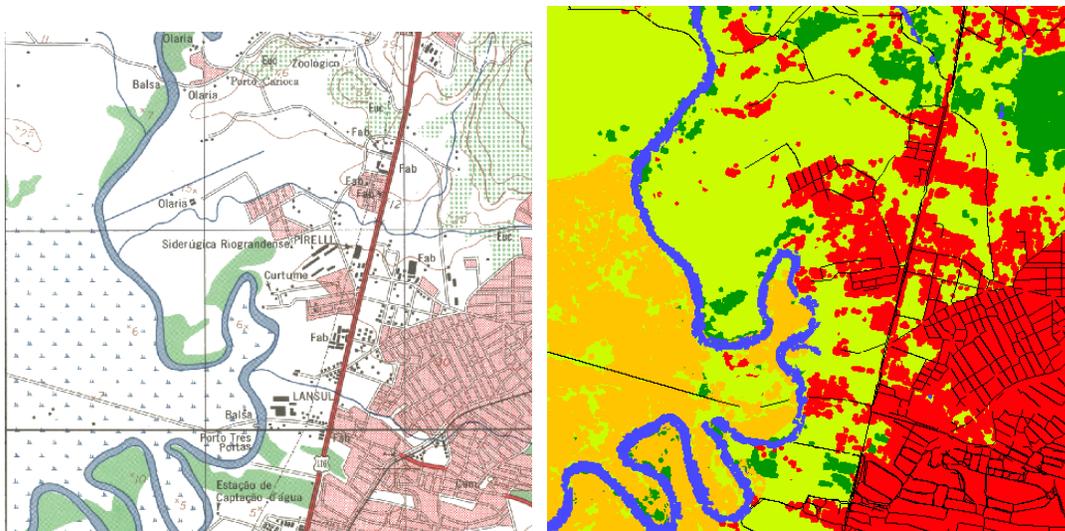
Ein Beispiel hierfür erkennt man am nördlichen Rand der Stadt. Diese Verbesserung tritt deutlicher bei anderen kleinen Siedlungsbereichen mitten im Bild hervor. Dort wird die Information aus der Karte

durch die Information aus den Satellitenbildern ergänzt. Dadurch entstehen größere Siedlungsbereiche um die Landstraße.

Die Klassifikation der Wälder und der Vegetation zeigt ebenfalls den Gewinn, der durch das Zusammenwirken beider Informationsquellen entsteht. Die Flächen sind kompakt (Eigenschaft der Karte) und wurden zugleich mit den Informationen der Satellitenbilder aktualisiert. Die Vegetationsflächen wurden kleiner, und vor allem am Ufer des Rio dos Sinos kann man große Veränderungen nachweisen. Fehler können in den bewaldeten Flächen gesehen werden (oben rechts im Abb. 61). Diese Flächen wurden entweder unterschätzt, nicht klassifiziert oder als landwirtschaftliche Flächen falsch klassifiziert (Abb. 61).

Der Übergang zwischen Klassen wird oft nicht klassifiziert. Dies läßt sich mit der Definition einiger Fuzzy-Funktionen erklären. Die Funktionen werden aus reinen Pixeln innerhalb der Flächen abgeleitet. Dort gibt es keine Mischung zwischen Klassen.

Die großen Lücken im Sumpfgebiet (Reisanbau) können durch die variierenden spektralen Eigenschaften der Reisanbauflächen im Landsat-Bild erklärt werden. Diese Klasse kann als grüne Fläche, Boden, feuchter Boden oder feuchte Vegetation auftreten.



(a) Topographischen Karte.

(b) Verbesserung des Ergebnisbildes mit der Maximum-Likelihood-Klassifikation. Rot: Siedlungsgebiete, grün: Wald, gelb: Landwirtschaft (Reisanbau), hell grün: Landwirtschaft, blau: Wasser, schwarz: Straßen.

Abbildung 62: Vergleich zwischen der ursprünglichen Karte (a) und dem Ergebnis (b).

8.4.3 Anwendung der genetischen Algorithmen zur Bestimmung der Fuzzy-Funktionen

Als zweite Option zur Bestimmung der unscharfen Funktionen bietet sich eine rechnerische Annäherung mittels eines numerischen Verfahrens. Im Abschnitt 7.1 wurde eine Optimierungsmethode zur Berechnung der Parameter dieser Funktionen beschrieben.

Das Ziel des Optimierungsverfahrens ist es, die Parameter der Funktionen zu schätzen, die die richtige Klassifikation der Pixel anhand der zur Verfügung stehenden Hinweise ermitteln. Diese Schätzung folgt aus einer Lernphase, in der die Ein- und Ausgabedaten analysiert werden und ein Modell vorgeschlagen wird. Dafür ist eine Trainingsmenge mit der wahren Klassifikation einiger Pixel nötig. Deswegen wurde ein kleines Gebiet im Bild gesucht, in dem die unterschiedlichen Klassen vorhanden sind und die

Variation der Merkmale deutlich geschätzt werden kann. Ein Sollbild entstand aufgrund einer visuellen Interpretation von Karte, Landsat-Bild und Spot-Bild. Alle zur Verfügung stehenden Merkmale werden benutzt, um Regionen mit einer relativ sicheren Klassifikation zu identifizieren. Trainingsgebiete für „Wald“ und „Landwirtschaft“ sind ziemlich leicht zu erkennen und zu begrenzen, vor allem wenn man die Daten aus dem Landsat-Bild und der Karte kombiniert. Bei „Siedlungsgebieten“ ist sogar die visuelle Interpretation des Landsat-Bildes schwierig. Trotzdem war es möglich, sichere Regionen aus der Karte und dem Spot-Bild abzuleiten. Es ist sehr wichtig, daß die Trainingsmenge nicht nur Pixel in homogenen Regionen einschließt, sondern auch Pixel, die sich auf Übergängen zwischen Klassen befinden. Diese Maßnahme spielt bei der Schätzung der Fuzzy-Mengen eine wichtige Rolle, denn der Zugehörigkeitsgrad wird nicht nur für die extremen Fälle ermittelt, sondern auch für die Übergänge zwischen „keine Zugehörigkeit“ und „volle Zugehörigkeit“. Diese Bereiche werden von den Seiten der trapezförmigen Funktionen dargestellt.

Die Kodierung der Punkte erfolgt gemäß Abschnitt 7.2. Für jedes Trapez sind 4 Punkte nötig und ausreichend. Da es sich um 16 Adjektive handelt, bilden insgesamt $64 \cdot 8$ Chromosomen die genetische binäre Kette.

Eine ziemlich kleine *Bevölkerung* wurde benutzt, vor allem weil die Analyse des Tauglichkeitsgrades jedes *Individuums* (d.h., die Analyse der Qualität des Ergebnisses der Klassifikation gemäß der in seiner genetischen Kette gespeicherte Parameter) sehr mühsam ist. Der *Tauglichkeitsgrad* wird aus der Klassifikation der ganzen Trainingsgebiete abgeleitet. Die Aufgabe ist rechenintensiv und hängt von der Größe der Trainingsgebiete ab. Tests mit *Bevölkerungen* von 30 und 40 *Individuen* wurden durchgeführt. Bei einer *Bevölkerung* von 40 *Individuen* wurden insgesamt 28 neue *Individuen* erzeugt und 12 *Eltern* in die Kindermenge kopiert (*Elitismus*).

Da die *Mutation* zufällig geschieht kann es passieren, daß eine Funktion außerhalb des Bereichs zwischen Null und 200 entsteht. In diesem Fall wird dieser Teil der *genetischen Kette* durch ein einfaches neutrales Trapez ersetzt (Tab. 11). Dieses neutrale Trapez liegt genau in der Mitte des Variationsbereiches und begünstigt damit keinen extremen Fall der linguistischen Terme. In der Praxis entspricht diese Entscheidung einer größeren Veränderung als die Umkehrung nur eines Bits.

Tabelle 11: Neutrales Trapez.

Funktion	p1	p2	p3	p4	p5
neutral	0. 0.	50. 0.	110. 1.	150. 1.	200. 0.

Ein Uniformitätstest wird nach jeder Iteration durchgeführt. Um den Prozeß anzuhalten, wird überprüft, ob 70 Prozent der besten *Individuen* ähnlich waren. In diesem Fall wird angenommen, daß der Prozeß konvergiert ist. Der mittlere Abstand zwischen den *Kettengliedern* aller *Individuen* wird als Kriterium gewählt. Die Tabelle 12 stellt eine Zusammenfassung der Parameter zur Steuerung des evolutiven Prozesses einer 40-Individuen-Bevölkerung dar.

Klassifikation mit den Ergebnissen der genetischen Algorithmen Der Datensatz wurde auch mit Hilfe der rechnerisch geschätzten Fuzzy-Funktionen klassifiziert. Die Schätzung der Funktionen erfolgt in zwei Schritten. Zuerst wurde ein Suchprozeß durchgeführt, in dem die erste Bevölkerung zufällig erzeugt wurde. Die Ergebnisse dieses Versuchs waren nicht besonders gut, aber konnten als eine erste Näherung für eine neue Suche benutzt werden.

Versuch 1 Im ersten Versuch werden 20 kleine Regionen des Bildes (10x10 Pixel) als Trainingsdaten zur Berechnung der Qualität der Funktionen ausgewählt. Diese Trainingsgebiete liegen in der Mitte der Objekte, und enthalten deswegen weniger Informationen über die Mischung der Klassen oder die Übergänge zwischen den Klassen. Abb. 63 stellt das Ergebnis dieser Klassifikation vor.

Tabelle 12: Zusammenfassung der Parameter der genetischen Algorithmen.

Parameter	Wert
Bevölkerung	40 Individuen
Anzahl der Chromosomen	64*8 Bits = 512
Auswahlmethode	Auswahl nach dem Tauglichkeitskriterium
Überlebende	30 Prozent
Größe der Elite-Menge	30 Prozent
Paarungsmethode	single point method
Cross point	Mitte der Kette
Anzahl der Mutanten	5
Mutation	zufällige Mutation
Uniformitätskriterium	mittlerer Abstand
Größe der Menge für den Uniformitätstest	70 Prozent
Maximales Alter	1500 Generationen

Im Vergleich zur Klassifikation mit den empirischen Fuzzy-Funktionen (Abb. 61) zeigt dieses Ergebnis einige Vorteile. Die Menge der unklassifizierte Flächen geht zurück, und die Objektkonturen sind besser definiert. Leider zeigt diese Lösung auch keine zufriedenstellende Qualität. Manche Klassen wurden anteilmäßig überschätzt, Wasserflächen sind überall im Bild zu finden, und „Wasser“ und „grüner Wald“ werden oft falsch klassifiziert. Auch die Grenzen der besiedelten Flächen sind dem Anteil nachüberschätzt. Der Verlauf ihrer Konturen erinnert an die Karte.

Versuch 2 Theoretisch ist der Startpunkt der genetischen Algorithmen eine Menge zufälliger Lösungen. Man kann aber auch eine nicht optimierte Lösung eines vorherigen Versuches als eines der *Startindividuen* einführen, um dem Prozeß eine bessere Ausgangssituation zu Verfügung zu stellen. Wenn diese Lösung sich nicht eignet, dann wird sie vom Verfahren aussortiert. Wenn sie aber eine gute Näherung ist, dann würde sie den Prozeß beschleunigen.

In einem zweiten Versuch wird die Lösung des ersten Tests als ein *Individuum* der ersten *Generation* kodiert, so daß die ersten Schritte übersprungen und der Prozeß beschleunigt werden konnte. Eine größere Region im Bild (512x512 Pixel), in der nicht nur alle Klassen vorkommen, sondern in der auch die Übergänge zwischen den Klassen deutlich zu sehen sind, wird als Trainingsmenge benutzt. Die Satellitenbilder und die Karte wurden analysiert und ein Sollbild für diese Region hergestellt. Ein direkter Vergleich zwischen dem Sollbild und den Zwischenergebnissen der genetischen Algorithmen ermöglicht die Berechnung der Anpassung jedes *Individuum*. Abb. 64 stellt das Ergebnis dieser Klassifikation vor.

Abb. 65 zeigt verschiedene Zwischenlösungen des genetischen Verfahrens. Es werden die besten Lösungen nach 10, 20, 100 und 200 Iterationen gezeigt. Nach 10 Iterationen erfährt die Klassifikation schon eine bemerkenswerte Verbesserung. Ein Vergleich zwischen Abb. 65a und Abb. 63 zeigt, daß die Siedlungsbereiche nicht mehr übertrieben und viele falsch klassifizierte Wasserflächen beseitigt wurden. Der Nachteil daran ist, daß mit der Reduktion der Wasserflächen auch richtig klassifizierte Pixel am Fluß verloren gehen.

Nach 20 Iterationen (Abb. 65b) werden die meisten Lücken geschlossen und die Klassifikation der landwirtschaftlichen Flächen verbessert. Die Siedlungsgebiete verbessern sich nach 100 Iterationen (Abb. 65c). Ihre Grenzen werden an dieser Stelle besser definiert, obwohl es noch viele nicht klassifizierte Pixel gibt, die aber nach 200 Iterationen gelöscht und durch die richtige Klassifikation ersetzt werden (Abb. 65d). Auch die grünen Flächen um den Fluß, die zuerst gelöscht wurden, werden zu diesem Zeitpunkt richtig klassifiziert.

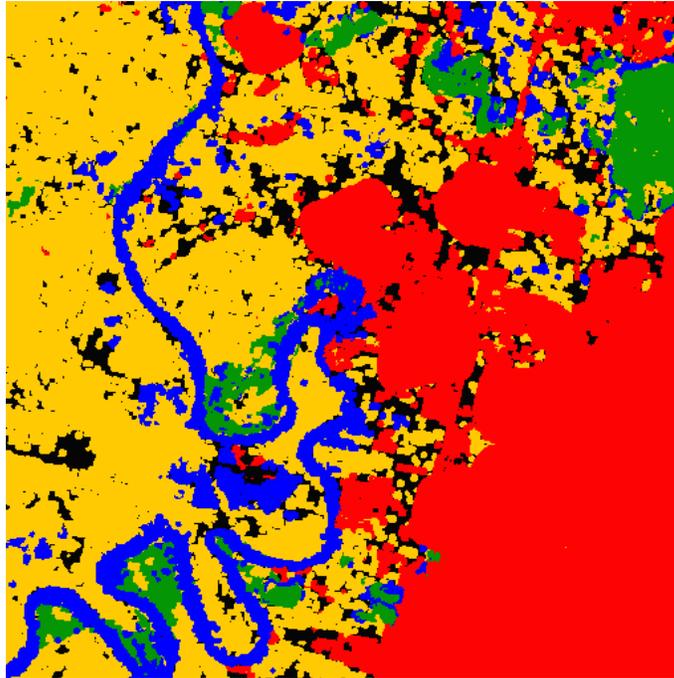


Abbildung 63: Ergebnis der Klassifikation. Zur Bestimmung der unscharfen Funktionen wurden 20 Trainingsgebiete (10x10) benutzt.

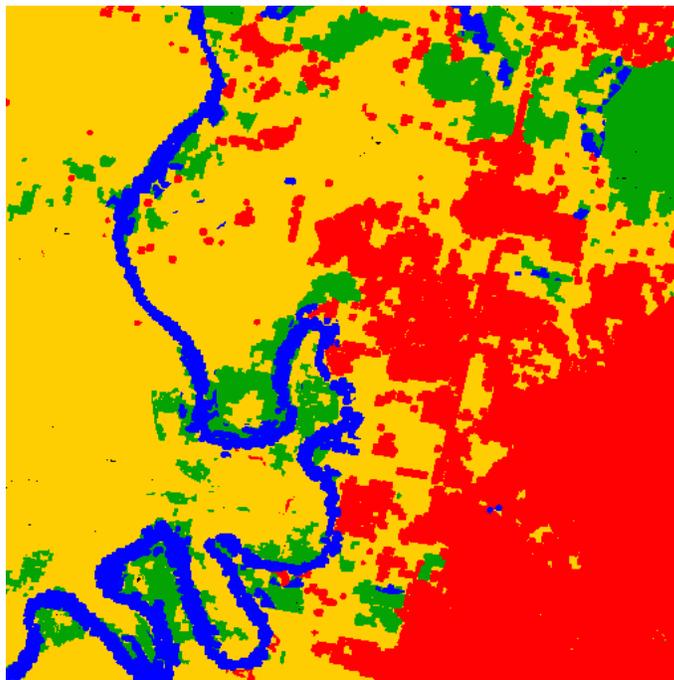
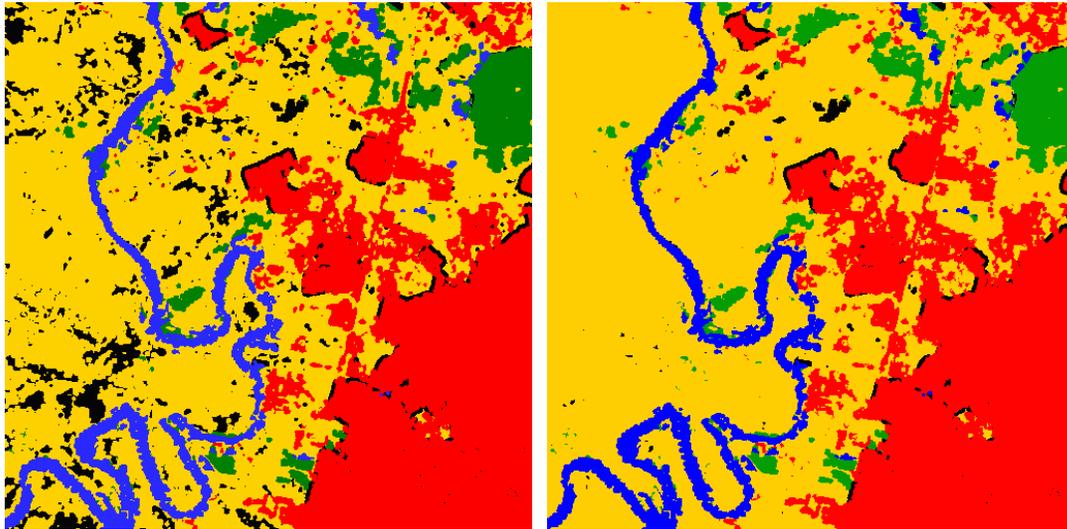
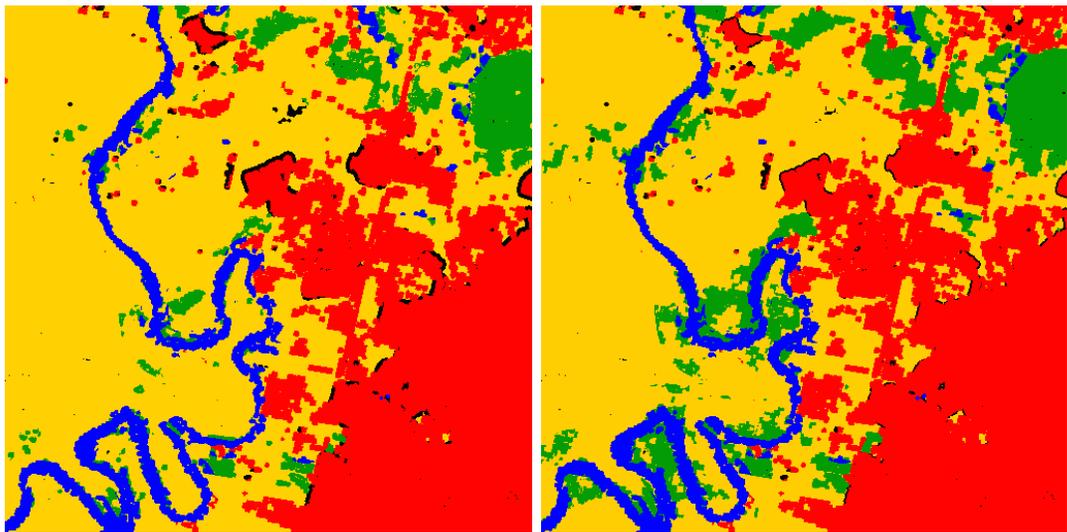


Abbildung 64: Klassifikation. Die Zugehörigkeitsfunktionen wurden mit Hilfe der genetischen Algorithmen geschätzt. Rot: Siedlungsgebiete, grün: Wald, gelb: Landwirtschaft, blau: Wasser.



(a) 10 Iterationen

(b) 20 Iterationen



(c) 100 Iterationen

(d) 200 Iterationen

Abbildung 65: Zwischenergebnis der automatischen Schätzung der Funktionen (a) 10 Iterationen, (b) 20 Iterationen (c) 100 Iterationen und (d) 200 Iterationen. Rot: Siedlungsgebiete, grün: Wald, gelb: Landwirtschaft, blau: Wasser, schwarz: nicht klassifizierte Pixel.

Typische Merkmale der iterativen Suche nach einer optimalen Lösung bei den genetischen Algorithmen ist die progressive Verbesserung der Qualität des Ergebnisses. Sie wird mit dem Tauglichkeitsgrad des besten Individuums, und der Konvergenz der genetischen Kette aller Individuen zu einem gemeinsamen Wert gemessen. Diese beiden Merkmale beschreiben den aktuellen Zustand des Prozesses und werden nach jeder Iteration überwacht.

Abb. 66 stellt die Entwicklung des Tauglichkeitsgrades des besten Individuums (beste Lösung) im letzten Intervall des Prozesses dar. Ein stetiges Wachstum kennzeichnet diese Graphik. Die Verbesserung des Ergebnisses erfolgt zufällig in unregelmäßigen Intervallen. Dieses Verhältnis wurde erwartet, denn die Optimierung der Lösung wird nicht von festen Regeln gesteuert. Wenn neues genetisches Material durch Mutationen in die genetischen Kette eingeführt wird, dann sind große Sprünge im Tauglichkeitsgrad zu beobachten, während die Paarung der besten Individuen kleinere Verbesserungen liefern. Wenn Mutationen und Paarung nicht zur Verbesserung des Ergebnisses beitragen, bleibt der Tauglichkeitsgrad wegen des Elitismus-Prinzips unverändert.

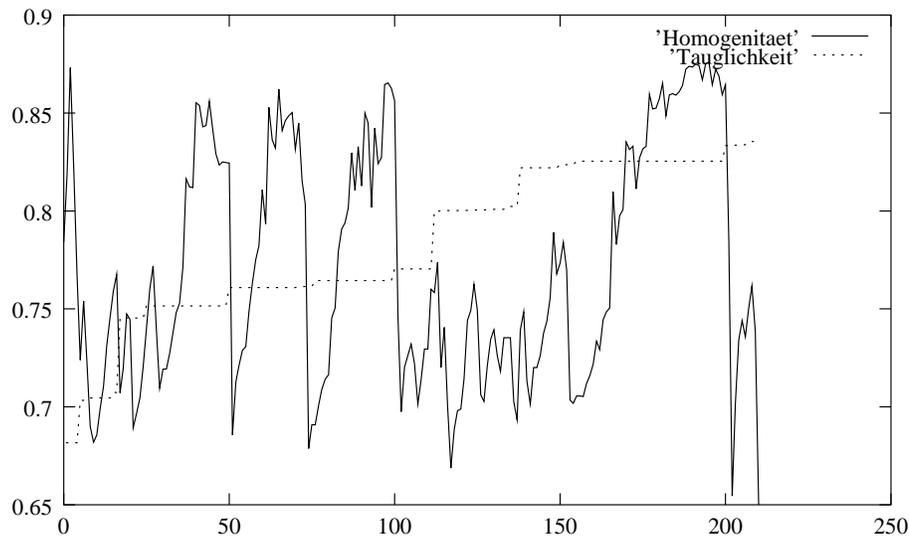


Abbildung 66: Entwicklung der Homogenität und der Bevölkerung.

Die Homogenität der Bevölkerung wächst mit jeder Iteration (Abb. 66). Obwohl die Bevölkerung des genetischen Prozesses sich in Richtung eines ähnlichen Zustandes entwickelt, wird die minimal erwartete Homogenität nicht erreicht und der Prozeß nach einer bestimmten Anzahl von Iterationen angehalten.

In Abb. 66 wird die Entwicklung der Homogenität dargestellt. Die Werte wurden skaliert, um einen Vergleich mit der Entwicklung des Tauglichkeitsgrades zu ermitteln. Der schwankende Verlauf dieser Kurve im Vergleich zur Tauglichkeitsfunktion folgt aus der konstanten Mutation. Nach jeder Iteration erfährt eine kleine Menge Individuen Mutationen, die die Homogenität der Bevölkerung beeinflussen. Wenn ein besseres Individuum durch eine Mutation erzeugt wird, dann sinkt die Homogenität und dieses neue genetische Material wird in den anderen Lösungen verbreitet. Durch die Paarung werden die Individuen die Merkmale des besseren Individuum erben und die Bevölkerung wird in den folgenden Iterationen homogener.

Der Prozeß entwickelt sich langsam und schrittweise, weil jede Operation auf der genetischen Kette (Mutation und Paarung) meist nur einen bestimmten Punkt der Funktionen verändert und weil die Mutation ganz zufällig stattfindet.

In den ersten Iterationen herrscht große Vielfalt an genetischen Ketten, und es gibt viel Platz für Mutationen, die zur Verbesserung der Qualität führen. Eine gute Lösung wird trotzdem nicht schnell erreicht, denn oft führen die zufälligen Mutationen zu schlechteren Optionen.

In einer zweiten Phase wird die Bevölkerung vom Profil eines Individuum dominiert. Viele Lösungen

sehen ähnlich aus, sind aber aufgrund der Paarung dieses Individuum mit anderen Individuen (mit geringer Qualität) weniger geeignet. Die Paarung führt nur zu kleinen Verbesserungen im Ergebnis, denn sie behält die Eigenschaften der besseren Lösung. Die Dominanz dieses Profils wird durch eine Mutation eines ihrer Söhne beendet, denn dadurch wird ein neues, besseres Profil erzeugt.

Das Lernverhalten ist sehr stark von der Tauglichkeitsfunktion und der Trainingsmenge abhängig. Mit einer größeren Menge wird das Lernverhalten verbessert, wobei aber auch der Rechenaufwand steigt. Das Erreichen einer guten Lösung für das Problem kann durch die Einbeziehung der Lösung eines ähnlichen Problems beschleunigt werden. Sie kann aus einem anderen Experiment mit derselben Datenbasis oder aus einem vorherigen Test mit einer kleineren Trainingsmenge als ein Individuum der ersten Generation eingefügt werden. Diese erste Schätzung würde die Rolle eines „Gedächtnisses“ spielen, denn dadurch kann sich der Prozeß an einen ähnlichen Fall „erinnern“ und die alte Lösung dem neuen Datensatz anpassen.

9 Diskussion

Die vorliegende Arbeit konzentriert sich auf die Integration unterschiedlicher Datensätze zur Klassifikation der Landnutzung in Gebieten, in denen nur wenig Daten verfügbar sind. Die Gesamtheit der Daten schließt Bilder der Fernerkundung und bereits existierende Karten ein, zwei Datenquellen, die relativ gut verfügbar sind. Ziel der Arbeit ist es, die Landnutzung anhand dieser Daten zu ermitteln.

Die Arbeit untersucht mehrere Aspekte, von der Erstellung einer passenden Beschreibung der Szene bis zur Klassifikation. Die Ergebnisse werden, nach den Themen gegliedert, auf den folgenden Seiten zusammengefaßt.

9.1 Kartenerkennung

Kartenerkennung ist ein bisher ungeklärtes Problem in der Bildverarbeitung. Low-level Bildverarbeitungsmethoden werden ständig optimiert und an bestimmte Probleme angepaßt, eine allgemeine Lösung steht jedoch noch nicht zur Verfügung. Jedes Bild hat besondere Merkmale, und jeder Benutzer eigene Interessen, welche die Auswahl der Erkennungsmethoden bestimmen. Die hier vorgestellte Kartenerkennung zielt nicht auf die volle Interpretation einer Karte, sondern nur auf die Ableitung von Hinweisen, die die Interpretation von Satellitenbildern unterstützen können. Dadurch wird die Aufgabe wesentlich vereinfacht.

Aus der Karte werden Informationen über ausgewählte Objekte abgeleitet. Manche Einzelheiten, die auf einer Karte sehr genau dargestellt werden, werden nicht berücksichtigt. Das betrifft zum Beispiel die Namen der Flüsse, Straßen oder Städte und die explizite Angabe der Funktion mancher Gebäude, wie z.B. Industriegebäude. Damit ist nicht gemeint, daß man auf diese Information verzichten soll. Da diese Art von Information in einem Satellitenbild nicht vorkommt, kann sie direkt aus der Karte abgelesen werden, und nimmt an der integrierten Verarbeitung nicht teil.

Aus dem gescannten Bild der Karte wird ein binäres Bild für jede Farbe extrahiert. Ein Optimierungsverfahren, basierend auf Fuzzy-Logik, wird angewendet um das RGB-Bild zu segmentieren. Nicht nur spektrale Merkmale (Homogenität und Kontrast der gefärbten Flächen) werden dabei berücksichtigt, sondern auch räumliche Merkmale, wie die Größe der gefärbten Flächen und ihre Gestalt.

Die Gewichtung der Merkmale wird in Form von Regeln und unscharfen Funktionen in den Segmentierungsprozeß eingeführt. Das Ziel der Segmentierung kann durch die Bestimmung der unscharfen Mengen, welche die gewünschten Eigenschaften der Farbflächen wiedergeben, festgelegt werden. Dieser Ansatz hat den Vorteil, daß die Segmentierung gesteuert und zur Lösung eines bestimmten Problems angepaßt werden kann.

Die Segmentierung der Karte zeigte eine gute Qualität im Bezug auf spektrale und räumliche Eigenschaften der Farbflächen. Die Pixel werden in homogene, kompakte Flächen zusammengefaßt, welche einen guten Kontrast zum Hintergrund oder zu anderen Flächen zeigen. Auf diese Weise werden aus der Karte Objekte erkannt.

Die volle Erkennung der Karten wurde nicht erforscht, sondern auf linien- und flächenhafte Elemente eingeschränkt. Auch die volle semantische Interpretation der Ergebnisse war nicht nötig, sondern nur eine Annäherung der möglichen Lage der Objekte in der Karte. Die Genauigkeitsansprüche wurden deswegen nicht sehr hoch gesetzt.

Die Erkennung von Linien und vollen Flächen bereitete weniger Aufwand als die Erkennung von parallelen Linien und Flächen mit Mustern. Die leeren Räume zwischen Flächen mit Mustern und parallelen Linien sind ein wichtiger Bestandteil dieser Elemente und werden mit zwei unterschiedlichen Methoden analysiert: der Delaunay Triangulation und dem Wasserscheide Verfahren. Beide zeigten Vor- und Nachteile.

Das Verfahren der Triangulation ergibt eine optimierte Schätzung der Konturen flächenhafter Elemente, obwohl falsch klassifizierte Pixel in der Nähe der äußeren Grenzen große Fehler verursachen können. Das Problem verschärft sich, wenn das Bild viel Rauschen enthält.

Die Berechnung der Triangulation erfordert eine große Rechenleistung. Da eine Karte mit einer kleinen Pixelegröße abgetastet wird, ist die Anzahl der Punkte in der Triangulation sehr groß und die Berechnung aufwendig.

Das Wasserscheide-Verfahren ist in der Lage, die Zwischenräume innerhalb der Muster oder Linien zu erkennen, wird aber von zwei Nachteilen gekennzeichnet: Es scheitert an den Grenzen der Flächen mit Mustern und kann keine dünnen Flächen erkennen. Das Verfahren kann nur leere Räume innerhalb der Regionen finden. Mit der Berechnung der Hülle des Ergebnisses wird diese Beschränkung teilweise aufgehoben. Trotzdem bleibt das Problem der sehr dünnen, gemusterten Flächen. Parallele Linien konnten erfolgreich mit dem Wasserscheide-Verfahren erkannt werden, da in diesem Fall das Problem an den Rändern selten vorkommt.

Eine gemischte Lösung ist die beste Option zur Erkennung von Regionen mit Mustern. Zuerst werden die Daten mit dem Wasserscheide-Verfahren verarbeitet, wobei ein großer Anteil der Regionen erkannt wird. Dabei wird die Anzahl von Vordergrundpixeln reduziert und die Berechnung der Triangulation einfacher. Die übrigen Regionen werden mit dem Verfahren der Triangulation erkannt.

Die räumliche Begrenzung der Regionen mit Mustern ist zufriedenstellend. Da solche Muster systematisch erzeugt werden, findet man oft Situationen an den Grenzen der Regionen, aus denen die Musterelemente nicht vollständig, sondern nur Teile dieser Elemente vorhanden sind. Diese Situationen werden von diesem Ansatz sehr gut gemeistert, denn er basiert auf dem Abstand zwischen Mustern und nicht auf der Erkennung einzelner Muster als Grundelemente. Tatsächlich werden in dieser Region nicht nur die Pixel zusammengefaßt, die zu einem ganzen Musterelement gehören, sondern auch die Pixel unvollständiger Symbole.

Die Überlagerung von mehreren kartographischen Elementen verursacht Lücken in den Regionen mit Mustern. Diese Lücken werden gefüllt, so daß am Ende gut definierte Regionen entstehen. Größere Lücken werden aber nicht geschlossen, was dazu führen kann, daß die Konturen der Region falsch geschätzt werden. Probleme kommen auch vor, wenn zwei Landnutzungsklassen mit derselben Farbe nah beieinander liegen. Trotz dieser Fehler hat der Ansatz einen großen Vorteil: Er ist nicht von der Erkennung der Musterelemente abhängig. Deswegen ist er ziemlich leicht zu implementieren und an die Verhältnisse anderer Karten zu adaptieren. Der Ansatz liefert keine optimalen Ergebnisse, wenn zwei Regionen mit unterschiedlichen Mustern ähnliche Abstände zwischen Symbolen haben.

Die Grenzen der Klassen werden ermittelt, aber da sie nicht unbedingt aktuell sind, werden sie später als Fuzzy-Mengen dargestellt, so daß ihr Wachstum oder ihre Verkleinerung modelliert werden kann. Diese Modellierung ist entscheidend, denn es besteht keine Sicherheit über die Gültigkeit alter Karten. Dadurch wird die Anwendung komplexer Analysemethoden vermieden und eine Beschreibung des Karteninhaltes oder Satellitenbildes auf Rasterbasis erzeugt.

9.2 Inferenzschema

Das hier vorgestellte Modell eignet sich zur Klassifikation der Landnutzung aus Daten der Fernerkundung und Karten. Das Fuzzy-Klassifikationsschema bietet ein offenes Modell, das an Eigenschaften der Szene angepaßt werden kann, denn es basiert auf Regeln und unscharfen Funktionen.

Die Bestimmung der Regeln verlangt Wissen über die Szene und ist in der Analyse entscheidend. Das ist nicht neu in der Bildverarbeitung, denn ein Bild kann nicht nur „*mehr als tausend Worte sagen*“, sondern auch mehrere Bedeutungen haben, je nach dem Standpunkt und dem Interesse des Benutzers. Durch die Formulierung der Regeln entsteht die Freiheit, die Analyse zu steuern und menschliches Wissen explizit darzustellen.

Das Inferenzschema besitzt die Fähigkeit, unterschiedliche Daten aus demselben Gebiet in die Analyse einzubeziehen. Die Daten selbst werden nicht analysiert, sondern Hinweise die daraus abgeleitet werden. Diese Hinweise behalten ihre wichtigen Merkmale und bei der Bewertung der Regeln greift das System direkt auf sie zu. Die Dimension des Problems wird dabei erhöht, denn aus einem Datensatz können mehrere Hinweise abgeleitet werden.

Die Modularität des Verfahrens erlaubt auch die Anpassung des Inferenzschemas an andere Datensätze mit neuen Informationen. Es ist dabei unbedingt nötig, die Informationen im passenden Format darzustellen und die Regeln neu zu definieren. Das ist eine schwere Aufgabe wenn die Beziehungen zwischen den Informationen und dem gewünschten Ergebnis nicht ausreichend bekannt sind.

9.3 Modellierung der unscharfen Mengen

Fuzzy-Mengen werden benutzt um Information darzustellen und sie zu kombinieren. Unscharfe Logik eignet sich gut für diesen Zweck, denn sie ist in der Lage, Unsicherheit über die Hinweise auszudrücken und zu verarbeiten. Diese Funktionen können durch die Übertragung der menschlichen Erfahrung und Wissen oder auch automatisch mit Hilfe von Trainingsgebieten bestimmt werden.

Ein rechnerunterstütztes Verfahren zur Bestimmung der Fuzzy-Funktionen, basierend auf Prinzipien der genetischen Algorithmen, wurde implementiert. In diesem Verfahren werden mehrere mögliche Lösungen mit Hilfe von Operatoren, die eine genetische Entwicklung simulieren, wie *Paarung* und *Mutation*, optimiert.

Der Prozeß besteht aus zwei Phasen. In der ersten Phase erfahren die möglichen Lösungen überwiegend große Veränderungen bis eine Lösung mit relativ guter Qualität gefunden wird. Danach wird diese Lösung langsam verändert und optimiert. Die erste Phase ist sehr mühsam, deswegen sind Alternativen, die die erste Suche beschleunigen, besonders interessant.

Das Verfahren erlaubt es, Ergebnisse anderer Experimente in den Suchprozeß zu importieren und die Suche zu optimieren. Die Einbeziehung dieser ersten Schätzung hat den Vorteil, daß viele Iterationen der ersten Phase übersprungen werden, und die Suche von einem ziemlich guten Startpunkt ausgeht, was mit einer signifikanten Reduktion des Rechenaufwands verbunden ist. Das ist ein positiver Beitrag, da die Methode sehr viele Iterationen braucht um zum richtigen Ergebnis zu konvergieren. Die Speicherung dieser Funktionen (die optimale Lösungen eines anderen Problems) spielt die Rolle eines Gedächtnisses, das als Startpunkt benutzt werden kann, wenn das Schema mit einer neuen Situation konfrontiert wird.

9.4 Klassifikation

Das Ergebnis der Klassifikation muß unter Berücksichtigung der unterschiedlichen Eigenschaften der benutzten Datensätze analysiert werden. Das Bild ist durch große homogene Regionen und weniger verstreute Pixel gekennzeichnet. Im Vergleich zur Karte enthält es eine aktualisierte Klassifikation des Gebietes mit einer Verbesserung der Konturen der Siedlungen und einer neuen Kartierung der grünen Flächen. Im allgemeinen ist die räumliche Definition der Objekte besser als im Ergebnis einer multispektralen Klassifikation, aber es gibt einen großen Verlust an Details wenn man die multispektrale Klassifikation als Vorbild nimmt, weil viele Klassen zusammengefaßt werden.

Die Beschreibung der Objekte, gemäß der Karte, ist sehr deutlich. Deswegen wird das Ergebnis von den geometrischen Eigenschaften der Karte an manchen Stellen dominiert, zum Beispiel bei der Definition der Grenzen von Siedlungsgebieten. Obwohl ein Spielraum aufgrund der Fuzzifizierung der Daten entsteht, werden viele Pixel mit dem maximalen Wert (volle Zugehörigkeit) gefunden. Diese Werte überwiegen bei der Bewertung der beiden Quellen.

Wenn es auf der Karte keine Informationen gibt (leere Regionen) dann wird das Ergebnis von Informationen aus den Satellitenbildern bestimmt. In diesem Fall entstehen öfter kleine homogene Regionen. Diese besondere Eigenschaft wird aus der spektralen Klassifikation übertragen.

Die Qualität der Klassifikation hängt ab von den verfügbaren Daten und der Ableitung der Hinweise aus allen Datensätzen, die eine gute Qualität haben müssen. Das gilt sowohl für die raumbezogenen, als auch für die spektralen Datensätze.

9.5 Schlußbetrachtungen und Ausblick

Das Hauptziel dieser Arbeit ist die Klassifikation der Landnutzung einer Region, von der keine ausreichenden Datenerhebungen zur Verfügung stehen und bei der Fernerkundung und raumbezogene Daten oft die einzigen Informationsquellen sind.

Die Verarbeitung von raumbezogenen Daten als Lösung der Probleme in Entwicklungsländern wird ständig vorgeschlagen. Oft wird nur die erste Phase solcher Projekte durchgeführt: die Datenerhebung oder die Digitalisierung bereits bestehender Karten. Nach dem Aufbau eines Geo-Informationssystems wird an dem Projekt nicht mehr gearbeitet, weil nicht ausreichende finanzielle Mittel zur Verfügung stehen. Mit dem Ende des Projekts werden die finanziellen Mittel gekürzt, und es entsteht das Risiko, daß diese Daten veralten, ohne daß sie effektiv genutzt werden. Trotzdem enthalten die Daten einen signifikanten Anteil an Information über das Gebiet, denn manche Merkmale erfahren mit der Zeit nur kleine Veränderungen. Das Hauptproblem ist die Trennung der gültigen Informationen von den veralteten. Dieses Problem wird hier mit Hilfe der Fuzzy-Logik gelöst. Die Informationen werden als Fuzzy-Funktionen dargestellt, mit einem langsamen Übergang zwischen „totaler“ und „keiner“ Zugehörigkeit eines Elements zu einer Menge.

Die Sensoren der Fernerkundung erfassen die Erdoberfläche und ermitteln eine ikonische Darstellung ihrer spektralen Eigenschaften. Satellitenbilder werden systematisch aufgenommen und sind für Entwicklungsregionen besonders interessant. Diese Daten können heutzutage sehr leicht verarbeitet und analysiert werden, denn die Entwicklung der Informatik erlaubt es, Bildverarbeitungssysteme auf billigen Rechnern laufen zu lassen. Daten wurden das teuerste Glied solcher Systeme, denn Software und Hardware werden immer preiswerter.

Neue Sensoren werden entwickelt und ihre Daten stehen zur Verfügung. Je besser die Auflösung, desto mehr Einzelheiten werden im Bild zu sehen sein. Die Analyse muß deswegen nicht nur auf spektralen Merkmalen basieren, sondern andere Merkmale, wie räumliche Eigenschaften und Beziehungen, müssen berücksichtigt werden.

Methoden zur Klassifikation der Landnutzung durch die Analyse unterschiedlicher Datensätze und neuer Merkmale gewinnen an Bedeutung mit der technischen Entwicklung von Fernerkundung und der Verfügbarkeit von raumbezogenen Daten. Die vollständige Integration dieser beiden Technologien verlangt nach Datenverarbeitungsmethoden, welche die Besonderheiten jedes einzelnen Datensatzes berücksichtigen.

Die Verfügbarkeit von Experten, die diese Daten und Systeme wirkungsvoll benutzen können, spielt jetzt die entscheidende Rolle. Erfahrung und Wissen wird gefragt sein. In Entwicklungsländern wird versucht durch Kooperationsprogramm mit anderen Ländern diese Lücke zu füllen. Die Anwendung von Methoden der künstlichen Intelligenz und von wissensbasierten Systemen, die Fakten und Prozesse für die Lösung des Problems speichern, kann eine günstige Alternative werden. Methoden wie die genetischen Algorithmen können dabei helfen, das Wissen von Experten zu übertragen und an andere Probleme anzupassen.

Literatur

- Ajayi, G. [1993], *Untersuchung der Einsatzmöglichkeiten von Satellitenaufnahmen für die topographische Kartenherstellung in Entwicklungsländern*, Dissertation, Fachbereich Bauingenieur- und Vermessungswesen der Universität Hannover.
- Bähr, H.-P. [1984], Abschätzung einiger geometrischer Fehlerkomponenten bei der multispektralen Klassifizierung, *Bildmessungen und Luftbildwesen*, **52**(1):23–30.
- Bähr, H.-P. [1998], From data to inference: Examples for knowledge representation in image understanding, in ISPRS (Hrsg.), *Proceedings of ISPRS, Commission 3, Columbus, Ohio*, pp. 23–30, ISPRS.
- Bähr, H.-P. & Vögtle, T. (Hrsg.) [1998], *Digitale Bildverarbeitung*, Wichmann, Heidelberg, 3. Aufl.
- Baltsavias, E. & Hahn, M. [1999], Integration of image analysis and GIS, in ISPRS (Hrsg.), *International Archives of Photogrammetry and Remote Sensing*, Bd. 32: 7-4-3 W6, Valladolid, Spain, ISPRS.
- Bässmann, H. & Besslich, P. [1989], *Konturorientierte Verfahren in der digitalen Bildverarbeitung*, Springer-Verlag, Berlin.
- Binaghi, E., Madella, P., M.G., M. & Rampini, A. [1997], Fuzzy contextual classification of multisource remote sensing images, *IEEE Transactions on Geoscience and Remote Sensing*, **35**(2):326–339.
- Brandenberger, A. & Ghosh, S. [1990], Status of world topographic and cadastral mapping, in *World Cartography*, Bd. XX, United Nations.
- Bronsveld, K., Chutirattanapan, S., Pattanakanok, B., Suwanwerakamtorn, R. & Trakooldir, P. [1994], The use of knowledge in land use/land cover mapping from satellite images, *ITC Journal*, pp. 349–358.
- Brügelmann, R. [1998], *Rasterbasierte Methoden zur Gebäudeextraktion aus gescannten Karten*, Dissertation, Univ. Karlsruhe.
- Burger, R. [1992], Einbeziehung von Geoinformation und Geowissen in der Klassifikation von Satellitenbildern mit Hilfe eines evidenztheoretischen Ansatzes, in Günther, O. & Rieker, W. (Hrsg.), *Wissensbasierte Methoden zur Fernerkundung der Umwelt*, pp. 111–145, Karlsruhe, Wichmann Verlag.
- Celenk, M. [1990], A color clustering technique for image segmentation, *Computer Vision, Graphics and Image Processing*, **52**:145–170.
- Chevres, M. *et al.* [1981], The SPOT satellite remote sensing mission, *Photogrammetric Engineering and Remote Sensing*, **47**(8):1163–1171.
- Chuvieco, E. [1990], *Fundamentos de teledetección espacial*, RIALP S.A., Madrid, 1. Aufl.
- Doermann, D. [1998], An introduction to vectorization and segmentation, in Tombre, K. & Chhabra, A. (Hrsg.), *Graphic recognition: algorithms and systems*, Bd. 1389 von *Lecture Notes in Computer Science*, pp. 1–7, Berlin, Springer.
- Ehlers, M. [1995a], Grundlagen von Geographischen Informationssystemen – GIS, in Ehlers, M. (Hrsg.), *Aktuelle Forschungsberichte der Abteilung Geographische Informationssysteme/ Fernerkundung*, Bd. 20A, pp. 7–13, Vechta, Institut für Strukturforchung und Planung in agrarischen Intensivgebieten, Hochschule Vechta.
- Ehlers, M. [1995b], Integrating remote sensing and GIS for environmental monitoring and modeling: where are we?, *Geo Info System*, pp. 36–43.

- Ehlers, M. & Shi, W. [1995], Quality indicators for integrated GIS datasets, in Foundation, E. S. (Hrsg.), *European Science Foundation GIS/DATA Specialist Meeting on Data Quality*.
- Förstner, W. [1992], Fusion information in remote sensing, in ISPRS (Hrsg.), *XVII ISPRS Congress, Washington, D.C.*, Washington, D.C., ISPRS, Invited Paper.
- Fu, K. & Mui, J. [1981], A survey on image segmentation, *Pattern Recognition*, **13**:3–13.
- Glemser, M. & Fritsch, D. [1998], Data uncertainty in a hybrid GIS, in *GIS – Between Vision and Applications*, Bd. 32 von *ISPRS Commission IV Symposium*, pp. 180–187, Stuttgart.
- Goldberg, D. [1989], *Genetic algorithms in search, optimization and machine learning*, Addison-Wesley, New York.
- Haefner, H. [1997], Fernerkundung und Geographie – thematische, methodische und technische Perspektiven, *Erdkunde*, **41**(3):169–182.
- Hagg, W. [1998], *Merkmalsbasierte Klassifikation von SAR-Satellitenbilddaten*, Dissertation, Universität Karlsruhe.
- Hake, G. & Grünreich, D. [1994], *Kartographie*, Springer Verlag, Berlin, New York, 7. Aufl.
- Haralick, R. & Shapiro, L. [1997], *Computer and robot vision*, Bd. 1, Addison-Wesley Publishing Company Inc.
- Hunt, R. [1992], *Measuring colour*, Ellis Horwood Ltd., Chichester.
- Hutchinson, C. [1982], Techniques for combining Landsat and ancillary data for digital classification improvement, *Photogrammetric Engineering and Remote Sensing*, **48**(1):123–130.
- IBGE [1993], *Manual de normas, especificações e procedimentos técnicos para a carta internacional do mundo, ao milionésimo - CIM 1:1000000*, IBGE, Instituto Brasileiro de Geografia e Estatística.
- Jain, R., Kasturi, R. & Schnuck, B. [1995], *Machine vision*, Mc.Graw-Hill, Singapore.
- Jenson, S. & Domingue, J. [1988], Extraction of topographic structure from digital elevation data for geographic information system analysis, *PERS*, **54**(11):1593–1600.
- Johnson, R. & Wichern, D. [1998], *Applied multivariate statistical analysis*, Prentice Hall, New York.
- Kanade, T. [1980], Region segmentation: Signal vs. semantics, *Computer Graphic and Image Processing*, **13**:279–297.
- Klauer, R. [1993], *Untersuchungen zur Optimierung von Verfahren der automatisierten Digitalisierung von Flurkarten*, Dissertation, Wissenschaftliche Arbeiten der Fachrichtung Vermessungswesen der Universität Hannover.
- Konecny, G. [1990], Review of the latest technology in satellite mapping, *World Cartography*, **20**:23–30.
- Kunz, D., Vögtle, T. & Schilling, K. [1998], Integrierte Verarbeitung von Satellitenbild- und vektorierter Karteninformation, in Bähr, H. & Vögtle, T. (Hrsg.), *Digitale Bildverarbeitung*, pp. 220–242, Wichmann, Heidelberg, 3. Aufl.
- Lauterbach, B. [1993], *Extraktion attributierter Strukturprimitiven aus kartographischen Vorlagen auf der Basis einer multispektralen Bildanalyse*, Dissertation, Bremer Beiträge zur Geographie und Raumplanung, Universität Bremen, Nr. 27.
- Lee, T., Richards, J. & Swain, P. [1987], Probabilistic and evidential approaches for multisource data analysis, *IEEE Transactions on Geoscience and Remote Sensing*, **25**(3):283–293.

- Levine, M. [1985], *Vision in man and machine*, Mc.Graw-Hill, New York.
- Lillesand, T. & Kiefer, R. [1994], *Remote sensing and image interpretation*, John Wiley and Sons, Inc., New York.
- Lin, W. [1994], *Ein Beitrag zur kartographischen Mustererkennung mittels Methoden der künstlichen Intelligenz*, Dissertation, Deutsche Geodätische Kommission, Reihe C, Nr. 419.
- Mark, D. [1983], Automated detection of drainage networks for digital elevation models, in *Proceedings of Auto-Carto 6*, Bd. 2, Ottawa Ontario, Canada.
- Moon, W. [1990], Integration of geophysical and geological data using evidential belief function, *IEEE Transactions on Geoscience and Remote Sensing*, **28**(4):711–719.
- Morrison, J. [1980], Computer technology and cartographic change, in Taylor, D. (Hrsg.), *The computer in contemporary cartography*, Kap. 2, John Wiley Press and Sons Ltd., London.
- Ohta, Y., Kanade, T. & Toshiyuki, S. [1980], Color information for region segmentation, *Computer Graphics and Image Processing*, **13**:222–241.
- Okabe, A., Boots, B. & Sugihara, K. [1992], *Spatial tessellations: concepts and application of Voronoi diagrams*, John Wiley Press, Chichester.
- Peddle, D. [1995], Knowledge formulation for supervised evidential classification, *Photogrammetric Engineering and Remote Sensing*, **61**(4):409–417.
- Perkal, J. [1956], On epsilon length, *Bulletin de l'Academie Polonaise des Sciences*, **4**:399–403.
- Pham, D. & Liu, X. [1995], *Neural networks for identification, prediction and control*, Springer-Verlag, Heidelberg.
- Pratt, W. [1974], *Digital image processing*, John Wiley and Sons, New York.
- Quint, F. [1997], *Zur automatischen kartengestützten Auswertung monokularer Luftbilder*, Dissertation, Univ. Karlsruhe.
- Rosenfeld, A. & de la Torre, P. [1988], Histogram concavity analysis as an aid in threshold selection, *IEEE Transactions on System, Man and Cybernetics*, **SMC-13**:62–66.
- Sahoo, P., Soltani, S., Wong, A. & Chen, Y. [1988], A survey of thresholding techniques, *Computer Vision, Graphics, and Image Processing*, **41**:233–260.
- Scheuring, R. [1995], *Zur Qualität der Basisdaten von Landinformationssystemen*, Dissertation, Univ. der Bundeswehr, München.
- Segl, K. [1996], *Integration von Form- und Spektralmerkmalen durch künstliche neuronale Netze bei der Satellitenbildklassifizierung*, Dissertation, Universität Karlsruhe.
- Sezan, M. [1990], A peak detection algorithm and its application to histogram-based image data reduction, *Computer Vision, Graphics, and Image Processing*, **49**:36–51.
- Shafer, G. [1976], *A mathematical theory of evidence*, Princeton University Press, Princetown, New Jersey, 297 Seiten.
- Shea, K. [1991], Design considerations for an artificially intelligent system, in Buttenfield, B. & McMaster, R. (Hrsg.), *map generalization: making rules for knowledge representation*, Kap. 1, S. 245, John Wiley and Sons, New York, 1. Aufl.

- Silva Centeno, J. [1998], Segmentation of thematic maps using colour and spatial attributes, in Tombre, K. & Chhabra, A. (Hrsg.), *Graphic recognition: algorithms and systems*, Bd. 1389 von *Lecture Notes in computer science*, pp. 221–230, Berlin, Springer.
- Silva Centeno, J., Kishi, R. & Bähr, H.-P. [1999], Recognition of buildings using scanned maps and laser scanner altitude data, *DGPF, Photogrammetrie, Fernerkundung, Geoinformation*, **99**(1):19–28.
- Solberg, A., Taxt, T. & Jain, A. [1996], A markov random field model for classification of multisource satellite imagery, *IEEE Transactions on Geoscience and Remote Sensing*, **34**(1):100–112.
- Stanley, L. & Zhu, Q. [1995], *Fuzzy and evidence reasoning*, Physica-Verlag, Heidelberg, 360 Seiten.
- Swain, P. & Davis, S. [1978], *Remote sensing: The quantitative approach*, Mc.Graw-Hill, New York.
- Ton, J., Stickeln, J. & Jain, A. [1991], Knowledge-based segmentation of Landsat images, *IEEE Transactions on Geoscience and Remote Sensing*, **29**(2):222–232.
- Vincent, L. & Soille, P. [1991], Watershed in digital spaces: An efficient algorithm on immersion simulation, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **13**(6):583–589.
- Vögtle [1998], Ladung und Ladungsverschiebung in Halbleiterelementen, in Bähr, H. & Vögtle, T. (Hrsg.), *Digitale Bildverarbeitung*, pp. 7–10, Wichmann, Heidelberg, 3. Aufl.
- Vögtle, T. & Schilling, K. [1995], Wissensbasierte Extraktion von Siedlungsbereichen in der Satellitenbildanalyse, *ZPF – Zeitschrift für Photogrammetrie und Fernerkundung*, **5**:199–207.
- Wahl, F. [1984], *Digitale Bildsignalverarbeitung*, Springer-Verlag, Berlin.
- Wald, L. [1999], Definitions and terms of reference in data fusion, in ISPRS (Hrsg.), *International Archives of Photogrammetry and Remote Sensing*, Bd. 32: 7-4-3 W6, Valladolid, Spain, ISPRS.
- Wang, F. [1990], Improving remote sensing image analysis through fuzzy information representation, *PERS*, **56**(8):1163–1169.
- Weindorf, M. [1994], *Segmentation von Siedlungsflächen aus Satellitenbilddaten*, Diploma thesis, Universität Karlsruhe.
- Whatmough, R. [1991], Threshold selection using the exponential hull, *CVGIP: Graphical Models and Image Processing*, **53**(6):592–600.
- Yang, J. [1989], *Automatische Digitalisierung von Deckfolien der deutschen Grundkarte 1:5000 - Bodenkarte*, Dissertation, Wissenschaftliche Arbeiten der Fachrichtung Vermessungswesen der Universität Hannover.
- Zhu, A.-X. [1997], Measuring uncertainty in class assignment for natural resource maps under fuzzy logic, *Photogrammetric Engineering and Remote Sensing*, **63**(10):1195–1202.

4.2 Erkennung von Landnutzungsgruppen

Nach Hake & Grünreich [1994] „gelten als kartographische Gestaltungsmittel die Grundelemente Punkt, Linie und Fläche sowie die zusammengesetzten Zeichen Signatur, Diagramm, Halbton und Schrift“. Die in einem Satellitenbild sichtbaren Landnutzungsklassen werden meist als Linien und Flächen in einer Karte dargestellt. Zu den linienhaften kartographischen Elementen in einer IBGE Karte (1:50.000) gehören Objekte wie Flüsse und Straßen. Kleine Flüsse werden als durchgehende/massive oder strichlierte blaue Linie dargestellt. Eine Straße ist als eine einzige Linie oder durch zwei parallele Linien in der Karte dargestellt. Andere Klassen wie Siedlungsgebiete, breitere Wasserflächen (größer als 4 mm im jeweiligen Maßstab, IBGE [1993]), Sumpfgebiete, Wälder oder Obstplantagen, werden als Flächen dargestellt. Dabei kennt man zwei Arten der Darstellung: geschlossene Flächen, wie zum Beispiel Wälder, und flächenhafte Signaturen, wie zum Beispiel Siedlungsgebiete.

4.2.1 Erfassung von Linien

Unter dem Begriff „Linien“ versteht man „alle nicht unterbrochenen Striche, die eine *Lage* angeben“ (Hake & Grünreich [1994]). Ihre graphische Variation ist hinsichtlich der Farbe oder der Strichbreite möglich. Linien werden benutzt, um Lageangaben über die Abgrenzung von flächenhaften Objekten (Umfang) anzugeben oder bei sehr schmalen Objekten zur Darstellung des Objekts selbst. In diesem Fall wird das Objekt durch eine Mittellinie dargestellt. Ein Spezialfall ist die Verwendung von zwei parallelen Linien, z.B. zur Repräsentation einer Straße; dabei ist der leere Zwischenraum ein wichtiger Bestandteil ihrer Darstellung.

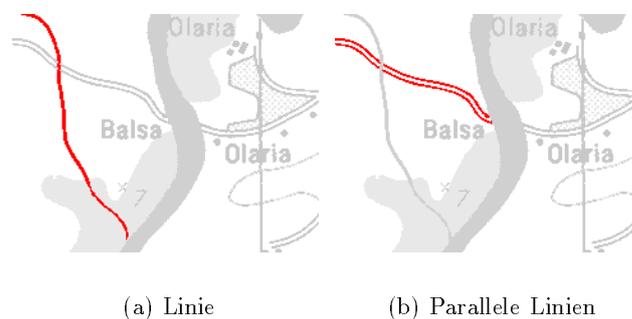


Abbildung 28: Zwei mögliche Darstellungen linienhafter kartographischer Elemente.

Die Erfassung von Linien aus binären Bildern ist im wesentlichen ein Problem der Vektorisierung und wird in Bässmann & Besslich [1989] sehr detailliert beschrieben. Ein Großteil der bekannten Ansätze fängt mit der Merkmalsextraktion in einer low-level Ebene (*Signalebene*) an und versucht dann später Linien anzunähern. Zur Schätzung der Mittelachse der Linien bieten sich zwei Möglichkeiten an. Die erste ist die iterative Verdünnung der erkannten Regionen bis zu dem Punkt, an dem die Flächen zu einer Kette von einzelnen Pixeln reduziert werden. Diese Punkte werden zu Linien verkettet (Linienverfolgung). Der alternative Ansatz berechnet die Mittelachse aus der Konturen der Regionen. Benachbarte Liniensegmente werden später zu größeren Linien aufgrund ihrer Abstände und Richtungen zusammengefaßt (Bässmann & Besslich [1989]).

4.2.2 Erfassung von Flächen

Als flächenhafte Elemente versteht man nicht nur die Regionen einer Karte, die in Form einer kompakten Menge gleichfarbiger Pixel vorliegen, sondern auch die Regionen, die durch Muster und Schraffuren