

Universität Osnabrück

Fachbereich Mathematik und Informatik

Lehrstuhl für Geoinformatik

**Auswirkung des Rauschens
und Rauschen vermindender Maßnahmen
auf ein fernerkundliches Segmentierungsverfahren**

Karl Gerhards

Inauguraldissertation zur Erlangung des akademischen Grades

Doktor der Naturwissenschaften (Dr. rer. nat.)

Danksagung

Herzlich danken möchte ich meinem Doktorvater Prof. Dr. Manfred Ehlers für seine Bereitschaft, dieses Projekt zu betreuen. Ihm verdanke ich Ermutigung und zahlreiche hilfreiche Hinweise besonders in der Endphase dieser Arbeit.

Hervorzuheben sind alle diejenigen, die den Mut hatten, mir selbst entwickelte Software zu überlassen: Prof. Dr. Herbert Jahn und Dr. Winfried Halle vom Institut für Luft- und Raumfahrt der TU Berlin für Code zum Graphennetzwerk, Dr. Jochen Bückner vom Institut für Informationsverarbeitung der Universität Hannover für die Überlassung des Programms geoAIDA, Dr. James C. Tilton von der NASA für das hseg-Programm sowie die vielen ungenannten Programmierer von Open Source- und Freeware-Programmen.

Danken möchte ich auch den Mitarbeitern des Institutes für Geoinformatik, insbesondere Herrn Ronald Janowsky und Herrn Dr. Ulrich Michel für die Überlassung von Daten.

Ein besonderer Dank gilt schließlich meiner Frau Barbara für den Freiraum, den sie mir gegeben hat, um dieses Projekt zu verwirklichen.

Abstract

A huge number of smoothing algorithms exist for the decrease of the noise of very high resolution satellite images. The effect of different low pass and edge preserving filters on the behavior of an object oriented segmentation procedure is examined based on two synthetic gray value images and an IKONOS image.

For noise measurement, a modified, originally from Baltasvias et al. [2001] proposed procedure provides evidence in which for each grey value only the standard deviations of the most uniform areas are taken into consideration. Nevertheless, in comparison with synthetically noised images appears that in this case noise in the image is systematically underestimated around two times.

Simple filters like mean filter and from mean filter derived filters make the precision of the object recognition dramatically worse, edge preserving filters can be advantageous with noisier data.

The modified EPOS filter, originally introduced by Haag and Sties [1994, 1996] proves to be the best filter that is to be used with claims to exact segment borders in the pixel domain. Besides, it can be steered with only one parameter. By this edge preserving filter over all image parameters, as standard deviation or histogram are only slightly influenced.

Zusammenfassung

Zur Verminderung des Rauschens sehr hochauflösender Satellitenbilder existieren eine Vielzahl von Glättungsalgorithmen. Die Wirkung verschiedener Tiefpaß- und kantenerhaltender Filter auf das Verhalten eines objektorientierten Segmentierungsverfahrens wird anhand zweier synthetischer Grauwertbilder und einer IKONOS-Aufnahme untersucht.

Als Rauschmaß hat sich ein modifiziertes, ursprünglich von Baltsavias et al. [2001] vorgeschlagenes Verfahren bewährt, in dem je Grauwert nur die Standardabweichungen der gleichförmigsten Gebiete berücksichtigt werden. In Vergleich mit synthetisch verrauschten Bildern zeigt sich jedoch, daß auf diese Weise das Rauschen im Bild systematisch um fast den Faktor zwei unterschätzt wird.

Einfache Filter wie Mittelwertfilter und davon abgeleitete Verfahren verschlechtern die Präzision der Objekterkennung dramatisch, kantenerhaltende Filter können bei stärker verrauschten Daten vorteilhaft sein.

Als bester Filter, der bei Ansprüchen an präzise Segmentgrenzen im Pixelbereich sinnvoll einzusetzen ist und dabei mit nur einem Parameter gesteuert werden kann, erweist sich der modifizierte EPOS-Filter, ursprünglich vorgestellt von Haag und Sties [1994, 1996]. Die generellen Bildparameter, wie Standardabweichung oder Histogramm werden durch diesen kantenerhaltenden Filter nur unwesentlich beeinflusst.

Schlagworte

Rauschen, Mittelwertfilter, Medianfilter, Bilateraler Filter, eCognition, multiresolution segmentation, noise estimation, mean filter,

Inhalt

1	Einleitung	7
1.1	Motivation.....	7
1.2	Ziele	9
1.3	Aufbau	10
2	Grundlagen	12
2.1	Definitionen.....	12
2.1.1	Bild.....	12
2.1.2	Kanal	13
2.1.3	Filter	13
2.1.4	Kante	13
2.1.5	Partition.....	13
2.1.6	Ground Truth.....	14
2.2	Verfahren.....	14
2.2.1	Abtastung.....	14
2.2.2	Rand.....	15
2.2.3	Faltung	15
2.3	Segmentierung	17
2.4	Verwendete Hard- und Software.....	18
3	Bildrauschen.....	19
3.1	Begriffe.....	19
3.1.1	Signal-Rausch-Verhältnis	19
3.1.2	Quantisierungsrauschen.....	22
3.1.3	weißes Rauschen.....	22
3.1.4	additives Rauschen.....	22
3.1.5	multiplikatives Rauschen.....	22
3.2	Ursachen	23
3.3	Addition des Rauschens.....	24
3.3.1	Rauschen zufügen.....	24
3.3.2	Rauschen messen.....	25
4	Segmentierung.....	28
4.1	Objektextraktion	28
4.2	Segmentierungsverfahren.....	28
4.2.1	Kantendetektierende Verfahren.....	29
4.2.2	Pixelorientierte Verfahren.....	29
4.2.3	Regionenorientierte Verfahren	30
4.3	eCognition.....	31
4.3.1	Algorithmus	31
4.3.2	Einstellbare Parameter.....	32
4.3.3	Benutzeroberfläche	34
4.4	Untersuchung von Bildern mit konstantem Grauwert	36
4.4.1	Segmentierung der weißen Fläche.....	37
4.4.2	Einfluß des Rauschens.....	39
4.4.3	Segmentierung eines Grauwertverlaufs.....	42

5	Verwendete Filter	45
5.1	Lineare Filter	45
5.1.1	Mittelwertfilter	45
5.1.2	Kantenextraktion	46
5.1.3	Kantenverstärkung.....	47
5.2	Medianfilter	48
5.3	Filter mit veränderlicher Filterfunktion	49
5.3.1	Anisotrope Diffusion	50
5.3.2	Graphen Netzwerk Filter.....	52
5.3.3	Edge Preserving Optimized Speckle Filter	55
5.3.4	Bilateraler Filter.....	58
6	Filterbewertung	61
6.1	Übersicht	61
6.1.1	Ground Truth.....	61
6.1.2	Parametereinstellung.....	63
6.1.3	Segmentierung und Klassifikation der Ground Truth	64
6.1.4	Ergebnisse	68
6.2	Schwache Kontraste	73
6.2.1	Segmentierung der erzeugten Bilder.....	77
6.2.2	Verrauschen und Filtern.....	78
6.2.3	Parameterwahl.....	81
6.3	Ergebnisse	82
6.3.1	Mittelwertfilter	83
6.3.2	Medianfilter.....	84
6.3.3	Filter mit anisotroper Diffusion	85
6.3.4	Kantenerhaltende Filter	86
6.3.5	Bilateraler Filter.....	87
7	Anwendung auf eine reale Szene	90
7.1	Bildanalyse	91
7.1.1	Topographie.....	91
7.1.2	Histogramm.....	92
7.1.3	Rauschen.....	92
7.2	Segmentierung	96
7.2.1	Kriterien für die Filterbewertung	96
7.2.2	Granularität.....	97
7.2.3	Segmentkanten	99
8	Zusammenfassung, Wertung und Ausblick	106
9	Literatur	109

1 Einleitung

Die Anforderungen an eine Segmentierung bewegen sich in dem Spannungsfeld, das durch die drei divergenten Ziele Präzision, Geschwindigkeit und minimierten Speicherbedarf (aufgrund der Größe der Bilddateien) aufgespannt wird.

Schnelle Programme versuchen, aus den vorhandenen Bilddaten durch Verringerung der Auflösung kleinere Dateien zu erhalten, auf denen Farb- oder Grauwert-ähnliche Gebiete als zusammenhängend gedeutet werden. Meistens ist eine Vorverarbeitungsstufe voran gesetzt, um Bildrauschen zu entfernen oder Kanten zu schärfen.

Andere Programme, besonders solche, die Texturen erkennen möchten, leiten aus den Bilddaten eine Vielzahl weiterer Daten ab, die Auskunft geben über lokale und globale Frequenzspektren, Richtungen und Gradienten. Dadurch werden die auszuwertenden Daten vervielfacht, in einfachen Fällen um Faktoren im zweistelligen Bereich, häufig auch in der Größenklasse $O(n^2)$. Damit steigen bei den heutigen Auflösungen die Datenmengen in nicht mehr zu bewältigende Dimensionen.

Lange bekannt sind die Probleme der unüberwachten gegenüber der überwachten Klassifikation: im ersteren Fall erhält man zwar für jeden Punkt eine Klasse, eine Vielzahl davon ist jedoch auf die Zielklassen, die als Ergebnis gefunden werden sollen, nicht abzubilden, im zweiten Fall bleibt eine Vielzahl von Punkten unklassifiziert.

Der beste Weg, aus einer Vielzahl von gefundenen Segmenten eine Zielklasse zu bilden, besteht in einer darauf aufbauenden wissensgesteuerten Klassifizierung. Diesen Ansatz verfolgen Programme wie GeoAida¹ und eCognition®². Die ursprüngliche Anzahl der Ausgangssegmente ist dabei offenbar unerheblich, solange keine Klassengrenzen verwischt werden und die Agglomeration automatisch erfolgen kann. Durch die Zusammenfassung multisensorieller Daten kann ein neuer Ansatz verfolgt werden: Aus Objekten, die durch ein Sensorsystem identifiziert werden können, kann eine Hypothese über Nachbarobjekte gebildet werden, der im weiteren Verlauf eine bestimmte Wahrscheinlichkeit zugeordnet wird. In einem folgenden Bewertungsschritt wird schließlich die wahrscheinlichste Hypothese als richtig angenommen. Diese Bewertungen sollen offengelegt und dokumentiert werden, damit bei Fehlinterpretationen die Ursache gefunden werden kann und die Erkennungsleistung verbessert werden kann. Alternativ ist auch ein Lernen nach dem neuronalen Modell möglich.

1.1 Motivation

Mit der Einführung hochauflösender Sensoren mit Bodenauflösungen im Dezimeterbereich ergeben sich neue Aufgaben in der maschinellen Auswertung von Fernerkundungsaufnahmen. Waren für Landsat TM³-Aufnahmen noch Klassifikationsverfahren mit aus Maximum-Likelihood-Klassifizierern abgeleiteten Verfahren sinnvoll und ausreichend, so erfordert das gesteigerte Auflösungsvermögen neuer Sensorsysteme objektorientierte und wissensbasierte Verfahren zur Klassifizierung der Bilddaten. Eine Übersicht über die aktuellen Sensorsysteme ist zu finden bei Ehlers [2004, 2005].

¹<http://tnt.uni-hannover.de>

²<http://www.definiens-imaging.com>

³Thematic Mapper

Die vorliegende Arbeit betritt Neuland in der Untersuchung der Folgen der Vorverarbeitung digitaler Bilddaten für ein nachfolgendes Segmentierungs- und Klassifikationsverfahren. Systematische Untersuchungen zu diesem Thema fehlen bislang.

Bisher wurde stillschweigend vorausgesetzt, daß Bilder, die visuell gut aussehen und gut differenzierbar sind, auch für das maschinelle Sehen gut geeignet wären. Andererseits kann der menschliche Betrachter Grauwertunterschiede am besten bei mittleren Helligkeiten unterscheiden. Für die Maschine ist es egal, ob Subtrahend und Diminuend groß oder klein sind, wenn die Differenz die gleiche ist. So gleicht beispielsweise der Wallis-Filter Helligkeitsunterschiede im Bild aus, fügt aber andererseits Rauschen hinzu. Dies erspart dem menschlichen Betrachter das Drehen an der Helligkeitssteuerung, ist für eine maschinelle Bildverarbeitung wegen des Informationsverlustes durch das zugefügte Rauschen aber nachteilig.



Abbildung 1.1

San Francisco aus 450 km Höhe, Satellit IKONOS. Das Bild wurde von der Firma Space Imaging ins Internet gestellt.

Diese Aufnahme (Abbildung 1.1) von San Francisco wird von der Firma Space Imaging⁴ ins Internet gestellt. Sie soll die Leistungsfähigkeit der Sensoren des Satelliten IKONOS darstellen, mit denen dieses Bild aus 450 km Höhe aufgenommen wurde. Das Bild ist wolkenfrei und von beeindruckender visueller Qualität.

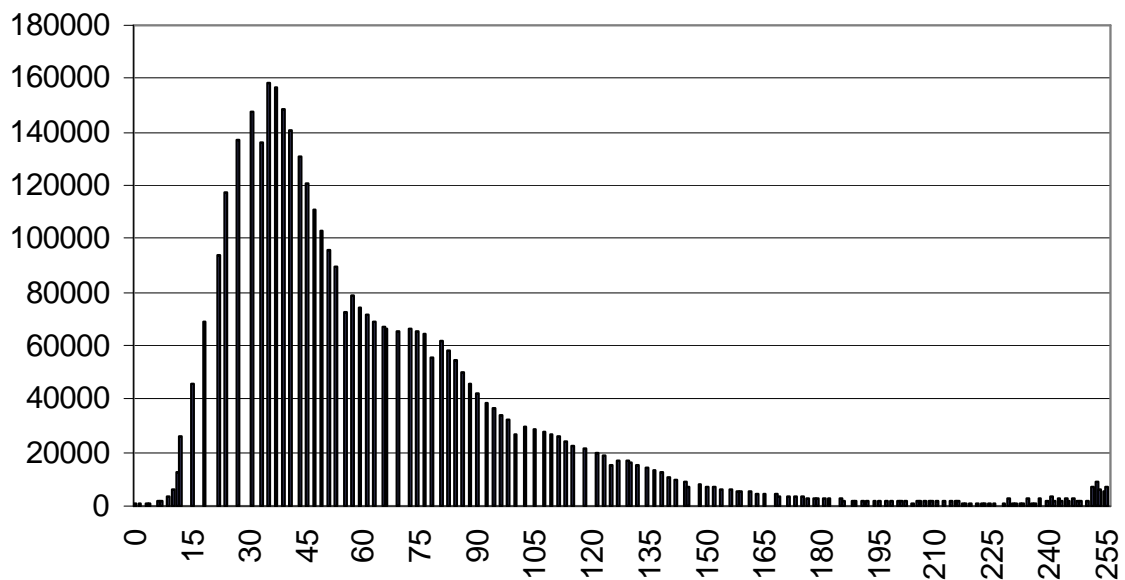


Abbildung 1.2

Histogramm zu Abbildung 1.1 von Seite 8. Das Histogramm zeigt die Spuren der Bildvorverarbeitung und des Resamplings.

Betrachtet man das zugehörige Histogramm (Abbildung 1.2), so fällt auf, daß von den 4096 möglichen Werten radiometrischer Auflösung, mit denen geworben wird, in diesem Bild lediglich 151 verwendet wurden. So stellt sich die Frage, ob diese Datenreduktion auf weniger als 4% der möglichen Werte auch zu einem merklichen Informationsverlust führt. Verbessern Verfahren, die dazu entwickelt wurden, dem menschlichen Auswerter die Arbeit zu erleichtern, auch die Erkennungsleistung eines maschinellen Systems?

1.2 Ziele

In dieser Arbeit soll untersucht werden, ob verschiedene Filter, die zur Verminderung des Rauschens benutzt werden, die Objekterkennungsleistung von eCognition verbessern. eCognition wurde als Analyseprogramm verwendet, weil es nach einer vergleichenden Untersuchung von Neubert und Meinel [2003] das für die Segmentierung hochauflösender Satellitendaten bestgeeignete Programm war und seit jener Untersuchung weiter entwickelt wurde. Der Segmentierungsalgorithmus von eCo-

⁴<http://www.space-imaging.com>

gnition ist in wesentlichen Teilen nicht dokumentiert. Deshalb kann das Verhalten des Programms auf Änderungen der Eingabedaten nicht theoretisch vorhergesagt werden.

Für dieses Ziel ist es erforderlich, zunächst das Rauschen in seinen verschiedenen Komponenten zu analysieren. Dann soll ein Verfahren entwickelt werden, Rauschen und die Wirkung rauschvermindernder Maßnahmen in Bezug auf das Rauschen zu messen. Erwünscht ist ein Verfahren, das für die infrage kommenden Bilder vollautomatisch arbeitet.

Das verwendete Segmentierungsprogramm eCognition bietet zahlreiche Parameter, die die Segmentierung beeinflussen, in ihren Auswirkungen aber kaum dokumentiert sind. Ein Ziel der Untersuchungen ist daher eine Darstellung der Auswirkungen von Änderungen der Parametereinstellungen auf das Segmentierungsergebnis.

In kommerziellen Bildverarbeitungsprogrammen werden eine Vielzahl von Algorithmen angeboten, die die Qualität von Bildern verbessern sollen. Es kann nicht davon ausgegangen werden, daß ein Segmentierungsprogramm die gleichen Anforderungen an die Bilddaten stellt wie ein menschlicher Untersucher. Mit dieser Arbeit soll deshalb dargestellt werden, welche Anforderungen das Segmentierungsprogramm an die Bilddaten stellt, um ein erwünschtes Ergebnis liefern zu können.

Hauptziel ist damit die systematische Analyse der Auswirkungen verschiedener Filter in der Bildvorverarbeitung sowohl auf die Objekterkennungsleistung von eCognition im allgemeinen wie auch auf die Präzision der Segmentkanten im speziellen.

1.3 Aufbau

Um die Übersichtlichkeit zu verbessern, wurden theoretische Überlegungen den folgenden praktischen Ausführungen möglichst unmittelbar vorausgestellt. Deshalb wechseln sich theoretische und praktische Teile in der Arbeit ab.

Die Arbeit hat folgenden Aufbau:

Im nächsten Kapitel werden die die wesentlichen Grundlagen bereitgestellt. Hier wird auch Auskunft gegeben über die verwendete Hard- und Software.

Das dritte Kapitel beschäftigt sich mit dem Bildrauschen, den Definitionen, dem künstlichen Verrauschen von Bildern und dem Problem der Messung.

Das vierte Kapitel über Segmentierung gliedert sich in einen theoretischen Teil und einen praktischen Teil. Zunächst werden die verschiedenen Segmentierungsverfahren kurz dargestellt und die Möglichkeiten zur Objektextraktion aus Kanten- und Segmentinformationen beschrieben. Anschließend wird der von eCognition verwendete Algorithmus vorgestellt, soweit er veröffentlicht wurde. Es folgen die für das Verständnis dieser Arbeit notwendige Angaben über die verwendeten Parameter zur Steuerung der Segmentierung und der Benutzeroberfläche. Im praktischen Teil werden die Ergebnisse vorgestellt, die durch die Variation der Segmentierungsparameter an Bildern ohne Objekte erzielt wurden. Diese Voruntersuchungen geben einen ersten Aufschluß über das Segmentierungsverhalten von eCognition in homogenen Gebieten.

Im fünften Kapitel werden die untersuchten Glättungsfilter dargestellt. Dies sind zunächst zwei weit verbreitete, die in jedem Bildverarbeitungsprogramm enthalten sind wie der Mittelwertfilter und der Medianfilter. Dann folgen mit den Anisotropic Diffusion Filtern, dem Graphen Netzwerk Filter und

dem Edge Preserving Optimized Speckle Filter verschiedene Filter, die Kanteninformationen direkt berücksichtigen, indem ein Homogenitätskriterium verwendet wird und der neue Wert eines Pixels nur aus solchen Gebieten in der Nachbarschaft des Pixels berechnet wird, die dieses Homogenitätskriterium nicht verletzen.

Auch der zuletzt beschriebene Filter, der Bilaterale Filter will Kanten erhalten und glättet Bilder nur in homogenen Regionen. Darüber hinaus kann er mehrere Bildebenen berücksichtigen und eignet sich deshalb besonders für Multispektralaufnahmen.

Nach diesen Vorbereitungen folgen im sechsten Kapitel zwei Versuche an synthetisch erzeugten Grauwertbildern. Im ersten Versuch geht es darum, einen Überblick zu gewinnen über das grundsätzliche Verhalten bei verschiedenen Objektgrößen und Grauwertsprüngen in Kantenbereichen. Um den Hintergrund des Bildes zu erzeugen, werden die im dritten Kapitel gewonnenen Erkenntnisse genutzt. Im zweiten Versuch werden die Filterwirkungen im Grenzbereich der spektralen Auflösung differenziert betrachtet.

Im siebten Kapitel werden die Untersuchungen an einem realen Bild fortgesetzt.

Das achte Kapitel faßt die erhaltenen Ergebnisse zusammen und zeigt weiteren Forschungsbedarf auf.

Das Literaturverzeichnis schließt die Arbeit ab.

2 Grundlagen

In diesem Kapitel werden einige grundlegende Definitionen gegeben und allgemeine Algorithmen beschrieben.

Die Bildbearbeitung ist jene Teildisziplin der digitalen Bildverarbeitung, welche Abbildungen von Bildern auf Bildern zum Gegenstand hat. Mit dem Oberbegriff digitale Bildverarbeitung wird im deutschsprachigen Bereich im allgemeinen jenes Fachgebiet bezeichnet, das im englischsprachigen Bereich als Computer Vision bezeichnet wird. Es geht hierbei allgemein um die rechnerbasierte Bearbeitung, Auswertung, Klassifizierung, Interpretation von bildhaften Informationen. Dabei ist der Bereich der Bildbearbeitung allgemein für eine erste Phase der digitalen Bildverarbeitung zuständig, in welcher „schönere Bilder“, besser analysierbare Bilder oder Bilder mit hervorgehobenen Bildmerkmalen produziert werden (Klette und Zamperoni [1995]).

Mit welchen Methoden ein Bild sinnvollerweise behandelt wird, hängt dementsprechend sowohl von den Eingabedaten als auch von den sich anschließenden Auswerteverfahren ab. Radar- und Satellitendaten müssen anders aufbereitet werden, als Daten hochauflösender Satellitenkameras. So kann es aber auch nicht verwundern, daß ein maschinelles Auswerteverfahren andere Ansprüche an die vorangegangene Bildbearbeitung stellt, als ein menschlicher Betrachter.

2.1 Definitionen

Je weiter verbreitet ein Begriff ist, desto unschärfer ist seine Definition. Für die wichtigsten Begriffe wird in diesem Abschnitt definiert, wie sie in dieser Arbeit verstanden werden.

2.1.1 Bild

Ein Bild ist eine zweidimensionale Anordnung f von skalaren oder vektoriellen Meßwerten $f(x, y)$. Bei einem fotografischen Bild erfolgt der „Meßprozeß“ beispielsweise optisch-analog mit dem Resultat unterschiedlicher Farbintensitäten in den Bildpunkten. In der digitalen Bildverarbeitung werden Bilder als Eingabeinformation verwendet, wobei für einen Bildpunkt (x, y) ein Bildwert $f(x, y)$ einer visuell zu erfassenden Eigenschaft entspricht, welche durch spezielle bildgebende Techniken (Mikroskopie, Lasersensor, Infrarotmeßtechnik, CCD-Sensoren, Computertomographie u.a.m.) realisiert wird. Ein Bild f ist eine Funktion, die gewissen Punkten (x, y) der reellen Ebene Bildwerte $f(x, y)$ zuordnet. Die Menge aller Punkte (x, y) , für die f definiert ist, wird Bildträger genannt. Falls der Bildträger ein Gebiet der reellen Ebene ist, so wird von analogen Bildern gesprochen. Falls der Bildträger ein Bildraster von $M \times N$ Bildpunkten ist, so spricht man von diskreten Bildern. Bei einer Betrachtung der Bilder als informationelle Einheit (Signaltheorie) werden die Sprechweisen kontinuierliches Bildsignal und diskontinuierliches Bildsignal bevorzugt. Bei einem digitalen Bild wird ein Bildraster und ein endlicher Wertebereich für die Bildwerte $f(x, y)$ vorausgesetzt (nach Klette und Zamperoni [1995]). Der Begriff Bildfunktion wird synonym gebraucht.

2.1.2 Kanal

Unter Kanal wird in diesem Zusammenhang eine Ebene eines vektoriiellen Bildes verstanden. So haben RGB⁵-Bilder drei Bildebenen, die auch als Kanäle bezeichnet werden. Dies ist ein Unterschied zu Bildbearbeitungsprogrammen, die sowohl Ebenen wie auch Kanäle anders definieren.

2.1.3 Filter

Ein Filter ist eine Abbildung eines Eingabebildes f auf ein Ausgabebild g . Dies ist die allgemeinste Definition. Im engeren Sinn betrachtet man die Bilddaten in ihrer ikonischen Form, d. h. im Unterschied zur Signalprozessierung in der Nachrichtentechnik wird die Bildgeometrie besonders berücksichtigt. Für das Resultatsbild h ist der Wert $h(x, y)$ von dem Bildwert $f(x, y)$ sowie von den Bildwerten in einer gewissen Umgebung $U(x, y)$ um den Bildpunkt (x, y) im Eingabebild f abhängig, wobei im allgemeinen der Wert $h(x, y)$ von dem Wert $f(x, y)$ besonders stark beeinflusst wird. Lineare 2D-Filter können durch Faltung mit einem Faltungskern g (Filterfunktion) realisiert werden. Bei globalen 2D-Filtern besteht für mindestens einen Wert $h(x, y)$ eine potentielle Abhängigkeit von allen Eingangswerten des Bildes (Klette und Zamperoni [1995]).

Filter werden unterschieden in lineare und nicht-lineare Filter oder auch in Transformationen, nicht-rekursive und rekursive Filter (Huang [1979]). Eine andere Unterscheidung ist die zwischen Filtern mit feststehender Filterfunktion und solchen mit variabler Filterfunktion. Zu den letztgenannten gehören insbesondere die sogenannten kantenerhaltenden Filter. Solche Filter analysieren an jeder Stelle die Pixelumgebung daraufhin, ob möglicherweise eine Kante, d.h. ein größerer Grauwertsprung vorliegt. Wenn dieser Fall eintritt, versuchen sie, das Pixel dem „richtigen“ Segment zuzuordnen, um so die Qualität einer anschließenden Segmentierung und Klassifizierung zu verbessern.

2.1.4 Kante

Im Grauwertbild ist eine Kante allgemein gesprochen durch eine starke Grauwertänderung zwischen zwei dicht beieinander liegenden Bildsegmenten charakterisierbar. In realen Grauwertbildern sind Kanten im allgemeinen unregelmäßig im Grauwertverlauf bzw. mit verschiedenen Störungen behaftet. Die durch die Kante getrennten Bildsegmente sind ebenfalls im allgemeinen nicht durch gleichförmige Grauwerte beschreibbar bzw. können ebenfalls gestört sein (Klette und Zamperoni [1995]). Bei wenig gestörten Flächen sind im mittleren Grauwertbereich Grauwertunterschiede von 5 Einheiten bei einem Definitionsbereich von 256 Graustufen visuell gut unterscheidbar.

2.1.5 Partition

Eine Partition im mathematischen Sinne ist die Aufteilung einer Menge M in Teilmengen M_i , so daß gilt

$$M = \bigcup_i M_i \text{ und } \bigcap_i M_i = \emptyset \quad (2.1)$$

Sie wird mit $\{M_i\}$ bezeichnet. Andere Bezeichnungen für Partition sind Klasseneinteilung oder Zerlegung (Behnke et al. [1964]). Eine Partition $\{M'_j\}$ heißt feiner als $\{M_i\}$ genau dann, wenn es zu jedem j ein i gibt derart, daß M'_j in M_i enthalten ist.

⁵Rot-Grün-Blau

2.1.6 Ground Truth

Das Wort stammt aus dem militärischen Bereich und bedeutet dort den schwer zu erreichenden Zustand vollständiger Information.⁶

Im GIS-Bereich heißt es „*Information collected at the same site and at the same time as a remote sensing system is collecting data. Ground Truth is considered more accurate, and is used to interpret and calibrate remotely sensed observations.*“⁷ oder in Kurzform “Geländeerkundung bzw. Feldvergleich, die zur Stützung und Validierung der Klassifikation von Bilddaten dient.“⁸

2.2 Verfahren

Nach den Begriffsdefinitionen sollen nun grundlegende Verfahren vorgestellt werden.

2.2.1 Abtastung

Die Digitalisierung einer natürlichen Szene, sei es durch eine Kamera oder durch einen Scanner, bedeutet eine Diskretisierung der ursprünglich analog vorliegenden Daten. Praktisch werden die Daten, z. B. die Intensität der reflektierten Strahlung, an den durch das Raster vorgegebenen Punkten gemessen und als diskrete Zahlen erfaßt. Das Nyquist-Shannonsche Abtasttheorem⁹, engl.: Nyquist scan theorem, in neuerer Literatur auch WKS-Sampling-Theorem (für Whittaker-Kotelnikow-Shannon) genannt, beschreibt die theoretischen Voraussetzungen für eine informationsverlustfreie Wiederherstellung des analogen Signals aus den digitalen Daten.

Es wurde von Shannon [1949] zum Ausgangspunkt seiner Theorie der maximalen Kanalkapazität, d.h. der maximalen Bitrate in einem frequenzbeschränkten, rauschbelasteten Übertragungskanal. Dabei stützte er sich auf Überlegungen von Nyquist [1928] zur Übertragung endlicher Zahlenfolgen mittels trigonometrischer Polynome und auf die Theorie der Kardinalfunktionen (Whittaker [1929]). Unabhängig davon wurde das Abtasttheorem von Kotelnikow [1933] in der sowjetischen Literatur eingeführt, was im Westen allerdings erst in den 1950er Jahren bekannt wurde.

Das Abtasttheorem besagt, daß ein kontinuierliches, bandbeschränktes Signal mit einer Minimalfrequenz von 0 Hz und einer Maximalfrequenz f_{\max} mit einer Frequenz größer als $2 \cdot f_{\max}$ abgetastet werden muß, damit man aus dem so erhaltenen zeitdiskreten Signal das Ursprungssignal ohne Informationsverlust (aber mit unendlich großem Aufwand) rekonstruieren bzw. (mit endlichem Aufwand) beliebig genau approximieren kann.

$$f_{\text{abtast}} > 2f_{\max} \quad (2.2)$$

$\frac{1}{2} f_{\text{abtast}}$ nennt man die *Nyquist-Frequenz*.

Weil scharfe Kanten in einem Bild zu beliebig hohen Ortsfrequenzen führen, führt die Digitalisierung einer natürlichen Szene stets zu Artefakten, weil die Voraussetzungen des WKS-Sampling-Theorems verletzt werden. In der akustischen Signalverarbeitung kann man sich durch einen vorgeschalteten

⁶ http://www.aec.at/de/archives/center_projekt_ausgabe.asp?iProjectID=11694

⁷ <http://www.gisdevelopment.net/glossary/g.htm>

⁸ <http://www.geoinformatik.uni-rostock.de>

⁹ <http://de.wikipedia.org/wiki/Nyquist-Shannon-Abtasttheorem>

Tiefpaßfilter helfen, um die hohen Frequenzanteile auszufiltern. Bei der Aufnahme eines Bildes ist jedoch ein solcher vorgeschalteter Weichzeichner unerwünscht, weil er optisch eine unzureichende Flankensteilheit der Filterung aufweist.

2.2.2 Rand

Alle Verfahren, die die Umgebung eines Pixels berücksichtigen, wie die Faltung oder Kanten erhaltende Rauschfilter, erfordern eine besondere Berücksichtigung des Bildrandes, falls Eingabe und Ausgabe die gleiche Größe behalten sollen.

Aus einer beliebigen Anzahl von Verfahren für diese Berechnungen im Randbereich seien beispielhaft die folgenden drei genannt (Klette und Zamperoni [1995]):

1. f wird außerhalb des Bildrasters R identisch einer fixierten Konstante angenommen, etwa $f(x, y) = 0$ für alle Gitterpunkte $p = (x, y)$ außerhalb von R , oder $f(x, y)$ gleich dem mittleren Grauwert des Bildes f für Gitterpunkte $p = (x, y)$ außerhalb von R .
2. f wird außerhalb des Bildrasters R identisch fortgesetzt, zum Beispiel ist dann $f(x + M, y) = f(x, y)$ und $f(x, y + N) = f(x, y)$ unter dieser Annahme.
3. f wird an den Rändern von R nach außen gespiegelt, zum Beispiel ist dann $f(M + c, y) = f(M - c, y)$ und $f(x, N + y) = f(x, N - y)$, wobei die Bedingungen $0 < c < M/2$ und $0 < b < N/2$ sinnvollerweise aufzustellen sind.

2.2.3 Faltung

In der Mathematik und besonders in der Funktionalanalysis beschreibt die Faltung¹⁰ (engl. Convolution) einen mathematischen Operator, welcher für zwei Funktionen f und w eine dritte Funktion liefert, die die „Überlappung“ zwischen f und einer gespiegelten verschobenen Version von w angibt.

Für zwei auf dem reellen Intervall D definierte Funktionen $f, w: D \rightarrow C$ wird die Faltung von f mit w als $f * w$ notiert und ist definiert als das Integral über das Produkt von f mit einer gespiegelten verschobenen Version von w :

$$(f * w)(t) := \int f(\tau)w(t - \tau)d\tau \quad (2.3)$$

Der Integrationsbereich ist der Definitionsbereich D beider Funktionen. Im Fall eines beschränkten Definitionsbereichs können f und w als periodisch fortgesetzt oder als konstant (siehe Abschnitt 2.2.2, Seite 15) angenommen werden, damit der Faktor $w(t - \tau)$ stets definiert ist. Oft werden auch f und w statt dessen durch Null fortgesetzt.

Die lineare Filterung eines elektronischen Signals (das kann außer einem Bild auch ein akustisches Signal sein) stellt die Faltung der Originalfunktion mit der Impulsantwort des Sensors dar.

Diffusions-Prozesse, wie sie in Abschnitt 5.3 beschrieben werden, lassen sich ebenfalls durch eine Faltung beschreiben.

¹⁰angelehnt an [http://de.wikipedia.org/wiki/Faltung_\(Mathematik\)](http://de.wikipedia.org/wiki/Faltung_(Mathematik)) und Schowengerdt[1997]

Eine anschauliche Deutung der Faltung ist die Gewichtung einer Funktion mit einer anderen. Der Funktionswert der Gewichtsfunktion an einer Stelle t gibt an, wie stark der um t zurückliegende Wert der gewichteten Funktion in den Wert der Ergebnisfunktion eingeht.

2.2.3.1 Diskrete Faltung

In der digitalen Signalverarbeitung und der digitalen Bildverarbeitung hat man es meist mit diskreten Funktionen zu tun. Die diskrete Faltung kann analog zu Gleichung (2.3) definiert werden:

$$(f * w)(n) = \sum_k f(k)w(n-k) \quad (2.4)$$

In abgekürzter Form schreibt man auch

$$g = f * w \quad (2.5)$$

Für das Ausgabepixel g_{ij} , das Eingabebild f und den bearbeitenden Filter w bedeutet dies in ausführlicher Form

$$g_{ij} = \sum_{m=i-\lfloor \frac{W_y}{2} \rfloor}^{i+\lfloor \frac{W_y}{2} \rfloor} \sum_{n=j-\lfloor \frac{W_x}{2} \rfloor}^{j+\lfloor \frac{W_x}{2} \rfloor} f_{mn} w_{i-m, j-n} \quad (2.6)$$

w sei auf $(0,0)$ zentriert und W_x, W_y die Anzahl der Zeilen bzw. Spalten in x - bzw. y -Richtung.

Die Gaußklammer $\lfloor \cdot \rfloor$ bedeutet die Abrundung nach unten (größte ganze Zahl, die nicht größer ist als ...), also $\lfloor \frac{5}{2} \rfloor = 2$.

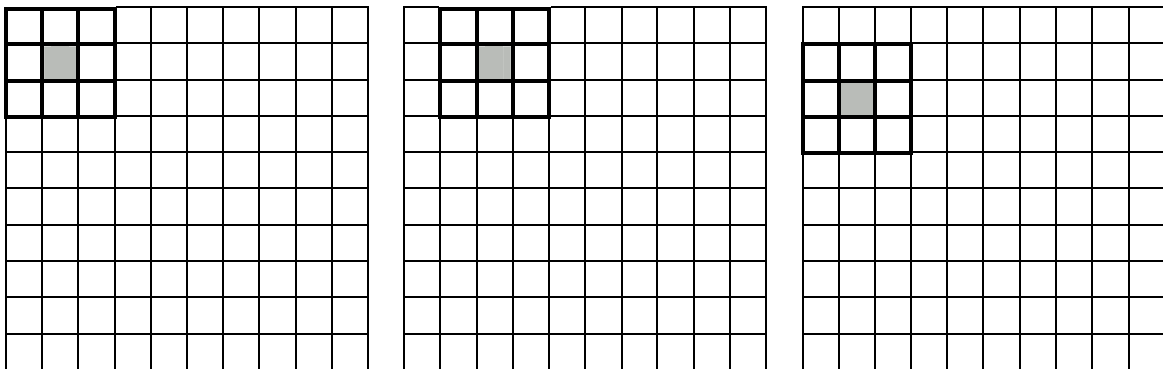


Abbildung 2.1

Das Ausgabepixel (grau) wird berechnet aus der Faltung der Filterfunktion w (kräftige Linien) auf dem Eingabebild f (groß). Anschließend wird die Filtermatrix um ein Pixel verschoben. Nachdem eine Zeile bearbeitet wurde, wird die nächste Zeile bearbeitet. Wie man sieht, ist für die Pixel im Randbereich eine besondere Berechnung erforderlich, falls Ausgabe und Eingabe die gleiche Größe behalten sollen.

Der gesamte Algorithmus zur Anwendung einer Faltung lautet damit nach Schowengerdt [1997]:

Drehe w um 180° .

Verschiebe w , so daß es auf f_{ij} zentriert ist.

Berechne g_{ij} nach Gleichung (2.6).

Wiederhole 2. und 3. für alle in Frage kommenden Pixel.

Auch dabei hängen die Summationsgrenzen von der Art der Funktionen $f, w: D \rightarrow C$ ab (dabei ist D Teilmenge von \mathbb{R}).

2.3 Segmentierung

Die Zahl möglicher Partitionen (vgl. Abschnitt 2.1.5) steigt größenordnungsmäßig mit der Fakultät der Anzahl der Elemente der Grundmenge an. Im Sinne der Bildverarbeitung sind fast alle dieser Partitionen nicht weiterführend. Es wird vielmehr verlangt, daß die zu einem Segment zusammengefaßten Pixel eines Bildes untereinander in einem gewissen Zusammenhang stehen. Rosenfeld und Kak [1982] bemerken dazu: „It should be emphasized that there is no single standard approach to segmentation.“ (zitiert nach Hanning [2001])

Alle Autoren sind sich darüber einig, daß die Segmentierung eines Bildes eine Partition darstellt. Differenzen bestehen in der Auffassung darüber, wie die Segmente zu definieren seien.

Zur Segmentierung von Rasterbildern können sehr unterschiedliche Kriterien herangezogen werden: Farb- und Helligkeitsunterschiede, Farb- und Helligkeitsverläufe, Texturen, Ergebnisse aus diskreten Objekterkennungsverfahren und externe Informationen. Für eine Segmentierung nach Helligkeitsunterschieden in einem Spektralkanal, d. h. nach Grauwerten, konnte eine vorherige Grauwertglättung das Ergebnis älterer Segmentierungsverfahren entscheidend verbessern. Die Auswertung mehrerer Merkmale führt zur Erfordernis, multivariate Verteilungen zu berücksichtigen.

Hanning [2001] gibt zahlreiche mathematische Definitionen für die zur Segmentierung benötigten Begriffe, stellt aber fest „Der Begriff „Segmentierung“ bzw. „Bildsegmentierung“ ist sicherlich einer der zentralen Begriffe der Bildverarbeitung überhaupt. Es gibt jedoch keine einheitliche Definition des Begriffes.“ um dann fortzufahren: „... gehen Robert Haralick und Linda Shapiro dem Problem der Definition des Begriffes „Segmentierung“ elegant aus dem Weg, indem sie beschreiben, welche Bedingungen die meisten Algorithmen, die sich „Segmentierungsalgorithmen“ nennen, erfüllen“:

1. Segmente einer Segmentierung sind uniform oder homogen bezüglich einer Charakteristik wie zum Beispiel dem Grau- oder Farbwert oder der Textur.
2. Adjazente Segmente haben signifikant unterschiedliche Werte in der Charakteristik, in der sie uniform sind.
3. Das Innere der Segmente ist einfach und hat, wenn überhaupt, nur wenige kleine „Löcher“.
4. Die Ränder von Segmenten sind ebenfalls einfach und nicht zerfranst. (Haralick und Shapiro [1992]).

Pavlidis [1977] und Wahl [1984] geben die Kriterien etwas formaler an:

1. Die Segmentierung M ist eine Partition im Sinne der Definition in Abschnitt 2.1.5 (Seite 13).
2. Die einzelnen Segmente M_i sind stets zusammenhängend.
3. Es existiert ein Einheitlichkeitskriterium E_k , das für alle $E_k(M_i)$ erfüllt ist.
4. für jede Vereinigungsmenge zweier benachbarter Teilbereiche M_i und M_j ist $E_k(M_i \cup M_j)$ nicht erfüllt.

2.4 Verwendete Hard- und Software

Bevor im nächsten Kapitel die Thematik des Rauschens diskutiert wird, soll hier eine Übersicht über die verwendete Hard- und Software gegeben werden.

Alle Programme wurden ausgeführt auf einem Rechner mit einem Prozessor AMD Athlon 2200+ mit 1,8 Ghz und 512 MB RAM. Als Betriebssystem wurde Microsoft® Windows® XP Home SP2 verwendet. Diese Arbeit wurde geschrieben mit OpenOffice 1.1.3¹¹ und Microsoft® Word 2002 SP3. Die mathematischen Formeln wurden mit dem FormelEditor Version 3 aus dem Microsoft Office-Paket gesetzt. Die erhaltenen Daten wurden mit Microsoft® Excel 2002 ausgewertet. Kleine Hilfsroutinen wurden mit Microsoft® QuickBasic® 4.0 aus dem Jahre 1988 programmiert. Zur Bildbetrachtung, -bearbeitung und Formatkonvertierung wurde Adobe® Photoshop® 6.0 benutzt.

Als Programmieroberfläche für C-Programme wurde die Professional Edition aus Microsoft® Visual Studio® 6.0 verwendet.

Das Programm zur Bestimmung des Rauschens wurde unter Verwendung der Routinen aus dem Insight Toolkit¹² nach der Originalarbeit von Baltsavias et al. [2001] neu geschrieben.

Der Mittelwert-, der Median-, beide Anisotropic Diffusion und der Bilaterale Filter wurden aus dem Insight Toolkit übernommen. Das Insight Toolkit liegt im C-Quelltext vor. Es wurde mit CMake¹³ Version 2 patch 6 für die Kompilierung vorbereitet.

Die Programme Graphennetzwerkfilter und der Edge Preserving Speckle Filter wurden nach den Originalarbeiten von Halle [1999] bzw. Haag [1998] neu programmiert.

Manche Insight Toolkit Programme sind in einem CVS Repository gespeichert. Zum Zugriff darauf war WinCvs¹⁴ 1.3.20.3 Beta 20 build 3 erforderlich. WinCvs ist angewiesen auf Python¹⁵ oder Tk/Tcl¹⁶, die in der Version 2.4 (Python) bzw. 8.3.4 (Tcl) verwendet wurden.

Nachdem das ursprünglich favorisierte GeoAida (<http://tnt.uni-hannover.de>) nicht mehr weiterentwickelt wird, wurde zur Segmentierung und Klassifikation die im Internet zur Verfügung gestellte Trial Version des kommerziellen Pakets eCognition®¹⁷ in der Version 4.0 verwendet.

¹¹<http://www.openoffice.org> (Freeware)

¹²<http://www.itk.org> (Freeware)

¹³<http://www.cmake.org> (Freeware)

¹⁴<http://www.wincvs.org> cvsgui.sourceforge.net (Freeware)

¹⁵<http://www.python.org> (Freeware)

¹⁶<http://tcl.sourceforge.net> (Freeware)

¹⁷<http://www.definiens-imaging.com>

3 Bildrauschen

Nach diesen Vorbereitungen folgt nun ein Einstieg in die Problematik des Bildrauschens. Zunächst wird wieder eine formale Definition gegeben, die später mit praxisorientierten Beispielen erläutert wird.

Das Eingabebild f wird verstanden als zusammengesetzt aus der eigentlichen Information f' und einer zusätzlichen Signalkomponente r :

$$f(\mathbf{x}) = f'(\mathbf{x}) + r(f, \mathbf{x}) \quad (3.1)$$

Die Bestimmung von r gelingt leicht für $f'=0$, einer Aufnahme mit abgedecktem Sensor. Allerdings ist diese Signalkomponente nicht konstant, sondern statistisch zufällig und je nach Aufnahmebedingungen und Sensor auch von f abhängig, so insbesondere auch beim multiplikativen Rauschen, wie es beispielsweise bei Radarbildern auftritt. Weil die digitalen Bildsensoren ein analoges Lichtsignal über einen bestimmten Zeitraum integrieren und dann einen digitalen Wert aus einer kleinen endlichen Menge möglicher Werte ausgeben, kommen Rundungsfehler hinzu. Ein Versuch, Rauschen meßtechnisch zu erfassen, wird im Abschnitt 3.3.2 auf Seite 25 beschrieben.

3.1 Begriffe

3.1.1 Signal-Rausch-Verhältnis

Das Signal-Rausch-Verhältnis, engl. *Signal-Noise-Ratio* (SNR) ist definiert durch das Verhältnis der Leistungen des Signals, der Nutzinformation, durch das Hintergrundrauschen.

$$SNR = \frac{P_{signal}}{P_{noise}} \quad (3.2)$$

Nach Bowen [2002] sind die bereitgestellten digitalen Zahlen der Pixelrepräsentation (digital numbers, DN) nahezu proportional der vom Sensor empfangenen Strahlungsleistung (vergl. Abbildung 3.1). Deshalb können zur Berechnung des Signal-Rauschverhältnisses die Digitalzahlen unmittelbar genutzt werden.

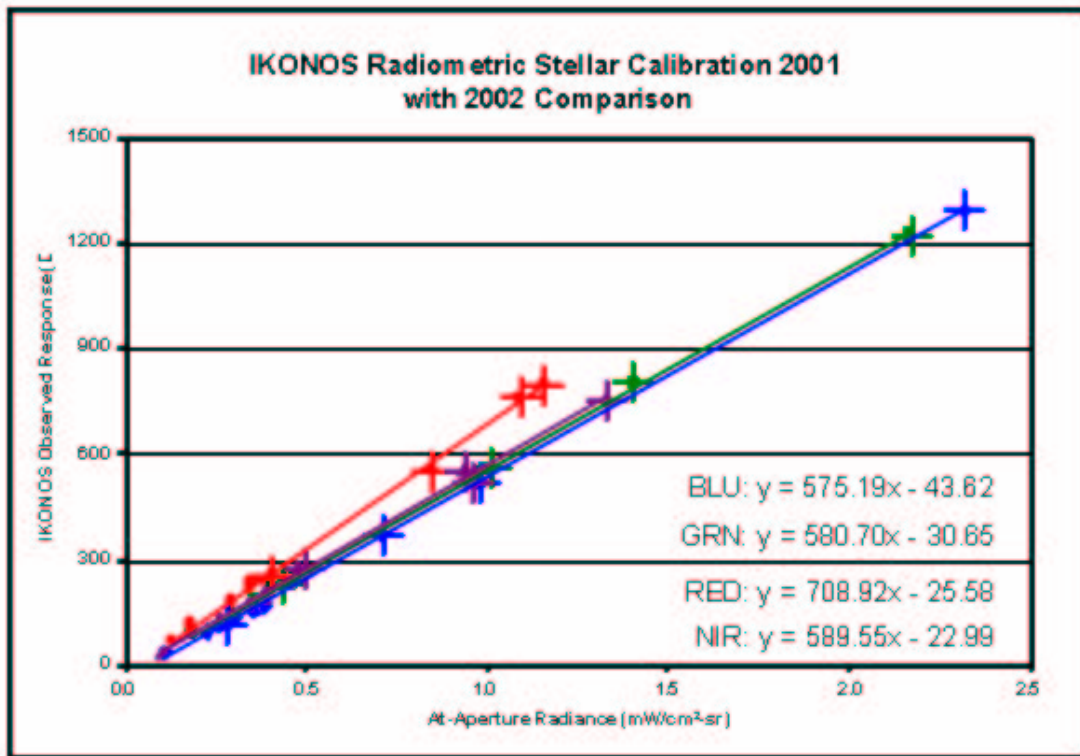


Abbildung 3.1

Die empfangene Strahlungsleistung ist der ausgegebenen Digitalzahl ungefähr proportional. Aus Bowen [2002]

Als Signalleistung können je nach Absicht der maximal mögliche Digitalwert, der größte im Bild vorhandene Digitalwert, die Differenz zwischen größtem und kleinstem Wert oder der Mittelwert verwendet werden.

Billingsley [1979] verwendet als Bezeichnungen

- S/σ für das Verhältnis zwischen der größtmöglichen Signaldifferenz und dem nach Gleichung (3.3) definierten Rauschmaß P_{noise} (root mean square (rms)),
- SNR für das Verhältnis zwischen dem aktuellen (allerdings auch nicht näher definierten) Signal und P_{noise} und
- β für das Verhältnis zwischen dem Quantisierungsschrittmaß und P_{noise} .

$$P_{noise} = \frac{1}{N} \sum_{i=1}^N (f'_i - f_i)^2 \quad (3.3)$$

Statt SNR findet sich häufig auch die Bezeichnung S/N und als Einheit wird dB angegeben. Bei elektrischen Signalen verwendet man die Formel

$$S/N = 20 \log_{10}(P_{signal}/P_{noise}) \text{ dB} \quad (3.4)$$

wenn das Verhältnis der Signalleistungen gebildet wird und

$$S/N = 10 \log_{10}(A_{signal}/A_{noise}), \text{ dB} \quad (3.5)$$

wenn die Spannungen (Amplituden) der Signale zueinander ins Verhältnis gesetzt werden.

Bei künstlich zugefügtem Rauschen kann die Differenz zwischen der Ground Truth und dem verrauschten Bild gebildet werden. Üblich ist die Berechnung des mittleren quadratischen Fehlers als die mittlere Differenz über die $i = 1 \dots N$ Punkte des Originalbildes f' und dem verrauschten Bild f , bei mehrkanaligen Bildern auch je Kanal.

Da normalerweise keine Ground Truth vorliegt, wird häufig die Varianz σ^2 aus den Differenzen zwischen Originalbild und dem Mittelwert \bar{f} berechnet. Dann sind jedoch sämtliche Bildinhalte in diesem Wert mit enthalten. Um das Rauschen abzuschätzen, sind daher andere Verfahren erforderlich, die im Abschnitt 3.3.2 (Seite 25) beschrieben werden. In den Kapiteln 6 und 7 wird gezeigt werden, daß gute, d. h. kantenerhaltende Filterverfahren die Varianz im Bild nicht wesentlich verändern.

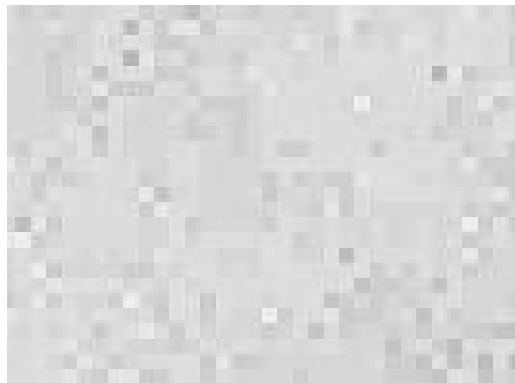


Abbildung 3.2

Ausschnitt aus einer IKONOS-Szene zur Bestimmung des Rauschens (vgl. Abbildung 7.7).

Zur Veranschaulichung der verschiedenen Maßzahlen soll der in Abbildung 3.2 dargestellte, wegen seiner Homogenität ausgewählte Ausschnitt aus einer IKONOS-Szene dienen: Die Original-Szene enthält Grauwerte von 0 bis 1083. Nach Spezifikation können IKONOS-Szenen Grauwerte von 0 bis 2047 enthalten. Der Ausschnitt selbst enthält Grauwerte von 189 bis 237 bei einem Medianwert von 213 und einem Mittelwert von 213,8. Mit diesem Mittelwert ergibt sich nach Gleichung (3.3) für diesen Ausschnitt ein Rauschmaß von

$$P_{noise} = 32.$$

Die folgende Tabelle 3.1 listet verschiedene Maßzahlen für diesen Ausschnitt auf:

P_{noise}		=	32
S/σ	= 2048/32	=	64
S/σ	= 1084/32	=	33,88
SNR	= (237-189)/32	=	1,5
S/N	= $20 \log_{10} 33,88$	=	30 dB
S/N	= $20 \log_{10} 1,5$	=	3,5 dB
σ	= $\sqrt{32}$	=	5,7

Tabelle 3.1

Vergleich der verschiedenen Maßzahlen für das Rauschen in
Abbildung 3.2

Das Signal-Rausch-Verhältnis als Zahlenwert wird im folgenden nicht verwendet werden, weil dazu in jedem Einzelfall die bezogene Signalleistung benannt werden müßte und statt dessen bei allen Bildern alle oben genannten Werte beziffert werden. In dieser Arbeit wird entsprechend den in der Literatur meist verbreiteten Angaben die Standardabweichung σ als die Quadratwurzel aus der Varianz σ^2 angegeben.

3.1.2 Quantisierungsrauschen

Das Quantisierungsrauschen N_Q entsteht durch die Umwandlung eines analogen Signals in ein digitales Signal. Es ist ein Rundungsfehler, es ist additiv und signalunabhängig, falls die Digitalzahl mehr als 4 bit hat.¹⁸ Es ist proportional dem Quadrat aus der Amplitude des Eingangssignals V_{in} dividiert durch die Anzahl der digital darstellbaren Zahlen 2^Q .

$$N_Q \approx \left(\frac{V_{in}}{2^Q} \right)^2 \quad (3.6)$$

3.1.3 weißes Rauschen

Unter weißem Rauschen wird ein Rauschen verstanden, das vom Signal unabhängig ist. Das Störsignal wird separat erzeugt, sein Mittelwert ist Null. Der Gegensatz ist ein vom Signal abhängiges Rauschen, das autokorreliertes oder farbiges Rauschen genannt wird (Hütte [1989]).

3.1.4 additives Rauschen

Additives Rauschen ist von der Amplitude des Signals unabhängig. Im Falle der Bildbearbeitung spricht man von additivem Rauschen, wenn die durch das Rauschen bedingte Grauwertvariation in jedem Grauwertbereich die gleiche ist.

3.1.5 multiplikatives Rauschen

Im Unterschied zum additiven Rauschen ist das multiplikative Rauschen proportional zur Amplitude des Eingangssignals. Multiplikatives Rauschen findet man insbesondere bei RADAR-Daten.

¹⁸ http://en.wikipedia.org/wiki/Quantization_noise

3.2 Ursachen

Nach diesen Definitionen werden jetzt die verschiedenen Ursachen des Rauschens beschrieben.

Folgende Faktoren bewirken ein höheres Rauschen: kleine Sensoren, hohe Packungsdichte, kurze Belichtungszeit, wenig Licht, hohe Sensortemperatur.

Zunächst sind die CCD-bedingten Rauschursachen zu berücksichtigen:

1. **Thermales Rauschen.** Thermales Rauschen verringert sich je 20K Temperaturabnahme im Verhältnis zur Raumtemperatur um den Faktor 10. Es ist besonders störend in der Astronomie mit langen Aufnahmezeiten. Bei einer Satellitenkamera und sehr kurzen Aufnahmezeiten wird es eher zu vernachlässigen sein.
2. **Variabilität der Sensitivität der Photodioden.** Die einzelnen Dioden eines Zeilensensors wandeln die empfangenen Photonen unterschiedlich effektiv in Ladung um. Das Ausgabesignal kann durch die Subtraktion der Aufnahme einer uniformen Graufäche kalibriert werden. Besser ist es, wenn mehrere Graufächenaufnahmen gemittelt werden und die Kalibrierung sehr frühzeitig in der Präprozessierungskette erfolgt.
3. **Elektronisches Rauschen.** Elektronisches Rauschen wird durch Streukapazitäten im elektrischen Aufbau und der Ausleseschaltung hervorgerufen. Elektronisches Rauschen wird durch eine Aufnahme mit einer Expositionszeit von 0 dargestellt und ist in einer Graufächenaufnahme enthalten.
4. **Photonenrauschen.** Photonen als diskrete Teilchen treffen zufällig verteilt auf dem Sensor auf. Das Rauschen folgt einer Poisson-Verteilung mit der Quadratwurzel der Intensität. Eine Verlängerung der Aufnahmezeit würde daher zu einer quadratischen Verbesserung des Rauschverhaltens führen, ist aber systembedingt nicht möglich. Dennoch ist zu berücksichtigen, daß in hellen Bereichen das Photonenrauschen mit der Quadratwurzel der Intensität steigt.
5. **Ausleserauschen.** Schließlich verursacht die Ausleseschaltung, bestehend aus dem Reset-Transistor, dem eigentlichen Auslesetransistor und dem sich anschließenden Signalverstärker ein Rauschen. Die folgende¹⁹ Abbildung 3.4 zeigt schematisch die Schaltung einer CCD-Ausabeelektronik. Der Sensor wandelt empfangenes Licht in Ladung um, die über den Ausgabeknotenpunkt und den Auslesetransistor in Spannung gewandelt wird, um durch den nachfolgenden Verstärker verstärkt zu werden, bevor sie von einem AD-Konverter digitalisiert wird. Jede einzelne Komponente fügt Rauschen zu.

¹⁹aus: http://www.mssl.ucl.ac.uk/www_detector/ccdgroup/opttheory/darkcurrent.html

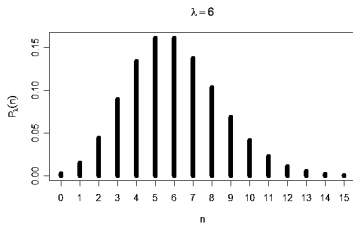


Abbildung 3.3

Poissonverteilung mit $\lambda = 6$
(<http://www.wikipedia.de>)

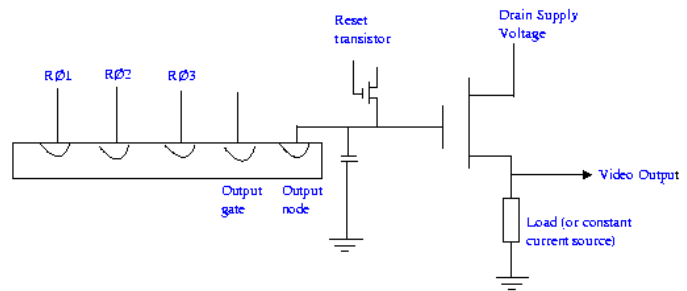


Abbildung 3.4

Ausleseschaltung eines CCD-Sensors. Jede einzelne Komponente trägt zum Rauschen bei.

3.3 Addition des Rauschens

In den folgenden beiden Abschnitten wird beschrieben, wie Rauschen zugefügt wurde und wie Rauschen in Fernerkundungsaufnahmen gemessen werden kann.

3.3.1 Rauschen zufügen

Additives Rauschen mit der Standardabweichung σ wird zugefügt, indem zu jedem Pixel ein Wert im Bereich von $(-4\sigma, +4\sigma)$ addiert wird, der sich durch die gewichtete Verteilungsfunktion ergibt. Die Intervallgrenzen wurden so gewählt, daß in der Normalverteilung weniger als 10^{-4} der Werte außerhalb dieses Intervall liegen. Als Verteilungsfunktion wurde eine Gaußverteilung nach Gleichung (3.7) gewählt.

$$r(t) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{t^2}{2\sigma^2}} \quad t \in \{-4\sigma, \dots, +4\sigma\} \quad (3.7)$$

Dazu wird eine Look-Up-Tabelle der Länge $8\sigma + 1$ vorbereitet, die eine Diskretisierung des Integrals der gewünschten Verteilungsfunktion, also der Dichtefunktion darstellt. Der Zufallszahlengenerator wird initialisiert. In der Verarbeitungsschleife wird eine Zufallszahl im Bereich $[0, 1]$ generiert und in der Look-Up-Tabelle die nächst größere Zahl gesucht.

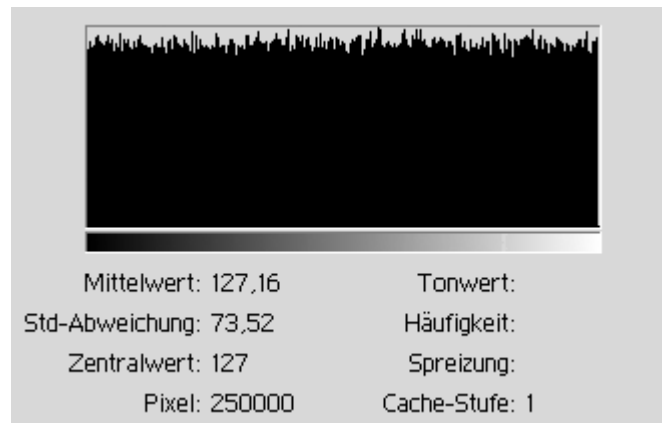


Abbildung 3.5

250000 Zufallszahlen aus dem Zufallszahlengenerator: Auch die Verteilung der Zufallszahlen ist verrauscht. Der Wert 255 kommt durch Rundungseffekte mit 14 am seltensten vor, während die Häufigkeiten der Werte von 0 bis 254 zwischen 940 und 1040 streuen.

Der Tabellenplatz, vermindert um die Hälfte der Zahl der Tabelleneinträge, ergibt den zu addierenden Wert. Bei Pixelwerten nahe Null oder im Bereich der Sättigung kann der erhaltene Wert aus dem Wertebereich der Darstellung herausfallen und muß durch den zulässigen Randwert ersetzt werden. Bei der Bestimmung des Rauschmaßes gelten diese Werte als Sättigungswerte und werden nicht berücksichtigt.

3.3.2 Rauschen messen

In Abschnitt 3.1.1 wurde erwähnt, daß die einfache Berechnung der Varianz σ^2 eines Bildes zur Bestimmung des Rauschens wenig aussagekräftig ist, weil Rauschen und Bildinhalte zusammengefaßt werden. Um sinnvolle Daten zu erhalten, ist es erforderlich, in einem Bild homogene Bereiche aufzusuchen und in diesen das Rauschen zu bestimmen. In der Regel markiert ein menschlicher Untersucher entsprechende Bildbereiche, um dann in diesen ausgewählten Bereichen die Varianz bzw. die daraus berechnete Standardabweichung σ zu ermitteln.

So bestimmte Schowengerdt [2002] das Rauschen in einer IKONOS-Aufnahme in einem dunklen Bildbereich, nämlich der Wasseroberfläche eines Sees in Arizona, mit $\sigma = 8,23$ bei elf bit Bilddaten. Nach den Anmerkungen zum Signal-Rauschverhältnis (Abschnitt 3.1.1 auf Seite 19) kann dieser Wert nicht ohne weiteres auf acht bit Bilddaten umgerechnet werden, weil nicht bekannt ist, wie groß die tatsächliche Signalamplitude im Bild war.

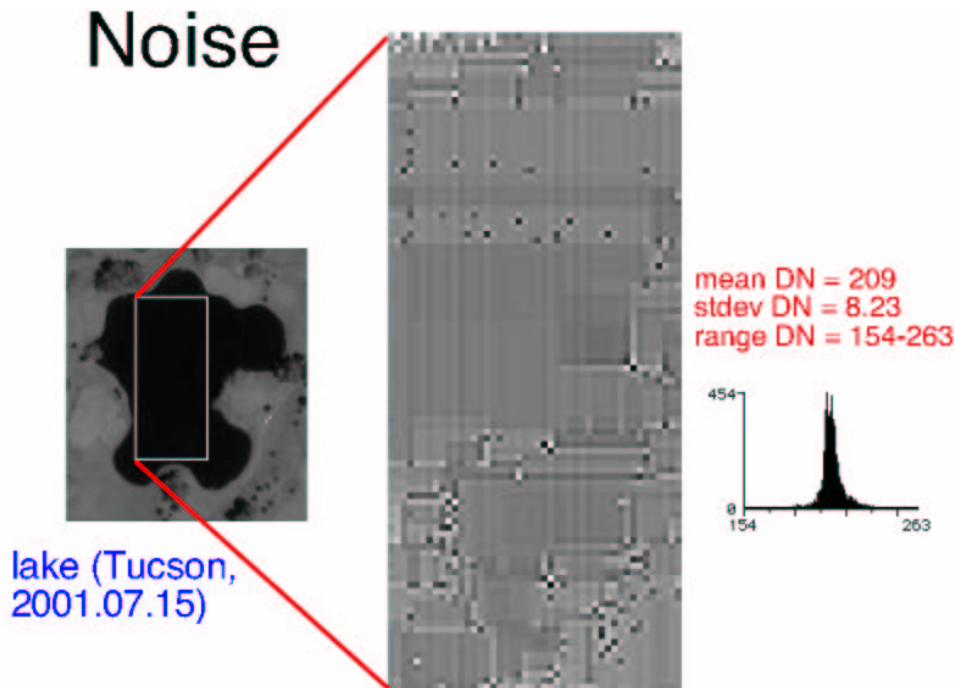


Abbildung 3.6

Darstellung des Rauschens in einer IKONOS-Aufnahme mit elf bit Grauwertumfang (aus Schowengerdt [2002])

Baltsavias et al. [2001] stellten fest, daß dieses Verfahren nicht für alle Szenen anwendbar ist, weil häufig im Bild große homogene Flächen fehlen.

So wurde das Bildrauschen mit folgendem Verfahren bestimmt: Zunächst wurde das Bild manuell unterteilt in homogene, nicht gesättigte Flächen wie Wasserflächen und inhomogene Bereiche wie der Rest des Bildes. Ein kleines Fenster wird über den Flächen verschoben und die Standardabweichung der Pixelwerte bestimmt. Für homogene Flächen werden die erhaltenen Standardabweichungen der Größe nach sortiert und die N% kleinsten zur Berechnung der durchschnittlichen Standardabweichung herangezogen. Typische Werte für N sind 80 bis 95. Die Fensterwerte der nichthomogenen Flächen werden nach ihren Grauwerten sortiert. Innerhalb eines Grauwertes werden die Ergebnisse wieder nach ihrer Standardabweichung sortiert und das Rauschen als Mittelwert der kleinsten 5% berechnet. Dabei werden jedoch nur die Werte verwendet, für die der Mittelwert aus einer ausreichenden Zahl (z. B. 100) von Werten berechnet werden kann. Dieses Vorgehen rechtfertigt sich aus der Annahme, daß auch in nicht homogenen Flächen kleine homogene Flächen enthalten sind. Dieses Verfahren ist universell einsetzbar und nicht von plattformsspezifischen Kalibrierungen oder speziellen Objektflächen abhängig.

Auf diese Art bestimmten die Autoren das Bildrauschen in den von ihnen verwandten IKONOS-Aufnahmen zu folgenden Werten: Die MTF²⁰-geschärfte panchromatische Aufnahmen rauschten bei niedrigen Intensitätswerten um etwa 1,5 Grauwertstufen, in den hohen Intensitätsstufen jedoch bis zu 10,5 Stufen. Die multispektralen Aufnahmen zeigten ein geringeres Rauschen als die panchromati-

²⁰ Modulation Transfer Function Correction: Die Modulations-Transfer-Funktion beschreibt das Verhältnis zwischen Eingangs- und Ausgangssignal als Funktion der räumlichen Frequenz. So geben dicht beieinander liegende Linienpaare in der Ausgabe weniger Kontrast als dicke Balken. www.schneideroptics.com/info/white_papers/quality_criteria_of_lenses.pdf. Die Korrektur soll diesem Mangel abhelfen.

schen Aufnahmen, das entgegen der Erwartung vom NIR²¹ zum blauen Kanal hin noch abnahm. Die mittlere rauschen-bedingte Standardabweichung betrug 4,5 bis 5 Stufen für die panchromatischen Bilder und zwei Stufen für die multispektralen Bilder. Solches Rauschen kann insbesondere nach Kontrastverstärkung irreführende Texturen erzeugen.

In den IKONOS-Aufnahmen fanden sie weitere Artefakte, die nicht alle zu erklären waren:

1. Für CCD-Sensoren typische helle und dunkle Streifen in Flugrichtung,
2. Signalsättigung aufgrund starker Reflexion mit Ausbreitung auf Nachbarsensorzellen im Read-Out-Prozess,
3. physikalisch nicht vorhandene helle Linien, die weder in noch quer zur Flugbahn verliefen.
4. Auch die Datenkompression auf 2,6 bit führte zu eher in homogenen Flächen sichtbaren Artefakten.

Leider wurde nicht angegeben, wie die Werte aus homogenen und nicht homogenen Flächen miteinander verrechnet werden und ob die Intensitätsabhängigkeit berücksichtigt wurde. Der Algorithmus wurde deshalb für die vorliegende Arbeit neu geschrieben. Homogene Flächen wurden nicht besonders berücksichtigt, weil Wasserflächen durch Wellenbildung bei den untersuchten Auflösungen stets texturiert sind und Asphaltflächen, wenn sie nicht völlig schwarz erscheinen, in der Realität üblicherweise scheckig aussehen.

Unter Verwendung des ITK-Toolkits²² wurde ein Programm geschrieben, das ein beliebiges Bild einliest und für jede 3 x 3-Umgebung eines Pixels Mittelwert und Standardabweichung berechnet. Randpixel werden dabei nicht berücksichtigt. Der gerundete Mittelwert wird als Zeiger in eine Liste verwendet, in der die zugehörigen Standardabweichungen gesammelt werden. Die Werte 0 und 255 gelten als Sättigungsbereiche für die keine Listen angelegt werden. Listen, die unterdurchschnittlich viele Standardabweichungen enthalten, werden nicht ausgewertet. Als unterdurchschnittlich besetzt gilt eine Liste, wenn sie weniger als die Gesamtzahl der Pixel im Bild geteilt durch die Anzahl der möglichen Grauwerte enthält. Für jede verbliebene Liste wird aus den 5% kleinsten Werten der Mittelwert gebildet und zusammen mit der Anzahl der Einträge in der Liste ausgegeben. Das Rauschmaß eines Bildes ist der Mittelwert aus allen so errechneten Standardabweichungen für die verschiedenen im Bild berücksichtigten mittleren Grauwerte.

Diese Berechnungsmethode benötigt keinen Benutzereingriff, um homogene Flächen in einem Bild auszuwählen. Sie geht davon aus, daß in einem Bild eine genügende Anzahl kleiner (3 x 3 Pixel große) homogener Flächen vorhanden sind. Die Methode mittelt über alle im Bild vorkommenden Intensitäten, d. h. bei einem multiplikativen Rauschen ist der erhaltene Wert abhängig vom mittleren Grauwert des Bildes²³. Mischpixel an Objektkanten tragen nicht zum Rauschmaß bei, wenn der mittlere Grauwert in anderen, homogenen Bildbereichen oft genug vorkommt. Sie tragen auch dann nicht zum Rauschmaß bei, wenn sie in ihrer Farbe selten vorkommende Linienobjekte bilden. Kammförmige Artefakte im Histogramm, wie sie durch Equalizing und Histogrammstretch entstehen, ändern das Rauschmaß nur innerhalb ihrer Rundungseffekte.

²¹ NIR. Nahes Infrarot, Wellenlängenbereich etwa 760 bis 900 nm.

²² <http://www.itk.org>

²³ Will man multiplikatives Rauschen besser berücksichtigen, wäre ein zweiter Parameter zu seiner Beschreibung erforderlich, was die Handhabbarkeit verschlechtert.

4 Segmentierung

In diesem Kapitel soll zunächst eine kurze Übersicht in die verschiedenen Segmentierungsverfahren in der Bildbearbeitung geben. Die Vielzahl der bis heute vorgestellten Verfahren kann nicht einzeln dargestellt werden. Auch sind andere Einteilungen als die hier gewählten möglich. So gibt z. B. Blaschke [2000] eine andere Zusammenstellung als Überblick.

In Abschnitt 4.1 werden Verfahren zusammengefaßt, die einzelne Objekte aus einem Bild herausfiltern. Dies entspricht einem wissensbasierten Top-Down-Ansatz. Der Gegenstand des Interesses ist vorab bekannt und wird im Bild gesucht. Die Verfahren werden auch modellbasiert genannt, weil vorab für die zu identifizierenden Objekte verallgemeinerte Beschreibungen, Modelle, erstellt werden müssen.

Im Abschnitt 4.2 werden Segmentierungsverfahren im engeren Sinne beschrieben. Dies entspricht einem Bottom-Up-Ansatz. Ausgehend vom gesamten Bild wird dieses in Segmente entsprechend der in Abschnitt 2.3 gegebenen Definition unterteilt. Anschließend muß dieses so gefundenen Segmenten eine Bedeutung zugewiesen werden. Dieser Schritt wird in der Regel als Klassifikation bezeichnet.

Schließlich wird die verwendete Segmentierungssoftware eCognition vorgestellt. Nach einer kurzen Übersicht über den verwendeten Algorithmus und die genutzten Parameter und ihre Einstellung wird in Abschnitt 4.4 untersucht, wie sich die Software unter verschiedenen Parametereinstellungen verhält, wenn das Bild keine Informationen enthält und nur eine weiße Fläche segmentiert werden soll. Im letzten Abschnitt wird das Segmentierungsergebnis dargestellt für den Hintergrund der Ground Truth, wie sie später in Kapitel 6 verwendet werden wird.

4.1 Objektextraktion

Neben dem Versuch, eine sinnvolle Segmentierung aller Pixel zu erreichen, können auch lediglich die interessierenden Objekte in einem Bild aufgesucht werden. Vey [1993] stellt ein Bündel verschiedener Möglichkeiten vor, wie einzelne Objekte in Grauwertbildern auch unabhängig von ihrer Lage erkannt werden können. In Range-Aufnahmen will man Vordergrund und Hintergrund voneinander unterscheiden. (Hoover et al. [1994, 1996]). Einen frühen Ansatz zur Erkennung von Bäumen in Fernerkundungsaufnahmen beschreibt Pollock [1994]. Neuere Verfahren zur Baumerkennung wurden von Straub [2003, 2004] vorgestellt. Verfahren zur Straßenextraktion stellen Gerke et al. [2001a, b] und Gerke [2005] dar. Hat man Objekte geortet, bieten sich Snakes zu Erfassung der Objektkanten an (Kerschner [2003]). Lin und Nevatia [1998] entwickeln mit „BUDDS“ (Buildings Detection and Description System) ein Verfahren, mit dem aus einem einzelnen Grauwertbild eine dreidimensionale Darstellung von Häusern rekonstruiert werden kann. Dieses Verfahren wurde von Maloof et al. [2003] mit Lernalgorithmen weiter verbessert.

Diese Verfahren sind modellbasiert: Der Gegenstand des Interesses ist vorab bekannt und es existiert ein geometrisches Modell des Objektes. Solche Verfahren sollen hier nicht weiter betrachtet werden, weil das im Wesentlichen untersuchte Segmentierungsverfahren eine echte Partitionierung im Sinne der in Abschnitt 2.1.5 (Seite 13) gegebenen Definition liefert.

4.2 Segmentierungsverfahren

In Erinnerung an Abschnitt 2.3 sollen noch einmal die allgemeinen Anforderungen an eine Segmentierung zusammen gefaßt werden:

- Vollständigkeit: Jedes Pixel wird mindestens einem Segment zugeordnet.
- Überdeckungsfreiheit: Kein Pixel wird mehr als einem Segment zugeordnet.
- Zusammenhang: Jedes Segment bildet ein zusammenhängendes Gebiet.

4.2.1 Kantendetektierende Verfahren

Ausgehend von der Idee des „Durchzeichnens“ versucht man, in einem Bild zusammenhängende Streckenzüge zu finden, die die Begrenzungen der Segmente bilden sollen. Sobel-, Kirsch- und Laplace²⁴-Operatoren dienen dazu, Kanten in einem Bild zu verstärken. Die Hough-Transformation²⁵ erzeugt aus diesen Kantenabschnitten einen Dualraum, der als Koordinaten die Parameter der interessierenden geometrischen Objekte wie Geraden oder Kreise enthält. Mit Kantenverfolgungsalgorithmen können fehlende Kantenstücke ergänzt werden. Beispielsweise bieten sich Snakes, „active contour models“ an, um fehlende Kanten zu ergänzen und aus den Kanten ein Objekt zu generieren. (Kass et al. [1987, 1988], ein schneller Algorithmus wird von Williams und Shah [1992] beschrieben.)

Diese Verfahren sind für die Auswertung von Fernerkundungsbildern mit eher schwachen Kontrasten und wenig zusammenhängenden Flächen seltener geeignet. Sie finden Verwendung in Spezialfällen wie der Erkennung von Straßenzügen oder Gebäudekanten.

4.2.2 Pixelorientierte Verfahren

Bei pixelorientierten Verfahren wird für jedes Pixel einzeln entschieden, zu welcher Klasse es gehören soll. Die Klasseneinteilung muß nicht vorab bekannt sein, sie kann auch durch ein Schwellwertermittlungsverfahren, z. B. durch eine Histogrammanalyse, erzeugt werden. Pixelorientierte Verfahren erzeugen vollständige und überdeckungsfreie Segmentierungsergebnisse, führen aber in der Regel zu nicht zusammenhängenden Segmentierungen. Der fehlende Zusammenhang führt zum sogenannten Salt-and-Pepper-Effekt²⁶. Das einfachste Verfahren ist die Binärisierung eines Bildes anhand eines Schwellwertes. Andere bekannte Verfahren sind das *k-mean*-Verfahren und der ISODATA²⁷-Algorithmus. Wegen ihrer weiten Verbreitung sollen sie an dieser Stelle vorgestellt werden. Beide Verfahren sind unüberwachte Klassifikationsalgorithmen, die Cluster bilden. Ein Cluster ist eine Anhäufung von Punkten, in diesem Fall von Vektoren in einem n -dimensionalen Merkmalsraum. Als Merkmale dienen die Intensitätswerte eines Pixels in den verschiedenen Spektralkanälen, aber auch abgeleitete Werte wie der NDVI²⁸ oder Texturmerkmale wie Grauwertdifferenzen zu den Nachbarpixeln. Um Anhäufungen zu beschreiben, ist ein Abstandsmaß erforderlich, mit dem die Abstände zwischen den Punkten im Merkmalsraum beschrieben werden. Viele solcher Abstandsmaße oder Distanzfunktionen wurden beschrieben²⁹. Als Beispiele seien genannt die L_p -Distanz oder Minkowski-Metrik:

²⁴ vergl. Abschnitt 5.1.3.1 (Seite 48)

²⁵ <http://de.wikipedia.org/wiki/Hough-Transformation>

²⁶ In an sich zusammenhängenden Gebieten finden sich sehr häufig kleinste Gebiete, Einsprengsel, die einer anderen Klasse zugeordnet wurden oder die Pixel in einem Gebiet werden durchgängig auf mehr als eine Klasse aufgeteilt: Salz- und Pfeffer-Effekt

²⁷ Iterative Self-Organizing Data Analysis Technique

²⁸ Normalized Differenced Vegetation Index

²⁹ <http://de.wikipedia.org/wiki/Distanzfunktion>

$$L_p(x, y) = \sqrt[p]{\sum_{i=1}^n |x_i - y_i|^p}, \quad x, y \in \mathbf{R}^n \quad (4.1)$$

Mit $P = 1$ erhält man die City-Block oder Manhattan-Distanz, $P = 2$ ergibt den Euklidischen Abstand und für $\lim P \rightarrow \infty$ erhält man die Maximum- oder Chebyshev-Distanz.

Häufig wird auch die Mahalanobis-Distanz³⁰ verwendet: Dazu muß eine $n \times n$ -Kovarianzmatrix S für die verschiedenen Vektoren x (die Punkte im n -dimensionalen Merkmalsraum) erstellt oder geschätzt werden. Die Kovarianzmatrix kann für jede Klasse verschieden oder für alle Klassen gleich sein. Die Verwendung der Kovarianzmatrix bewirkt, daß die Abstandsmaße in den verschiedenen Dimensionen unterschiedlich gewichtet werden können. Die Mahalanobis-Distanz $d(x, y)$ zwischen den Vektoren x und y lautet dann:

$$d(\mathbf{x}, \mathbf{y}) = (\mathbf{x} - \mathbf{y})^T \mathbf{S}^{-1} (\mathbf{x} - \mathbf{y}) \quad (4.2)$$

Gelegentlich wird wegen der Vergleichbarkeit mit anderen Abstandsmaßen die Quadratwurzel $\sqrt{d(\mathbf{x}, \mathbf{y})}$ als Mahalanobis-Distanz bezeichnet. Ist die Kovarianzmatrix die Einheitsmatrix, dann ist die Quadratwurzel aus der Mahalanobis-Distanz der Euklidische Abstand.

4.2.2.1 *k-mean-Algorithmus*

Der Untersucher gibt die Anzahl k der gewünschten Klassen vor, die ungefähren Mittelpunkte der Cluster im Merkmalsraum und die Anzahl der Iterationen. Der Algorithmus berechnet die Abstände der Vektoren x zu den Mittelpunkten und ordnet die Vektoren den nächstgelegenen Mittelpunkten zu. In einem iterativen Verfahren werden die Mittelpunkte der Cluster neu berechnet und so das Ergebnis optimiert. Werden in einem Schritt Cluster nicht mit Vektoren besetzt, dann entfallen sie im nächsten. Das Verfahren ist schnell, aber stark von der Separierbarkeit der Cluster und der ursprünglichen Verteilung der geschätzten Mittelpunkte abhängig. Werden keine Vorgaben gemacht, können diese auch vom Programm geschätzt werden.

4.2.2.2 ISODATA-Algorithmus

Der ISODATA-Algorithmus entspricht dem *k-mean-Algorithmus*. Allerdings ist hier die Anzahl der Cluster nicht festgelegt, sondern der Algorithmus kann Cluster teilen, wenn die zugehörige Punktwolke besser durch zwei Zentren zu beschreiben ist. Cluster werden zusammengelegt, wenn der Abstand der Zentren und die Verteilung der Punktwolken dies nahelegen und Cluster werden gelöscht, wenn sie mit zu wenigen Punkten besetzt sind.

4.2.3 Regionenorientierte Verfahren

Im Gegensatz zu den pixelorientierten Verfahren betrachten regionenorientierte Verfahren von vornherein zusammenhängende Gebiete. Dabei kann vom kleinen zum großen (bottom-up) oder vom großen zum kleinen (top-down) vorgegangen werden.

³⁰ Eine ausführliche Herleitung und Begründung der Vorteile findet sich unter <http://www.wu-wien.ac.at/usr/h99c/h9951826/distance.pdf>

4.2.3.1 Region Growing Verfahren

Region Growing Verfahren starten bei einzelnen Pixeln, den so genannten Seedpixel, und vergleichen sie mit ihren Nachbarn. Ist die Ähnlichkeit hinreichend groß, dann wird das Nachbarpixel in das vom Seedpixel ausgehende Segment aufgenommen. Die Seedpixel können dabei regelmäßig oder wahllos über das Bild verteilt werden, sie können vom Bearbeiter von Hand vorgegeben werden, oder sie können vom Programm nach einer Suche homogener Bereiche vorgegeben werden. Um texturierte Flächen segmentieren zu können, kann statt auf den Originaldaten auch auf abgeleiteten Daten gearbeitet werden.

4.2.3.2 Watershed Verfahren

Ein erfolgreich angewandtes Verfahren ist die Wasserscheiden- (*watershed*-) Transformation. Die Grauwerte eines Bildes werden als Höhendaten eines Geländes aufgefaßt. Der Algorithmus sucht anhand der Gradienten die Täler in diesem Gelände auf. Diese Täler werden so lange mit Wasser aufgefüllt, bis sie ihr Nachbartal berühren. Die so gefundenen Wasserscheiden werden als die Segmentgrenzen aufgefaßt. Dieses Verfahren wurde vielfältig optimiert und vermeidet den in Abschnitt 4.2.2 angesprochenen Salt-and-Pepper-Effekt der pixelorientierten Verfahren. Ein ausgereiftes Verfahren, Täler bzw. Gipfel zu finden, stellen Cheng [1995], Comaniciu und Meer [1997, 2002], und Comaniciu [2000] mit dem Mean-Shift-Verfahren vor. Das ganze Segmentierungsprogramm wurde von Georgescu unter der Bezeichnung EDISON (Code for the Edge Detection and Image SegmentatiON system) zur Verfügung gestellt (<http://www.caip.rutgers.edu/riul/riul.html>). Weitere Verbesserungen wurden von Wang und Suter [2003] vorgeschlagen, um den Algorithmus noch robuster zu machen.

4.2.3.3 Split and Merge Verfahren

Split and Merge Verfahren sind häufig blockorientiert. Das ganze Bild wird analysiert und so lange in Blöcke aufgespalten, bis diese einem Homogenitätskriterium genügen. Anschließend können benachbarte Blöcke aus vorangehenden Aufteilungen wieder zusammengefügt werden. Verfeinerungen ergeben sich, wenn im Anschluß die treppenförmigen Blockgrenzen wieder geglättet werden.

4.3 eCognition

eCognition beschreibt sich als Werkzeug zu objektorientierten Klassifizierung. Objektorientiert bedeutet in der Informatik, daß die früher übliche Trennung zwischen Daten und den zu ihrer Verarbeitung erforderlichen Algorithmen aufgehoben wird. In eCognition wird darunter verstanden, daß innerhalb einer Analyse eine Datenbank erstellt wird, in der zu jedem Objekt, sei es Segment oder Klasse, ein Katalog von Merkmalen gespeichert wird. Außerdem werden die topologischen Beziehungen zu Nachbarn und in der Hierarchie unter- und übergeordneten Objekten so aufgezeichnet, daß sie für weitere Analysen ausgewertet werden können (de Kok [2001]).

Im folgenden Abschnitt werden einige wenige Aspekte beschrieben, die beim Erstellen dieser Arbeit wichtig waren. Diese Beschreibungen sollen und können nicht das 486 Seiten starke Benutzerhandbuch (Baatz et al. [2004]) ersetzen. Auf die vielfältigen Möglichkeiten, Objekteigenschaften zur Klassenbeschreibung zu nutzen, soll nicht eingegangen werden.

4.3.1 Algorithmus

Die Klassifikation erfolgt bei eCognition in mehreren Schritten. Zunächst gibt der Benutzer die Parameter für die Segmentierung vor. Eine Segmentierung kann auch mehrfach in verschiedenen Auflö-

sungsstufen erfolgen, die dann hierarchisch strukturiert werden. Anschließend wird klassifiziert, wobei die Klassenbeschreibung ebenfalls hierarchisch strukturiert sein kann. In der Klassenbeschreibung kann auch angegeben werden, daß nur Objektbegrenzungen einer bestimmten Segmentierungsebene verwendet werden sollen. Das Klassifizierungssystem wurde in dieser Arbeit nur genutzt, um die Segmentierungsergebnisse darzustellen, deshalb wird darauf nicht weiter eingegangen. Für die Segmentierung wichtig und verschieden von bisherigen Ansätzen ist der Fractal Net Evolution Approach (Baatz und Schäpe [1999]).

4.3.1.1 Fractal Net Evolution Approach

Im Prinzip gehört der Segmentierungsansatz zu den in Abschnitt 4.2.3.3 beschriebenen Split- und Merge-Verfahren. Ausgehend von der Betrachtung des einzelnen Pixels wird dieses mit seinen Nachbarn verglichen und anhand des Homogenitätskriteriums entschieden, ob es mit zusammen mit seinem Nachbarn zu einem Segment gehören soll. Der wesentliche Unterschied zu den bisher verwendeten Verfahren liegt in der Art und Weise, wie dies geschieht. Das für eCognition entwickelte Unähnlichkeitsmaß lautet nach Baatz und Schäpe [2000]:

$$h_{diff} = \sum_c w_c (n_1 (h_{mc} - h_{1c}) + n_2 (h_{mc} - h_{2c})) \quad (4.3)$$

Hier werden für jedes Layer c die Unähnlichkeitsmaße der beiden Segmente h_1 und h_2 sowie des vereinigten Segments h_m berechnet und die Differenzen $(h_{mc} - h_{1c})$ bzw. $(h_{mc} - h_{2c})$ mit der Zahl der Pixel n_1 bzw. n_2 der Ausgangssegmente multipliziert. Schließlich wird die vom Benutzer vorgegebene Gewichtung w_c für jedes Layer berücksichtigt.

Die meisten Segmentierungsverfahren versuchen, im Sinne eines *global mutual best fitting* das Segmentierungsergebnis über das ganze Bild zu optimieren. Deshalb können sie Segmente sehr unterschiedlicher Größe erzeugen. Im Gegensatz dazu betrachtet der Algorithmus in eCognition nur die direkten Nachbarn im Sinne eines *local mutual best fitting*. Nach einem entsprechend dem Scale-Parameter (für den eine rationale Zahl akzeptiert wird) gewählten n sucht er in festgelegter Reihenfolge jedes n -te Pixel für einen weiteren *merge*-Schritt auf. Als geeignete Heterogenitätskriterien werden in erster Linie die Farbe in Form der Varianz des mittleren Grauwerts oder der Standardabweichung des mittleren Grauwerts des Objektes genannt. In zweiter Linie sind dies die unter Abschnitt 4.3.2.3 genannten Formkriterien Compactness und Smoothness.

4.3.2 Einstellbare Parameter

Entscheidend für das Segmentierungsergebnis ist die richtige Einstellung der wählbaren Parameter. Als erstes ist an dieser Stelle die Auswahl und Wichtung der zur Segmentierung verwendeten Kanäle zu nennen. Kanäle, die die differenzierenden Informationen nicht enthalten, können ausgeblendet werden. Andere Daten als Bilddaten können hinzugenommen werden. Welche Kanäle verwendet werden, hängt von den zur Verfügung stehenden Daten und den gesuchten Informationen ab.

Anschließend müssen die Homogenitätskriterien festgelegt werden. Entsprechend dem Ansatz von eCognition, daß unterschiedliche Dinge in verschiedenen Auflösungen am Besten zu erkennen sind, ist das erste einzustellende Kriterium der Scale-Parameter.

4.3.2.1 Scale

Der Scale-Parameter ist ein abstrakter Parameter, der die zulässige Heterogenität der gebildeten Segmente begrenzt. In heterogener Umgebung wird die Größe der gebildeten Segmente kleiner sein, als in bei homogenen Daten. Die Wirkung bei völlig homogenen Daten wird in Abschnitt 4.4 näher untersucht werden. Die bei Programmstart vorgegebene Standardeinstellung zeigt einen Wert von zehn.

4.3.2.2 Color and Shape

Beide Begriffe bezeichnen die Pole eines Kriteriums. Der Benutzer kann eine Gewichtung vornehmen im Intervall zwischen einem Minimum von 0 für Shape entsprechend 1 für Color bis 0,9 für Shape bzw. 0,1 für Color. Die vom Programm vorgegebene Standardeinstellung ist 0,1 für den Shape-Parameter.

Color kann nicht kleiner gewählt werden, weil ohne spektrale Information keine auswertbare Information dem Bild entnommen werden kann. Andererseits bewirkt eine stärkere Wichtung des Shape-Kriteriums, daß stark texturierte Daten zusammengefaßt werden können.

Das Shape-Kriterium wiederum wird näher definiert durch die Pole Compactness und Smoothness, die die Form der gebildeten Segmente beeinflussen.

4.3.2.3 Compactness and Smoothness

Diese beiden Pole zur Beschreibung der gewünschten Form der Segmente sind nicht antagonistisch. Dennoch hat der Nutzer die Wahl der Wichtung im Intervall 0 ... 1 der Kriterien gegeneinander. Voreingestellt ist ein Wert von 0,5, also eine gleichmäßige Wichtung. Die Wahl eines höheren Wertes der Compactness wird für RADAR-Daten empfohlen. Die Heterogenität der Compactness wird berechnet als das Verhältnis zwischen der effektiven Länge der Begrenzung eines Segmentes und der Wurzel aus der Zahl seiner Pixel:

$$h_{compact} = \frac{l}{\sqrt{n}} \quad (4.4)$$

Das kompakteste Objekt in diesem Sinne ist demnach ein Quadrat.

Die Heterogenität der Smoothness wird berechnet aus der Länge der Begrenzung eines Segmentes und der Länge des Umfangs des kleinstmöglichen umschließenden Rechtecks parallel zu den Bildkanten (Bounding Box).

$$h_{smooth} = \frac{l}{b} \quad (4.5)$$

Das glatteste Objekt in diesem Sinne ist also ein Rechteck parallel zu den Bildkanten, unabhängig vom Verhältnis der Seitenlängen, während ein längliches Rechteck diagonal zu den Bildkanten als wenig *smooth* verstanden wird.

Im Falle eines Quadrates sind also Compactness und Smoothness maximal. Die Untersuchungsergebnisse in Abschnitt 4.4 werden zeigen, daß diese Angaben die Entscheidungskriterien nicht vollständig beschreiben.

4.3.3 Benutzeroberfläche

In diesem Abschnitt sollen nur die Elemente der Benutzeroberfläche beschrieben werden, die für das Verständnis dieser Untersuchung wichtig sind. Das Programm ist komplex und bietet zahllose weitere Dialoge und Einstellmöglichkeiten.

4.3.3.1 Ebenenwichtung

Nachdem alle für eine Segmentierung eventuell erforderlichen Bilder in das Projekt aufgenommen wurden, kann vor der Segmentierung ausgewählt werden, welche Layer mit welcher Gewichtung w_c aus Gleichung (4.3) in die Auswertung eingehen sollen.

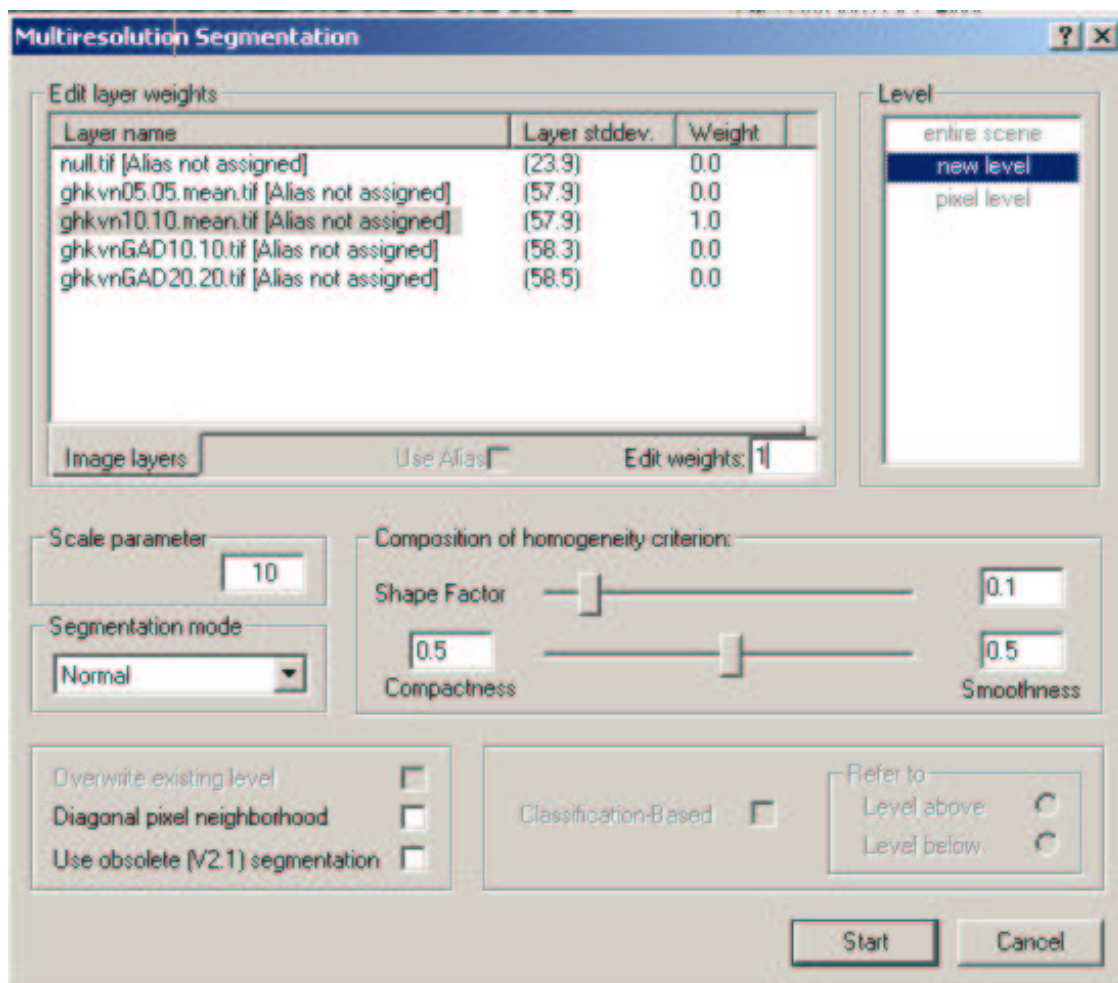


Abbildung 4.1

Der Benutzerdialog in eCognition erlaubt zunächst die Auswahl der für die Segmentierung erforderlichen Bildebenen auch aus verschiedenen Dateien. Anschließend werden der Scale-Parameter und die Parameter für das Homogenitätskriterium eingestellt.

4.3.3.2 Parameter

Anschließend werden der Scale-Parameter und die Parameter für das Homogenitätskriterium eingestellt. Nachdem eine Segmentierung auf Pixelebene stattgefunden hat, kann in weiteren Segmentierungsschritten im Feld Level festgelegt werden, ob eine weitere Segmentierungsebene in der Hierarchie ober- oder unterhalb der bestehenden eingefügt werden soll oder ob die bestehende überschrieben werden soll. Im Feld Segmentation mode kann für folgende Segmentierungsschritte außer der normalen Segmentierung eingestellt werden, ob unterhalb der Pixelebene alternativ nach einer *line analysis* weiter segmentiert werden soll. Oberhalb der Pixelebene kann auch nach spektraler Differenz segmentiert werden. Nach einer Klassifizierung können auch deren Ergebnisse für eine weitere (größere) Segmentierung verwendet werden.

4.3.3.3 Membershipfunktion

eCognition bietet vom Benutzer veränderbare Membershipfunktionen an. Damit kann innerhalb der Klassenbeschreibung festgelegt werden, welche Segmente im Rahmen einer fuzzy-Zuordnung einer Klasse zugeordnet werden sollen.

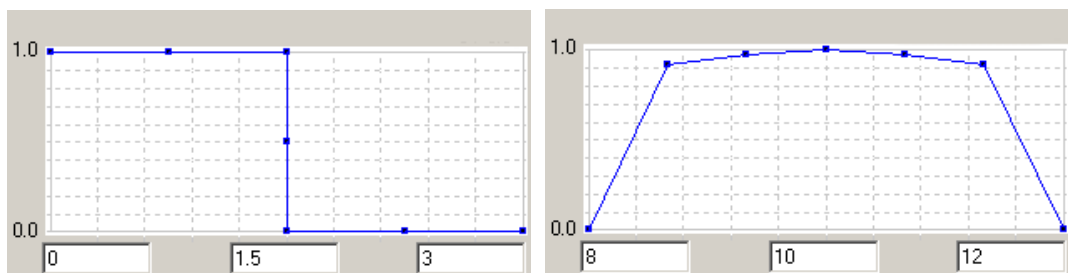


Abbildung 4.2

Einstellung der Membershipfunktion in eCognition.

Die Abbildung zeigt links eine treppenförmige Membershipfunktion. So werden nur Segmente mit Werten im Bereich 0 ... 1 einer bestimmten Klasse zugeordnet. Rechts ist eine wannenförmige Membershipfunktion abgebildet, mit der die Klassenzugehörigkeit auf Segmente mit Werten zwischen 9 und 11 beschränkt wird.

Abbildung 4.2 stellt zwei der möglichen Ausprägungen vor, wie sie in dieser Arbeit Verwendung finden. Die Ordinate gibt die Wahrscheinlichkeit der Klassenzugehörigkeit des Segmentes nach diesem Auswahlkriterium im Rahmen der Rechenregeln der fuzzy-Mengen an. In der Abszisse gibt der Benutzer die linke und rechte Begrenzung des Bereichs vor. Die Größe der zulässigen Werte hängt dabei vom Auswahlkriterium ab. Soll die Größe eines Objektes beschrieben werden, kann der Maximalwert die Zahl der Pixel im Bild sein. Wenn in den Bildinformationen ein Bildmaßstab angegeben ist, können Objektgrößen auch metrisch angegeben werden. Wird der Winkel eines Objektes beschrieben, liegt der Maximalwert bei 180 (Grad).

4.3.3.4 Klassenbeschreibung

In der Klassenbeschreibung (Abbildung 4.3) wird die Klasse definiert. Die Klasse erhält einen eindeutigen Namen. Ihr wird eine Farbe für die Darstellung zugewiesen. Die Klasse kann von einer oder mehreren Elternklassen abgeleitet werden und so deren Eigenschaften erben. Wird sie von mehreren Elternklassen abgeleitet, muß angegeben werden, welcher Elternklasse sie bei der Darstellung der

Gruppenhierarchie zugeordnet werden soll. Sie kann als abstrakte Klasse gekennzeichnet werden. Solche, wie auch inaktive Klassen werden im Klassifikationsprozeß nicht ausgewertet. Dann werden die Membershipfunktionen und deren Verknüpfungen angegeben, die die Klasse bestimmen. Die Icons vor den Membershipfunktionen stellen diese nicht dar, sondern geben einen Hinweis auf die Art der Funktion.

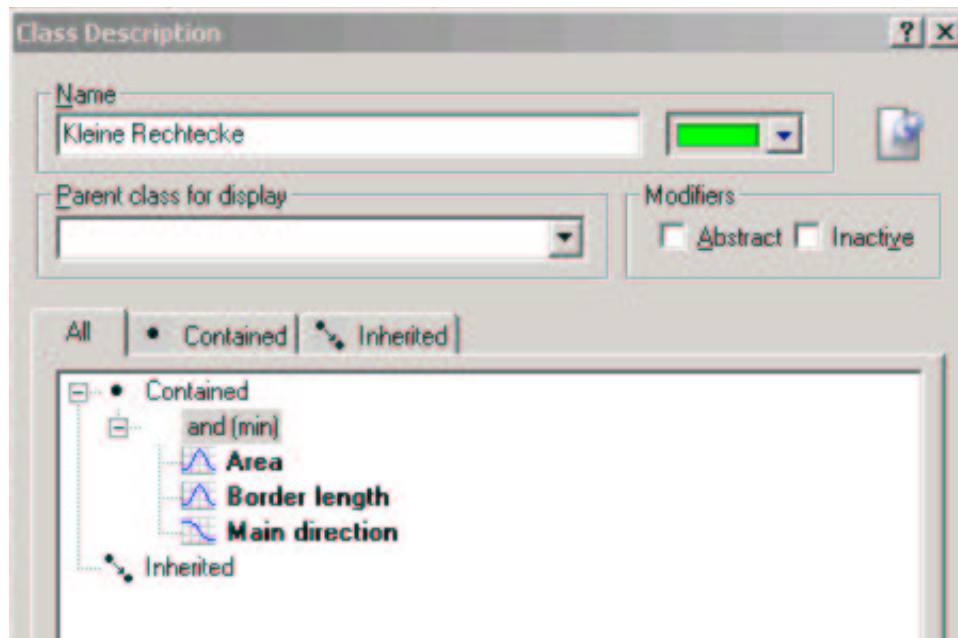


Abbildung 4.3

Klassenbeschreibung in eCognition.

In der Klassenbeschreibung werden die Membershipfunktionen und deren Verknüpfungen angegeben, die die Klasse bestimmen. Die Icons vor den Membershipfunktionen stellen diese nicht dar, sondern geben nur einen Hinweis auf die Art der Funktion.

Nach diesen einleitenden Erläuterungen folgt nun ein Abschnitt mit den Ergebnissen einer ersten orientierenden Untersuchung der Wirkung der erwähnten Parameter.

4.4 Untersuchung von Bildern mit konstantem Grauwert

In diesem Abschnitt wird zunächst die Wirkung der Parametereinstellung des Scale-, Color-, Shape-, Compactness- und Smoothness-Parameters für den einfachsten Fall, nämlich die Segmentierung einer weißen Fläche untersucht. Von Anwendern war beklagt worden, daß die Segmentgrenzen sehr zufällig und regellos verliefen und minimale Änderungen am Bild wie das Entfernen einer Pixelreihe zu erheblich veränderten Segmentierungsergebnissen führten. Diese Probleme im Segmentierungsalgorithmus sollen mit der Version 4.0 behoben sein.

Bislang existieren keine Veröffentlichungen zum Verhalten des Programms auf die Eingabe einfach strukturierter Flächen. Als einfachster Fall wurde deshalb eine uniforme weiße Fläche gewählt und die Ausgabe bei systematischer Variation der vom Benutzer vorzugebenden Parameter dokumentiert.

4.4.1 Segmentierung der weißen Fläche

Es wurde ein Testbild generiert mit einer Kantenlänge von 500 x 500 Pixel. Dieses Testbild wurde mit eCognition 4.0 segmentiert. Der Segmentierungsalgorithmus wurde im Vergleich zur Vorgängerversion so verändert, daß bei sehr geringen Bildänderungen die Segmentierung ein gleiches Ergebnis bringen soll. Dennoch sind die gebildeten Segmente nicht nur Quadrate, und deutliche Moiré-Muster sind erkennbar.

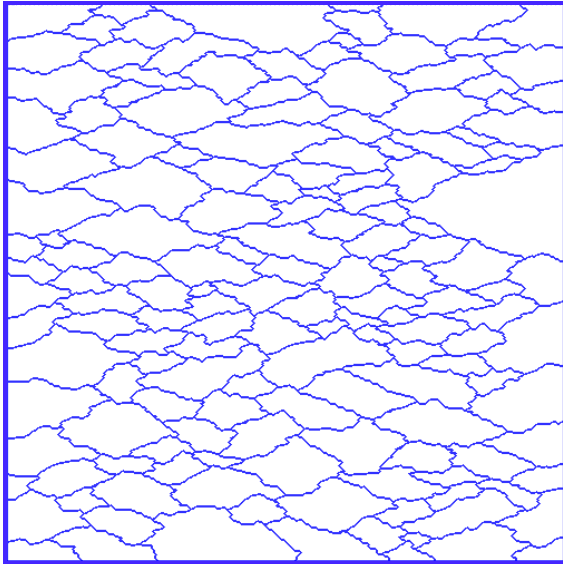


Abbildung 4.4

Segmentierung einer weißen Fläche mit den im Programm voreingestellten Parametern (Scale=10, Shape=0,1, Compactness=0,5)
alter Algorithmus der Versionen 2 und 3

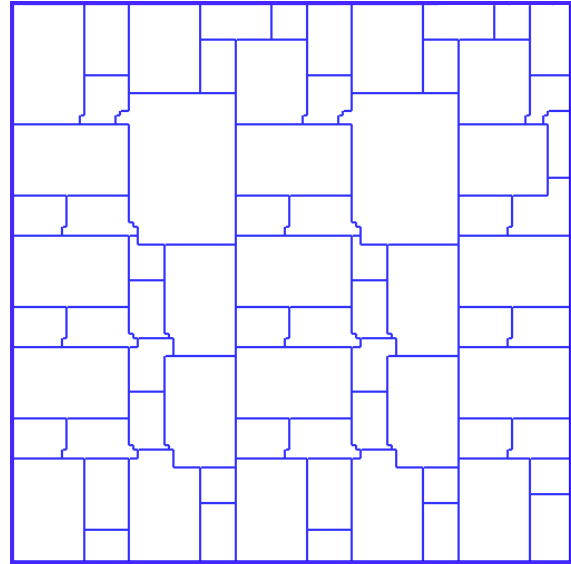


Abbildung 4.5

neuer Algorithmus der Version 4

Abbildung 4.4 zeigt die Segmentierung nach dem alten Algorithmus der Versionen 2 und 3, Abbildung 4.5 zeigt die Segmentierung nach dem Algorithmus der Version 4. Die kleinen Verschiebungen am Rand vieler Rechtecke sind typisch und kommen bei fast jeder Parameterwahl vor.

Nach den Empfehlungen soll man erstens den größtmöglichen Scale-Parameter verwenden und zweitens das Color-Kriterium so groß wie möglich wählen. Nach diesen Empfehlungen wurde das Testbild immer dann richtig mit einem einzigen, die gesamte Fläche umfassenden Segment segmentiert, wenn das Color-Kriterium auf 1 gestellt wurde. Um jedoch einen Eindruck zu erhalten von den internen Abläufen, die nach wie vor nicht offengelegt sind, wurden die Segmentierungsparameter variiert:

Der Scale-Parameter wurde eingestellt auf die Werte 2, 5, 10, 20, 50, der Shape-Parameter (gegen das Color-Kriterium) wurde variiert mit den Werten 0, 0,25, 0,5, 0,75 und dem Maximum von 0,9, schließlich die Compactness (gegen die Smoothness) mit den Werten 0, 0,25, 0,5, 0,75, 1. Bei einem Shape-Parameter von 0 läßt sich dieser Wert nicht variieren.

Insgesamt ergaben sich so 105 Segmentierungsergebnisse. Die kleinste Anzahl von Segmenten war erwartungsgemäß 1 mit einer Segmentgröße von 250.000 Pixeln. Dieses Ergebnis wurde immer dann erzielt, wenn bei beliebiger Werten für den Scale-Parameter die Segmentierung nur nach der Farbe vorgenommen wurde (Shape-Parameter = 0). Die größte Anzahl von Segmenten betrug 62.500, hier war jedes einzelne Segment ein Quadrat mit der Kantenlänge 2 Pixel. Dieses Ergebnis wurde erzielt mit

einem Scale-Parameter von zwei und Shape-Parameter von 0,9 und 0,75. Die Verteilung der Größen der Segmentflächen war recht unterschiedlich. Häufig hatte die Hälfte oder mehr Flächen die Größe des Medians, und besonders kleine Flächen waren mehr als „Reststücke“ entstanden, in anderen Fällen waren die Größen der Segmentflächen ziemlich gleichmäßig über alle Klassen verteilt. Das Verhältnis zwischen Maximalwert der Fläche und dem Medianwert betrug maximal 5,5, das Verhältnis zwischen Medianwert und Minimum lag maximal bei 16, bedingt durch wenige kleine „Reststücke“. Abbildung 4.6 zeigt ein Beispiel einer eher gleichmäßigen Verteilung, hier finden sich im Bereich zwischen 2000 und 3000 Pixeln höchstens 8 mit der gleichen Pixelzahl. Das zugehörige Histogramm ist in Abbildung 4.8 dargestellt. Die Abszisse wurde logarithmisch im Sinne der Normreihe R10 mit dem Streckfaktor 200 geteilt.

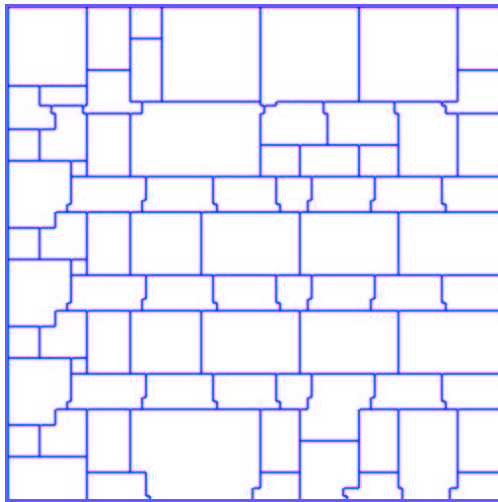


Abbildung 4.6

Segmentierungsergebnis eCognition

Eingestellte Parameter:

Scale=50, Shape=0,9, Compactness=1

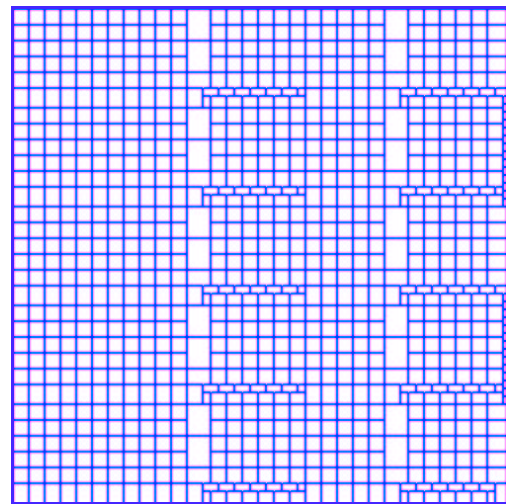


Abbildung 4.7

Scale=5, Shape=0,25, Compactness=1

oder auch

Scale=5, Shape=0,5, Compactness=0,5. Beide Einstellungen ergeben eine identische Segmentierung.

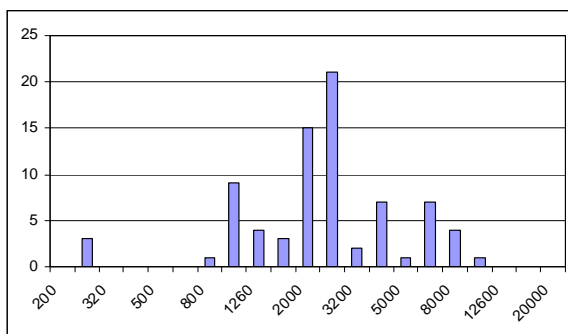


Abbildung 4.8

Histogramm zu Abbildung 4.6

Abszisse: Segmentgröße in Pixel,

Ordinate: Anzahl der Segmente.

Gleichmäßige Verteilung.

Im Bereich zwischen 2000 und 3000 Pixel finden sich höchstens 8 Segmente mit der gleichen Pixelzahl.

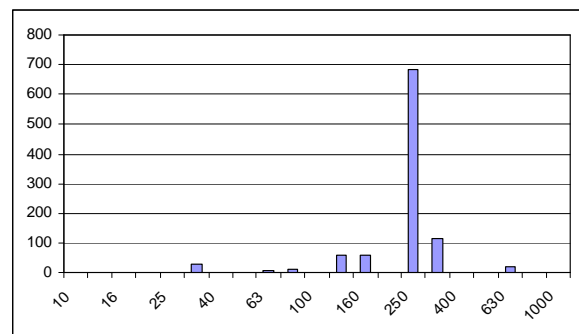


Abbildung 4.9

Histogramm zu Abbildung 4.7

Abszisse: Segmentgröße in Pixel,

Ordinate: Anzahl der Segmente.

Medianzentrierte Verteilung mit einem Anteil von 70% 256 Pixel großer Segmente.

Abbildung 4.10 zeigt den bei mittlerer Parameterwahl von Shape und Compactness bei verschiedener Wahl des Scale-Parameters recht gleichmäßigen Quotienten des Medians der Anzahl der gebildeten Segmente durch das Quadrat des Scale-Parameters. Die Abszisse gibt den Scale-Parameter an, die Kurvenbeschriftungen bezeichnen die Werte für (Shape / Compactness).

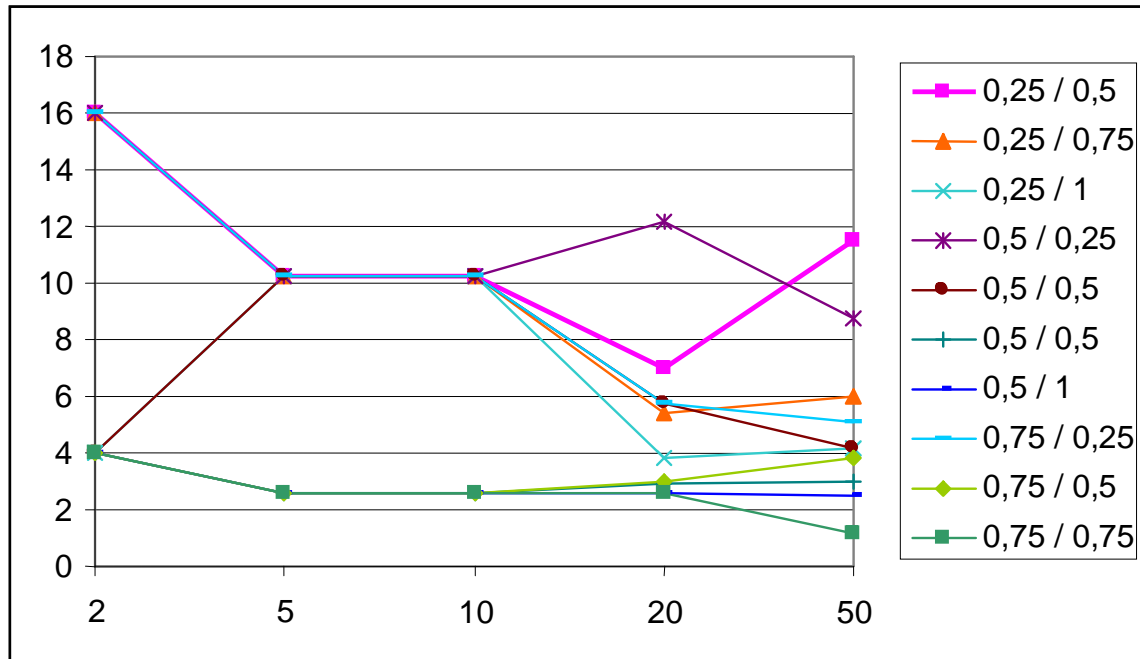


Abbildung 4.10

Die Abszisse gibt die Werte des Scale-Parameters an, die Kurvenbeschriftungen bezeichnen die Werte für (Shape / Compactness).

Die Abbildung zeigt auf der Ordinate den bei mittlerer Parameterwahl von Shape und Compactness bei verschiedener Wahl des Scale-Parameters recht gleichmäßigen Quotienten Median/Scale².

4.4.2 Einfluß des Rauschens

In den folgenden Abbildungen wird gezeigt, wie sich das Segmentierungsergebnis von eCognition bei gleichbleibender Parametereinstellung in Abhängigkeit vom zugefügten Rauschen verändert. Alle Bilder haben die Größe 500 x 500 Pixel und zunächst einen einheitlichen Grauwert von 127. Für die in Abbildung 4.11 bis Abbildung 4.14 vorgestellten Segmentierungsergebnisse wurden folgende Parameter eingestellt: Scale=10, Shape=0,1, Compactness=0,5.

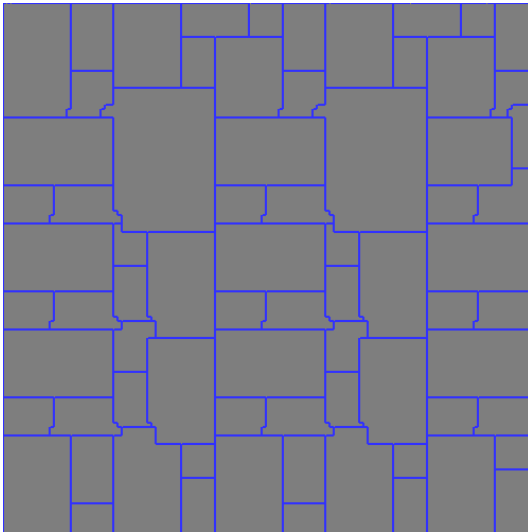


Abbildung 4.11

Rauschen = 0, identische Segmentierung wie in Abbildung 4.5. Es wurden 74 Segmente gebildet.

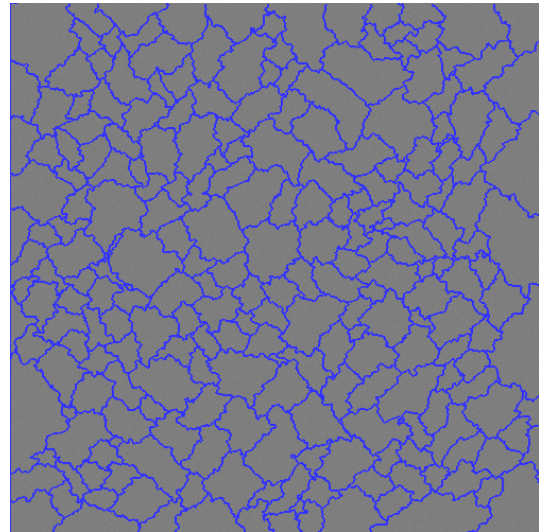


Abbildung 4.12

Rauschen mit $\sigma = 2$ zugefügt, überraschende Ähnlichkeit mit Abbildung 4.4. Es wurden 239 Segmente gebildet.

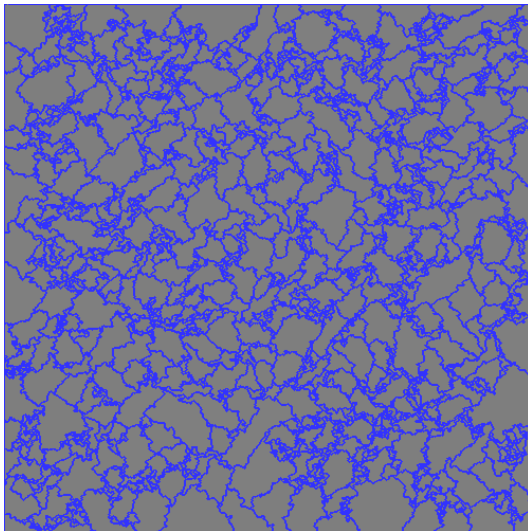


Abbildung 4.13

Rauschen mit $\sigma = 5$ zugefügt, 784 Segmente

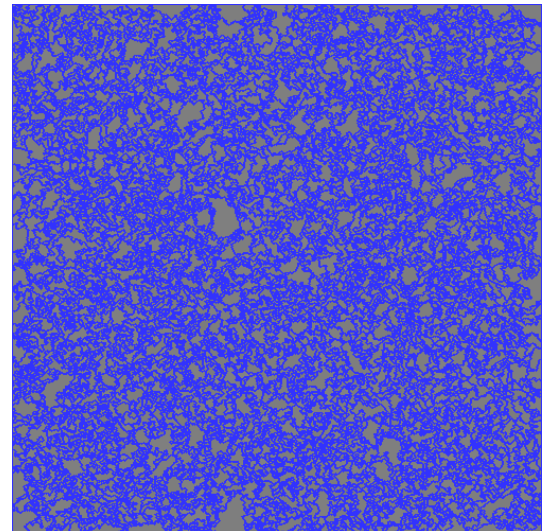


Abbildung 4.14

Rauschen mit $\sigma = 10$ zugefügt, 2821 Segmente

Aufgrund der niedrigen Einstellung des Shape-Parameters, der einer starken Color-Gewichtung entspricht, wirkt sich die Varianz der Grauwerte im Bild stark aus. Die größtmögliche Einstellung des Shape-Parameters von 0,9 bei ansonsten unverändert belassenen Parametern (Scale = 10, Compactness = 0,5) führt dazu, daß verschieden starkes Rauschen, sobald es überhaupt vorhanden ist, sich nicht mehr sichtbar auf die Segmentierung auswirkt. Dieser Befund wird durch die in Abbildung 4.15 bis Abbildung 4.18 dargestellten Segmentierungsergebnisse dokumentiert.

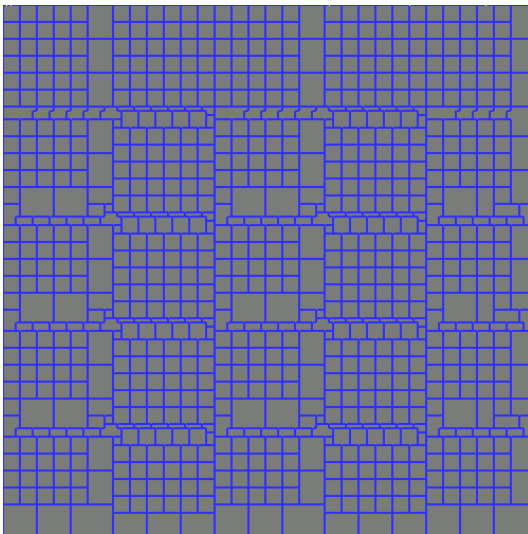


Abbildung 4.15

Unverraushtes Bild, 915 Segmente

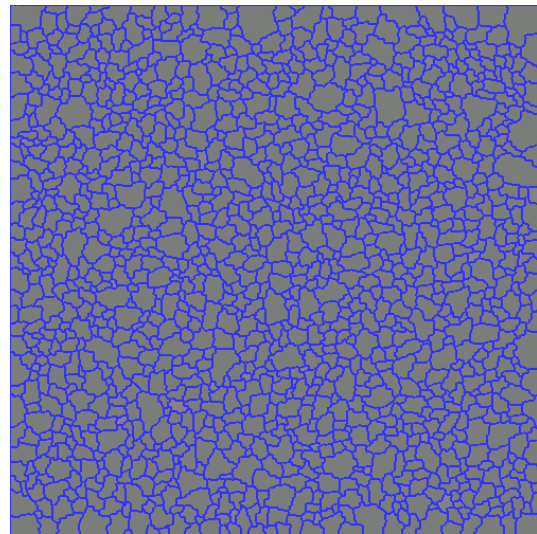


Abbildung 4.16

Rauschen mit $\sigma = 2$ zugefügt, 1303 Segmente

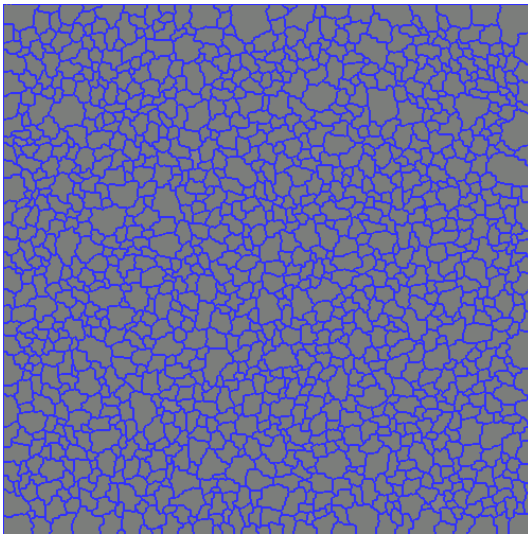


Abbildung 4.17

Rauschen mit $\sigma = 5$ zugefügt, 1286 Segmente

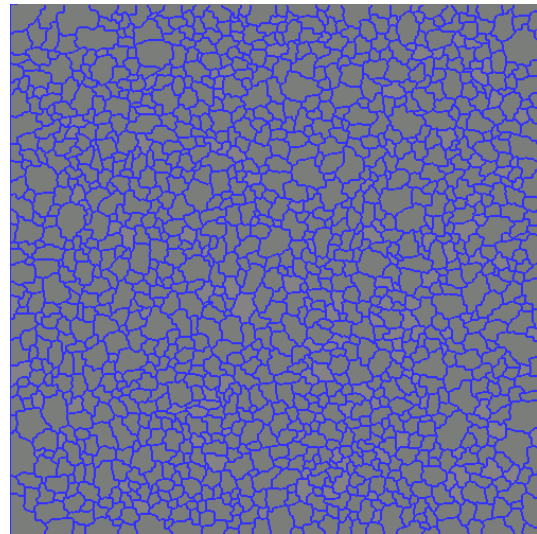


Abbildung 4.18

Rauschen mit $\sigma = 10$ zugefügt, 1277 Segmente

Segmentierungen unterschiedlich stark verrauschter Flächen unterscheiden sich untereinander nicht stärker, als solche mit anderen Initialisierungen des Zufallszahlengenerators. Abbildung 4.19 und Abbildung 4.20 stellen mit ihren logarithmischen Skalen den Zusammenhang zwischen starker Rauschabhängigkeit bei niedrig eingestelltem Shape-Parameter bzw. fehlendem Zusammenhang bei hoch eingestelltem Shape-Parameter noch einmal graphisch dar. In Abbildung 4.19 wurde eine Trendlinie eingefügt mit der Formel $S = 59,1\sigma^{1,54}$. Diese Potenzfunktion ergibt für die dargestellten fünf Werte ein Bestimmtheitsmaß von $r^2 = 0,9848$.

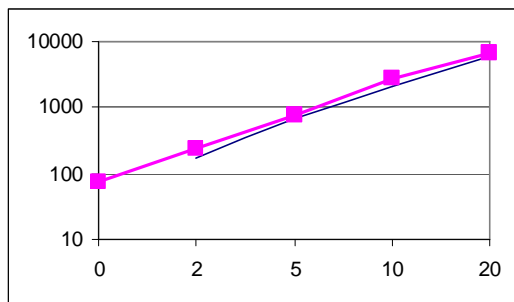


Abbildung 4.19

Abszisse: Standardabweichung σ im Bild, Ordinate: Anzahl S der gebildeten Segmente.

Eine starke Color-Gewichtung führt zu einer potenziell zunehmenden Anzahl S von Segmenten. Die eingefügte Trendlinie mit $S = 59,1\sigma^{1,54}$ hat ein Bestimmtheitsmaß von $r^2 = 0,9848$

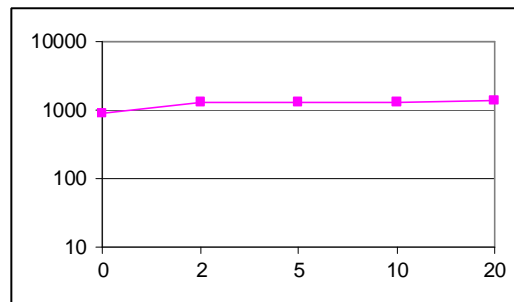


Abbildung 4.20

Abszisse: Standardabweichung im Bild, Ordinate: Zahl der gebildeten Segmente.

Eine starke Shape-Gewichtung führt zu einer konstanten Anzahl von Segmenten trotz zunehmendem Rauschen.

4.4.3 Segmentierung eines Grauwertverlaufs

Im nächsten Versuch soll eine Fläche mit einem gleichmäßigen Grauwertverlauf segmentiert werden. Beabsichtigt ist ein diagonaler Verlauf, so daß die mittleren Grauwerte in einer größeren Fläche vertreten sind als die Extremwerte.

4.4.3.1 Erzeugung eines gleichmäßigen Grauwertverlaufs

Die Aufgabe, einen gleichmäßigen Grauwertverlauf von schwarz nach weiß zu erzeugen, erweist sich als nicht trivial. Das zunächst aus Photoshop® herangezogene Hilfsmittel Farbverlauf³¹ ergibt zwar einen optisch gleichmäßigen Farbverlauf, das Histogramm Abbildung 4.21 zeigt jedoch einen nicht-linearen Verlauf mit ungleich verteilten Grauwerten.



Abbildung 4.21

Darauf hin wurde das Bild mit einem Quick-Basic-Programm direkt generiert. Bei einer Bildgröße von 500 x 500 Pixeln ergaben sich unerwünschte rundungsbedingte Sprünge in der Pixelverteilung, wie das Histogramm in Abbildung 4.22 ausweist.



Abbildung 4.22

³¹ (Ebene → Neue Füllfläche → Verlauf, anschließend Ebene → Mit darunterliegender auf eine Ebene reduzieren)

Die rundungsbedingten Sprünge konnten nach einer Änderung des Rundungsverhaltens entsprechend einem Vorschlag von Kender [1976] deutlich verbessert werden, siehe Abbildung 4.23. Hier wurde erst nach Addition einer kleinen Zufallszahl aus dem Intervall $[-0,5 \dots 0,5]$ gerundet.



Abbildung 4.23

Ein wirklich gleichmäßiges Ergebnis entsprechend dem Histogramm in Abbildung 4.24 war jedoch nur mit einer Bildvergrößerung auf 512×512 Pixel zu erreichen.



Abbildung 4.24

4.4.3.2 Ergebnis

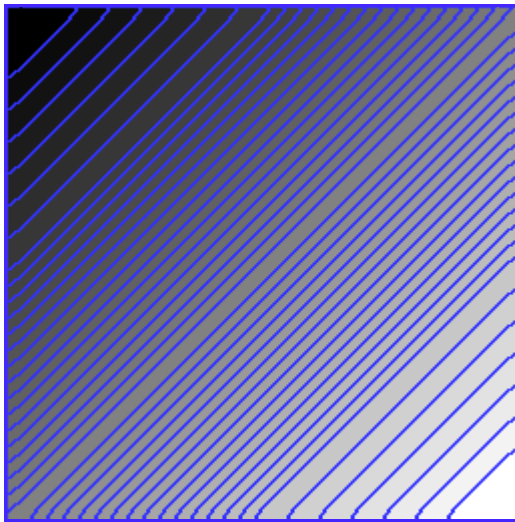


Abbildung 4.25

Segmentierung des homogenen Grauwertverlaufs entsprechend dem in Abbildung 4.24 dargestellten Histogramm

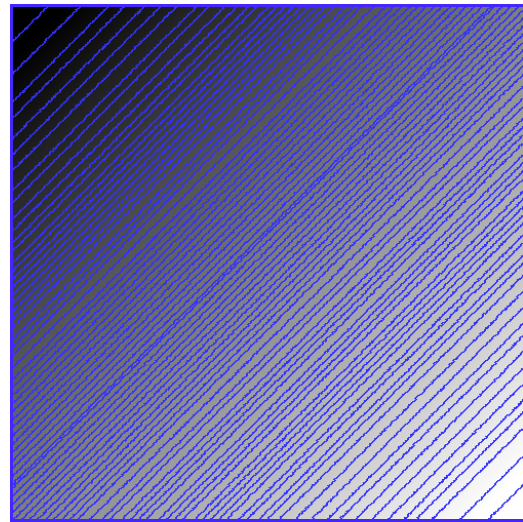


Abbildung 4.26

Segmentierung des stochastisch verbesserten Grauwertverlaufs entsprechend dem in Abbildung 4.23 dargestellten Histogramm

Abbildung 4.25 zeigt die Segmentierung des Verlaufsbildes mit einer Einstellung des Shape-Parameters von 0, d. h. $\text{Color} = 1$ und einer Einstellung des Scale-Parameters von 100. Die meisten Streifen umfassen vier verschiedene Grauwerte, der mittlere breite Streifen enthält 6 Grauwerte. Bei einer Einstellung des Shape-Parameters von 50 sind die Streifen etwa halb so breit, statt 44 werden 91 Segmente gebildet. Segmentierungen mit Einstellungen von 10 und 20 unterscheiden sich im mittleren Bereich nicht. Die Streifen enthalten nur je einen Grauwert, nur die in der dunklen bzw. hellen Ecke enthalten zwei bzw. neun Segmente enthalten zwei bis vier Grauwerte.

Zum Vergleich wird in Abbildung 4.26 die zum Histogramm Abbildung 4.23 gehörige Fläche segmentiert dargestellt. Die Parameter wurden genauso wie für die Segmentierung in Abbildung 4.25 eingestellt. Es ergeben sich 87 statt 44 Segmente. Die optisch nicht wahrnehmbaren Unregelmäßigkeiten führen zu starken Störungen in der Segmentierung trotz des hohen Wertes für den Scale-Parameter.

Dies führt zur Motivation dieser Untersuchung: Wie weit kann das Segmentierungsergebnis verbessert werden, indem das Ausgangsbild vorprozessiert wird, Bildrauschen entfernt wird und Kanten geschärft werden?

Dieser Grauwertverlauf wird später in Kapitel 6 den Hintergrund für die weiteren Untersuchungen bilden.

Zuvor werden jedoch im folgenden Kapitel die verwendeten Filter vorgestellt.

5 Verwendete Filter

Ausgangspunkt der Überlegungen zur Filterung ist die Feststellung, daß sich ein zufällig verteiltes Rauschen durch Mittelwertbildung mehrerer Pixel ausgleichen läßt. Die Methoden der Mittelwertbildung haben sich erheblich verfeinert, allerdings wurden damit auch die erforderlichen Rechenzeiten verlängert. Die Fortschritte in der Computertechnik konnten damit Schritt halten, so daß heutige Verfahren mit heutiger Technik ähnliche Laufzeiten erfordern, wie frühere Verfahren mit älterer.

Filter lassen sich in verschiedener Weise einteilen, so kann man unterscheiden zwischen lokalen Filtern, die ihre Eingangsdaten je Antwortpixel nur aus einem Teil des Bildes beziehen, und globalen Filtern, bei denen potentiell jedes Pixel im Bild ein Ausgabepixel beeinflussen kann. Eine andere Unterscheidung ist die zwischen linearen Filtern und nichtlinearen. Huang [1979] teilt die Filter ein in Transformationen, nichtrekursive und rekursive Filter (vgl. Abschnitt 2.1.3, Seite 13). Schließlich kann man nicht-iterative Filter von iterativ arbeitenden Filtern unterscheiden. Aus dieser großen Auswahl wurden einige Filter ausgewählt, einerseits die gebräuchlichsten, wie Mittelwert- und Medianfilter, andererseits darüber hinaus sinnvoll erscheinende wie sie im folgenden beschrieben werden.

5.1 Lineare Filter

Lineare Filter sind mathematisch durch die Faltung (Abschnitt 2.2.3, Seite 15) beschrieben. Nur der folgende Mittelwertfilter dient unmittelbar der Rauschverminderung, die anschließend folgenden linearen Filter zur Kantenschärfung sollen Nachteile der Mittelwertfilterung wieder ausgleichen.

5.1.1 Mittelwertfilter

Der einfachste Filter zur Glättung entsteht durch eine ungewichtete Mittelwertbildung über die Nach-

barschaft eines Pixels. Für eine 3 x 3-Umgebung nimmt man die Matrix $\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$. Diese Matrix wird

über dem Pixel zentriert und die Grauwerte aller neun darunterliegenden Pixel mit den Koeffizienten der Matrix multipliziert. Die Summe der Werte wird normalisiert, in diesem Fall also durch neun geteilt und ergibt den Grauwert des berechneten Pixels.

Dieser Filter erzeugt eine Grauwertglättung, jeder Punkt wird durch den Mittelwert seiner Nachbarschaft ersetzt, einzelne „Ausreißer“ werden auf diese Art eliminiert. In einem Programm wie Imagine^{®32} findet man speziell diese Matrix unter Tiefpaßfilter, weil hohe Frequenzen, d. h. abwechselnde große und kleine Grauwerte von Pixeln ausgeglichen werden und nur die großflächigen Grauwertänderungen übrig bleiben. In der Photographie bezeichnet man diese Art der Bildbearbeitung als Weichzeichner. In Modifikationen des Verfahrens werden die Pixel der Umgebung unterschiedlich stark gewichtet, z. B. entsprechend einer Gaußglocke. Der Gaußsche Weichzeichner wichtet das Zentralpixel stärker als weiter entfernt liegende Pixel. Die Koeffizienten der Matrix werden vorab berechnet. Im ITK-Toolkit ist der Mittelwertfilter (engl. *mean filter*) mit der oben beschriebenen Matrix realisiert. Programmintern läßt sich auch eine andere Größe der Nachbarschaft einstellen.

Der Nachteil dieser Art der Grauwertglättung ist der Verlust von Kanteninformation: Das Bild wird unscharf, Strukturen werden verwischt, vorher scharfen Begrenzungslinien verlieren ihre Ortsinformation. Man handelt sich also durch dieses Verfahren erhebliche räumliche Unschärfe ein.

³² www.erdas.com

5.1.2 Kantenextraktion

Wegen des Verlustes von Kanteninformation wird nach einer Faltung mit einem Mittelwertfilter gerne versucht, die Kanten wieder zu verstärken. Das geschieht durch Operatoren, die die zweite Ableitung der Bildfunktion darstellen. Sollen ausschließlich Kanten dargestellt werden, muß die Summe der Koeffizienten des Faltungskerns null betragen. Dadurch werden homogene Gebiete gelöscht. Bekannte Vertreter sind die richtungsorientierten Sobel- oder Kirsch-Operatoren.

5.1.2.1 Prewitt-Operator

Der Prewitt-Operator sucht die erste Ableitung in horizontaler und vertikaler Richtung. Dazu verwendet er eine Faltung nach Gleichung (2.6) (Seite 16) mit den Matrizen $w_h = \begin{pmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{pmatrix}$ und $w_v = \begin{pmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix}$.

Eine Weiterentwicklung ist der folgende Sobel-Operator.

5.1.2.2 Sobel-Operator

Der Sobel-Operator entspricht dem Prewitt-Operator, allerdings verwendet er als Faltungskerne $w_h = \begin{pmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{pmatrix}$ und $w_v = \begin{pmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{pmatrix}$, um das Zentralpixel besser zu betonen. Will man alle Kanten darstellen,

kann man die Bildpunkte g_{xy} der Ergebnisbilder $g_h = f * w_h$ und $g_v = f * w_v$ mit $g_{xy} = \sqrt{g_{h,xy}^2 + g_{v,xy}^2}$ geometrisch mitteln.

Diese Operation wird in Abbildung 5.2 für das in Abbildung 5.1 gezeigte Ausgangsbild dargestellt.



Abbildung 5.1

Ausschnitt aus einer Szene in Bremen-Walle
(vgl. Abbildung 7.1, S. 90)



Abbildung 5.2

Sobel-Operator, angewandt auf Abbildung 5.1

5.1.2.3 Kirsch-Operator

Der Kirsch-Operator liefert den größten Gradienten in den acht Richtungen (parallel zu den Bildrändern und diagonal). Er ist definiert in Gleichung (5.1) mit den in den Gleichungen

(5.2) angegebenen Faltungskernen.

$$g_{nm} = \max_{z=1,\dots,8} \sum_{i=-1}^1 \sum_{j=-1}^1 w_{ij}^{(z)} \cdot f_{n+i,m+j} \quad (5.1)$$

$$w^{(1)} = \begin{pmatrix} 5 & 5 & 5 \\ -3 & 0 & -3 \\ -3 & -3 & -3 \end{pmatrix} \quad w^{(2)} = \begin{pmatrix} -3 & 5 & 5 \\ -3 & 0 & 5 \\ -3 & -3 & -3 \end{pmatrix} \quad w^{(3)} = \begin{pmatrix} -3 & -3 & 5 \\ -3 & 0 & 5 \\ -3 & -3 & 5 \end{pmatrix} \quad w^{(4)} = \begin{pmatrix} -3 & -3 & -3 \\ -3 & 0 & 5 \\ -3 & 5 & 5 \end{pmatrix} \quad (5.2)$$

$w^{(1)}$ bis $w^{(4)}$, die anderen vier Faltungskerne sind dazu spiegelsymmetrisch.



Abbildung 5.3

wie Abbildung 5.1

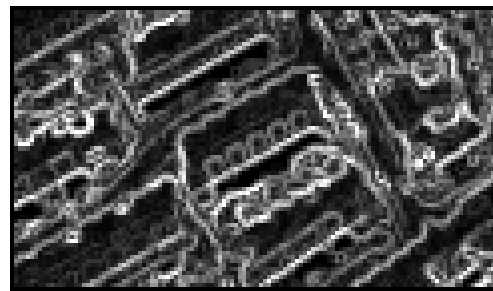


Abbildung 5.4

Kirsch-Operator, angewandt auf Abbildung 5.1. Der Operator ergibt einen maximalen Pixelwert von 2706, die Pixelwerte wurden deshalb mit 256/2706 multipliziert, anschließend wurde ein Histogrammausgleich durchgeführt.

Diese Filter enthalten eine rauschunterdrückende Glättungswirkung (Klette und Zamperoni [1995]).

5.1.3 Kantenverstärkung

Zur Kantenverstärkung mit Erhalt der generellen Bildinformation muß die Summe der Koeffizienten des Faltungskerns eins betragen, damit in homogenen Gebieten der ursprüngliche Grauwert erhalten bleibt. Hier sind richtungsunabhängige Operatoren erwünscht. Ein einfacher Vertreter ist der folgende Laplace-Filter.

5.1.3.1 Laplace-Filter

Von Jensen [1986] und Möller [2003] wird als Faltungskern $\begin{bmatrix} 0 & -1 & 0 \\ -1 & 5 & -1 \\ 0 & -1 & 0 \end{bmatrix}$ verwendet. Dieser

Filter betont Kanten, ihm fehlt aber die rauschunterdrückende Tiefpaßwirkung der zuvor beschriebenen Operatoren. Er ist deshalb rauschempfindlich und verstärkt das Rauschen, weil er jedes einzelne Pixel aus seiner Umgebung hervorhebt. In dieser Arbeit wurde er versuchsweise nach dem Mittelwertfilter verwendet, um die unscharf gewordenen Kanten wieder zu schärfen.

5.1.3.2 Ford-Filter

Wie der Laplace-Filter wirkt der von Ford et al. [1983] beschriebene Filter kantenverstärkend. Die Autoren entwickelten die Koeffizienten experimentell für LandSat-Bilder und binärisierten die gefilterten Bilder mit einem Schwellwert von 90% der Dichtefunktion. In der ursprünglich vorgestellten Form (Abbildung 5.5) ergibt die Summe der Koeffizienten 2,5, d. h. in einem homogenen Gebiet wird der ursprünglich vorhandene Grauwert mit 2,5 multipliziert. Dies war wegen der anschließenden Binärisierung unerheblich. Weil diese generelle Grauwertänderung in dieser Untersuchung aber unerwünscht ist, wurde in dieser Arbeit der in Abbildung 5.6 notierte Kernel verwendet.

$$\begin{bmatrix} -0,3 & -0,3 & -0,3 & -0,3 & -0,3 \\ -0,3 & -0,3 & -0,3 & -0,3 & -0,3 \\ -0,3 & -0,3 & 9,7 & -0,3 & -0,3 \\ -0,3 & -0,3 & -0,3 & -0,3 & -0,3 \\ -0,3 & -0,3 & -0,3 & -0,3 & -0,3 \end{bmatrix} \qquad \frac{1}{73} \begin{bmatrix} -1 & -1 & -1 & -1 & -1 \\ -1 & -1 & -1 & -1 & -1 \\ -1 & -1 & 97 & -1 & -1 \\ -1 & -1 & -1 & -1 & -1 \\ -1 & -1 & -1 & -1 & -1 \end{bmatrix}$$

Abbildung 5.5

Abbildung 5.6

Das größere Verhältnis zwischen der Wichtung des Zentralpixels und seiner Nachbarschaft bewirkt eine Abschwächung der Filterwirkung. Aufgrund seiner Größe ist dieser Kernel weniger rauschempfindlich als der Laplace-Operator.

Die bisher dargestellten Filter gehören zu den linearen Filtern, die sich durch eine Faltung beschreiben lassen. Der nachfolgend vorgestellte Medianfilter ist ein nichtlinearer Filter. Mit ihm versucht man, nur die „Ausreißer“ im Bild zu eliminieren.

5.2 Medianfilter

Um zu vermeiden, daß einzelne „Ausreißer“ in der Nachbarschaft eines Pixels den zu berechnenden Wert des Pixels zu stark beeinflussen und vor allem, um zu vermeiden, daß neue Farben (Grauwerte) in das Bild eingefügt werden, die im Original nicht vorhanden waren, wurde der Medianfilter entwickelt. Hier wird nicht mehr der Mittelwert der Nachbarschaft berechnet, sondern die Werte der Nachbarschaft sortiert und der Zentralwert (Median) als Ausgabe verwendet. Dieser Filter wirkt besser Kanten erhaltend, wie die Graphik in Abbildung 5.7 zeigt.

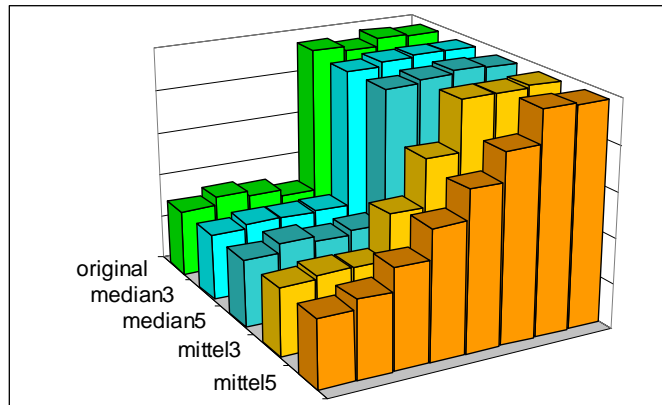


Abbildung 5.7

Kantenverhalten der Filter. Ordinate: Grauwert eines Pixels
Mittelwertfilter verwischen Kanten im Gegensatz zum Medianfilter. Die Ziffer
bedeutet die Kantenlänge des Filters.

Besonders stark verwischt der Tiefpaßfilter der Breite 5 (mittel5). Wie weit Ausreißer geglättet werden, hängt von der Filtergröße ab. Bei einem einzelnen Ausreißer genügt eine Filterbreite von 3 im linearen Fall, bei zwei nebeneinander liegenden Pixeln ist eine Filterbreite von 5 erforderlich (Abbildung 5.8 und Abbildung 5.9).

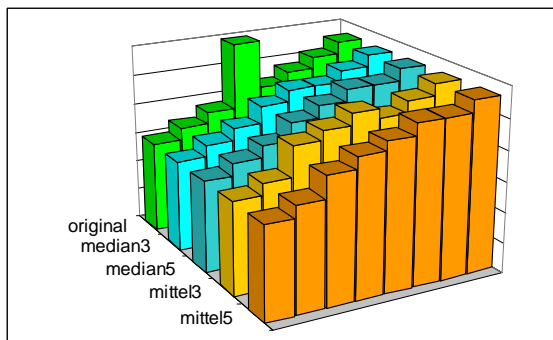


Abbildung 5.8

Einzelne Ausreißer werden durch alle Filter geglättet.

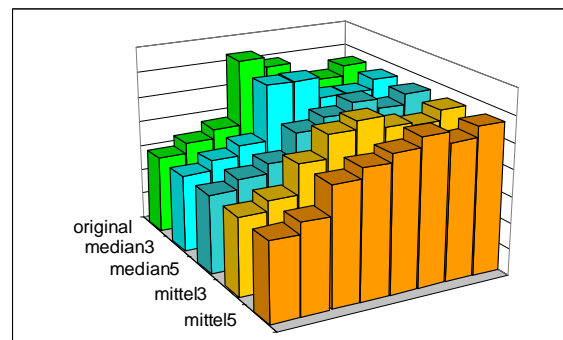


Abbildung 5.9

Bei zwei nebeneinander liegenden Pixeln ist für den Medianfilter eine Filterbreite von 5 erforderlich.

5.3 Filter mit veränderlicher Filterfunktion

Bisher lag die Filterfunktion fest. Bei den nachfolgend beschriebenen Filtern steuern die Bildinhalte nicht nur das Ergebnis, sondern beeinflussen vorab auch die Filterfunktion selbst. Die Filterfunktion wird hier durch Parameter gesteuert, die die Filterwirkung beeinflussen. So kann der Filter in einem inhomogenen Gebiet beispielsweise eine andere Fläche zur Berechnung der Filterwirkung benutzen als in einem homogenen Gebiet. Für jedes Pixel muß also nicht nur der Ausgabewert berechnet werden, sondern zunächst der Filter selbst. Dies beansprucht erhebliche Rechenzeiten. Für die vorgenannten Filter lagen die Ergebnisse nahezu in Echtzeit vor, während für die meisten im folgenden beschriebenen Filter Rechenzeiten um die sieben Sekunden erforderlich waren. Diese Filter arbeiten entweder iterativ, wie die auf dem Prinzip der anisotropen Diffusion beruhenden Filter (Abschnitt 5.3.1) und der Gra-

phennetzwerkfilter (Abschnitt 5.3.2) oder sie benötigen nur einen einzigen Durchgang, wie der EPOS-Filter (Abschnitt 5.3.3) oder der Bilaterale Filter (Abschnitt 5.3.4). Dabei benötigen die iterativen Filter nicht notwendig mehr Rechenzeit. Der Bilaterale Filter benötigte hier in seiner nicht optimierten Form mit 15 Minuten die 130fache Rechenzeit gegenüber den iterativen Filtern.

5.3.1 Anisotrope Diffusion

Vor der Erläuterung der Filter eine kurze Begriffserklärung: Wärmediffusion, auch Wärmeleitung oder Konduktion genannt, bezeichnet den Übergang der Wärme in festen Stoffen oder strömungslosen Flüssigkeiten von wärmeren zu kühleren Regionen.

Der Begriff wird verallgemeinert auf die iterative Änderung von Grauwerten in einem Bild in Abhängigkeit von benachbarten Grauwerten. Anisotrop bedeutet richtungsabhängig. Das Attribut anisotrop wurde von Perona und Malik [1990] eingeführt, um die Richtungsabhängigkeit des Diffusionsprozesses zu betonen. Das Verfahren wird auch *nonuniform* oder *variable conductance diffusion* genannt.

Der Operator ∇ (Nabla) liefert sämtliche partiellen Ableitungen in den Richtungen der Koordinatenachsen und damit den Gradienten, in den hier betrachteten Fällen also Betrag und Richtung der größten Grauwertdifferenz in einem Bild.

Das Zeichen * soll hier für die im Abschnitt 2.2.3 (Seite 15) beschriebene Faltung stehen.

Die isotrope Filterung mit einem Gaußschen Weichzeichner kann aufgefaßt werden als eine Lösung der Wärmediffusionsgleichung

$$\frac{\partial g(x, y, t)}{\partial t} = \nabla \cdot \nabla g(x, y, t) \quad (5.3)$$

mit $g(x, y, 0) = f(x, y)$ als dem Eingabebild und $g(x, y, t) = G(\sqrt{2t}) * f(x, y)$ als die Lösung zum Zeitpunkt t mit $G(\sigma)$ als Gaußfunktion mit der Standardabweichung σ .

Anisotrope Diffusion erweitert die Formel um eine variable Wärmeleitung, die von der Differentialstruktur des Bildes abhängen soll. So kann das Weichzeichnen an Kanten begrenzt werden, die durch den Betrag ihrer Ableitung definiert werden:

$$g_t = \nabla \cdot c(|\nabla g|) \nabla g \quad (5.4)$$

Die unabhängigen Parameter x und y wurden weggelassen, t bezeichnet den Parameter, nach dem partiell abgeleitet wird.

Mit $c(|\nabla g|)$ soll die Wärmeleitung für große Gradienten begrenzt werden. Nach Whitaker [1993] hat sich

$$c(|\nabla g|) = e^{-\frac{|\nabla g|^2}{2k^2}} \quad (5.5)$$

als effektiv erwiesen. k wird als freier Parameter eingeführt, um die Sensitivität an Kanten kontrollieren zu können. Anisotrope Diffusion erfordert also zwei freie Parameter, den Leitfähigkeitsparameter k und den Zeitparameter t , die analog zu σ die Reichweite des Filters bestimmen.

Gleichung (5.5) ist eine nichtlineare Partialgleichung, die im diskreten Gitternetz durch eine endliche Zahl von Differenzen gelöst werden kann. Das geglättete Bild läßt sich deshalb nur iterativ berechnen, die Rechenzeiten sind aber auch für große Bilder erträglich. Das Verfahren wurde später in verschiedenen Richtungen verbessert und erweitert (vgl. Weikert et al. [1996] und Whitaker und Xue [2001]). Insbesondere läßt es sich auch für vektorielle Eingabebilder wie Farb- oder Multispektralaufnahmen einrichten:

$$F : U \rightarrow \mathfrak{R}^m \quad F_t = \nabla \cdot c(\mathcal{D}F)F \quad (5.6)$$

Dabei ist $\mathcal{D}\vec{F}$ ein allgemeines Unähnlichkeitsmaß, eine Verallgemeinerung der Gradientensteilheit für Vektorbilder, das beliebige Berechnungen über die verschiedenen Vektorkoordinaten bzw. Bildebenen hinweg erlaubt. Der Parameter k verbindet die Glättung in den verschiedenen Kanälen. Rauschen und Texturen können geglättet werden, Kanten bleiben erhalten und können sogar verstärkt werden.

Der im folgenden noch benötigte TimeStep-Parameter t beeinflusst die Stärke der Veränderung eines Pixels im Bild von einer Iteration zur nächsten nach folgender Gleichung (5.7):

$$f_i^{n+1} = f_i^n + \Delta f_i^n t \quad (5.7)$$

Gleichung (5.7) beschreibt die Veränderung des Pixels i im Bild f von der n -ten zur $n+1$ -ten Iteration, mit Δf ist die durch die Filterfunktion berechnete Veränderung bezeichnet. Ein zu groß gewählter TimeStep-Parameter kann zu Schwingungen führen und beeinträchtigt deshalb die Stabilität des Algorithmus (Read et al. [1979]).

5.3.1.1 Gradient Anisotropic Diffusion Filter

Die Implementierung wurde dem Insight Toolkit entnommen. Das allgemeine Verfahren wurde bereits auf Seite 49 dargestellt. Der Algorithmus ist eine Verallgemeinerung des von Perona und Malik [1990] vorgeschlagenen Verfahrens auf den N -dimensionalen Raum. Als Wärmeleitungsfunktion wird

$$C(\mathbf{x}) = e^{-\left(\frac{\|\nabla U(\mathbf{x})\|}{k}\right)^2} \quad (5.8)$$

verwendet. Die Pixelwerte können wahlweise ganzzahlig oder Fließkomma-Werte sein. Als Parameter sind anzugeben die Anzahl der Iterationen, vorgeschlagen werden und hier verwendet wurden fünf Iterationen. Für den TimeStep-Parameter t werden für 2-D-Bilder 0,25 vorgeschlagen (für 3-D-Bilder 0,125) und für den Parameter k wurde der Wert 3,0 gewählt.

Je nach Bilddaten und Parametereinstellung kann dieser Filter auch Kanten verstärkend wirken, was zu Artefakten führt.

Die Laufzeit betrug mit diesen Parametern je 512 x 512 Pixel großes Grauwertbild mit dem auf Seite 18 beschriebenen System 7 Sekunden.

5.3.1.2 Curvature Anisotropic Diffusion Filter

Im Unterschied zum Gradient Anisotropic Diffusion Filter verwendet dieser Filter eine „modified curvature diffusion equation“ (MCDE). MCDE kann nicht kantenverstärkend wirken. Dieser Filter erhält feinere Strukturen und ist weniger kontrastempfindlich. Der Nachteil ist eine verlängerte Laufzeit. Die MCDE-Gleichung lautet

$$f_t = |\nabla f| \nabla \cdot c(|\nabla f|) \frac{\nabla f}{|\nabla f|} \quad (5.9)$$

Als Eingabeparameter werden vorgeschlagen und wurden verwendet fünf Iterationen, $t = 0,125$ (für 3-D-Bilder wird hier 0,0625 empfohlen) und $k = 3,0$ wie beim Gradient Anisotropic Diffusion Filter.

5.3.2 Graphen Netzwerk Filter

Von Jahn [1996] inauguriert und in Halle [1999] näher ausgeführt gibt es ein anderes Verfahren der Grauwertglättung. Das Verfahren verwendet ein „*layered graph network*“, ein Graphennetzwerk und ist nicht nur zur Grauwertglättung, sondern primär zur Segmentierung von homogenen und texturierten Flächen vorgesehen. Der Algorithmus benötigt keine Vorabinformationen und keine *seedpoints* und ist zur Realisierung auf Parallelrechnern vorgesehen. Er ist schnell und robust und erzielt bei guter Konturerhaltung eine starke Grauwertglättung. Wie im Namen angekündigt, arbeitet das Verfahren in verschiedenen Schichten oder Fortschritten der Segmentierung.

Im ersten Schritt werden die Pixel als Knoten des zu erstellenden Graphen betrachtet, in nachfolgenden Schritten können dies zusammenhängende Flächen mit gleichem oder ähnlichem Grauwert sein oder in späteren Schritten auch Texturen.

Aus den Pixeln wird ein Nachbarschaftsgraph erzeugt. Ein Pixel kann null bis vier Nachbarn haben, mit denen es durch eine Kante verbunden wird, wenn es sich in seinem Grauwert nur gering von den Werten seiner Nachbarn unterscheidet.

Der besondere Vorteil einer Darstellung mit einem Graphen ist die Tatsache, daß den die Pixel verbindenden Kanten Gewichte zugeordnet werden können. Dadurch braucht in einem Durchgang nicht sofort entschieden werden, ob zwei Pixel $g_{i,j}$ und $g_{i',j'}$ zusammen gehören. Das wird der Fall sein, wenn sie sich in ihren Grauwerten nicht unterscheiden. Haben sie sehr unterschiedliche Grauwerte, werden sie mit hoher Wahrscheinlichkeit zu verschiedenen Segmenten gehören. Dazwischen sind jedoch beliebige unscharfe Werte möglich, die durch eine Funktion μ :

$$0 \leq \mu(g_{i,j}, g_{i',j'}) \leq 1 \quad (5.10)$$

dargestellt werden können. Für die Grauwertglättung können nun die Grauwerte des Pixels mit denen seiner vier Nachbarn gemittelt werden. Dabei wird der Grad der Nachbarschaft durch die Funktion μ beschrieben. Weil größere Grauwertdifferenzen geringer zu Buche schlagen, wirkt dieser Algorithmus bei geeigneter Wahl der sonstigen noch zu beschreibenden Parameter Kanten erhaltend.

Des weiteren hat sich ein rekursives Vorgehen bewährt. Weil es im Falle eines rekursiven Vorgehens zu Schwingungen durch Rückkoppelung kommen kann, wählt man bei lernenden Algorithmen gern einen Konvergenzfaktor α (l) mit Abhängigkeit von der Anzahl l der stattgefundenen Iterationen, mit dem die Änderungsrate sukzessive reduziert wird.

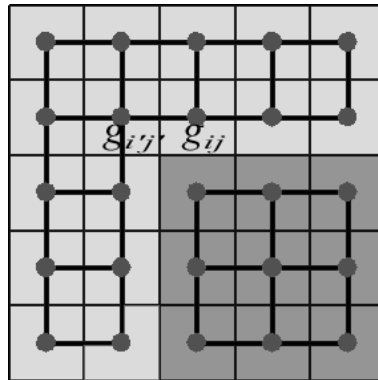


Abbildung 5.10

Der Ähnlichkeitsgraph. Nur Pixel mit ähnlichen Grauwerten werden durch eine Kante verbunden.

Die beiden benachbarten Pixel $p_{i'j'}$ und p_{ij} mit den Grauwerten $g_{i'j'}$ und g_{ij} werden miteinander durch eine Kante im Ähnlichkeitsgraphen verbunden, wenn sie der Ähnlichkeitsbedingung

$$\left| g_{i'j'} - g_{ij} \right| \leq t_{ij} \quad (5.11)$$

genügen. Dabei ist t_{ij} ein bestimmter Schwellwert, der von der Grauwertvariation innerhalb einer gewissen Umgebung der Pixel $p_{i'j'}$ und p_{ij} abhängt. Weil eine harte Klassifikation der Kanten des Ähnlichkeitsgraphen nachteilig ist, wird eine „fuzzy“-Funktion μ eingeführt:

$$\mu(x) = \frac{t^2}{t^2 + x^2} \quad \text{mit} \quad x = \left| g_{i+r \ j+s} - g_{ij} \right| \quad (5.12)$$

Die Funktion μ kann je nach Erfordernis in den verschiedenen Unteralgorithmen noch variiert werden. Der Schwellwert t wird in den Unteralgorithmen dynamisch berechnet. Weil die Frequenz des Rauschens in verschiedenen Bildern sehr unterschiedlich sein kann, ist die Berechnung des Schwellwertes nicht trivial.

Als optimal für verschiedene Bilder wurde der Mittelwert $\overline{d3_{ij}}$ der drei kleinsten Werte von $\left| g_{i+r \ j+s}(l) - g_{ij}(l) \right|$ aus einem quadratischen Fenster mit 11 x 11 Bildpunkten als stabile Schwelle t_{ij} der l -ten Iteration gefunden:

$$t_{ij} =: \overline{d3_{ij}} = \overline{MIN}_{120\text{-Umgebung}}^3 \{ |g_{i+r \ j+s}(l) - g_{ij}(l)| \} \quad (5.13)$$

Um Rechenzeit zu sparen, wird diese Schwelle nur für die erste Iteration auf diese Weise berechnet und für die nachfolgenden Iterationen nach der folgenden Funktion abgeschwächt:

$$\overline{d3_{ij}}(l+1) = \frac{\overline{d3_{ij}}(l)}{l} \quad (5.14)$$

Gleichzeitig stabilisiert diese Abschwächung das Konvergenzverhalten dieses nichtlinearen, rekursiven Filters mit dem Konvergenzfaktor $\alpha(l)$, der konstant 4 gewählt werden kann.

$$g_{ij}(l+1) = g_{ij}(l) + \alpha(l) \cdot \frac{\sum_n \mu(|g_{i+r,j+s}(l) - g_{ij}(l)|) \cdot [g_{i+r,j+s}(l) - g_{ij}(l)]}{1 + \sum_n \mu(|g_{i+r,j+s}(l) - g_{ij}(l)|)} \quad (5.15)$$

Die in Gleichung (5.12) beschriebene Funktion μ ergibt sich damit zu

$$\mu(|g_{i+r,j+s}(l) - g_{ij}(l)|) = \frac{\overline{d3}_{ij}^2}{\overline{d3}_{ij}^2 + |g_{i+r,j+s}(l) - g_{ij}(l)|^2} \quad (5.16)$$

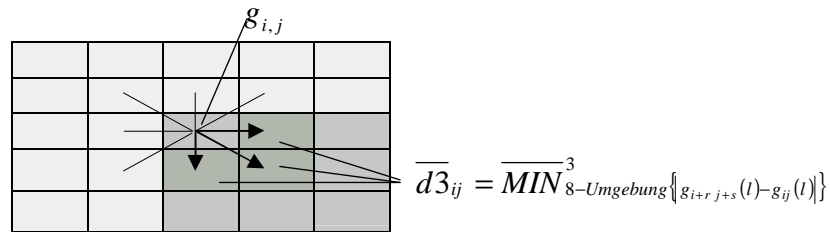


Abbildung 5.11

Bestimmung des Mittelwertes $\overline{d3}_{ij}$ der drei kleinsten Grauwertdifferenzen für eine 8-Umgebung: $\overline{d3}_{ij} = \overline{MIN}^3_{8-Umgebung} \{g_{i+r,j+s}(l) - g_{ij}(l)\}$, zur Berechnung wird eine 120-Umgebung verwendet (siehe Text)

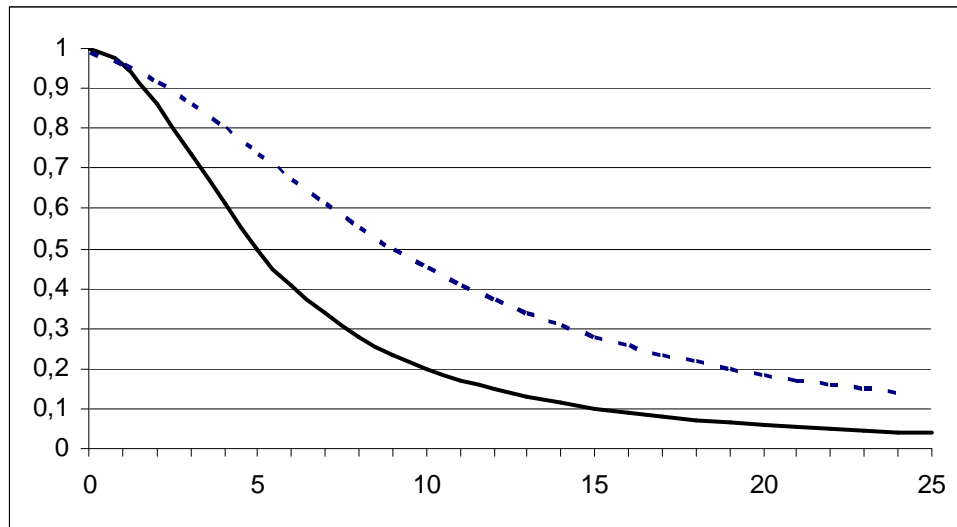


Abbildung 5.12

Die Funktion μ aus Gleichung (5.16) für $d_{ij}=10$ (punktierter Linie) und $d_{ij}=5$ (durchgezogene Linie). Auf der Abszisse ist die Grauwertdifferenz eingetragen.

Die Iterationen werden abgebrochen, sobald die maximale Änderung der Grauwerte aller Pixel einer Szene unter einem festzulegenden Schwellwert liegt. Beispielsweise kann als Schwellwert $0,03 * 256 = 7,68$ bei 8-bit unsigned Daten gewählt werden.

5.3.3 Edge Preserving Optimized Speckle Filter

Der Edge Preserving Optimized Speckle Filter (EPOS) wurde von Haag und Sties [1994, 1996] entwickelt und in Haag [1998] optimiert und ausführlich dargestellt und mit den bis dahin bekannten Verfahren verglichen. EPOS wurde entwickelt, um Speckle³³ aus SAR³⁴-Bildern zu entfernen. Das Verfahren läßt sich einfach modifizieren, so daß es auch auf beliebige andere verrauschte Rasterbilder angewendet werden kann. In dieser Untersuchung wurde es für die Verwendung mit HRSC-A³⁵- und IKONOS-Daten angepaßt.

Das Verfahren benötigt für Bilder mit 8 Bit Auflösung in Abhängigkeit von der internen Implementierung und Darstellung der Zahlen etwa die 12-fache Bildgröße als Speicher. Zunächst wird ein Bildmodell zugrundegelegt, bei dem die Signalkomponente der Bildfunktion aus in sich homogenen Flächen besteht. Größe und Form der Fläche sollen beliebig sein. Aufgabe ist es, eine Fläche zu finden, in der die Signalkomponente tatsächlich konstant ist und den Speckle bzw. das Rauschen durch Mittelwertbildung zu eliminieren.

Dazu wird ein Homogenitätskriterium benannt, nach dem die mutmaßlich homogene Fläche überprüft wird. Die Fläche wird im Rahmen der Filtergröße zunächst maximal gewählt. Wird das Homogenitätskriterium nicht erfüllt, wird die Fläche sukzessive verkleinert, bis das Homogenitätskriterium erfüllt wird. Anschließend kann das Zentralpixel durch einen erwartungstreuen Schätzwert ersetzt werden.

³³ Speckle: Sprengel, multiplikative Störung des SAR-Signals, d.h. die Signalvariation wächst mit der Intensität des Signals an im Gegensatz zum additiven Rauschen

³⁴ Synthetic Aperture RADAR

³⁵ High Resolution Stereo Camera-Airborne

Haag [1998] nennt folgende Forderungen an ein das Zentralpixel enthaltendes Segment:

Das Zentralpixel muß in jeder Stichprobe enthalten sein, um den Bezug zu seiner Signalkomponente zu gewährleisten.

Die Bereiche sollten möglichst groß sein, da die Zuverlässigkeit eines Homogenitätskriteriums mit der Stichprobengröße anwächst.

Die Form eines Segmentes sollte möglichst kompakt sein, da die Wahrscheinlichkeit für ein instationäres Signal mit zunehmender Ausdehnung größer wird.

Es müssen Meßbereiche existieren, bei denen sich das Zentralpixel am Rand des Segmentes befindet. Nur so können homogene Bereiche in unmittelbarer Umgebung von Bildkanten identifiziert werden.

Die Form der Segmente sollte eine möglichst gute Anpassung an vorkommende Geometrien erlauben. So dürfen nicht nur gerade Kanten zwischen homogenen Flächen betrachtet werden, vielmehr muß die Begrenzung der gesamten Fläche in Betracht gezogen werden. Faßt man diese als Polygonzug auf, so ergibt sich die Geometrie im einzelnen durch den Winkel zwischen den Polygonkanten.

Die Größe der Bereiche muß skalierbar sein, so daß auch kleinere homogene Flächen im Bild einer Verarbeitung zugeführt werden können.

Die Anordnung aufeinander folgender Messungen sollte so erfolgen, daß stets das größtmögliche verbleibende Segment als nächstes herangezogen wird. Eine Skalierung hin zu kleineren Meßbereichen sollte erst nach Ausschöpfung möglicher geometrischer Konstellationen erfolgen.

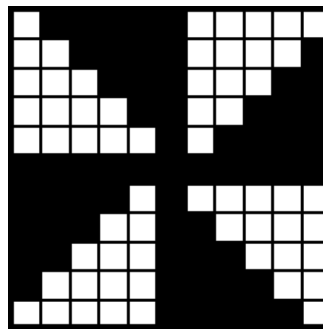


Abbildung 5.13

Segmente des EPOS-Filters

Für einen Filter der Größe 11×11 ist die Einteilung der Segmente in Abbildung 5.13 beispielhaft dargestellt. Das Zentralpixel soll in jedem Fall enthalten sein. Falls das Pixel am Rand der angenommenen homogenen Fläche liegt, wird der nicht zur Fläche gehörige Bereich ausgeblendet. Die Flächen sollen kompakt sein. Das hat allerdings den Nachteil, daß linienförmige Strukturen nicht erfaßt und homogenisiert werden können. Um möglichst große Meßbereiche für das Homogenitätskriterium zu erhalten, wird zunächst das gesamte Operatorfenster untersucht. Wird das geforderte Kriterium dort nicht erfüllt, dann werden sukzessive diejenigen Segmente vom untersuchten Gebiet ausgeschlossen, welche in sich die größte Inhomogenität aufweisen und somit das geforderte Kriterium am wenigsten erfüllen. Ist das Kriterium in keinem der Segmente erfüllt, so wird die Seitenlänge des Operatorfensters um zwei Pixel reduziert und mit der Berechnung in gleicher Weise fortgefahren. In einer heterogenen

Umgebung führt dies schließlich zur Reduktion des Operatorfensters auf das Zentralpixel: Es findet keine Mittelwertbildung und deshalb auch keine Veränderung des Grauwerts an dieser Stelle statt.

Das Verfahren wird weiter optimiert, in dem nicht einfach das Operatorfenster untersucht und ungeeignete Segmente ausgeschlossen werden und dann das Operatorfenster verkleinert wird, sondern die tatsächliche Anzahl der für die Mittelwertbildung berücksichtigten Pixel zugrunde gelegt wird. Anhand einer Tabelle wird so zum Beispiel ein Operatorfenster der Kantenlänge neun Pixel mit fünf Segmenten (51 Pixel) vor einem Operatorfenster der Kantenlänge sieben Pixel mit acht Segmenten (49 Pixel) vor einem Operatorfenster der Kantenlänge elf Pixel mit drei Segmenten (46 Pixel) untersucht.

5.3.3.1 Das Homogenitätskriterium

Ausschlaggebend für die Qualität des Filters ist das Homogenitätskriterium, nach dem entschieden wird, aufgrund welcher Datenbasis das Zentralpixel angepaßt werden soll, bzw. welche Nachbarpixel zur Mittelwertbildung herangezogen werden.

Dieses Homogenitätskriterium ist mehrfach abhängig von den Bilddaten: Multiplikative Störungen wie Speckle bei SAR-Bildern haben einen anderen Einfluß als additives Rauschen bei HRSC-A- oder IKONOS-Aufnahmen, optischen Grauwertbildern oder auch Farbbildern. Außerdem spielt das Maß des Rauschens eine Rolle: in stärker verrauschten Bildern ist naturgemäß die Varianz des erhaltenen Signals in homogenen Bildbereichen höher als in schwächer verrauschten Bildern.

Haag [1998] beschreibt ausführlich verschiedene auch von anderen Autoren verwendete Verfahren, um einen optimalen Wert zu erhalten. In einem homogenen Gebiet ist die Varianz der Signalkomponente null. Ist die Varianz in einem beobachteten Gebiet von null verschieden, so kann dies am Signal liegen, dann sollte keine Mittelwertbildung stattfinden, oder durch die Störkomponente bedingt sein. Im letzteren Fall ist eine Mittelwertbildung angezeigt. Da in der Regel, außer bei Multi-Look³⁶-Verfahren, nicht entschieden werden kann, ob die Varianz der Stör- oder der Signalkomponente zuzuordnen ist, verwendet man eine kontinuierliche, von der Varianz σ_z^2 gesteuerte Funktion um darüber zu entscheiden, ob bei hoher Signalvarianz der ursprüngliche Pixelwert z oder bei geringer Varianz der Mittelwert \bar{z} eingesetzt wird. Schließlich kann man noch zwischen rein statistischen Verfahren unterscheiden und solchen, die neben der Statistik auch die Geometrie berücksichtigen.

Lee [1980] verwendet eine Linearkombination aus dem Mittelwert \bar{x} des beobachteten Gebietes und dem ursprünglichen Wert z im Pixel,

$$\hat{x} = \bar{x} + w(z - \bar{x}) \quad \text{mit} \quad w = \frac{\sigma_x^2}{\sigma_z^2} \quad (5.18)$$

wobei sich w aus der Minimierung des mittleren quadratischen Fehlers ergibt.

Ein derartiger Filter wird in der Literatur häufig als MMSE (Minimum Mean Square Error) -Filter bezeichnet.

Der Filter kann nach Lopes et. al. [1990a] verfeinert werden, indem Grenzwerte C_v , C_{\max} für den erlaubten Variationskoeffizienten $C_z = \sigma / \bar{z}$ in homogenen Gebieten, in heterogenen Gebieten und in

³⁶ Aus einem Datensatz werden mit entsprechend verringerter geometrischer Auflösung mehrere unabhängige Aufnahmen des gleichen Gebiets erzeugt, um Störungen durch pixelweise Mittelung herauszufiltern.

Gebieten, in denen im Fall von SAR-Daten Punktstreuer erhalten werden sollen, angegeben werden, so daß

$$\hat{x} = \begin{cases} \bar{x} & \text{für } C_z \leq C_v \\ \bar{x} + w(z - \bar{x}) & \text{für } C_v < C_z < C_{\max} \\ z & \text{für } C_z \geq C_{\max} \end{cases} \quad (5.19)$$

Eine weitere Verfeinerung ergibt sich, wenn statt des Terms $\bar{x} + w(z - \bar{x})$ aufwendiger zu berechnende Terme wie von Lopes et al. [1990b] angegeben, verwendet werden.

In Ergänzung zu den bisher beschriebenen Verfahren erscheint es sinnvoll, bei der Mittelwertberechnung die räumliche Distanz vom Zentralpixel zu berücksichtigen, um nicht in Kantennähe unerwünschte Verwischungseffekte zu erhalten. Dazu gibt es wieder die einfachen Verfahren nach Frost et al. [1982] und Lopes et al. [1990b], die die Pixel in der Umgebung unabhängig von ihrer Richtung nur nach ihrem Abstand statistisch auswerten und einen optimierten Schätzwert berechnen.

Bereits Lee [1981] versucht jedoch, durch Betrachtung einer 3×3 Maske in unmittelbarer Nähe des Zentralpixels einen Gradienten zu berechnen. Senkrecht zum größten Gradienten in acht möglichen Richtungen wird eine Kante angenommen und nur die Pixel, die nicht jenseits der vermeintlichen Kante liegen, werden für die weitere Berechnung berücksichtigt.

Eine ganz andere Vorgehensweise beschreibt Crimmins [1985a, b]. Hier wird der Grauwert als geometrische Größe in z -Richtung betrachtet. Nach der Idee der konvexen Hülle³⁷ wird iterativ in vier Richtungen (waagrecht, vertikal und in den Diagonalen) der Grauwert jedes Pixels mit denen seiner Nachbarn verglichen und im Falle von das Konvexitätskriterium verletzenden Werten angepaßt.

Dieses geometrische Filter liefert eine gute Glättungswirkung und erhält Kanten, hat jedoch den Nachteil, daß dünne Strukturen wie Linien stark eingeebnet werden.

In den folgenden Untersuchungen wird die Varianz σ^2 , d. h.

$$\sigma^2 = \frac{1}{n} (g_{x,y} - \bar{g})^2 \quad (5.20)$$

mit $g_{x,y}$ als dem Grauwert an der Stelle $(x|y)$, \bar{g} als dem zugehörigen mittleren Grauwert und n als der Zahl der betrachteten Grauwerte als Homogenitätsmaß verwendet. Dabei wird die Varianz für ein im Rahmen der gewählten Filtergröße maximales n getestet und dann das Testgebiet nach dem EPOS-Verfahren soweit verkleinert, bis die Varianz das vorgegebene Maß von z. B. 100 unterschreitet. Wird die geforderte Varianz nicht unterschritten, dann wird das Zentralpixel nicht verändert, ansonsten wird es durch den berechneten Mittelwert ersetzt.

5.3.4 Bilateraler Filter

Der Bilaterale Filter wurde von Tomasi und Manduchi [1998] zunächst zur Kanten erhaltenden Glättung von Farbbildern entwickelt. Werden die einzelnen Kanäle eines RGB- oder CMYK-Bildes ge-

³⁷ Eine konvexe Hülle einer Menge von Punkten in der zweidimensionalen Ebene ist definiert als der Durchschnitt aller Halbebenen, die diese Punkte enthalten (als wäre um die Punkte ein Gummiband gespannt).

trennt gefiltert, dann entstehen bei der anschließenden Synthese der Kanäle zu einem Bild an den Kanten Farbsäume durch unterschiedliche Mittelwertbildung in der Nähe der Kanten. Die Idee des Bilateralen Filters ist, neben der Mittelung durch Nachbarschaft (*spatial domain*), d. h. in der Ebene eines Kanals, zusätzlich die anderen Kanäle einzubeziehen (*range domain*) in der Weise, daß eine beliebige Abstandsfunktion für die Farben angegeben werden kann. Im Falle von Grauwertbildern werden nur die Pixel zur Glättung herangezogen, die ausreichend kleine Differenzen in ihren Grauwerten haben. Im Falle von Farbbildern bietet sich als Abstandsfunktion der CIE-Lab-Farbraum an, der eine der menschlichen Wahrnehmung angenäherte Metrik bietet, für Mehrkanalaufnahmen die spektrale Nähe.

Die Autoren beschreiben den Bilateralen Filter wie folgt:

Ein Tiefpaßfilter, angewandt auf ein Bild $f(\mathbf{x})$ ergibt ein Ergebnisbild $h(\mathbf{x})$ mit dem Filterkernel $k_d(\mathbf{x})$

$$\mathbf{h}(\mathbf{x}) = k_d^{-1}(\mathbf{x}) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathbf{f}(\xi) c(\xi, \mathbf{x}) d\xi \quad (5.21)$$

Die Funktion $c(\xi, \mathbf{x})$ beschreibt die geometrische Nähe zwischen dem Zentralpixel \mathbf{x} und einem benachbarten Punkt ξ . Die fettgedruckten Zeichen h und f sollen betonen, daß sowohl das Eingabebild als auch das Ergebnisbild Mehrkanalbilder sein können. Wenn die Tiefpaßfilterung den Mittelwert nicht verändern soll, gilt

$$k_d(\mathbf{x}) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} c(\xi, \mathbf{x}) d\xi \quad (5.22)$$

Ist, wie z. B. beim Gauß-Filter der Filter ortsunabhängig, so ist $c(\xi, \mathbf{x})$ nur eine Funktion der Vektordifferenz $\xi - \mathbf{x}$ und k_d ist konstant.

Auf ganz ähnliche Weise kann die Ähnlichkeit definiert werden:

$$\mathbf{h}(\mathbf{x}) = k_r^{-1}(\mathbf{x}) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathbf{f}(\xi) s(\mathbf{f}(\xi), \mathbf{f}(\mathbf{x})) d\xi \quad (5.23)$$

und die Funktion $s(\mathbf{f}(\xi), \mathbf{f}(\mathbf{x}))$ beschreibt hier die photometrische Ähnlichkeit zwischen dem Zentralpixel \mathbf{x} und dem benachbarten Punkt ξ . Ebenso ergibt sich der Kernel

$$k_r(\mathbf{x}) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} s(\mathbf{f}(\xi), \mathbf{f}(\mathbf{x})) d\xi \quad (5.24)$$

Dieser kann jetzt jedoch nicht mehr ortsunabhängig sein, sondern soll gerade von den Kanten im Bild abhängen. Die Kombination beider Prinzipien kann dann mit folgender Formel zusammengefaßt werden:

$$\mathbf{h}(\mathbf{x}) = k^{-1}(\mathbf{x}) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathbf{f}(\xi) c(\xi, \mathbf{x}) s(\mathbf{f}(\xi), \mathbf{f}(\mathbf{x})) d\xi \quad (5.25)$$

mit dem Kernel

$$k(\mathbf{x}) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} c(\xi, \mathbf{x}) s(\mathbf{f}(\xi), \mathbf{f}(\mathbf{x})) d\xi \quad (5.26)$$

Dieser sogenannte Bilaterale Filter ersetzt also den Wert des Pixels x mit einem gewichteten Mittel ähnlicher und nahegelegener Pixelwerte. In homogenen Bildbereichen haben Pixel innerhalb einer engen Nachbarschaft ähnliche Grauwerte oder spektrale Signaturen und die normalisierte Ähnlichkeitsfunktion k^1s liegt nahe 1. In diesen Bereichen wirkt der Bilaterale Filter wie ein gewöhnlicher Gauß-Tiefpaß. Sobald eine Kante in den Nachbarschaftsbereich des Filters kommt, werden wegen der Ähnlichkeitsfunktion s nur noch diejenigen Pixel für die Filterung berücksichtigt, die einen ähnlichen Grauwert oder eine ähnliche spektrale Signatur haben.

Als Parameter werden empfohlen und wurden verwendet 5,0 für die *spatial domain* und 6,0 für die *range domain*.

Der Bilaterale Filter ist wegen seiner Ortsabhängigkeit in der Filterfunktion nicht separierbar und deshalb langsam. Er kommt zwar im Gegensatz zu manch anderen Filtern ohne Iterationen aus. In der verwendeten Implementierung von ITK³⁸ benötigte er mit den vorgeschlagenen Parametern jeweils 15 Minuten je Grauwertbild der Größe 512 x 512 Pixel, während die einfachen Filter Ergebnisse in Echtzeit und die aufwendigeren, aber zeitlich stark optimierten Filter wie der Graphennetzwerk- und der EPOS-Filter sieben Sekunden benötigten. Elad [2002] untersucht algebraisch die theoretische Basis des Filters, um Wege zur Beschleunigung zu finden. Er findet Gemeinsamkeiten mit den Filtern, die anisotrope Diffusion nutzen. Durand und Dorsey [2002] machen praktische Vorschläge, um mit linearer Approximation und unter Verwendung der schnellen Fouriertransformation (FFT) den Filter um Größenordnungen zu beschleunigen.

In diesem Kapitel wurden die untersuchten Filter beschrieben. Zahllose andere Filter wurden in der Literatur beschrieben. Abschnitt 5.3, Filter mit veränderlicher Filterfunktion, hätte als Überschrift auch tragen können „Kantenerhaltende Filter“, weil die Algorithmen sich darum bemühen, vorhandene Kanteninformationen nicht zu verwischen. Mit dieser Bezeichnung werden im nächsten Kapitel jedoch nur der Graphennetzwerkfilter und der EPOS-Filter zusammengefaßt werden.

³⁸<http://www.itk.org> Insight Toolkit

6 Filterbewertung

In diesem Kapitel wird die Komplexität der Bewertung von Filteralgorithmen dargestellt. Abgesehen davon, daß die Anzahl der verschiedenen veröffentlichten Filteralgorithmen nicht mehr überschaubar ist, haben die meisten Filter Parameter, die ebenfalls das Filterergebnis stark beeinflussen. Schließlich verändern die eingesetzten Filter die Bildinhalte in einer Weise, die eine Änderung der Segmentierungsparameter der hier verwendeten Segmentierungssoftware erforderlich macht.

6.1 Übersicht

Angesichts der Vielzahl der einstellbaren Parameter ist eine perfekte Auswahl der Vorverarbeitungsschritte und der anschließenden Parameterwahl für die Segmentierung nicht möglich. Gesucht wird daher ein Verfahren, wie diese beiden Schritte optimiert werden können. Wie bei einer Maximumsuche in einem unbekanntem Gebiet werden in dieser Arbeit gebräuchliche und mögliche Verfahren der Bildvorverarbeitung mit von den Nutzern empfohlenen Standardparametern an ausgewählten Bildbeispielen überprüft. Anschließend werden die einstellbaren Parameter der vorteilhaftesten Verfahren variiert und in ihrer Wirkung auf das Segmentierungsergebnis optimiert. Diese Vorgehensweise wurde übrigens von Comaniciu und Meer [2002] für ein anderes Segmentierungsverfahren benutzt und algorithmisch ausgefeilt (siehe auch Abschnitt 4.2.3.2).

6.1.1 Ground Truth

Entsprechend der Zielsetzung der vorliegenden Arbeit war eine exakte Ground Truth wichtig. Die Ermittlung der Ground Truth aus natürlichen Bildern ist ein mit Unsicherheiten behaftetes Problem. Im vorliegenden Fall geht es um die Segmentierung urbaner Objekte. Die Abgrenzung von Häusern mit Vorsprüngen und überdeckender Vegetation im Grenzbereich der Auflösung beinhaltet zwangsläufig Meßungenauigkeiten. Kraftfahrzeuge sind für den menschlichen Betrachter meist aus dem Kontext zu erschließen, eine eindeutige Signatur ist nicht gegeben, zumal sie aus verschiedenen Farbflächen zusammengesetzt erscheinen (Fensterscheiben, andersfarbige Dachflächen) und die Abgrenzungen auf jeden Fall im Subpixelbereich liegen.

Um die Zahl der Freiheitsgrade für die statistische Auswertung einzuschränken, wurde eine Maske erstellt (Abbildung 6.1), die „Häuser“ mit den Maßen 10 x 20 Pixel und „Kraftfahrzeuge“ mit den Maßen 2 x 5 Pixel enthielt. Diese Objekte wurden in 30°-Schritten rotiert. So entstanden bei den Objekten in den Bildachsen scharfkantige Objektbegrenzungen, bei den Objekten in 30° und 60° Richtung eine Mittelung mit Grauwerten entsprechend dem Histogramm in Abbildung 6.1b. Die Randwerte 0 und 255 wurden hier der Übersichtlichkeit halber weggelassen, sie betragen für 0 (schwarz) 76510 und für 255 (weiß) 164914. Die Maske enthält also die typischen Subpixeleffekte. Sie enthält je 420 größere und kleinere Objekte in insgesamt sechs Richtungen.

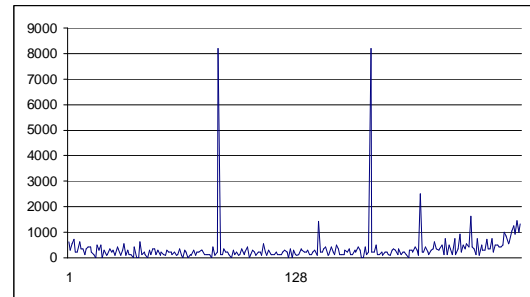
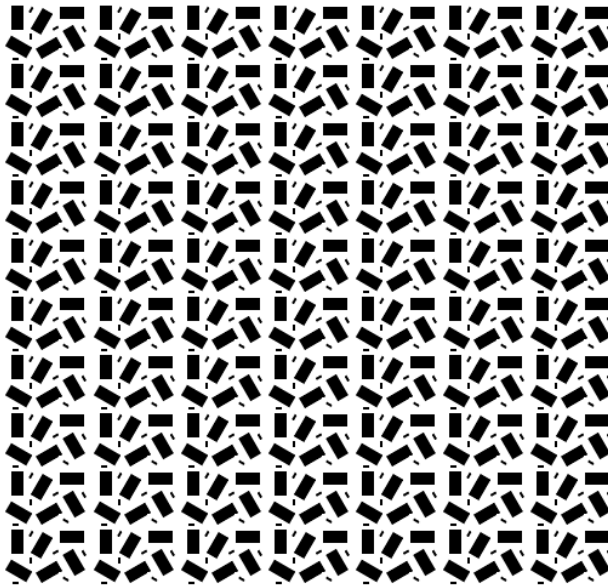


Abbildung 6.1a, b

Wenn man bei dieser Grafik das Rauschmaß nach dem auf Seite 25 (Abschnitt 3.3.2) beschriebenen Verfahren bestimmt, erhält man die in Abbildung 6.2 gezeigte Verteilung.

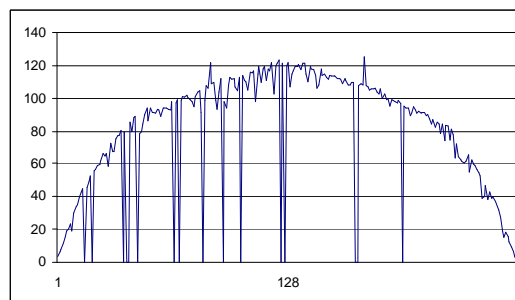


Abbildung 6.2

Rauschmaß für die Ground Truth aus Abbildung 6.1

Diese extremen Standardabweichungen sind natürlich nicht durch Bildrauschen bedingt, sondern dadurch, daß es in der Grafik außer für die Werte 0 und 255 keine homogenen Bereiche gibt, sondern alle Zwischenwerte Ausdruck der steilen Gradienten an den Rändern der Objekte sind. Für eine derartige Grafik ist die oben genannte Bestimmung des Rauschmaßes also nicht angezeigt. Dies ändert sich, sobald ein homogener Verlauf hinterlegt wird. Das auf Seite 43 in Abbildung 4.25 segmentiert gezeigte Bild wurde als Hintergrund verwendet und ein Summenbild nach folgender Formel erzeugt:

$$p = b + (f - b) \frac{255 - c}{255} \quad (6.1)$$

Dabei bezeichnet c den Grauwert des Pixels aus der Vordergrundmaske. An den Objekträndern sind bei der Rasterung durch die Subpixeleffekte Werte im Bereich zwischen 0 und 255 entstanden. f ist der gewünschte Sollwert des Vordergrundpixels mit seinem Verlauf von 0 bis 255 über das ganze Bild. b

ist der Grauwert des Hintergrundes aus Abbildung 4.25 und p ergibt schließlich den Grauwert des einzutragenden Pixels.

Auf diese Art sind alle Kontraste zwischen 0 und ± 127 Grauwerten im Bereich von 5 bis 250 vertreten und die Grenzen der maschinellen Segmentierung können präzise ausgelotet werden.



Abbildung 6.3

Die Ground Truth umfaßt alle Kontraste zwischen Vorder- und Hintergrund zwischen 0 und 127. Die großen Objekte sind 10 x 20 Pixel groß und entsprechen stilisierten Häusern bei einer Auflösung von 1m/Pixel, die kleinen Objekte sind 2 x 5 Pixel groß und entsprechen stilisierten PKWs.

6.1.2 Parametereinstellung

Zunächst sollte die optimale Parametereinstellung für die Segmentierung gefunden werden.

Entsprechend der Empfehlung aus dem eCognition User Guide³⁹ (S. 205) wird die erste Segmentierung für die kleinsten Objekte optimiert. Die Standardeinstellung mit einer Einstellung des Scale-Parameters von 10, einem Wert von 0,1 für den Shape-Parameter und einer Compactness von 0,5 stellt die kleinen Objekte überzeugend dar. Die Formen der schräg verlaufenden Objekte weichen mitunter deutlich von der rechteckigen ab, aber geringe Variationen der Parametereinstellungen erbringen keine sichtbaren Verbesserungen.

³⁹<http://www.definiens-imaging.de/documents/index.htm>

Die größeren Objekte sind aufgrund des inneren Grauwertverlaufs von bis zu 10 Grauwerten in bis zu drei Segmente unterteilt. Da diese Unterteilung auf dem Grauwertverlauf beruht, wurde für die Klassifizierung dieser Objekte ein weiteres *Object Level* mit veränderten Parametern erzeugt.

Eine Veränderung des Scale-Parameters war weniger effektiv als die Veränderung des Shape-Parameterfaktors auf seinen Grenzwert von 0,9. Eine weitere Verbesserung erbrachte die Einstellung der Smoothness auf 0,8. So wurden diese Einstellungen für die weiteren Versuche beibehalten.

6.1.3 Segmentierung und Klassifikation der Ground Truth

In Abbildung 6.4 und Abbildung 6.5 sind die Segmentierungsergebnisse nach dem neuen Algorithmus ab Version 3.0 und dem vom Hersteller als obsolet bezeichneten Algorithmus der Version 2.1 für die Einstellungen des ersten Levels gegenübergestellt. Nennenswerte Unterschiede sind im Gegensatz zur Segmentierung der weißen Fläche nicht erkennbar.

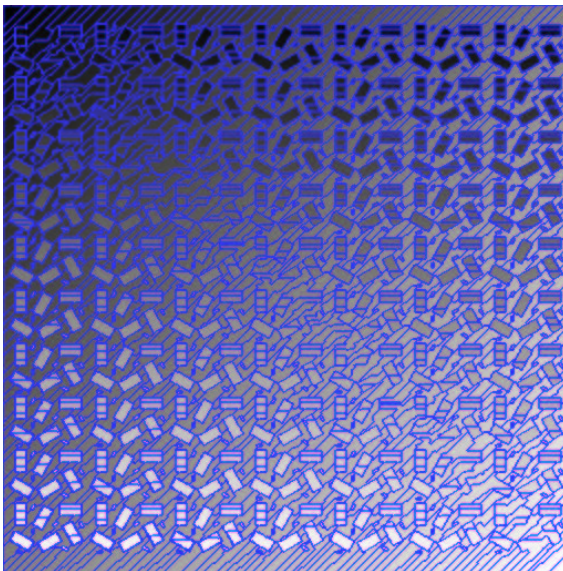


Abbildung 6.4

Segmentierung nach dem Algorithmus ab Version 4.0
(Scale = 10, Shape = 0,1, Smoothness = 0,5)

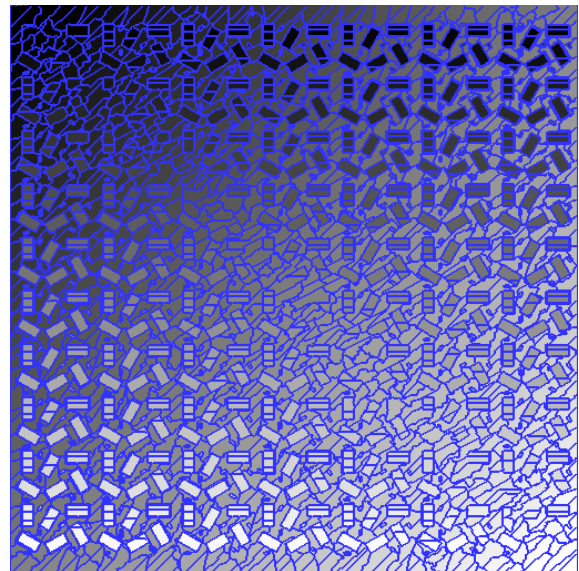


Abbildung 6.5

Segmentierung mit dem obsoleten Algorithmus der
Version 2.1, sonst gleiche Einstellungen wie in
Abbildung 6.4

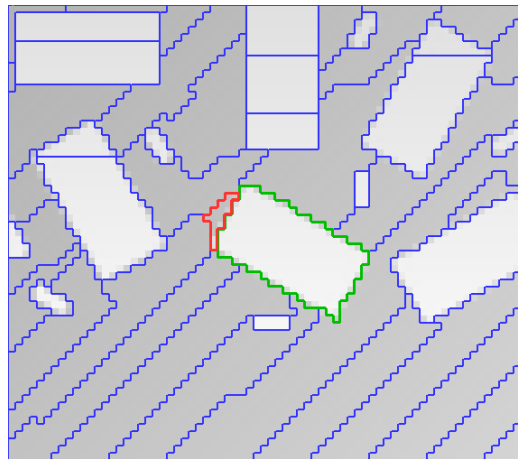


Abbildung 6.6

Ausschnitt aus dem unteren Bildrand aus Abbildung 6.4. Bei unvertauschten Daten reicht eine Grauwertdifferenz von nur einem Grauwert zur sicheren Segmentierung. Das rot markierte Segment wird falsch klassifiziert werden.

Für die Klassifikation der erhaltenen Segmente wurden folgende Einstellungen gewählt:

Die kleinen Objekte wurden in drei Gruppen eingeordnet: Eine Gruppe für die parallel zu den Bildkanten verlaufenden, und zwei Gruppen der mit Winkeldifferenzen von 30° zu den Bildkanten ausgerichteten Objekte. Die eine Gruppe wurde aus den eher in Richtung des diagonalen Grauwertverlaufs orientierten Objekten gebildet, die andere aus den quer dazu liegenden. Die Motivation für ein solches Vorgehen ergab sich daraus, daß die primäre Segmentierung der in Richtung des Grauwertverlaufs orientierten Objekte schwieriger ist, weil die Farbe innerhalb des Objektes mit zwei bis drei Grauwerten stärker changiert.

Nach den im vorigen Abschnitt gemachten Feststellungen waren unterschiedliche Segmentierungsebenen für ein optimales Klassifikationsergebnis erforderlich. Die kleinen Objekte wurden deshalb schließlich in der ersten Segmentierungsebene aufgesucht, was durch eine treppenförmige Membershipfunktion nach Abbildung 6.7 für die auszuwählende Ebene eingestellt werden kann.

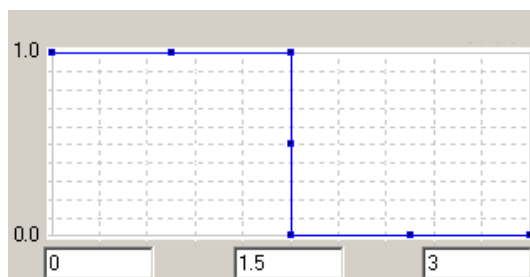


Abbildung 6.7

Treppenförmige Membershipfunktion zur „harten“ Abgrenzung

Die Membershipfunktionen können mit Hilfe der sieben Kontrollpunkte beliebig gezogen werden.

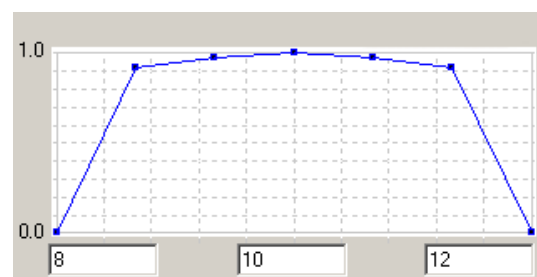


Abbildung 6.8

Wannenförmige Membershipfunktion zur „fuzzy“-Klassifizierung

Die Fläche (Area) wurde mit der wannenförmigen Membershipfunktion nach Abbildung 6.8 im Bereich 8 bis 12 Pixel für die orthogonalen Objekte und 8 bis 15 Pixel für die schrägen Objekte einge-

stellt. Letzteres war deshalb erforderlich, weil durch die Subpixeleffekte die Segmentgrößen variierten. Schließlich wurde die Achsenrichtung als Selektionskriterium herangezogen. Dieses Merkmal wäre in einem Bild einer realen Szene zur Erkennung von Kraftfahrzeugen natürlich nicht geeignet, war für die hier betrachteten Merkmale jedoch ideal. Die Achsenrichtung kann im Bereich $0^\circ \dots 180^\circ$ eingestellt werden. Die waagerechten Objekte werden durch eine Membershipfunktion in Rechteckform im Intervall [89, 91] erfaßt, für die lotrechten Objekte berechnet eCognition eine Achsenrichtung von einheitlich $179,97^\circ$, so daß diese durch eine Membershipfunktion spiegelsymmetrisch zu Abbildung 6.7 mit einem Intervall von [178, 180] erfaßt werden konnten. Beide Membershipfunktionen wurden mit der Funktion $\text{or}(\max)$ eingefügt. Zur Diskriminierung der schrägen Objekte wurden für die Achsenrichtungen die wannenförmigen Funktionen nach Abbildung 6.8 mit Intervallgrenzen von [20, 40] $\text{or}(\max)$ [50, 70] bzw. [110, 130] $\text{or}(\max)$ [140, 160] gewählt.

Die größeren Objekte wurden ausschließlich im zweiten Level klassifiziert. Verwendet wurde die Fläche mit der wannenförmigen Membershipfunktion mit den Intervallgrenzen [180, 230] und der Border length [56, 64] $\text{or}(\max)$ [74, 84] zur Erfassung der orthogonalen bzw. schrägen Objekte.

Abbildung 6.9 zeigt die Klassifikationsergebnisse für die kleinen Objekte. Vier Fehlzuordnungen an Kanten der größeren schrägen Objekte wurden mit gelben Kreisen hinterlegt. Die orangefarbenen Kreise markieren kleinere Objekte im Bereich der geringen Grauwertdifferenzen, die visuell deutlich erkennbar waren, mit den vorliegenden Parametereinstellungen aber nicht mehr segmentiert wurden.

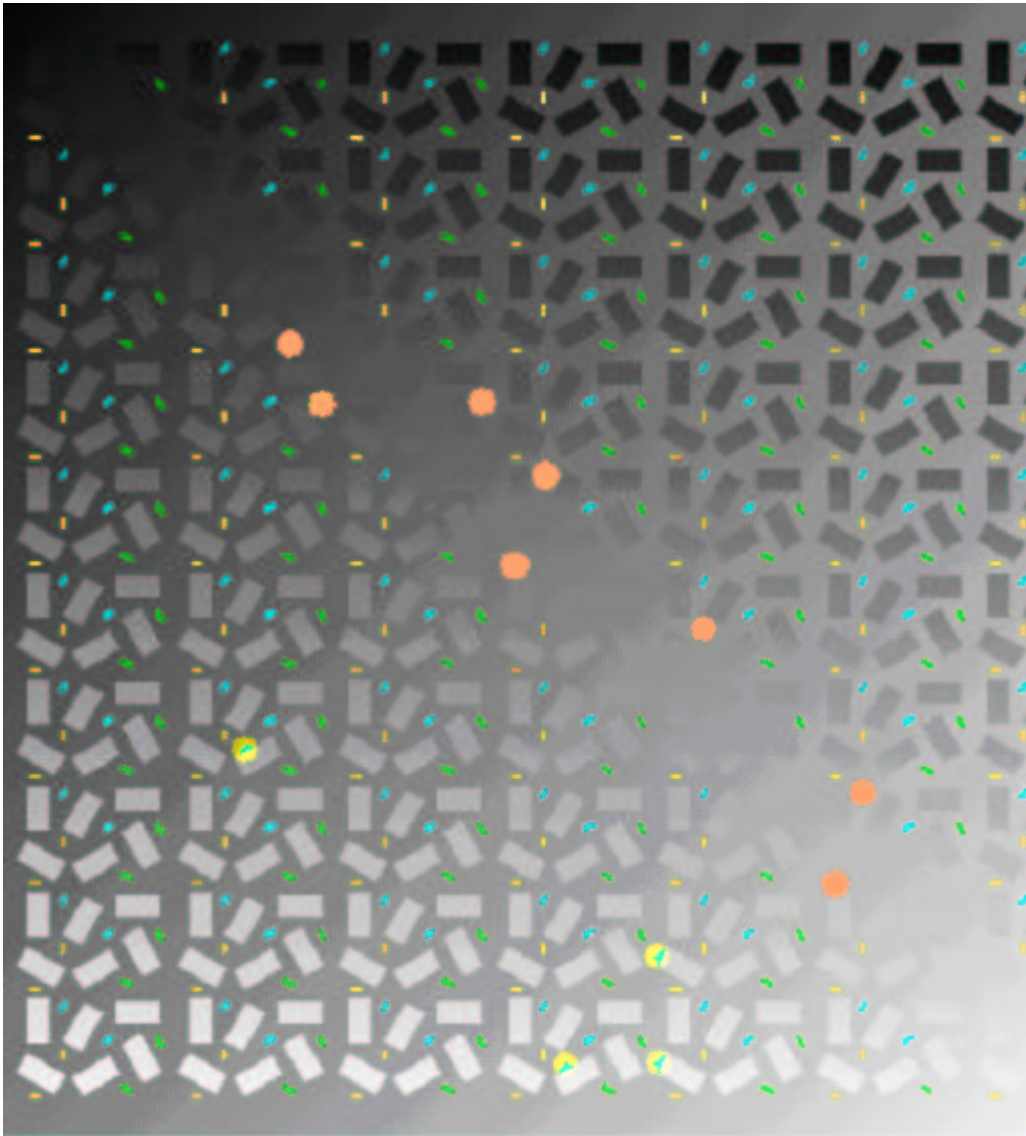


Abbildung 6.9

Abbildung 6.9 zeigt das Klassifikationsergebnis für die kleinen Objekte in den unverschmutzten Originaldaten. Vier Fehlzugeordnungen an Kanten der größeren schrägen Objekte wurden mit gelben Kreisen hinterlegt. Die orangefarbenen Kreise markieren kleinere Objekte im Bereich der geringen Grauwertdifferenzen, die visuell deutlich erkennbar waren, mit den vorliegenden Parametereinstellungen aber nicht mehr segmentiert werden konnten.

Inneres	Umgebung	Differenz
75	70	-5
90	82	-8
90	97	7
106	109	3
126	120	-6
141	148	7
178	185	8
198	192	-6

Tabelle 6.1

Die Grauwertdifferenzen der nicht mehr erkannten kleinen Objekte. (In Abbildung 6.9 orange hinterlegt.)

Zusammenfassend bleibt das Ergebnis, daß auch bei unverrauschten Daten nicht sicher vorhergesagt werden kann, ob und wie Flächen segmentiert werden. Einerseits kann eine Grauwertdifferenz von nur einem Wert ausreichend sein, andererseits kann es vorkommen, daß visuell gut unterscheidbare Gebiete mit einer Grauwertdifferenz von bis zu 8 nicht separiert werden. Generell sind größere Gebiete besser zu segmentieren. Die gewählten Objekte mit einer Größe von 2 x 5 Pixeln bewegen sich tatsächlich an der Grenze des Auflösungsvermögens.

Im folgenden Abschnitt wird die Wirkung der verschiedenen Filter auf das Segmentierungsergebnis gezeigt. Dabei interessieren folgende Fragen:

- Was passiert bei zunehmendem Rauschanteil im Bild?
- Wie stark darf ein Bild verrauscht sein, um noch auswertbar zu sein?
- Verändern Filter das Segmentierungsergebnis?
 - Verbessern Filter das Ergebnis?
 - Oder können Filter das Ergebnis verschlechtern?

Diese Fragen will der nachfolgende Abschnitt für die beschriebene Ground Truth beantworten.

6.1.4 Ergebnisse

Die Ground Truth wurde nach dem in Abschnitt 3.3.1 (Seite 24) beschriebenen Verfahren verrauscht. Entsprechend den von Schowengerdt [2002] und Baltsavias et al. [2001] beschriebenen Daten aus IKONOS-Aufnahmen (vgl. Abschnitt 3.3.2) wurde als Standardabweichung für das zugefügte Rauschen ein Wert aus der Menge {2, 3, 4, 5, 10, 20, 40} gewählt.

Für die kleinen Objekte ändert sich das Klassifizierungsergebnis in Abhängigkeit vom Grad des zugefügten Rauschens zunächst kaum. Bei mehr als mit $\sigma = 10$ zugefügtem Rauschen fällt die Anzahl der richtig segmentierten Flächen stark ab. Außerdem explodiert die Anzahl der falsch klassifizierten Flächen, so daß eine sinnvolle Klassifizierung nicht mehr möglich ist. Bei den größeren Flächen ist das Ergebnis ähnlich, wenn auch bei weitem nicht so ausgeprägt (Abbildung 6.10 a – c).

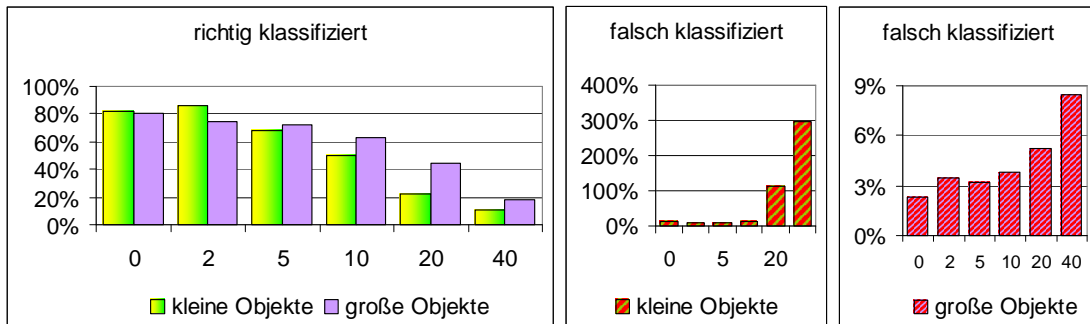


Abbildung 6.10 a – c

Anzahl der richtig und falsch klassifizierten Pixel in den ungefilterten Daten in Abhängigkeit von der Menge des zugefügten Rauschens.

Abszisse: Standardabweichung σ des zugefügten Rauschens.

Ordinate: Anteil der richtig bzw. falsch klassifizierten Pixel bezogen auf die Zahl der Pixel der zugehörigen Ground Truth.

Bei einem Rauschmaß von mehr als $\sigma = 10$ ist für die kleinen Objekte, bei mehr als $\sigma = 20$ auch für die großen Objekte keine sinnvolle Segmentierung mehr möglich. Etwas Rauschen kann die Erkennungsleistung verbessern.

In Abbildung 6.1a auf Seite 62 wurde der Vordergrund für die Klassifikation dargestellt. Abbildung 6.10a listet nun den Prozentsatz der richtig klassifizierten Pixel in den ungefilterten Bildern auf. 100% konnten nicht erreicht werden, weil in einer Diagonalen die Grauwerte von Vorder- und Hintergrund übereinstimmen. Es wird aber deutlich, wie bei den kleinen Objekten ein bißchen Rauschen die Erkennungsleistung verbessert, um dann bei Werten über $\sigma = 10$ stark abzusinken. Umgekehrt ist die Rate der Fehlklassifikationen bis $\sigma = 10$ ungefähr konstant, um dann stark anzusteigen. Bei den größeren Objekten ist dieser Effekt nicht so ausgeprägt, bei $\sigma = 20$ wird jedoch weniger als die Hälfte der Objekte erkannt und auch die Fehlklassifikationen nehmen zu.

Die so verrauschten Ausgangsdaten wurden nun mit den verschiedenen im Kapitel 5 beschriebenen Algorithmen gefiltert. Abbildung 6.11 dokumentiert das überraschende Ergebnis: Die verschiedenen Filter konnten das Ergebnis wenn überhaupt nur unwesentlich verbessern. Alle Filter, die nicht betont kantenerhaltend arbeiten, verschlechterten das Klassifikationsergebnis zum Teil dramatisch. Insbesondere der gern genutzte Mittelwertfilter reduzierte die Zahl der richtig klassifizierten Pixel auf 20%. Diese Ergebnisse galten für alle auswertbaren Bilder bis zu einem zugefügten Rauschen von $\sigma = 10$.

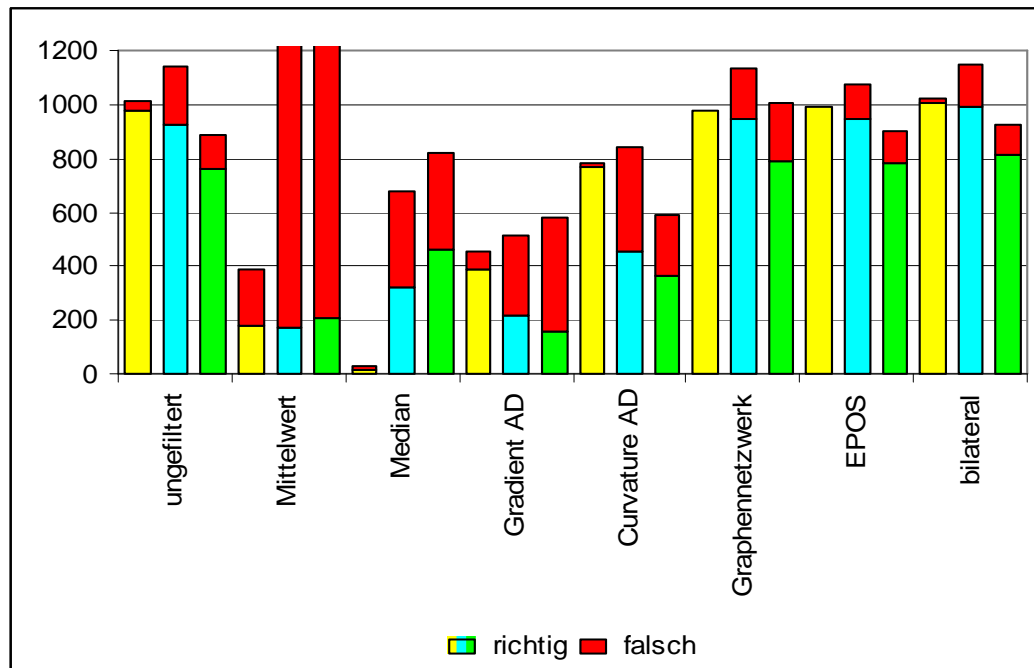


Abbildung 6.11

Das Bild wurde mit $\sigma = 5$ verrauscht. Dargestellt ist die Anzahl der klassifizierten Pixel der kleinen Rechtecke in den verschiedenen Richtungen aus Abbildung 6.9. Rot eingefärbt ist die Anzahl der falsch klassifizierten Pixel. Bei mit $\sigma = 10$ verrauschten Daten nimmt die Zahl der richtig klassifizierten Pixel insgesamt um etwa 30 % ab (vgl. Abbildung 6.10a).

Kein Filter kann das Ergebnis merklich verbessern, manche Filter aber erheblich verschlechtern.

Die Farben Gelb, Cyan, Grün in Abbildung 6.11 entsprechen den Farben der klassifizierten Objekte in Abbildung 6.9. Die Rechtecke der grün markierten Klasse verlaufen in Richtung der Grauwertabstufung des Hintergrundes und sind deshalb schwieriger zu klassifizieren. Die Kanten der gelb markierten Rechtecke verlaufen in Richtung der Koordinatenachsen. Sie haben an den Rändern keine Mischpixel und sind deshalb am besten zu segmentieren, obwohl hier die Segmentierungskriterien bezüglich der Fläche am strengsten waren. Nach Mittelwertfilterung nahm die Zahl der richtig klassifizierten Pixel stark ab. Dies liegt daran, daß die Flächen durch die Mittelung der Randpixel wesentlich verkleinert wurden. Veränderte Klassifizierungskriterien hätten dem abhelfen können, obwohl man damit in Kauf hätte nehmen müssen, daß die Grenzen der Segmente nicht mehr korrekt gezogen worden wären. Durch die Mittelwertfilterung wurden jedoch auch vorher kleinräumig zufällige Grauwertverteilungen großräumiger. So wurden viele die Klassenmerkmale erfüllenden Segmente gebildet, die eine Verbesserung des Klassifikationsergebnisses durch veränderte Klassenmerkmale unmöglich machten.

Bei stärkerem Rauschen wie in Abbildung 6.12 mit $\sigma = 20$ dargestellt, übersteigt die Anzahl der falsch klassifizierten Pixel die Anzahl der richtig klassifizierten bei weitem, so daß die Ergebnisse in einem automatisierten Prozeß nicht mehr sinnvoll weiter verwendet werden können.

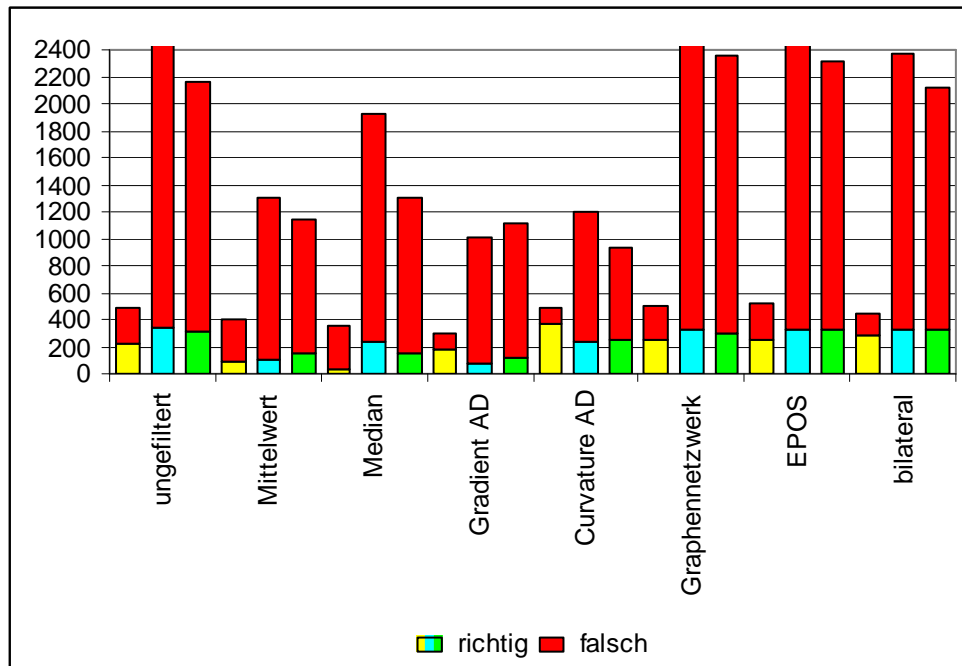


Abbildung 6.12

Bei mit $\sigma = 20$ verrauschten Daten ist ein sinnvolles Ergebnis für die kleinsten Objekte weder mit noch ohne Filterung zu erhalten.

In Abbildung 6.9 wurde eine Gruppe diagonalen Rechtecke grün markiert. Die nachfolgende Abbildung 6.13 gibt noch einmal eine Übersicht über die Anzahl der richtig bzw. falsch erkannten Rechtecke. Ausgangsdaten waren hier unterschiedlich verrauschten Daten nach Bearbeitung mit den besprochenen verschiedenen Filtern.

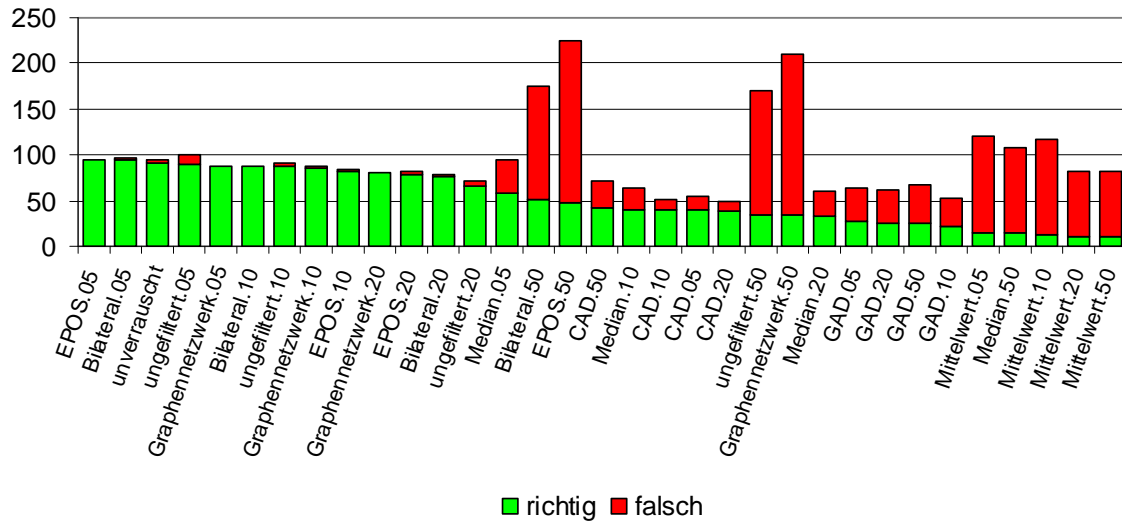


Abbildung 6.13

Zusammenfassung der Leistung verschiedener Filter für die in Abbildung 6.9 grün dargestellten kleinen Rechtecke. Die Ziffern geben an, wie weit die Daten verrauscht wurden. Die Ordinate gibt die Anzahl der richtig bzw. falsch klassifizierten Rechtecke an.

Nach diesen Ergebnissen scheinen Filter generell wenig nützlich zu sein für die Verbesserung der Segmentierungsergebnisse von eCognition. Etwas Rauschen kann unter Umständen sogar die Qualität der Segmentierung verbessern, wie Abbildung 6.10 a im Falle der kleinen Objekte nahelegt.

Im nächsten Kapitel werden die Grenzbereiche der Segmentierungsmöglichkeiten näher betrachtet, es wird untersucht, wie groß der Grauwertunterschied für eine sichere Kantenerkennung sein muß. Aus der Vielzahl möglicher Kombinationen von Filtern werden die geläufigsten ausgesucht und näher analysiert, insbesondere soll geprüft werden, wie sich Verfahren zur Kantenschärfung auswirken, die die Unschärfe des Mittelwertfilters wieder ausgleichen sollen.

6.2 Schwache Kontraste

Im vorigen Kapitel wurde die Auswirkungen des Rauschens und der verwendeten Filter auf Kanten mit relativ starken Kontrasten untersucht. In diesem Kapitel sollen die Untersuchungen ausgedehnt werden auf Objekte, die sich nur wenig vom Hintergrund abheben. Dazu wurde ein neues Testbild entwickelt, bei dem sich die Testobjekte nur wenig vom Hintergrund abheben. Ausgangspunkt waren neunzehn Rechtecke der Größe 5 x 11 Pixel mit Grauwerten von 129 bis 147 vor einem Hintergrund mit dem Grauwert 128. Diese Rechtecke wurden in Schritten von 10° gedreht. Bei dieser Drehung entstehen durch die Rasterung der vektoriell erzeugten Rechtecke an den Rändern Mischpixel, wie sie auch in Fernerkundungsaufnahmen an Objektgrenzen zu finden sind.

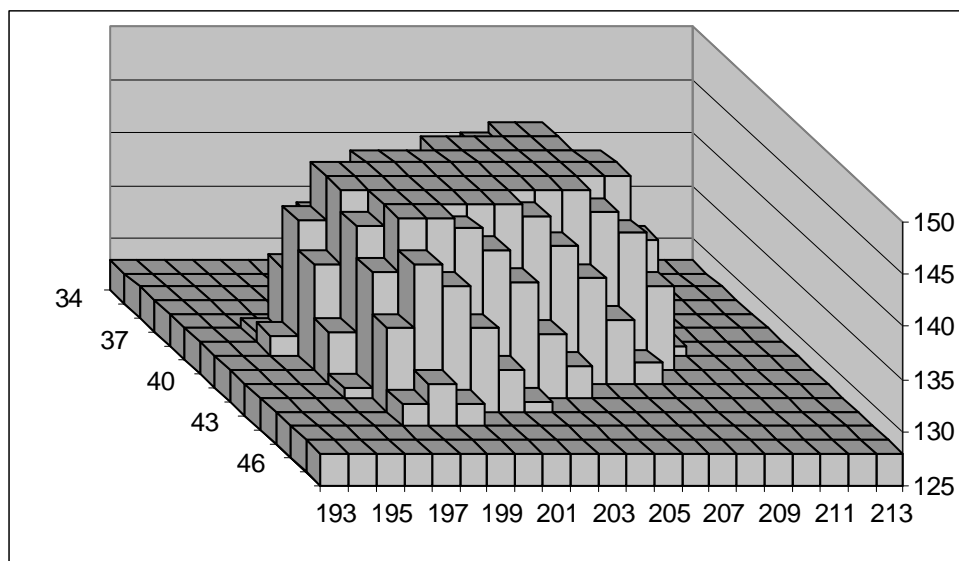


Abbildung 6.14

Die Abbildung zeigt eine 3-D-Ansicht eines um 70° gedrehten Rechtecks aus der unverrauschten Ground Truth. x- und y-Achse sind mit den Bildkoordinaten beschriftet, die senkrechte z-Achse gibt den Grauwert an. Die gestuften Ränder ergeben sich durch die Mischpixel bei der Umwandlung des Vektorobjektes Rechteck in die Rasterdarstellung.

Abbildung 6.15 zeigt die Ground Truth für die Untersuchungen verschiedener Filter bei Objekten mit geringen Kontrasten zum Hintergrund. Der Kontrast reicht von einer Graustufe (unterste Reihe, visuell nicht erkennbar) bis 19 Graustufen in der obersten Reihe.

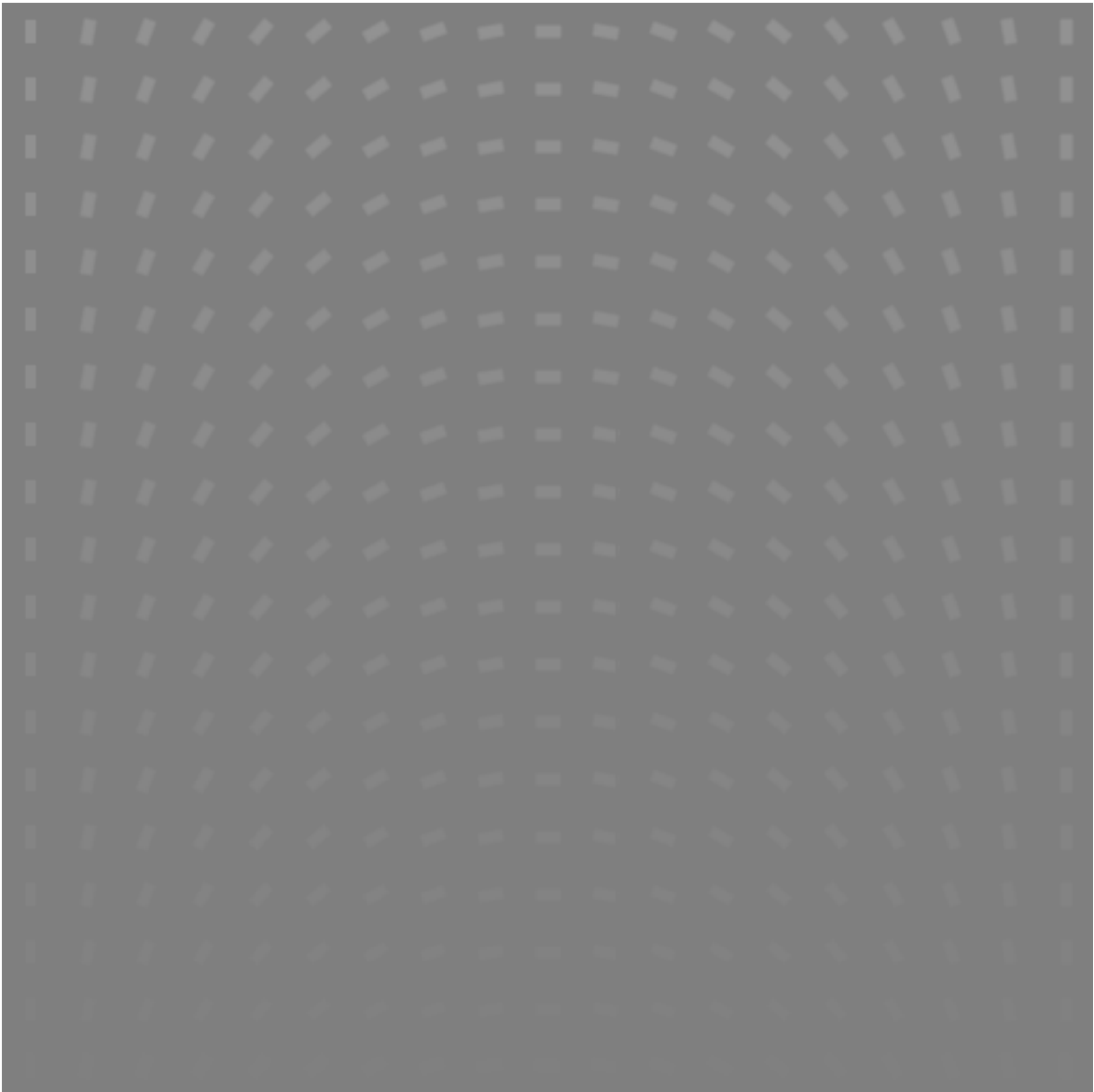


Abbildung 6.15

Ground Truth für die Filterwirkungen bei schwachen Kontrasten. Der Hintergrund hat einen einheitlichen Grauwert von 128, die Rechtecke im Vordergrund sind 5 x 11 Pixel groß und wurden jeweils um 10° weiter gedreht. Der Kontrast reicht von einer Graustufe (unterste Reihe, visuell nicht erkennbar) bis zu 19 Graustufen in der obersten Reihe.

Die so erzeugte Ground Truth wurde mit normalverteiltem Rauschen verrauscht. Als Maß für das zugefügte Rauschen wurden als Verteilungsparameter $\sigma \in \{1, 2, 3, 4, 5, 10, 20\}$ gewählt. Ein Verteilungsparameter $\sigma = 5$ gibt dabei an, daß die Standardabweichung einer Fläche mit vorher konstantem Grauwert nach Addition des Rauschens 5 beträgt. (Strukturierte Flächen ergeben andere Standardabweichungen, weil die Objektstrukturen die Standardabweichung erhöhen, vergl. Kapitel 3.3.2, Seite 25.)

Das Rauschen in IKONOS-Aufnahmen ist intensitätsabhängig. Diese Intensitätsabhängigkeit muß im Testbild nicht simuliert werden, weil die Intensitäten im Bild nur wenig differieren und verschiedene Rauschintensitäten untersucht werden. So können die Ergebnisse mit stärkeren Rauschanteilen die Auswirkungen in helleren und deshalb stärker verrauschten Bildbereichen darstellen und Ergebnisse mit geringeren Rauschanteilen solche in dunkleren Bildbereichen.

Für die visuelle Interpretation werden im Rahmen der Bildvorverarbeitung gern Weichzeichner und / oder kantenverstärkende Filter eingesetzt. Möller [2003] benutzte einen von Ford et al. [1983] vorgeschlagenen modifizierten Gauß-Filter als Tiefpaß und einen abgeschwächten Laplace-Filter in der von Bernstein [1983] verwendeten Form zur Kantenverstärkung. Diese Filter wurden allein oder zusätzlich zu den rauschvermindernden modernen Filtern eingesetzt.

Eine Standardeinstellung zur Segmentierung und Klassifizierung wird der Filterwirkung nicht gerecht und ist nicht sinnvoll, weil der Auswerter die Segmentierungsparameter und die Klassenbeschreibungen auf dem auszuwertenden, fertig vorverarbeiteten Bild optimiert oder optimieren sollte. Deshalb wurden die Segmentierungsparameter und Klassenbeschreibungen nach Filterung des Bildes individuell angepaßt, um optimale Erkennungsergebnisse zu erhalten. Beispielsweise ergibt die Segmentierung bei einem mit $\sigma = 5$ verrauschten ungefilterten Bild recht fraktale Segmentgrenzen wie im Ausschnitt in Abbildung 6.16 zu sehen ist.



Abbildung 6.16

Bei mit $\sigma=5$ -verrauschten Daten ergeben sich recht fraktale Segmentgrenzen.

Größere Werte für den Shape-Parameter-Index verbessern in diesem Fall das Klassifikationsergebnis. Die Filterung mit dem einfachen Mittelwert-Filter führt dazu, daß die Segmentgrenzen innerhalb der Rechtecke liegen, während nach Laplace- und anschließender Gradient Anisotropic Diffusion Filterung die Segmentgrenzen den visuellen Erwartungen entsprechen und deshalb eine Klassifizierung mit größeren Werten für Area und Borderlength bessere Ergebnisse zeitigt.

Eine Optimierung der Parametereinstellungen kann das beste Ergebnis nicht garantieren. Wegen der Vielzahl der möglichen Einstellungen ist es in menschlich verfügbarer Zeit nicht möglich, alle Einstellungen zu testen, deshalb ist eine heuristische Suche entsprechend den für den Mean-Shift-Algorithmus (vgl. Abschnitt 4.2.3.2, Seite 31) entwickelten Prinzipien erforderlich.

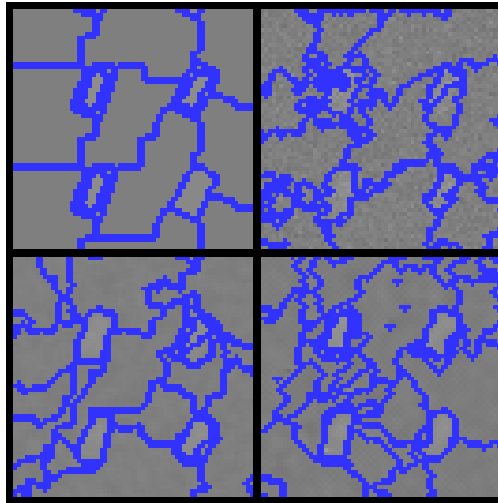


Abbildung 6.17 a – d

Wirkung der Filterung auf die Segmentierung:

- a) oben links: unverrauscht und ungefiltert, guter Kantenerhalt, jedoch ab einer Grauwertdifferenz > 10 separate Segmente für die Mischpixel an den Objektgrenzen.
- b) oben rechts: mit $\sigma = 5$ verrauscht, ungefiltert, sehr schlechter Kantenerhalt.
- c) unten links: mit $\sigma = 5$ verrauscht, Mittelwert-Filter, in der Regel guter Kantenerhalt.
- d) unten rechts: mit $\sigma = 5$ verrauscht, Laplace- und GAD-Filter. Stärkere Segmentierung, häufig bessere Objektdarstellung.

6.2.1 Segmentierung der erzeugten Bilder

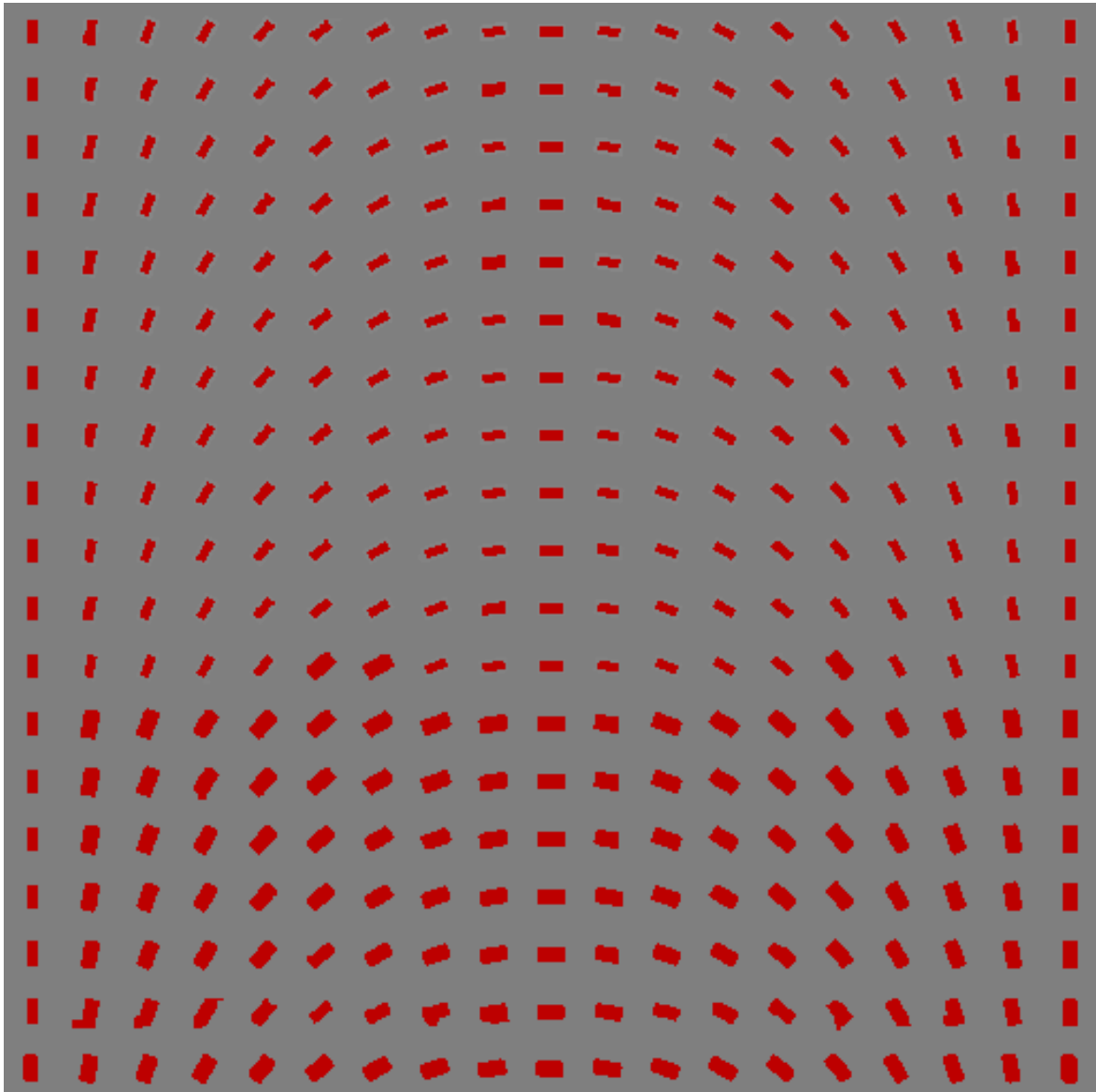


Abbildung 6.18

Klassifikationsergebnis der unverrauschten Ground Truth. Eine Grauwertdifferenz von 1 reicht für perfekte Segmentgrenzen aus.

Die Segmentierung der unverrauschten Ground Truth gelang mit den vorgegebenen Standardeinstellungen fast perfekt. Für die Klassifikation wurde Brightness $> 75,5$, Area zwischen 39 und 120 und Borderlength < 60 eingestellt. Das Kriterium Brightness allein wäre nahezu ausreichend gewesen. Besonders in der oberen Hälfte wurden jedoch um die Rechtecke weitere ringförmige Segmente gebildet, bedingt durch die unscharfe Kante aufgrund der Mischpixel wie auf Seite 73 besprochen. Nach einer Reduktion des Scale-Parameters von standardmäßig 10 auf 8 konnte das in Abbildung 6.18 gezeigte Ergebnis erhalten werden. Es sei noch einmal darauf hingewiesen, daß sich die Rechtecke in der untersten Reihe nur um einen Grauwert von ihrer Umgebung unterscheiden.

Nach Vorverarbeitung mit dem Mittelwert-Filter wurden 21 Rechtecke weniger erkannt. Dies lag im wesentlichen daran, daß durch die Mittelung die Rechtecke vor der Segmentierung verkleinert wurden, so daß das Area-Kriterium nicht mehr erfüllt wurde. Eine Verminderung der unteren Schranke führte jedoch andererseits zu einer Zunahme fehlerhafter Zuordnungen. Insbesondere wurden durch die glättungsbedingte Vergrößerung der Randbereiche der Rechtecke in der oberen Hälfte diese Randbereiche als eigene Objekte segmentiert.

Nach Filterung mit dem Median-Filter wurden im Vergleich zum Mittelwert-Filter noch einmal 14 Rechtecke weniger gefunden.

Der bilaterale Filter führte zu einer kräftigen Histogrammspreizung. Der vorherige minimale Grauwert von 128 wurde auf Null abgesenkt, der vorherige Maximalwert von 149 auf 240 angehoben. Dadurch verschwanden auch die Rechtecke in den beiden unteren Reihen. In der dritten Reihe wurden sie auf weniger als 26 Pixel verkleinert. Mit einem Scale-Parameter von 8 wurden die Rechtecke im oberen Teil in drei Segmente aufgeteilt. Mit einer Anpassung der Segmentierungsparameter an das Filterergebnis konnten ab der vierten Reihe alle 309 Rechtecke richtig klassifiziert werden. Dazu war ein Scale-Parameter von 25 erforderlich und eine Untergrenze von Area von 31.

6.2.2 Verrauschen und Filtern

Dieses Ausgangsbild (Abbildung 6.15) zeigte bei stärker verrauschten Daten insbesondere für die schlechter geeigneten Filter nur noch sehr geringe Differenzierungen. Zur besseren Darstellung wurde dieses Bild so modifiziert, daß die Rechtecke einen größeren Kontrast zum Hintergrund erhielten. So wurde der Grauwert des Hintergrundes von 160 am oberen Rand bis 130 am unteren Rand gleichmäßig variiert und die Rechtecke erhielten Grauwerte von 95 in der obersten Reihe bis 124 in der untersten Reihe. Das Ausgangsrechteck in der oberen linken Ecke ist 10 x 20 Pixel groß und wurde in Schritten von jeweils 12° von einer Spalte zur nächsten Spalte gedreht. Abbildung 6.19 zeigt die kontrastverstärkte Ground Truth für die folgende Analyse.

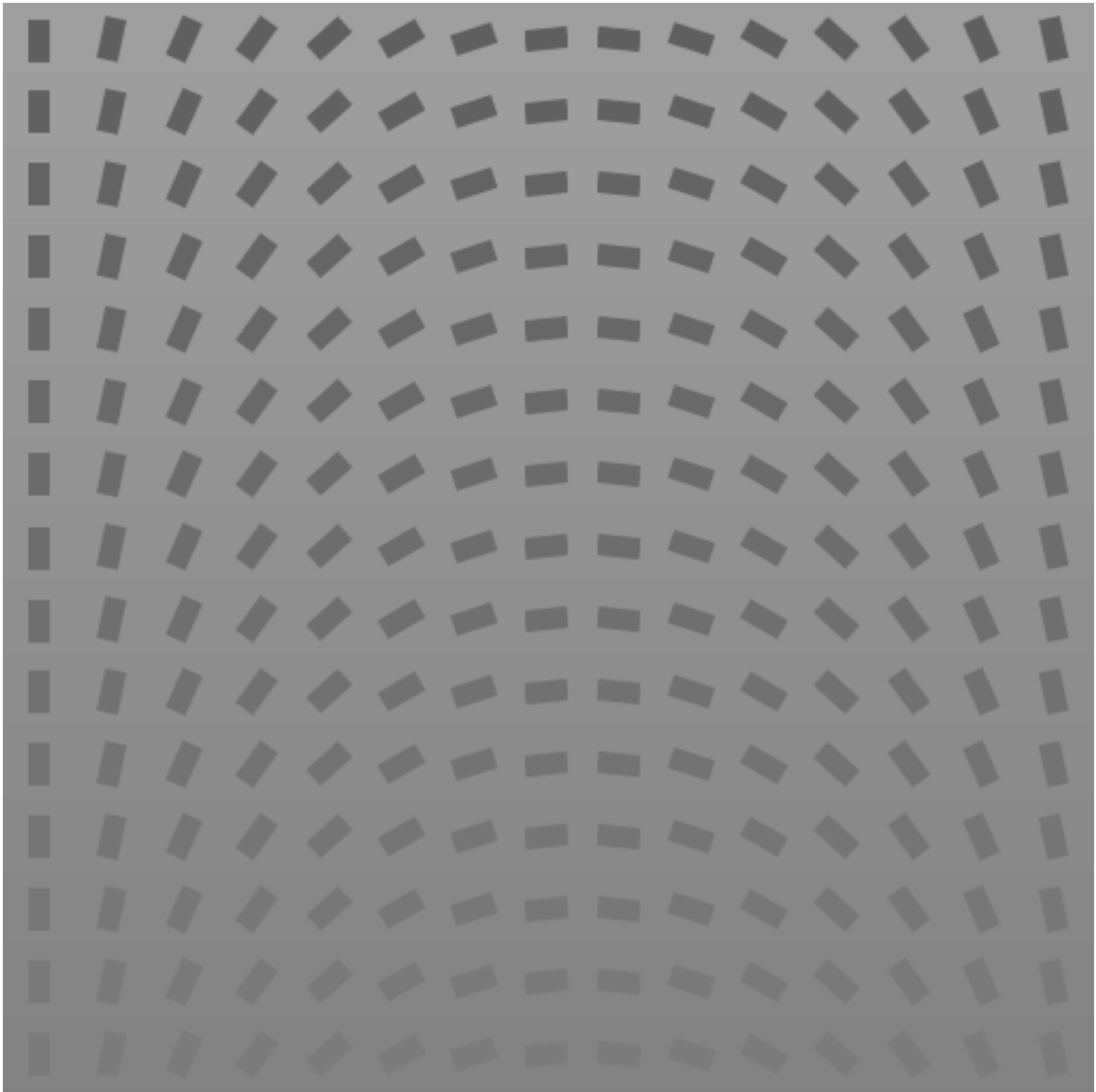


Abbildung 6.19

Kontrastverstärkte Version der Ground Truth

Diese Ground Truth wurde mit dem in Abschnitt 3.3.1 (Seite 24) beschriebenen Verfahren mit verschiedenen Werten für σ verrauscht. Anschließend wurde das Rauschen mit dem in Abschnitt 3.3.2 (Seite 25) beschriebenen Verfahren gemessen. Tabelle 6.2 beschreibt das Maß des zugefügten Rauschens, die Standardabweichung über das gesamte Bild, wie sie von eCognition berechnet wird und das gemessene Rauschen.

σ zugefügt	Standardabweichung	gemessen
0	15,7	0,46
1	15,7	0,69
2	15,8	1,17
3	16,0	1,67
4	16,2	2,18
5	16,5	2,69
7	17,2	3,73
10	18,6	5,29
15	21,7	7,86
20	25,4	10,40
25	29,5	12,92

Tabelle 6.2

Angegeben ist das Maß des zugefügten Rauschens nach Abschnitt 3.3.1, die Standardabweichung über das gesamte Bild, wie sie von eCognition berechnet wird und das nach dem in Abschnitt 3.3.2 beschriebenen Verfahren gemessene Rauschen.

Vergleicht man zugefügtes und gemessenes Rauschen, so fällt auf, daß der Zahlenwert des zugefügten Rauschens etwa doppelt so hoch ist wie der des gemessenen.

6.2.3 Parameterwahl

Sämtliche verwendete Filter haben Parameter, die ihre Wirkung beeinflussen. Bei der Vielzahl der möglichen Einstellungen ist es daher in endlicher Zeit nicht möglich, den „besten“ Wert für alle Filter und alle Bildinhalte zu finden. In der Regel wurde deshalb auf die in der Literatur jeweils vorgeschlagenen Einstellungen zurückgegriffen und diese, wo es sinnvoll erschien, variiert. Interessanter als einen absolut besten Filter zu finden, ist darum die Frage, wie sich die verschiedenen Filter auf den bisher ja noch nicht offengelegten Segmentierungsalgorithmus von eCognition auswirken, um für eine gegebene Situation einen sinnvollen Filter mit sinnvollen Einstellungen auswählen zu können. Aus diesem Grund werden im folgenden die Wirkungen der verschiedenen Filter an ausgewählten Beispielen näher beleuchtet. In den folgenden Beispielen werden die in Abschnitt 6.1 beschriebenen Objekte klassifiziert. Die erzeugende Fläche hat die Maße 10 x 20 Pixel und damit eine Fläche von 200 Pixeln und einen Umfang von 60 Pixeln. Für die Klassifizierung werden die in Tabelle 6.3 notierten Kriterien und Grenzen verwendet.

Kriterium	ist	von	bis
Area	200	176	232 (270)
Borderlength	60	57	91
Elliptic fit	0,82	0,69	0,88
length/width	2	1,57	2,2

Tabelle 6.3

Klassenbeschreibung für die Rechtecke

Je nach vorangegangener Filterung muß auch der Scale-Parameter von eCognition unterschiedlich eingestellt werden. Für die meisten Filter war ein Scale-Parameter von 20 eine gute Wahl. Speziell nach Filterung mit dem Mittelwert-Filter, auch mit nachfolgender Kantenschärfung mit dem modifizierten Ford- oder dem Laplace-Filter erzielte ein Scale-Parameter von 50 erheblich bessere Ergebnisse.

6.3 Ergebnisse

Die folgende Abbildung stellt Segmentierungsergebnisse für die ungefilterte, aber mit $\sigma = 10$ verrauschte Ground Truth dar.

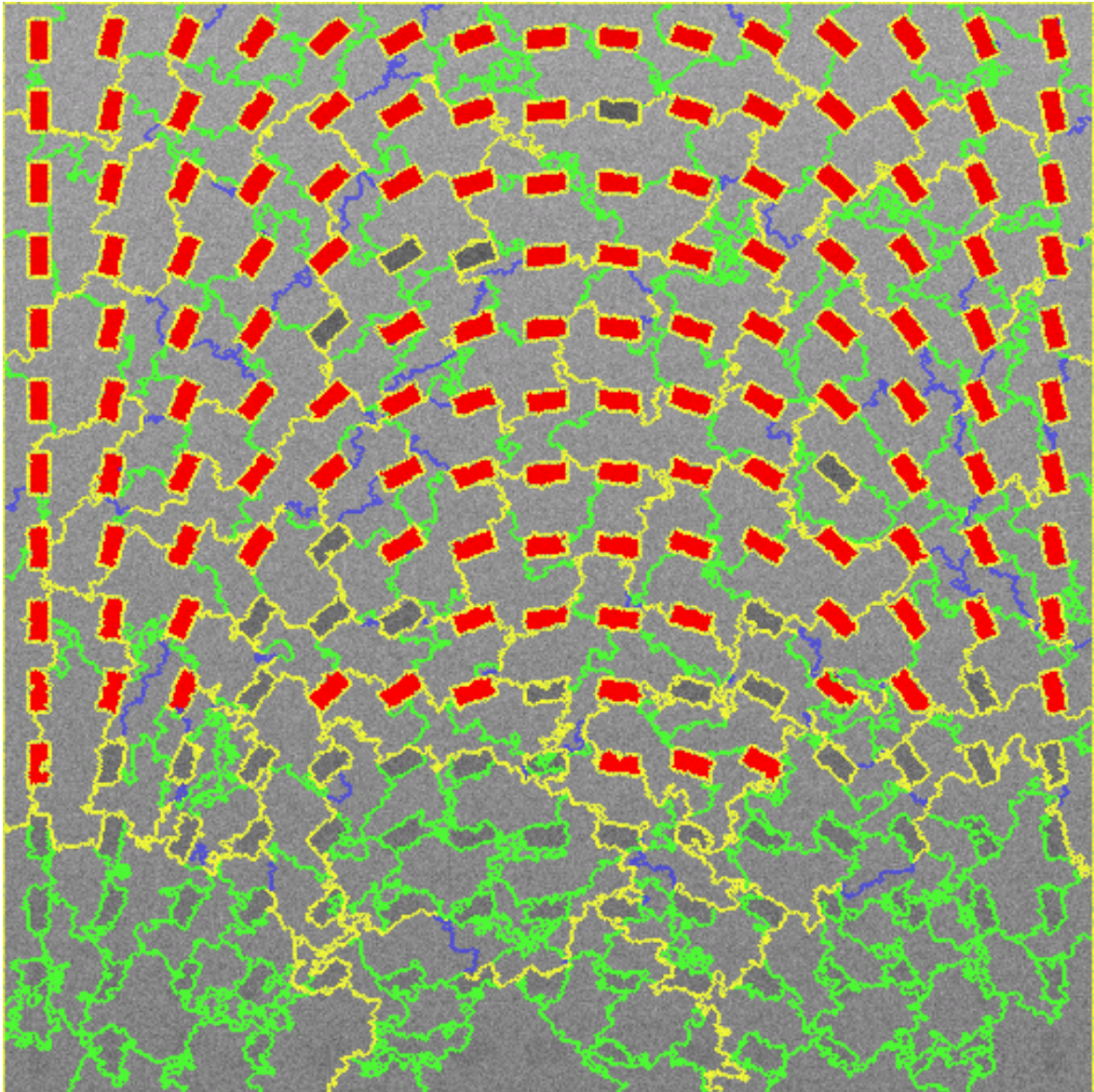


Abbildung 6.20

Mit $\sigma = 10$ verrauschte Ground Truth,

Blau und Gelb: Segmentgrenzen bei einem Scale-Parameter von 50

Grün und Gelb: Segmentgrenzen bei einem Scale-Parameter von 20

Rot: Rechtecke entsprechend der Klassenbeschreibung in Tabelle 6.3

In der Abbildung wurden die mit einer Scale-Parametereinstellung von 50 (Blau) und einer Scale-Parametereinstellung von 20 (Hellgrün) erzeugten Segmentgrenzen dargestellt. In beiden Einstellungen vorhandene Segmentgrenzen wurden gelb eingefärbt. Beide Läufe zur Segmentierung erfolgten im ersten Level, d.h. die Ergebnisse der jeweils anderen Segmentierung wurden nicht verwendet.

Der größere Teil der Segmentgrenzen ist für beide Scale-Parametereinstellungen identisch. Abbildung 6.21 zeigt die relativen Längen der Segmentgrenzen aus Abbildung 6.20, gemessen als Anzahl der Pixel in der entsprechenden Farbe.

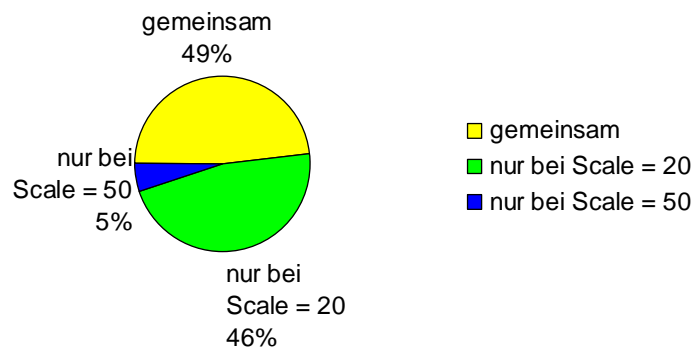


Abbildung 6.21

Der größere Teil der Segmentgrenzen ist für beide Scale-Parametereinstellungen identisch.

Die Wahl der Klassenbeschreibung bestimmt die Anzahl der klassifizierten Rechtecke. Von besonderem Interesse ist jedoch, wie die vorhandenen Kanten im Bild dargestellt werden. Daß einige Rechtecke nicht klassifiziert wurden, lag bei den gut erkennbaren Rechtecken im oberen Teil an der Auswahl durch die Formparameter *elliptic fit* und *length/width*, die auf Grund von Störungen im Randbereich als Ausschlußkriterien wirkten. Der untere Teil des Bildes mit seinen relativ schwachen Kontrasten wurde bei dem grob eingestellten Scale-Parameter sehr großflächig zusammengefaßt.

Im folgenden werden die Wirkungen der verschiedenen Filter speziell auf die Kantenverläufe näher analysiert.

6.3.1 Mittelwertfilter

Der in der Literatur offenbar beliebteste Filter ist der Mittelwert-Filter. Er ist, wie im folgenden gezeigt werden wird, jedoch für eine nachfolgende Segmentierung mit eCognition der am wenigsten geeignete Filter. Die Probleme entstehen dadurch, daß einerseits eCognition auf Rauschen im Bild sehr robust reagiert und andererseits durch die Filterung vorhandene Kanten unscharf werden. Durch die Unschärfe werden bei kleinen Werten für den Scale-Parameter neue Segmente um die Objektgrenzen gebildet oder bei größer gewähltem Scale-Parameter Randbereiche, die ursprünglich nicht zum Objekt gehörten, diesem zugeschlagen. Statt der 139 klassifizierten Rechtecke in Abbildung 6.20 wurde bei einem Scale-Parameter von 20 und unveränderter Klassenbeschreibung nach Mittelwert-Filterung nur noch 41 Rechtecke klassifiziert. Nach einer Vergrößerung des Scale-Parameters auf 50 und insbesondere einer Änderung der Klassenbeschreibung (die Obergrenze von Area wurde auf 270 erhöht), konnten wieder 124 Rechtecke klassifiziert werden.

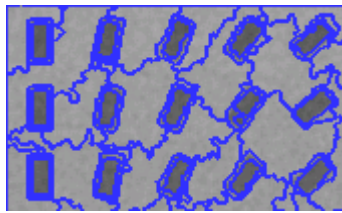


Abbildung 6.22

Segmentierung und Klassifizierung nach Mittelwert-Filterung, Scale = 20

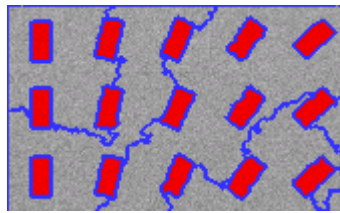


Abbildung 6.23

Segmentierung und Klassifizierung ohne Filterung, Scale = 50

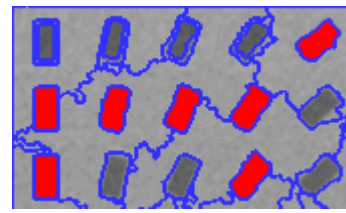


Abbildung 6.24

Segmentierung und Klassifizierung nach Mittelwert-Filterung, Scale = 50

Die durch die Mittelwert-Filterung erst entstehenden Probleme werden in den oben gezeigten Ausschnitten aus der linken oberen Ecke der Ground Truth verdeutlicht: Durch den Verlust von Kanteninformation wird das in seinem Grauwert gut abgrenzbare Objekt mit einem Hof segmentiert, so daß das Innere das Flächenkriterium nicht mehr erfüllt (Abbildung 6.22). Bei größerer Segmentierung kann ein verbessertes Klassifikationsergebnis ebenfalls nur durch eine Änderung des Flächenkriteriums erreicht werden, weil Randbereiche fälschlicherweise dem Rechteck zugewiesen wurden (gut erkennbar in der untersten Reihe in Abbildung 6.24).

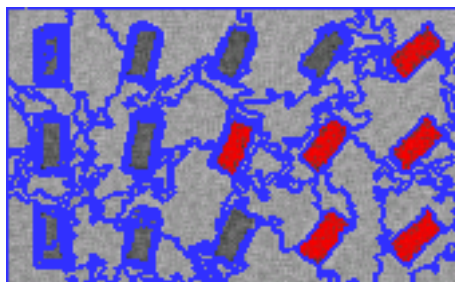


Abbildung 6.25

Mittelwert- und anschließende Laplace-Filterung erzeugen sehr unregelmäßige Objektbegrenzungen.



Abbildung 6.26

Auch eine Ford-Filterung nach Mittelwert-Filterung erzeugt separate Kantensegmente und ist wenig hilfreich.

Auch die nachfolgende Kantenschärfung mit einem Laplace⁴⁰- (Abbildung 6.25) oder Ford⁴¹-Filter (Abbildung 6.26) führt zwar dazu, daß einige Rechtecke mehr erkannt werden als nach der Mittelwert-Filterung, das Ergebnis bleibt aber erheblich schlechter als ohne Filterung.

6.3.2 Medianfilter

Die Ergebnisse des Median-Filters müssen differenziert betrachtet werden. Formal werden mit den gleichen Parametereinstellungen nach Median-Filterung (Abbildung 6.28) mehr Rechtecke richtig klassifiziert (bei einem Scale-Parameter von 50 in den Abbildungen 146 statt 139, mit einem Scale-Parameter von 20 sind es 151 statt 140).

⁴⁰ Beschrieben in Abschnitt 5.1.3.1, Seite 48

⁴¹ Beschrieben in Abschnitt 5.1.3.2, Seite 48

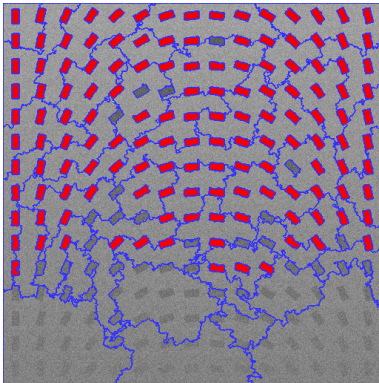


Abbildung 6.27

Ohne Filterung,
Scale = 50

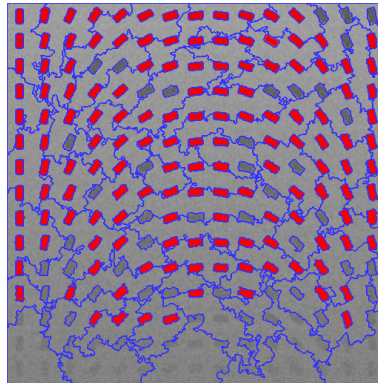


Abbildung 6.28

Median-Filterung,
Scale = 50

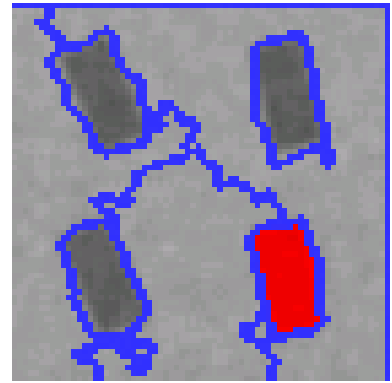


Abbildung 6.29

Ausschnitt aus Abbildung 6.28
rechts oben

Dies kommt daher, daß nach Medianfilterung Objekte mit schwachem Kontrast, wie in den unteren Bildabschnitten, besser segmentiert werden. Andererseits ist das Segmentierungsergebnis generell unzuverlässiger. Die gut abgrenzbaren Rechtecke in der oberen rechten Ecke werden nicht mehr klassifiziert, weil nach Filterung die Objektgrenzen stärker verändert werden und damit z. B. das Kriterium *Borderlength* nicht mehr erfüllt wird.

6.3.3 Filter mit anisotroper Diffusion

Von den verschiedenen kantenerhaltenden Glättungsfiltern, die im ITK-Toolkit zur Verfügung gestellt werden, wurden im Abschnitt 5.3.1 der Gradient Anisotropic Diffusion Filter und der Curvature Anisotropic Diffusion Filter vorgestellt. Beide Filter haben drei Parameter, mit denen sie gesteuert werden: die Leitfähigkeit (*conductance*), das zeitliche Intervall (*time step*) und die Zahl der Iterationen. Sie unterscheiden sich darin, daß der Gradient Anisotropic Diffusion Filter eine positive Kantenverstärkung aufgrund „negativer Diffusion“ erzeugen kann, während der Curvature Anisotropic Diffusion Filter ausschließlich glättend wirkt. In Vorversuchen hatte sich der Curvature Anisotropic Diffusion Filter als deutlich überlegen gezeigt, was er in dieser Versuchsreihe nicht bestätigen konnte. Beide Filter bewirken je nach Einstellung mehr oder weniger die auch vom Mittelwertfilter bekannte Tiefpaßfilterung mit Verlust von Kanteninformation und können in keinem Fall bessere Ergebnisse als die Originaldaten produzieren.

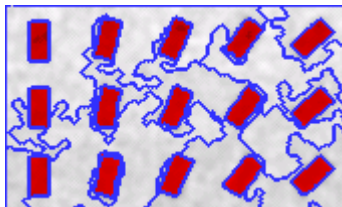


Abbildung 6.30

GradientAnisotropicDiffusion-Filter
5 0,25 3
Scale = 50

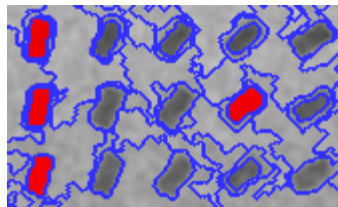


Abbildung 6.31

wie Abbildung 6.30, jedoch anderer
Ausschnitt

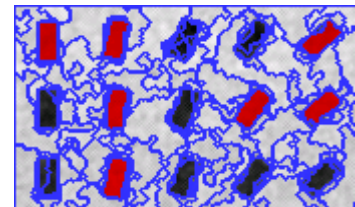


Abbildung 6.32

CurvatureAnisotropicDiffusion-
Filter 3 0,125 5
Scale = 20

Beide Filter unterscheiden sich von allen anderen untersuchten Filtern in der Weise, daß sie die durchschnittliche Standardabweichung im Bild fast verdreifachen. Ursache dafür ist eine starke Histo-

grammspreizung, die in Tabelle 6.4 und in den Histogrammen in Abbildung 6.33 a, b ausgewiesen wird. Dem menschlichen Betrachter erscheinen so gefilterte Bilder also kontrastreicher. Für die maschinelle Segmentierung ist dieser Vorzug jedoch ohne Bedeutung.

		unverrauscht	ungefiltert, $\sigma = 10$	Mittelwert Filter	GAD Filter	CAD Filter
	Standardabw.	15,7	18,6	15,2	45,5	47,7
Mittl. Grauwert oben	Vordergrund	95,0	93,7	93,7	31,0	21,0
	Hintergrund	159,0	159,0	158,2	228,0	219,4
Mittl. Grauwert unten	Vordergrund	123,0	124,4	122,4	139,7	167,5
	Hintergrund	131,0	130,9	129,9	141,5	157,6

Tabelle 6.4

Nach Filterung mit dem GAD- oder dem CAD-Filter steigt die Standardabweichung im Bild durch die starke Histogrammspreizung. Dadurch steigt der optisch wahrnehmbare Kontrast. Dies ist kein Vorteil für die maschinelle Segmentierung.

Zur Veranschaulichung werden die Histogramme des ungefilterten, mit $\sigma = 10$ verrauschten Bildes und des mit dem Gradient Anisotropic Diffusion Filter bearbeiteten Bildes in den folgenden Abbildungen gezeigt.

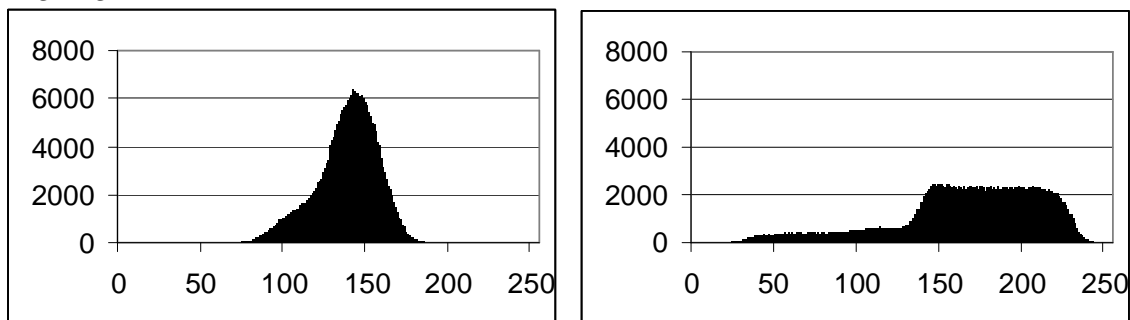


Abbildung 6.33 a, b

Die Histogramme des ungefilterten und des mit dem GAD-Filter bearbeiteten Bildes belegen die Histogrammspreizung, die den optisch wahrnehmbaren Kontrast deutlich verbessert.

6.3.4 Kantenerhaltende Filter

Sowohl der Graphennetzwerkfilter als auch der Edge Preserving Optimized Speckle Filter bemühen sich in besonderer Weise um Kantenerhalt, indem sie das Zentralpixel nur mit ähnlichen benachbarten Pixeln vergleichen und diesen angleichen.

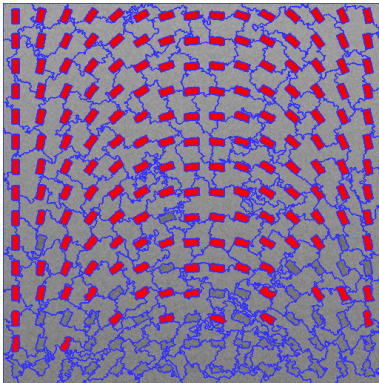


Abbildung 6.34

Der Edge Preserving Optimized Speckle Filter hat nur einen Parameter und erzeugt das beste Ergebnis.

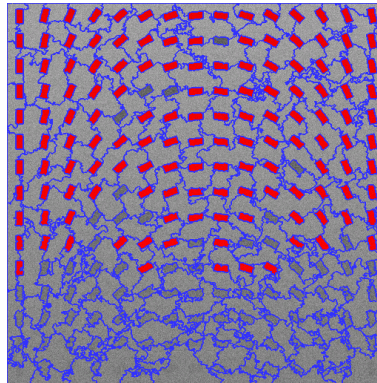


Abbildung 6.35

Zum Vergleich das ungefilterte Bild Scale = 20.

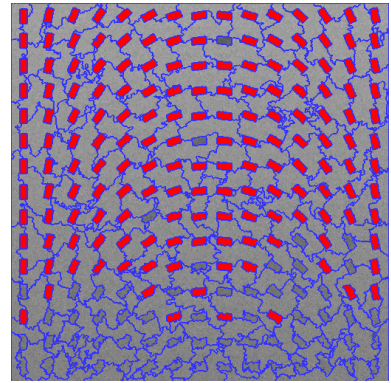


Abbildung 6.36

Der Graphennetzwerkfilter erhält die Kanten und kann ein um 10% verbessertes Klassifizierungsergebnis erzeugen.

Beide Filter lieferten die besten Ergebnisse in verschiedenen Versuchsreihen. Der Edge Preserving Optimized Speckle Filter hat darüber hinaus den Vorteil, nur einen steuernden Parameter zu verwenden. Mit diesem Parameter wird die maximal zulässige Varianz eingestellt, bis zu der das Zentralpixel (vergl. Abbildung 5.13 auf Seite 56) noch verändert wird. Ein höherer Wert führt also zu einer stärker glättenden Wirkung.

maximale Varianz	Standardabweichung	Erkannte Rechtecke
5	18,5	140
10	18,4	135
20	18,1	143
50	17,3	162
100	16,5	168
200	16,2	158

Tabelle 6.5

Die maximal zulässige Varianz als freier Parameter des Edge Preserving Optimized Speckle Filter beeinflusst das Segmentierungsergebnis.

Tabelle 6.5 belegt darüber hinaus, daß die Standardabweichung als Rauschmaß auch bei sonst unveränderten Parametern kein Qualitätsprädiktor für die nachfolgende Segmentierung ist. Bei hoch eingestellter Varianz und damit starker Glättungswirkung kann die Zahl der erkannten Rechtecke wieder sinken, aber auch bei niedrig eingestellter Varianz ist der Zusammenhang nicht gleichförmig.

6.3.5 Bilateraler Filter

Der Bilaterale Filter wurde entwickelt zur kanalübergreifenden Glättung mehrkanaliger Bilder. Für den Filter gibt es verschiedene Optionen zur Optimierung, insbesondere kann der benutzte Filterkern verändert werden. In dieser Untersuchung wurden jedoch nur die freien Parameter $\sigma_{(\text{domain})}$ und $\sigma_{(\text{range})}$ variiert, wobei die Änderung des $\sigma_{(\text{range})}$ im hier diskutierten einkanaligen Fall keine Auswirkung hat.

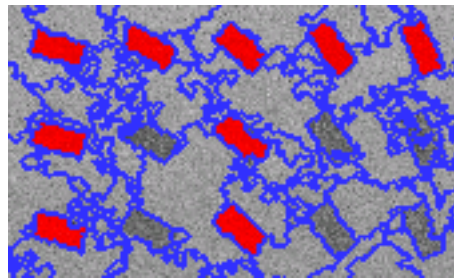


Abbildung 6.37

In Bereichen mit geringer Grauwertdifferenz führt die Filterung mit dem Bilateralen Filter zu sehr fraktalen Segmentgrenzen.

Die Filterung mit dem Bilateralen Filter läßt das Bild im visuellen Vergleich körniger aussehen. Die Standardabweichung über das gesamte Bild wird auch bei verschiedenen Parametereinstellungen kaum verändert (Tabelle 6.6). Das Segmentierungsergebnis entspricht den 140 erkannten Rechtecken, wenn nicht gefiltert wurde.

Parameter $\sigma_{(\text{domain})}$ $= \sigma_{(\text{range})}$	Standard- abweichung	Erkannte Recht- ecke
2	33,9	137
3	33,5	140
4	33,0	136
5	32,7	140
7	32,1	148
9	34,4	148

Tabelle 6.6

Beim Bilateralen Filter läßt sich kein Zusammenhang zwischen eingestellten Parametern, Standardabweichung im gesamten Bild und erkannten Objekten erkennen.

In Bereichen mit geringer Grauwertdifferenz führt die Filterung mit dem Bilateralen Filter zu sehr fraktalen Segmentgrenzen (Abbildung 6.37).

Auffällig ist im Histogramm, daß aufgrund von Rundungsartefakten ungefähr jeder zweite Grauwert nicht mehr besetzt ist.

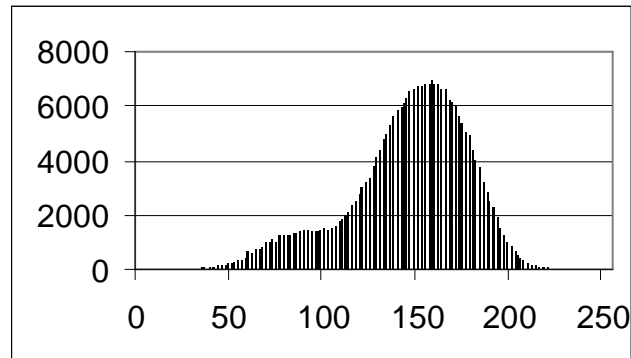


Abbildung 6.38

Histogramm nach bilateraler Filterung.

Im Vergleich zu den in Abbildung 6.33 a, b gezeigten Histogrammen ist nur noch ungefähr jeder zweite Grauwert besetzt. Auch hier verbessert die Grauwertspreizung den visuellen Kontrast, das verbessert jedoch nicht die Segmentierungsqualität.

In diesem Kapitel wurde die Wirkung der verschiedenen Filter bei unterschiedlichen Parametereinstellungen vorgestellt. Als Testobjekt dienten synthetische Bilder mit bekannter Ground Truth. Die Darstellung der Filterwirkung ist nicht trivial, weil nicht nur die Parameter der Filter variiert werden können, sondern für ein optimales Ergebnis anschließend auch die Parametrierung des verwendeten Segmentierungsprogramms optimiert werden muß.

Das folgende Kapitel verwendet die erhaltenen Erkenntnisse, um die Filterwirkungen an einer realen Szene zu überprüfen.

7 Anwendung auf eine reale Szene

Die gefundenen Erkenntnisse sollen nun an einer realen Szene überprüft werden.



Abbildung 7.1

Ausschnitt aus IKONOS-Szene Bremen vom 18.8.2001, 1 m panchromatisch, geocodiert, 16 bit. Der oben rot eingezeichnete und im folgenden betrachtete Ausschnitt hat die Koordinaten (3485363|5884726) bis (3485486|5884801) und liegt am Erlingplatz in Bremen-Walle.

Zur Verfügung stand eine 1024 x 1024 Pixel IKONOS-Szene von Bremen vom 18.8.2001. Das Bild ist bereits geocodiert und auf eine Pixelkantenlänge von 1 m transformiert. Die Daten liegen als TIFF-Datei mit elf bit spektraler Auflösung im 16-bit Format vor. Alle Untersuchungen wurden mit den Originaldaten durchgeführt. Die Filterprogramme waren für acht-bit-Daten eingerichtet und wurden für die Untersuchung für 16-bit-Daten umgeschrieben. So war keine Datenreduktion in der Vorverarbeitung erforderlich.

Manche Bildbearbeitungsprogramme haben mit solchen Daten Darstellungsprobleme, die Bilder sehen zu dunkel aus, weil die möglichen 16 bit nicht gleichmäßig besetzt sind. Die visuell beste Darstellung in eCognition erzielt man mit einem Histogramm-Ausgleich⁴². Für die Darstellung der obenstehenden Abbildung 7.1 waren Anpassungen in der Gradationskurve erforderlich.

7.1 Bildanalyse

Vor der Analyse der Wirkungen verschiedener Filter werden die allgemeinen Bilddaten beschrieben und zusammengefaßt.

7.1.1 Topographie

Anhand des in der Szene abgebildeten Fernmeldeturms Bremen-Walle wurden Aufnahmeuhrzeit und Blickwinkel ermittelt:



Abbildung 7.2

Anhand des Fernmeldeturms Bremen-Walle können Blickwinkel, Aufnahmeuhrzeit und Sonnenstand ermittelt werden.

Der Turm hat eine Höhe von 235,70 m, seine Länge im Bild beträgt 88,4 m. Daraus ergibt sich ein Blickwinkel von $\arctan(88,4/235,7) = 20,6^\circ$. Da das Bild geocodiert und in Nord-Süd-Richtung ausgerichtet ist, kann mit dem Winkel des Schattenwurfs von 21° zur Nordrichtung die Aufnahmezeit von 10:36 Uhr bezogen auf die Ortszeit bestimmt werden. Die Länge des Schattens beträgt 204,4 m und weist darauf hin, daß Schatten erhebliche Bildanteile beeinflussen werden.

⁴² Menu View → Edit Layer Mixing → Equalizing → Histogram

7.1.2 Histogramm

Die Szene hat Grauwerte im Bereich von 0 bis 1083. Die Grauwerte sind sehr ungleichmäßig verteilt.

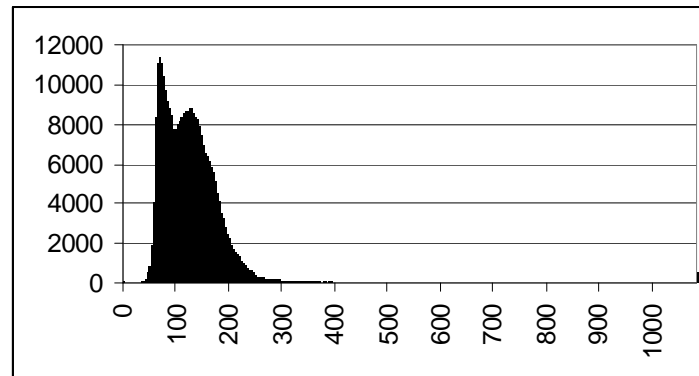


Abbildung 7.3

Histogramm der Szene aus Abbildung 7.1

Im 256 Grauwerte umfassenden Bereich von 41 bis 296 liegen 98,5% der Grauwerte, Erweitert man den Bereich auf eine neun-bit-Datendarstellung, dann liegen nur noch 3% der Grauwerte außerhalb des Bereichs von 23 bis 534.

7.1.3 Rauschen

Mit dem in Abschnitt 3.3.2 (Seiten 25ff.) beschriebenen modifizierten Verfahren nach Baltsavias wurde das Bildrauschen berechnet. Die folgende Abbildung 7.4 zeigt blau eingetragen mit der Häufung im linken Bereich die Anzahl der ausgewerteten 8-Nachbarschaften je zentralem Grauwert. Rot eingetragen ist die mittlere Standardabweichung der 5% kleinsten Standardabweichungen je Grauwert.

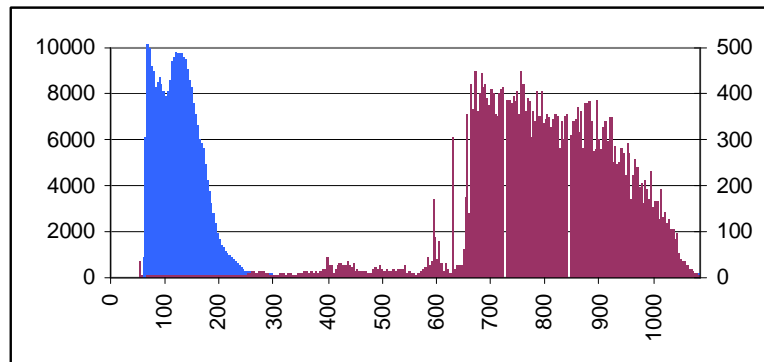


Abbildung 7.4

Rauschmaß in einem 3×3 – Fenster.

Abszisse: mittlerer Grauwert der 16-bit – Daten,

linke Ordinatenachse mit blau gezeichneten Daten:

Anzahl der ausgewerteten Fenster

rechte Ordinatenachse mit rot gezeichneten Daten:

mittlere Standardabweichung σ der 5% kleinsten Standardabweichungen je Grauwert.

Die Verteilung der ausgewerteten quadratischen Fenster mit jeweils 3 Pixeln Seitenlänge entspricht der Histogrammverteilung aus Abbildung 7.3. Die Abszisse ist mit den zentralen Grauwerten der Fenster beschriftet, die Häufigkeiten sind auf der linken Ordinatenachse aufgetragen. Im Bereich zwischen 61 und 193 konnten je Grauwert mehr als 2000 Quadrate ausgewertet werden. Aus den je Fenster berechneten Standardabweichungen wurde für jeden Grauwert der Mittelwert aus den 5% kleinsten gebildet. Diese gemittelten Standardabweichungen sind in roter Farbe dargestellt, die zugehörige Ordinatenachse rechts. Die hohen Standardabweichungen für größere Grauwerte ergeben sich, weil dort nur sehr wenige Fenster ausgewertet werden konnten. Wenn mehr als 30 Fenster ausgewertet werden konnten, lag die berechnete mittlere Standardabweichung stets unter 20.

Der Bereich, in dem mindestens 100 Fenster ausgewertet werden konnten, ist in Abbildung 7.5 noch einmal vergrößert dargestellt. Daraus ergibt sich ein mittleres Rauschmaß von 5,09.

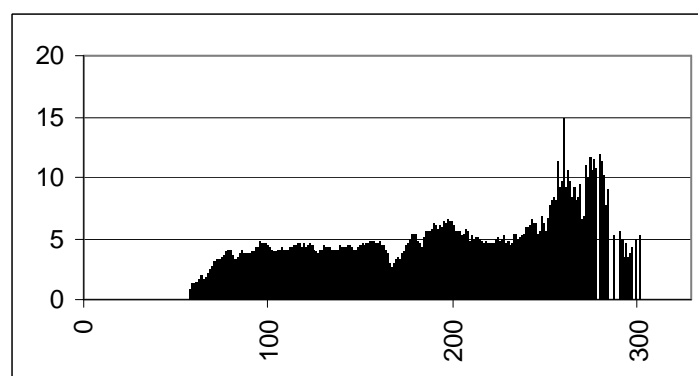


Abbildung 7.5

Mittelwert der Standardabweichung σ , wenn mindestens hundert 3×3 -Fenster ausgewertet werden konnten.

Diese Bestimmung des Rauschens kann vollautomatisch erfolgen. Allerdings wird das tatsächliche Rauschen systematisch unterschätzt. In der Original-Szene liegt am oberen Bildrand ein Sportplatz, auf dem ein hellerer Bereich und zwei Fußballtore zu erkennen sind. Ohne diese Bereiche kann ein homogen erscheinendes Gebiet von 2050 m² begrenzt werden. Dieses Gebiet hat einen mittleren Grauwert von 215 mit einer Standardabweichung von 7,54. Die nach dem beschriebenen Verfahren ermittelte Standardabweichung zum Grauwert von 215 beträgt jedoch 4,73. Offenbar gibt es viele „Ausreißer“ in den Daten, die durch das Verfahren ausgeschlossen werden, weil genügend „bessere“ Nachbarschaften zur Berechnung zur Verfügung stehen.



Abbildung 7.6

Zur vergleichenden Berechnung des Rauschens verwendeter Ausschnitt aus einem Sportplatz. 2050 m², mittlerer Grauwert 215 (16-bit-Daten), $\sigma = 7,54$.

In Abbildung 7.7a ist ein Teil des in Abbildung 7.6 rot markierten Ausschnitts vergrößert wiedergegeben. Zum Vergleich ist in Abbildung 7.7b ein mit einer Gaußverteilung synthetisch verrauschtes Bild dargestellt. Beide Bilder haben den gleichen mittleren Grauwert (215) und die gleiche Standardabweichung von 7,54. Dennoch sind die Bilder auffallend unterschiedlich. Bei den IKONOS-Aufnahmen fallen wie auch in Abbildung 3.6 (Seite 26) homogene Flecken auf, andererseits wirken die vom Zentralwert abweichenden Pixel wie durch ein Kantenschärfungsverfahren verstärkt. Offenbar entspricht das Modell einer Gaußverteilung des Rauschens nicht der realen Situation bei IKONOS-Aufnahmen. Um ein besseres Modell zu entwickeln, erscheinen jedoch Insiderkenntnisse der Bilderzeugungskette von den Sensoren über die Präprozessierung bis zur Art der Georeferenzierung erforderlich.

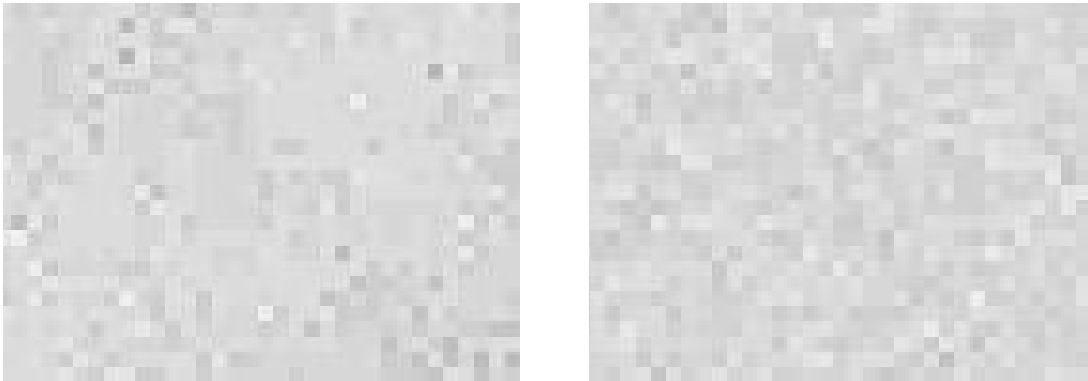


Abbildung 7.7 a, b

Links Detail aus dem Sportplatz Abbildung 7.6, rechts synthetisch verrauschtes Bild mit gleichem mittleren Grauwert (215) und gleichem Rauschmaß ($\sigma = 7,54$). Bei IKONOS-Aufnahmen fallen die Flecken mit homogener Färbung auf, siehe auch Abbildung 3.6 (Seite 26).

Betrachtet man reale Szenen aus IKONOS-Bildern wie in Abbildung 3.6 (Seite 26) oder Abbildung 7.7 a (einem Ausschnitt aus der in Abbildung 7.6 (Seite 94) gezeigten Szene), dann sieht man, daß das Rauschen in IKONOS-Bildern anscheinend nicht gleichmäßig verteilt ist. Im Unterschied zu dem in Abbildung 7.7 b gezeigten synthetisch verrauschten Ausschnitt aus der Ground Truth wechseln sich im kleinräumigen Maßstab mit sechs bis zehn Pixeln relativ homogene Bereiche mit stärker strukturierten ab.

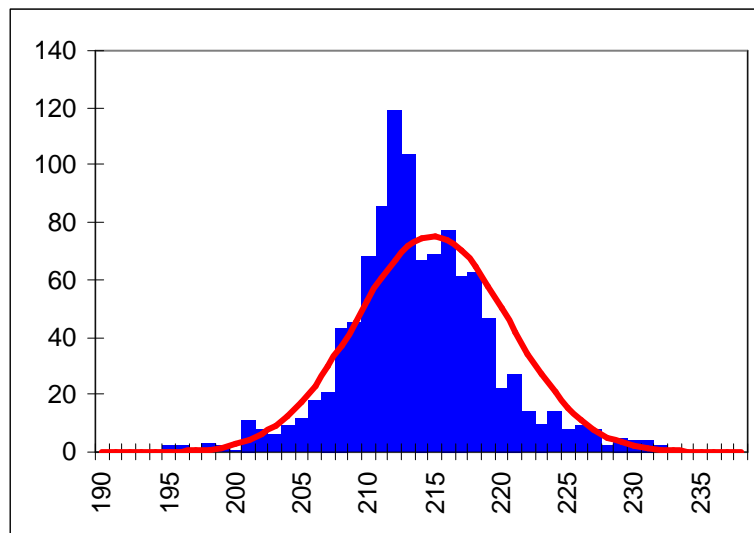


Abbildung 7.8

Histogramm zu Abbildung 7.7 a (blau) mit eingezeichneter Gaußverteilung (rot)

Andererseits zeigt das in Abbildung 7.8 dargestellte Histogramm mit überlagerter Gaußverteilung, daß eine solche anscheinend doch vorliegt, wenn man größere Bereiche betrachtet wie den Ausschnitt in Abbildung 7.7 a mit 30 x 36 Pixeln.

7.2 Segmentierung

7.2.1 Kriterien für die Filterbewertung

Bevor man die Qualität von Filtern beurteilen kann, muß man sich über die Kriterien zu ihrer Bewertung verständigt haben. Eine Methodik zur Bewertung von Qualitätskriterien haben Udupa et al. [2002] entwickelt. So reicht es offensichtlich nicht aus, Pixelsummen zu zählen, wenn etwa wichtige Objekte im Ergebnis unterschlagen werden, weil sie klein an Fläche sind. Außerdem erscheint es wünschenswert, zunächst die Segmentierungsparameter am ungefilterten Bild zu optimieren. Nach Filterung kann es jedoch erforderlich werden, die Segmentierungsparameter erneut anzupassen, um wieder zu optimalen Ergebnissen zu gelangen. Im Ergebnis ist es sicherlich für die meisten Verwendungszwecke sinnvoll, wenn der in Abbildung 7.6 zur Berechnung des Rauschens verwendete Sportplatz als ein Segment beschrieben wird. Andererseits enthält er ein größeres Stück mit verändertem, heller erscheinendem Bewuchs, das unter anderer Fragestellung getrennt segmentiert werden sollte. Schließlich handelt es sich um eine im Vergleich zu Siedlungshäusern und Gärten große Fläche, die im Sinne des Multiresolution-Ansatzes von eCognition erst auf einer höheren Ebene zusammengefaßt werden kann.

Erwünscht sind Filter, die Kanten auch bei möglichst groß gewähltem Scale-Parameter erhalten. Schließlich sollen bei der Segmentierung so wenige Segmente wie möglich, aber so viele wie nötig entstehen.

Nicht erwünscht sind Filter, die Kanten so verändern, daß an Stelle der vorher klar erkennbaren Kanten nach Filterung bei der Segmentierung neue Segmente entstehen, die den Kantenbereich umgrenzen. Solche Segmente lassen sich nämlich auch bei einem mehrstufigen Segmentierungsprozeß schlecht wieder trennen.

In diesem Sinne wurde besonderes Augenmerk gelegt auf Schattengrenzen, Dachflächen, Baumkronengrenzen auf der Sonnen- und Schattenseite, aber auch auf das Segmentierungsverhalten flächenhafter Objekte wie Straßenbelag und Grünflächen.

Wenn man den Shape-Parameter auf Null setzt, die Segmente also ausschließlich nach Grauwert ohne Berücksichtigung der Form gebildet werden, dann können sehr große, weit verzweigte Segmente entstehen. Würde es sich nur um Straßen handeln, wäre das durchaus angemessen. Im urbanen Raum trifft man jedoch auf viele, eher kompakte Flächen wie Dachflächen, die sich im ungünstigen Fall in der panchromatischen Aufnahme von der umgebenden Hoffläche kaum oder gar nicht unterscheiden lassen. Deshalb wurde bei den nachfolgenden Beispielen der Wert des Shape-Parameters auf 0,5 gesetzt.

Der Scale-Parameter wurde für jedes Filterergebnis variiert, um eine bessere Vergleichbarkeit zu gewährleisten. Weil die Filter die mittlere Varianz im Bild unterschiedlich stark beeinflussen (der Gradient Anisotropic Diffusion Image Filter senkt mit den hier gewählten Parametern die Varianz am stärksten, der bilaterale Filter am wenigsten) führt eine gleiche Einstellung des Scale-Parameters zu sehr unterschiedlichen Anzahlen von gebildeten Segmenten. Wenn Segmente zusammengefaßt werden, fallen Kanten weg. Die erwünschte Anzahl von Segmenten hängt jedoch nicht nur vom Bildinhalt ab, sondern hängt auch mit den Absichten des Auswerters zusammen. Sollen einzeln stehende Sträucher oder Pflützen auf Flachdächern getrennt segmentiert werden oder mit ihrer Umgebung zusammengefaßt werden?

Der Scale-Parameter ist eine Fließkommazahl und wurde deshalb für jedes Filterergebnis iterativ mit einer Genauigkeit von 1/100 so eingestellt, daß über die gesamte Szene 6500 Segmente gebildet wurden. In Tabelle 7.1 sind die Anzahlen der gebildeten Segmente bei einer Einstellung des Scale-Parameters von 20 und 25 sowie die für 6500 zu bildende Segmente erforderliche Einstellung zusammengefaßt.

Filter	Standard- abweichung	Segmente bei Scale = 20	Scale für 6500 Segmente	Segmente bei Scale = 25
ungefiltert	63,8	8641	23,85	6057
Mittelwert Filter	55,5	7398	21,50	5020
Median Filter	57,3	7448	21,72	5174
Gradient Anisotropic Diffusion	52,7	6744	20,46	4820
Graphennetzwerk Filter	61,1	9070	24,45	6252
Edge Preserving Optimized Speckle	63,2	8899	24,25	6219
Bilateraler Filter	63,4	8680	24,00	6104

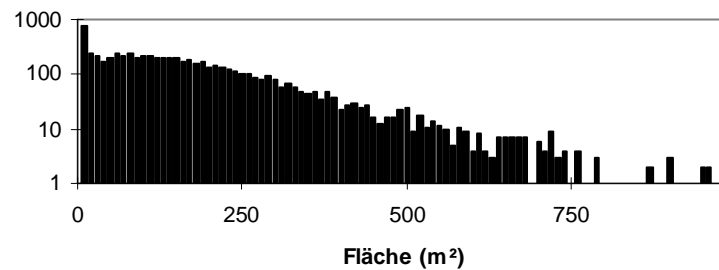
Tabelle 7.1

Der Scale-Parameter wurde mit einer Genauigkeit von 1/100 so eingestellt, daß möglichst 6500 Segmente gebildet wurden. Tatsächlich wurden 6498 bis 6505 Segmente gebildet.

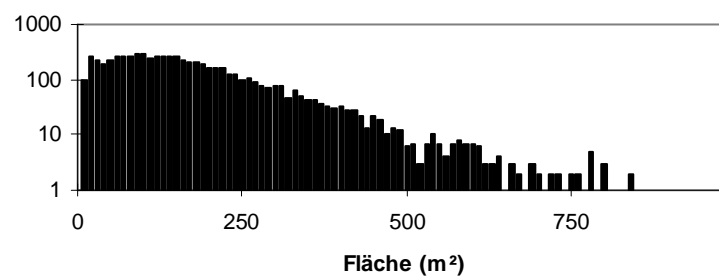
7.2.2 Granularität

Wie verändert ein Filter die Größenverteilung der gebildeten Segmente? Auch wenn durch Variation des Scale-Parameters eine gleiche Anzahl von Segmenten erzeugt wird, kann die Größenverteilung ähnlich wie auf Seite 38 beschrieben variieren.

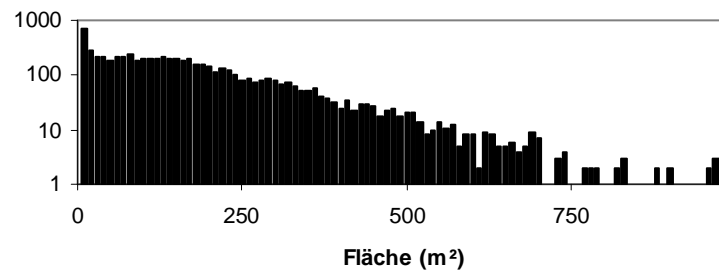
Verteilung der Segmentgrößen



ungefiltert, 6498 Segmente. Bis zu 10 Pixel (=m²) haben 752 Segmente, das größte ist 2082 m² groß.



mittelwertgefiltert, 6494 Segmente. Bis zu 10 Pixel haben 100 Segmente, das größte ist 1971 m² groß.



EPOS-gefiltert, 6499 Segmente. Bis zu 10 Pixel haben 733 Segmente, das größte ist 2906 m² groß

Tabelle 7.2

Die Ordinate gibt die Zahl der gebildeten Segmente je auf der Abszisse abgetragener Fläche an. Sie wurde der besseren Übersicht wegen logarithmisch geteilt.

Der EPOS-Filter ändert die Verteilung der Segmentgrößen nur unwesentlich, wenn der Scale-Parameter so eingestellt wird, daß gleich viele Segmente gebildet werden. Der Mittelwertfilter vermindert die Zahl sehr kleiner Segmente deutlich.

Tabelle 7.2 zeigt beispielhaft die Änderungen in der Verteilung der Segmentgrößen. Der Mittelwertfilter glättet stark über kleine Distanzen. Dadurch werden kleine Flächen mit zuvor hohem Kontrast mit der Umgebung verschmolzen, die anderenfalls das Homogenitätskriterium von eCogniton verletzen und deshalb sowohl im ungefilterten Original wie auch im EPOS-gefilterten Bild als separate Segmente erscheinen.

7.2.3 Segmentkanten

In diesem Abschnitt wird die Auswirkung der verschiedenen Filter auf Segmentkanten in einer realen Szene näher beschrieben. Als Demonstrationsobjekt dient der in Abbildung 7.1 (Seite 90) markierte Bildausschnitt (Abbildung 7.9).

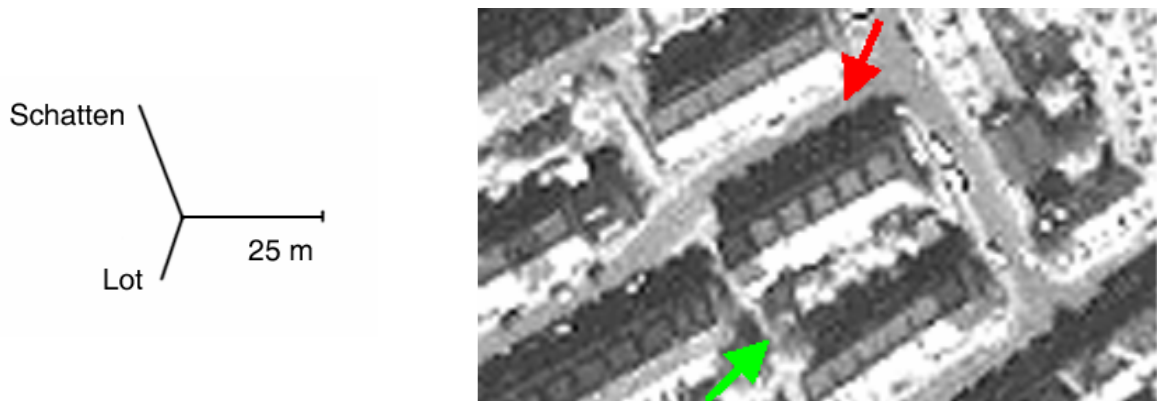


Abbildung 7.9

Der in Abbildung 7.1 markierte Bildausschnitt. Das Gebäude sind sechs für Bremen typische Reihenhäuser. Der Maßstab links mit Vertikaler und Schattenwurf für eine Höhe von 25 m wurde nach dem Funkturm Bremen-Walle ermittelt.

Der rote Pfeil markiert die Aufnahmeperspektive für das Photo in Abbildung 7.10, der grüne Pfeil zeigt die Perspektive für das in Abbildung 7.11 abgebildete Photo mit dem Blick in die Gärten.



Abbildung 7.10

Die in Abbildung 7.9 mit dem roten Pfeil markierte Perspektive. Die Aufnahmezeiten beider Aufnahmen stimmen überein, das Photo wurde allerdings vier Jahre und drei Monate später aufgenommen (im November 2005), deshalb sind die Schatten länger.







Abbildung 7.11

Der Blick in die Gärten, aus der in Abbildung 7.9 mit einem grünen Pfeil markierten Perspektive.

Zunächst ist festzuhalten, daß die Farbe der Dachflächen in der panchromatischen Satellitenaufnahme in erster Linie von ihrem Winkel zur Sonne abhängt. Die Abweichung der Aufnahmeperspektive vom Nadir von $20,6^\circ$ bewirkt, daß je Meter Höhendifferenz ein Versatz von 0,38 Metern in der Horizontalen beobachtet werden kann. Die Länge des Schattens beträgt 87 % der Gegenstandshöhe.

Die folgenden Abbildungen dokumentiert die Wirkungen der verschiedenen Filter an dem in Abbildung 7.1 (Seite 90) markierten Bildausschnitt. In diesen Ausschnitten wurde als Hintergrund jeweils das ungefilterte Bild verwendet. Darauf wurden in rot die ausschließlich im gefilterten Bild erzeugten Segmentgrenzen eingetragen, in grüner Farbe die sowohl im ungefilterten wie im gefilterten Bild gleich verlaufenden Grenzen und in gelb zur Kontrolle die nur im ungefilterten Bild vorhandenen Segmentgrenzen.

Im Ergebnis besagen also

-  grüne Kanten, das auf eine Filterung verzichtet werden kann. Die
-  roten Linien müssen bei einem nützlichen Filter die Segmentgrenzen im Bild
-  besser darstellen als die gelben. Sinnvolle
-  gelbe Grenzen sollten nicht verloren gehen.

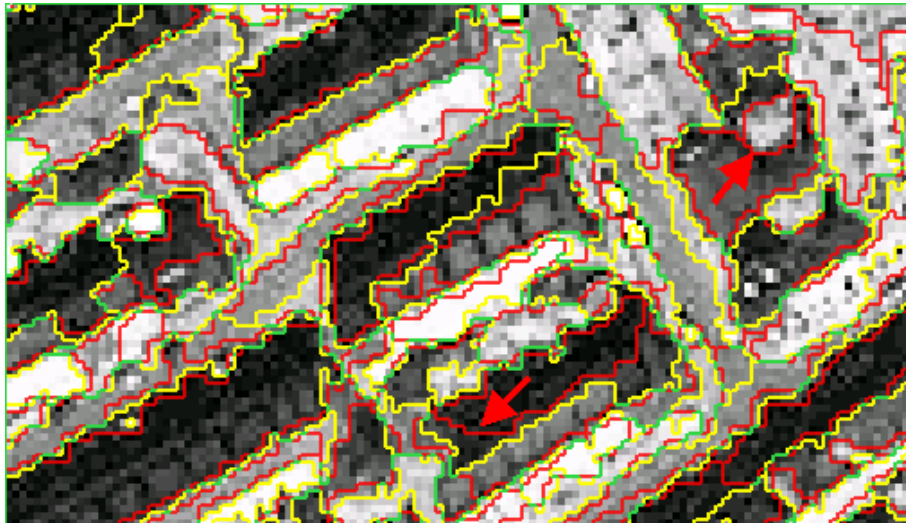


Abbildung 7.12

Filterung mit dem 3 x 3 Mittelwertfilter. Nur wenige Segmentgrenzen bleiben erhalten (grün gezeichnet). Die gelb eingetragene Segmentgrenzen sind nur im ungefilterten Bild, die rot eingetragene Grenzen nur im gefilterten Bild. Typisch für den Mittelwertfilter sind verschobene Segmentgrenzen, die nicht mehr dem ursprünglichen Kantenverlauf folgen (rote Pfeile). Kleine Segmente werden unterdrückt.

Im Unterschied zum ungefilterten Bild bewirken alle Filter, daß der Innenhof rechts oben im Bild segmentiert wird, wobei nach Mittelwertfilterung Teile des Wandschattens mit eingeschlossen werden (roter Pfeil rechts oben in Abbildung 7.12). Der rote Pfeil unten im Bild verweist auf eine neu gebildete Segmentgrenze, die sich dadurch ergibt, daß sich die ursprünglich recht gut dargestellte segmentierte Dachfläche (gelbe und grüne umfassende Kante) nach Filterung nur noch verschwommen darstellte. Der Mittelwertfilter faßt einerseits viele kleine Segmente zusammen, andererseits entstehen gerade an Kanten mit hohem Kontrast neue Segmente, die die ursprünglich scharf gezeichnete Kante zu einem mehrere Pixel breiten Saum geraten lassen. Dieser Effekt ist gut erkennbar an dem Häuserblock links unten im Bild, dort wird die gelbe Kante durch zwei rote Kanten gesäumt. Dies entspricht den Befunden in den synthetischen Bildern (vgl. Abbildung 6.22, Seite 84).

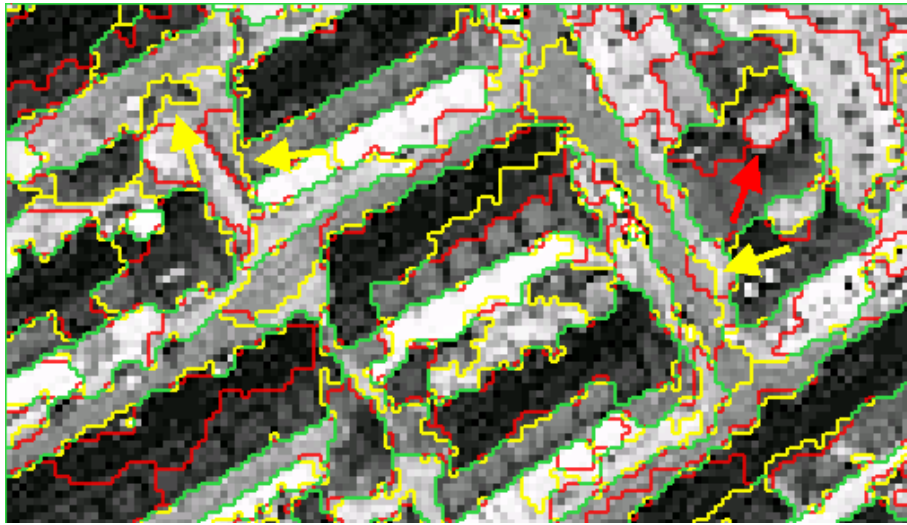


Abbildung 7.13

Filterung mit dem 3 x 3 Medianfilter. Der Kantenerhalt ist besser als nach Mittelwertfilterung. An manchen Kanten werden sinnvolle Grenzen gezogen (roter Pfeil), dafür verschwinden andere (gelbe Pfeile).

Sowohl der Median- wie auch in besonderem Maße der Mittelwert-Filter verändern die ursprünglichen Segmentgrenzen sehr stark. In den Abbildungen sind nur wenige Kanten unverändert geblieben (grün eingefärbt). Viele der ursprünglich vorhandenen Kanten (gelb eingefärbt) sind nach Filterung nicht mehr vorhanden und die hinzugekommenen (rot eingefärbten) Kanten entsprechen nur in wenigen Fällen den gesuchten und gewünschten Objektgrenzen. Übereinstimmend mit den im vorigen Kapitel gemachten Feststellungen verschlechtern beide Filter das Segmentierungsergebnis.

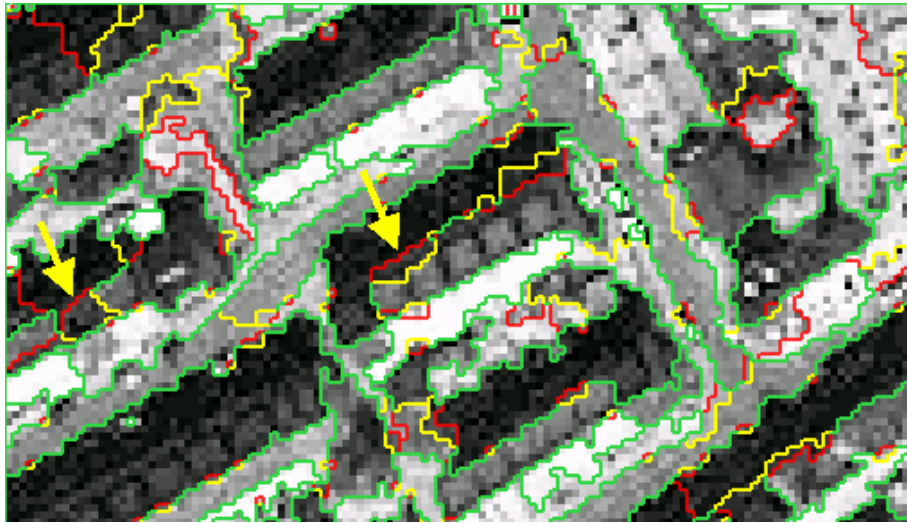


Abbildung 7.14

Filterung mit dem Bilateralen Filter ($\sigma_{\text{domain}}=5$, $\sigma_{\text{range}}=6$).

Der Bilaterale Filter trennt als einziger die wenig Kontrast gebenden Trauf- und Firstkanten (gelbe Pfeile). Neu erzeugt wurden wie in den vorhergehenden Abbildungen die roten Segmentgrenzen. Die gelben Segmentgrenzen stammen aus dem ungefilterten Bild und sind im gefilterten Bild nicht mehr vorhanden.

Der Bilaterale Filter trennt am besten die wenig Kontrast gebenden Trauf- und Firstkanten (Gelbe Pfeile in Abbildung 7.14). Die meisten im ungefilterten Bild bereits vorhandenen Segmentgrenzen bleiben erhalten (grüne Linien). An wenigen Stellen werden Segmentgrenzen verändert und die veränderten Segmentgrenzen erscheinen meist sinnvoller und sachlich gerechtfertigt. Der Filter verbessert das Segmentierungsergebnis.

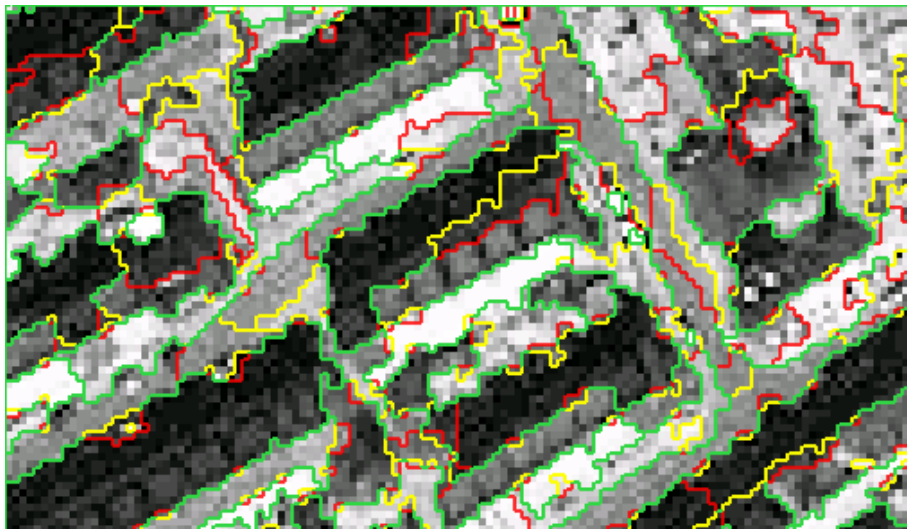


Abbildung 7.15

Filterung mit dem Graphennetzwerk-Filter ($\mu=0,12, 60$ Iterationen)

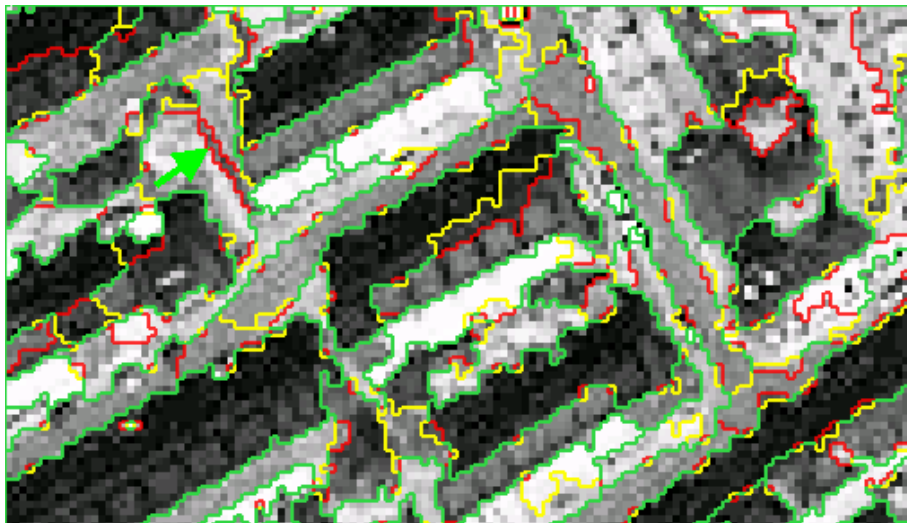


Abbildung 7.16

Filterung mit dem EPOS-Filter (Varianz=100). Der Eposfilter bewirkt als einziger, daß der Heckenschatten (grüner Pfeil rechts oben) segmentiert wird.

Die Ergebnisse nach Filterung mit dem Graphennetzwerkfilter und nach EPOS-Filterung gleichen sich sehr. Der EPOS-Filter bewirkt als einziger eine gute Darstellung der Schattenkante einer Hecke (grüner Pfeil in Abbildung 7.16). Die Firstkanten erzeugen im Gegensatz zum Bilateralen Filter keine Segmentgrenzen. Einige Kanten werden pixelbreit verschoben (vorwiegend rechts unten im Bild), wobei visuell kein Qualitätsunterschied in der Grenzziehung erkennbar ist.

Insgesamt fällt auf, daß nach EPOS-Filterung wie auch nach Filterung mit dem Bilateralen Filter oder dem Graphennetzwerk-Filter ein großer Teil der Segmentgrenzen identisch ist mit dem ungefilterten Original (grüne Grenzen). Dagegen gibt es kaum Übereinstimmungen mit den Segmentgrenzen nach Mittelwert-Filterung und wenige Übereinstimmungen nach Median-Filterung.

Der Graphennetzwerk-Filter hat mehrere Parameter, die in ihren Auswirkungen nicht alle einfach zu verstehen sind. Einige Parameter wurden empirisch festgelegt und im Algorithmus fixiert. Das Verfahren hat also zahlreiche Freiheitsgrade, die schwierig zu optimieren sind.

Ähnliches gilt für den Bilateralen Filter. Der Bilaterale Filter ist jedoch Gegenstand aktueller Forschung, so wurde insbesondere sein schlechtes Laufzeitverhalten (er benötigte die 130fache Rechenzeit der beiden anderen Algorithmen) erheblich verbessert. Der Bilaterale Filter ermöglicht die Verwendung unterschiedlicher Faltungskerne und den Einbezug mehrerer Ebenen bei Mehrkanalaufnahmen. Diesbezügliche Optimierungen wurden jedoch nicht untersucht.

8 Zusammenfassung, Wertung und Ausblick

Wir haben gezeigt,

- daß ein gewisses Maß an Rauschen vorteilhaft ist, insbesondere zur Behandlung von Rundungsfehlern,
- daß das Messen von Rauschen keine triviale Aufgabe ist,
- daß die Art des Rauschens mit dem einfachen Modell einer Gauß-Verteilung nicht hinreichend beschrieben ist,
- daß die bisher verwendeten Algorithmen zur Rauschverminderung, wie sie für die Bildverbesserung zur nachfolgenden visuellen Interpretation angemessen waren und sinnvoll genutzt wurden, für die Ergebnisqualität des untersuchten maschinellen Segmentierungsverfahren nachteilig sind,
- daß es Algorithmen gibt, die die Ergebnisqualität des untersuchten maschinellen Segmentierungsverfahren verbessern können,
- daß diese Algorithmen jedoch durch einen oder mehrere Parameter gesteuert werden, die die Ergebnisqualität entscheidend beeinflussen.

Rundungsfehler in Algorithmen führen häufig zu systematischen Fehlern (vgl. die Kammartefakte in dem in Abbildung 4.22 dargestellten Histogramm). Auch die in Abbildung 4.11 gezeigten Abweichungen von den zu erwartenden Rechteckformen im Segmentierungsergebnis der weißen oder grauen Fläche lassen auf Rundungsartefakte schließen. Ein geringes Maß an zugefügtem Rauschen führt dazu, daß diese systematischen Fehler durch stochastische Fehler ersetzt werden, die besser verteilt sind.

Bei jeder Bildauswertung ist das Wissen um die im Bild möglichen Artefakte von größter Bedeutung. Weil bei der Auswertung von Fernerkundungsaufnahmen im Grenzbereich der Auflösung der verwendeten Sensorsysteme gearbeitet wird, kann die Auswertbarkeit und die Qualität der erzielten Ergebnisse in besonders hohem Maße vom vorhandenen Rauschen abhängen. Um Rauschen vermindern zu können, wünscht man sich ein einfaches Verfahren, um das Rauschen messen zu können. Standardmäßig wird die Varianz über das gesamte Bild als Rauschmaß angeboten. Die Varianz über das gesamte Bild hängt jedoch im Wesentlichen von den Bildinhalten und den Kontrasteinstellungen ab und ist deshalb zur Bestimmung des Rauschens ungeeignet. Außerdem kann die Farbabweichung eines Pixels in einer ansonsten homogenen Fläche zwar ihre Ursache im Rauschen haben. Sie kann jedoch auch in einer Inhomogenität in der Fläche begründet sein, wie zum Beispiel einem Ölfleck auf dem Asphalt oder einer lokalen Wellenformation auf einem See. Ohne bekannte Ground Truth läßt sich das nicht verifizieren.

Als Rauschmaß hat sich ein ursprünglich von Baltsavias et al. [2001] vorgeschlagenes Verfahren bewährt. Das Verfahren wurde hier für die vollautomatische Nutzung modifiziert. Es mißt je Grauwert nur die Standardabweichungen der gleichförmigsten Gebiete berücksichtigt werden. In Vergleich mit synthetisch verrauschten Bildern zeigt sich jedoch, daß auf diese Weise das Rauschen im Bild systematisch um fast den Faktor zwei unterschätzt wird. Das Verfahren ist verwendbar, wenn es weniger auf das absolute Maß des Rauschens ankommt, sondern eher um den direkten Vergleich eines unbearbeiteten mit einem bearbeiteten Bild.

Der visuelle Vergleich einer homogenen Fläche in einer IKONOS-Aufnahme zeigt im Verhältnis zu einer artifiziell verrauschten Fläche trotz identischer makroskopischer Parameter ein anderes Erscheinungsbild. Die IKONOS-Aufnahme zeigt größere homogene Abschnitte und eher einzelne, vom Zentralwert stärker abweichende Pixel (Abbildung 7.7 a, b), obwohl schon über ein Gebiet von wenigen hundert Pixeln die Verteilungen übereinstimmen. Wahrscheinlich wird dieses Verhalten nicht durch die Sensoren, sondern durch die Art der Auslesung oder durch die Schärfungsverfahren des

Betreibers Space Imaging verursacht. Hier wäre ein besseres Modell als die Gaußverteilung zur Darstellung des Rauschens wünschenswert.

Zur Bildvorverarbeitung gehören für IKONOS- und andere Aufnahmen ähnlicher Qualität eine nicht näher beschriebene atmosphärische Korrektur sowie die Orthorektifizierung. Eine weitere Vorverarbeitung zur Rauschminderung und Kontrastverstärkung wird vor der Segmentierung mit eCognition offenbar nur noch selten angewandt. Es gibt jedoch Ausnahmen. So verwendete Hájek [2005] noch Medianfilter der Breiten 3 und 5. Fraser, Baltsavias und Gruen [2002] verwendeten einen nicht näher definierten anisotropen Gauß-Filter sowie einen Wallisfilter in der Vorverarbeitungskette. Der hier vorgestellte Gradient-Anisotropic-Diffusion-Filter bewirkt eine dem Gaußfilter entsprechende Tiefpaßfilterung in homogenen verrauschten Gebieten, spart jedoch die Kanten aus.

Der allgemein am häufigsten zur Verminderung des Rauschens genutzte Filter, der Mittelwertfilter, ist gleichzeitig auch der am wenigsten geeignete Filter für die Zwecke der maschinellen Segmentierung. Der Mittelwertfilter bewirkt eine starke Abnahme der Varianz im Bild. Betrachtet man die Helligkeitsinformation als Höheninformation, so würde man die Varianz als Rauigkeit der Oberfläche übersetzen können. Der Mittelwertfilter ebnet diese Werte ein, im extremen Fall bleibt eine plane Fläche zurück. Diese gibt jedoch keine Informationen mehr her. Insbesondere verschwimmen Kanteninformationen und je nach Parametereinstellungen des Segmentierungsprogramms werden um vorher klar abgrenzbare Objekte zusätzliche ringförmige, den Kantenbereich darstellende Segmente gebildet, oder Objekte mit schwachem Kontrast zum Hintergrund werden vom Segmentierer mit dem Hintergrund verschmolzen. Auch vorangehende oder nachfolgende kantenschärfende Maßnahmen können diesen Effekt nicht ausgleichen.

Ähnlich nachteilig wirken sich die meisten anderen untersuchten Filter aus. Die Funktion des Medianfilters ist inhärent im Segmentierungsalgorithmus integriert. In älteren Segmentierungsalgorithmen wie dem ISODATA-Algorithmus verbesserte der Medianfilter das Segmentierungsergebnis, weil er den Salt-and-Pepper-Effekt reduzierte. Im untersuchten Algorithmus von eCognition wird *cum grano salis* die gewünschte Segmentgröße vorgegeben, und einzelne Ausreißer, die der Medianfilter besser in Segmente integrieren würde, werden nicht separiert, sondern führen durch ihre Vergrößerung der Varianz im Segment nur zu einer lokal kleineren Segmentgröße.

Die untersuchten rekursiven Filter, die mit anisotroper Diffusion arbeiten, verbessern die visuelle Interpretierbarkeit, indem sie Kanten schärfen und eine kontrastverstärkende Histogrammspreizung applizieren. Dies kommt den menschlichen Sehbedürfnissen entgegen. Für die Maschine ist die Lage des mittleren Grauwertes jedoch egal, und da, wie erwähnt, eher die gewünschte Größe der Segmente eingestellt wird, als die mittlere Grauwertdifferenz in einem Segment, ist eine Histogrammspreizung für das maschinelle Verfahren nicht erforderlich.

Unsere Untersuchungen haben gezeigt, daß solche Filter das Segmentierungsergebnis in keinem Fall verbessern konnten, in der Regel aber deutlich verschlechterten. Die Standardabweichung über das gesamte Bild hat als Qualitätsprädiktor keine Aussagekraft, ebenso wie sich generalisierte Maßnahmen zur Rauschminderung nicht vorteilhaft auswirken. Von ihrer Verwendung muß daher nach dem Ergebnissen dieser Untersuchung im Rahmen der Vorverarbeitung für eine Segmentierung mit eCognition dringend abgeraten werden. Diese Feststellung deckt sich mit den Beobachtungen von Schiewe und Tuftte [2000], die ebenfalls keine Verbesserungen durch den Einsatz des Mittelwertfilters feststellen konnten.

Es gibt einige wenige Algorithmen, die in speziellen Situationen das Segmentierungsergebnis von eCognition erkennbar verbessern können. Sie zeichnen sich sämtlich dadurch aus, daß sie die Varianz über das gesamte Bild nicht verändern. Nach Literaturangaben wurden der Graphennetzwerkfilter nach

Jahn [1996] und Halle [1999] und der EPOS-Filter nach Haag [1998] neu programmiert und für die vorgesehene Aufgabe optimiert. Beide Filter wirken sehr gut kantenerhaltend, indem sie für jedes Pixel mögliche Segmente aus der näheren Umgebung untersuchen und den Grauwert des Pixels nach einzustellenden Kriterien dem mittleren Grauwert des am besten geeigneten Segments annähern. Darüber hinaus ist die Anwendung des EPOS-Filters relativ einfach, weil er nur einen einzustellenden Parameter hat und dieser Parameter, nämlich die Varianz im beurteilten Segment, ein recht anschauliches Kriterium darstellt (Gleichung (5.20)).

Das größte Entwicklungspotential hat der Bilaterale Filter nach Tomasi und Manduchi [1998]. Er wurde hier lediglich mit verschiedenen Einstellungen im wesentlichen eines Parameters getestet. Tatsächlich erlaubt er die Verwendung verschiedener Kernel, und er erlaubt die Berücksichtigung mehrerer Bildebenen bei Mehrkanalaufnahmen, die hier nicht untersucht wurden. Die hier verwendete Version verbesserte das Segmentierungsergebnis bereits in ähnlicher Weise wie der EPOS-Filter.

Durch die Verwendung eines optimierten Kernels und eines angepaßten Rundungsverhaltens sollte es möglich sein, die störende Eigenschaften wie grobkörniges Aussehen der Ergebnisbilder und Histogrammartefakte zu eliminieren. Elad [2002] und Durand und Dorsey [2002] haben bereits untersucht, wie sich der im Vergleich zu den anderen Algorithmen extrem hohe Rechenzeitbedarf um Größenordnungen beschleunigen läßt. Weitere Forschungen sollen klären, ob der Bilaterale Filter im mehrkanaligen Fall weitere Vorteile ausspielen kann und ob seine Parametrierung optimiert werden kann.

Schließlich sollte untersucht werden, wie sich die von Ehlers [2005] entwickelte Schärfung der Multispektralaufnahmen durch den panchromatischen Kanal auswirkt.

Bis auf den Bilateralen Filter, der in seiner verbesserungsbedürftigen Originalversion im Insight Toolkit implementiert und damit für C⁺⁺-Programmierer öffentlich zugänglich ist, liegen die sinnvoll einsetzbaren Filter bislang nicht in anwenderfreundlichen, nutzbaren Versionen vor.

Als bester Filter, der bei Ansprüchen an präzise Segmentgrenzen im Pixelbereich sinnvoll einzusetzen ist und dabei mit nur einem Parameter gesteuert werden kann, erwies sich der modifizierte EPOS-Filter, ursprünglich vorgestellt von Haag und Sties [1994, 1996]. Weitere Untersuchungen könnten zeigen, ob auch dieser Parameter automatisch optimiert werden kann. Die generellen Bildparameter, wie Standardabweichung oder Histogramm, werden durch diesen kantenerhaltenden Filter nur unwesentlich beeinflusst.

9 Literatur

Angegebene Internetadressen entsprechen dem Stand von Ende 2005 und können sich erfahrungsgemäß ändern.

- Baatz, M. und Schäpe, A. [1999]. Object-Oriented and Multi-Scale Image Analysis in Semantic Networks. Object-Oriented and Multi-Scale Image Analysis in Semantic Networks. In: Proc. of the 2nd International Symposium on Operationalization of Remote Sensing, Enschede, ITC, August 16–20, 1999.
<http://www.definiens-imaging.com/documents/publications/itc1999.pdf>
- Baatz, M. und Schäpe, A. [2000]. Multiresolution Segmentation – an optimization approach for high quality multi-scale image segmentation. In: Strobl, J. et al. (Hrsg.): Angewandte Geographische Informationsverarbeitung XII. Beiträge zum AGIT-Symposium Salzburg 2000, Karlsruhe, Herbert Wichmann Verlag: 12-23, 2000.
http://www.definiens-imaging.de/documents/publications/baatz_FP_12.pdf
- Baatz, M., Benz, U., Dehghani, S., Heynen, M., Hölzje, A., Hofmann, P., Lingenfelder, I., Mimler, M., Sohlbach, M., Weber, M. und Willhauck, G. [2004].
 eCognition User Guide 4. Definiens. München, 2004.
<http://www.definiens-imaging.com>
- Baltsavias, E., Pateraki, M. und Zhang, L. [2001]. Radiometric and geometric evaluation of Ikonos GEO images and their use for 3D building modelling. Proc. Joint ISPRS Workshop High Resolution Mapping from Space 2001, Hannover, 2001.
http://www.ipi.uni-hannover.de/html/publikationen/2001/isprs_cd.pdf
- Behnke, H. Remmert, R., Steiner, H.-G. und Tietz, H. [1964]. Mathematik I, Frankfurt/M., 1964.
- Bernstein, R. [1983]. Image Geometry and Rectification. In The Manual of Remote Sensing (Ch. 21). R.N. Colwell, Ed. Falls Church, Virginia. The American Society of Photogrammetry and Remote Sensing 1, 887-888, 1983.
- Billingsley, F.C. [1979]. Noise Considerations in Digital Image Processing Hardware. In Huang, T.S. (Ed.): Picture Processing and Digital Filtering. Topics in Applied Physics Band 6. Springer. Berlin, 1979.
- Blaschke, T. [2000]. Objektextraktion und regelbasierte Klassifikation von Fernerkundungsdaten: Neue Möglichkeiten für GIS-Anwender und Planer. In: Schrenk, M. (ed.), CORP'2000: Computergestützte Raumplanung, Wien, Band 1, 153-162, 2000.
http://213.47.127.15/corp/archiv/papers/2000/CORP2000_blaschke.pdf
- Bowen, H.S. [2002]. Absolute Radiometric Calibration of the Ikonos Sensor using Radiometrically Characterized Stellar Sources. 2002.
[http://www.spaceimaging.com/whitepapers_pdfs/2002/Absolute Radiometric Calibration of the IKONOS Sensor-ISPRS 2002.pdf](http://www.spaceimaging.com/whitepapers_pdfs/2002/Absolute%20Radiometric%20Calibration%20of%20the%20IKONOS%20Sensor-ISPRS%202002.pdf)
- Cheng, Y. [1995]. Mean shift, mode seeking, and clustering, IEEE Trans. Pattern Anal. Machine Intell., 17, 790-799, 1995.
<http://doi.ieeecomputersociety.org/10.1109/34.400568>
- Comaniciu, D. und Meer, P. [1997]. Robust Analysis of Feature Spaces: Color Image Segmentation. IEEE Conf. Comp. Vision and Pattern Recogn., 750-755. June 1997.
<http://doi.ieeecomputersociety.org/10.1109/CVPR.1997.609410>
- Comaniciu, D. [2000]. Nonparametric Robust Methods for Computer Vision. Dissertation. New Brunswick, New Jersey, 2000.
http://www.caip.rutgers.edu/riul/research/theses/abstract/comaniciu_thesis.html
- Comaniciu, D. und Meer, P. [2002]. Mean Shift: A robust approach toward feature space analysis. IEEE Trans. Pattern Anal. Machine Intell. 24 (5), 603-619, 2002.
www.cse.msu.edu/~cse902/S03/MeanShif.pdf
- Crimmins, T. [1985a]. Geometric Filter for Reducing Speckle. SPIE, International Conference on Speckle Bd. 556, 213-222, 1985.
- Crimmins, T. [1985b]. Geometric Filter for Speckle Reduction. Applied Optics Bd. 24 (10), 1438-1443, 1985.

- de Kok, R. [2001]. Objektorientierte Bildanalyse. Ein Lösungsansatz für den automatisierten Einsatz sehr hoch auflösender Satellitendaten für forstliche Fragestellungen. Dissertation an der Fakultät für Forstwissenschaft der Technischen Universität München, 2001.
<http://tumbI.biblio.tu-muenchen.de/publ/diss/ww/2001/kok.pdf>
- Durand, F. und Dorsey, J. [2002]. Fast Bilateral Filtering for the Display of High-Dynamic-Range Images. International Conference on Computer Graphics and Interactive Techniques. Proceedings of the 29th annual conference on Computer graphics and interactive techniques. San Antonio, Texas. 257-266, 2002.
<http://people.csail.mit.edu/fredo/PUBLI/Siggraph2002/DurandBilateral.pdf>
- Ehlers, M. [2004]. Remote Sensing for GIS Applikationen: New Sensors and Analysis Methods, in: Ehlers, M., Kaufmann, H.J. und Michel, U. (Eds.) Remote Sensing for Environmental Monitoring, GIS Applications, and Geology IV, Proceedings of SPIE 5574, 1-13, 2004.
- Ehlers, M. [2005]. Urban Remote Sensing: New Developments and Trends, Proceedings, 5th International Symposium Remote Sensing of Urban Areas (URS 2005)
- Elad, M. [2002]. On the bilateral filter and ways to improve it. IEEE Transactions On Image Processing, Vol. 11, No. 10, 1141-1151, 2002.
http://www-sccm.stanford.edu/~elad/Conferences/I3_Bilateral_Asilomar2002.pdf
- Ford, G. E., Algazi, V.R. und Meyer, D.I. [1983]. A Noninteractive Procedure for Land-Use Determination. Remote Sensing of Environment 13, 1-16, 1983.
- Fraser, C.S., Baltsavias, E. und Gruen, A. [2002]. Processing of IKONOS imagery for submetre 3D positioning and building extraction. ISPRS Journal of Photogrammetry and Remote Sensing 56 (2002) 177-194, 2002.
www.photogrammetry.ethz.ch/general/persons/AG_pub/Fraser_Baltsavias__Processing.pdf
- Frost, V., Stiles, J., Shanmugan, K., Holtzman, J. [1982]. A model for radar images and its application to adaptive digital filtering of multiplicative noise. IEEE Transactions on Pattern Analysis and Machine Intelligence 4 (1982), 2, 157-166, 1982.
- Gerke, M. [2005]. Automatic Quality Assessment of GIS Road Data Using Aerial Imagery - A Comparison between Bayesian and Evidential Reasoning: IntArchPhRS. Band XXXVI 3/W24. Wien, 171-177, 2005.
http://www.ipi.uni-hannover.de/html/publikationen/2005/paper/CMRT05_Gerke.pdf
- Gerke, M., Heipke, C., und Straub, B.-M. [2001a]. Building Extraction From Aerial Imagery Using a Generic Scene Model and Invariant Geometric Moments, Proceedings of the IEEE/ISPRS Joint Workshop on Remote Sensing and Data Fusion over Urban Areas, IEEE Piscataway, ISBN 0-7803-7059-7, 85-89, 2001.
http://www.ipi.uni-hannover.de/html/publikationen/2001/gerke/gerke_heipke_straub_ieee_isprs.pdf
- Gerke, M., Straub B.-M. und Koch, A. [2001b]. Automatic Detection of Buildings and Trees from Aerial Imagery Using Different Levels of Abstraction, Publikationen der Deutschen Gesellschaft für Photogrammetrie und Fernerkundung, Band 10, Eckhardt Seyfert (Hrsg.), 273-280, 2001.
http://www.ipi.uni-hannover.de/html/publikationen/2001/gerke/gerke_straub_koch_dgpf.pdf
- Hájek, F. [2005]. Object-Oriented Classification of Remote Sensing Data for the Identification of Tree Species Composition. eingereicht für ForwardFORESTS - International Conference, Prag, 2006.
<http://www.definiens-imaging.com/documents/publications/Forestsat2005 - Filip Hajek.pdf>
- Haag, W. und Sties, M. [1994]. Efficient Specklefiltering of SAR Images. Proceedings IEEE International Geoscience and Remote Sensing Symposium. 2140-2142, 1994.
- Haag, W. und Sties, M. [1996]. The Epos Specklefilter: A Comparison with some Well-Known Speckle Reduction Techniques. Proceedings ISPRS, Commission II, Vienna, Austria 135-140, 1996.
- Haag, W. [1998]. Merkmalbasierte Klassifikation von SAR-Satellitenbilddaten. Dissertation. Karlsruhe Halle, W. [1999]. Ausgewählte Algorithmen der Segmentierung und Klassifikation zur thematischen Bildverarbeitung in der Fernerkundung. Dissertation. TU Berlin, 1999.
- Hanning, T. [2001]. Vektorielle Mehrniveaupassung. Dissertation. Passau, 2001.
www.forwiss.uni-passau.de/~hanning/diss.pdf

- Haralick, R.M. und Shapiro, L.G. [1992]: Computer and Robot Vision. Bd. 1. Addison-Wesley, 1992.
- Hoover, A., Jean-Baptiste, G., Goldgof, D. und Bowyer, K. [1994]. A Methodology for Evaluating Range Image Segmentation Techniques. IEEE Workshop on Applications of Computer Vision, 264-271, Sarasota, Florida, 1994.
<ftp://figment.csee.usf.edu/pub/segmentation-comparison/WACV-final.ps.Z>
- Hoover, A., Jean-Baptiste, G., Jiang, X., Flynn, P.J., Bunke, H., Goldgof, D., Bowyer, K., Eggert, D., Fitzgibbon, A. und Fisher, R. [1996]. An Experimental Comparison of Range Image Segmentation Algorithms. IEEE PAMI, 18 (7):673-689, July 1996.
<ftp://figment.csee.usf.edu/pub/segmentation-comparison/PAMI-segcomp.ps.Z>
- Huang, T.S. [1979]. Picture Processing and Digital Filtering. Topics in Applied Physics Band 6. Springer. Berlin, 1979.
- Hütte [1989]. Die Grundlagen der Ingenieurwissenschaften. Springer, Berlin, 1989.
- Jahn, H. [1996]. Image Segmentation with a Layered Graph Network. Proceedings of the Society of Photo-Optical Instrumentation Engineers, 2662, 217-228, 1996.
- Jensen, J.R. [1986]. Introductory Digital Image Processing – A Remote Sensing Perspective. Prentice-Hall. Englewood Cliffs, New Jersey, 1986.
- Kass, M., Witkin, A. und Terzopoulos, D. [1987]. Snakes: active contours models. Proceedings of the First International Conference on Computer Vision, ICCV 1987, 259-268, 1987.
- Kass, M., Witkin, A. und Terzopoulos, D. [1988]. Snakes: Active Contour Models. International Journal of Computer Vision, 1 (4) 321-331, 1988.
- Kender, J. [1976]. Saturation, Hue and Normalized Color: Calculation, Digitization Effects, and Use. Technical Report. Department of Computer Science, Carnegie-Mellon University 1976.
- Kerschner, M. [2003]. Snakes für Aufgaben der digitalen Photogrammetrie und Topographie. Dissertation. Institut für Photogrammetrie und Fernerkundung der Technischen Universität Wien, 2003.
http://www.ipf.tuwien.ac.at/phdtheses/mk/diss_mk.pdf
- Klette, R. und Zamperoni, P. [1995]. Handbuch der Operatoren für die Bildbearbeitung. Vieweg. Braunschweig, 1995.
- Kotelnikow, V.A. [1933]. On the transmission capacity of "ether" and wire in electrocommunications, Izd. Red. Upr. Svyazzi RSKA. 1933
- Lee, J.-S. [1980]. Digital Image Enhancement and Noise Filtering by Use of Local Statistics. IEEE Transactions on Pattern Analysis and Machine Intelligence PAMI-2, 2, 165-168, 1980.
- Lee, J. [1981]. Refined filtering of image noise using local statistics. Computer Vision, Graphics and Image Processing 15, 380-389, 1981.
- Lin, C. und Nevatia, R. [1998]. Building detection and description from a single intensity image. Computer Vision and Image Understanding, 72, 101-121, 1998.
<http://iris.usc.edu/Outlines/papers/1998/lin-cviu98.pdf>
- Lopes, A., Touzi, R., Nezry, E. [1990a]. Adaptive specklefilters and scene homogeneity. IEE Transactions on Geoscience and Remote Sensing 28 (6), 992-1000.
- Lopez, A., Nezry, E., Touzi, R. [1990b]. Maximum a posteriori filtering and first order texture models in SAR images. Proc. IEEE International Geoscience and Remote Sensing Symposium, , 2409-2412, 1990.
- Malooof, M.A., Langley, P., Binford, T.O., Nevatia, R. und Sage, S. [2003]. Improved Rooftop Detection in Aerial Images with Machine Learning. Mach. Learn. 53, 157-191, 2003
www.cs.georgetown.edu/~malooof/pubs/mlj03.pdf
- Möller, M. [2003]. Urbanes Umweltmonitoring mit digitalen Flugzeugscannerdaten. Wichmann. Heidelberg, 2003. Auch: Untersuchungen an extrem hochauflösenden Flugzeugscannerdaten für urbane Fragestellungen und deren Integration in eine GIS-Umgebung. Dissertation. Vechta, 2002.
- Neubert, M. und Meinel, G. [2003]. Evaluation of Segmentation Programs for High Resolution Remote Sensing Applications. Joint Workshop of ISPRS WG I/2,I/5 and IC WG II/IV and

- EARSeL Special Interest Group 3D Remote Sensing. Hannover, 2003.
www.ipi.uni-hannover.de/html/publikationen/2003/workshop/neubert.pdf
- Nyquist, H. [1928]. Certain topics in telegraph transmission theory, *Trans. Amer. Inst. Elect. Eng.* 47, 1928.
- Pavlidis, T. [1977]. *Structural Pattern Recognition*. Springer-Verlag, Berlin, 1977.
- Perona, P. und Malik, J. [1990]. Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis Machine Intelligence*, 12, 629–639, 1990.
- Pollock, R.J. [1994]. A model-based approach to automatically locating tree crowns in high spatial resolution images. *Image and Signal Processing for Remote Sensing*, (eds.) Desachy, SPIE, Vol. 2315, 526-537, 1994.
- Read, R.R., Shanks, J.L. und Treitel, S. [1979]. Two-Dimensional Recursive Filtering. In Huang, T.S. (Ed.): *Picture Processing and Digital Filtering*. Topics in Applied Physics Band 6. Springer. Berlin, 1979.
- Rosenfeld, A. und Kak, A. [1982]. *Digital Picture Processing Vol. 2 ; 2. Aufl.*; Academic Press, 1982.
- Sahoo, P.K., Soltani, S., und Wong, A.K.C. [1988]. A survey of thresholding techniques, *Comp. Vis., Graph. and Imag. Proc.*, 41, 233-260, 1988.
- Schiewe, J. und Tufte, L. [2001]. Multi-skalige Klassifizierung von hochauflösenden Bilddaten. *Publikationen der DGPF (Jahrestagung Berlin)*. 9: 402-409, 2001.
- Schowengerdt, R.A. [1997]. *Remote Sensing: Models and Methods for Image Processing*, Academic Press. New York 1997.
- Schowengerdt, R.A. [2002]. Spatial Characterisation of IKONOS
http://ldcm.gsfc.nasa.gov/library/HSRCIW02/SpatialChar_IKONOS_Schowengerdt.pdf
- Shannon, C.E. [1949]: Communication in the Presence of Noise. *Proceedings of the IEEE*, 86, (2), February 1998, 447- 457 Nachdruck der Originalarbeit aus *Proceedings of the IRE*, 37, (1), 10–21, Jan. 1949. <http://www.stanford.edu/class/ee104/shannonpaper.pdf>
- Straub, B.M. [2003]. Automatische Extraktion von Bäumen aus Fernerkundungsdaten. Dissertation Hannover, 2004.
<http://www.ipi.uni-hannover.de/html/publikationen/2003/paper/diss-straub-hannoverreihe-249.pdf>
- Straub, B.M. [2004] A Multiscale Approach for the Automatic Extraction of Tree Tops from Remote Sensing Data. *IntArchPhRS*. Band XXXV, Teil B3. Istanbul, 418-423, 2004.
www.isprs.org/istanbul2004/comm3/papers/306.pdf
- Tomasi, C. und Manduchi, R. [1998]. Bilateral filtering for gray and color images. *Proc. IEEE Int. Conf. on Computer Vision*, 836-846, 1998. www.cse.ucsc.edu/~manduchi/Papers/ICCV98.pdf
- Udupa, J.K., LeBlanc, V.R., Schmidt, H., Imielinska, C., Saha, P.K., Grevera, G.J., Zhuge, Y., Currie, L.M., Molholt, P. und Jin, Y. [2002]. A Methodology for Evaluating Image Segmentation Algorithms. *Medical Imaging 2002: Image Processing*. Milan Sonka, J. Michael Fitzpatrick (Eds.), *Proceedings of SPIE 4684*, 299-277, 2002.
- Vey, S. [1993]. Objekterkennung in Grauwertbildern. Dissertation TH Darmstadt. 1993.
- Wahl, F.M. [1984]. *Digitale Bildsignalverarbeitung*. Nachrichtentechnik 13, Springer, Berlin, 1984.
- Wang, H. und Suter, D. [2003]. False-peaks-avoiding mean shift method for unsupervised peak-valley sliding image segmentation. In *Proceedings 7th International Conference on Digital Image Computing: Techniques and Applications (DICTA'03)*, Sydney, 581-590, 2003.
http://www.ds.eng.monash.edu.au/suter_publications/0581.pdf
- Weickert, J., ter Haar Romeny, B.M., und Viergever, M.A. [1996]. Conservative image transformations with restoration and scale-space properties. In *Proc. 1996 IEEE International Conference on Image Processing (ICIP-96, Lausanne, Sept. 16-19, 1996)*, 465-468, 1996. <http://www.mia.uni-saarland.de/weickert/Papers/icip96.ps.gz>
- Williams, D. J. und Shah, M. [1992]. A Fast Algorithm for Active Contours and Curvature Estimation. *CVGIP: Image Understanding* 55 (1) 14-26, 1992

- Whitaker, R. T. [1993]. Geometry-Limited Diffusion, Ph.D. Dissertation, The University of North Carolina, Department of Computer Science 1993.
<ftp://ftp.ecrc.de/ross/dissertation>
- Whitaker, R.T. und Xue, X. [2001]. Variable-Conductance, Level-Set Curvature for Image Processing. In International Conference on Image Processing. IEEE 2001.
http://www.cs.utah.edu/~xwxue/curvature_paper.pdf
- Whittaker, J.M. [1929]. The Fourier theory of the cardinal functions, Proc. Edinburgh Math. Soc. 1 (1929)

