

INSTITUT FÜR PHOTOGRAMMETRIE UND INGENIEURVERMESSUNGEN
UNIVERSITÄT HANNOVER

Diplomarbeit

**Verwendung von Texturparametern bei der Klassifizierung von
hochauflösenden Satellitendaten**

eingereicht im: Januar 2000

von: Martin Weis

geboren am 18. September 1972

in Detmold

Einleitung

In dieser Arbeit wird die Nutzung von Texturinformationen für die Auswertung von multispektralen Luftbildern untersucht. Dazu wird in Kapitel 1 ein Überblick über die Verfahren gegeben, mit denen sich Texturinformationen aus Bildern extrahieren lassen. Den mehrfachauflösenden Verfahren, die in der Bildverarbeitung häufig eingesetzt werden, ist Kapitel 2 gewidmet. Es wird das Verhältnis von Orts- und Frequenzraum beschrieben. Eine allgemeine Definition von Mehrfachauflösungen wird gegeben und einige mehrfachauflösende Verfahren auf ihre Tauglichkeit hin überprüft. In Kapitel 3 wird die Nutzung von Gaborfiltern zur Texturanalyse untersucht. Die Auswahl des Verfahrens fällt auf ein mehrfachauflösendes Verfahren, das Orientierungsinformationen bereitstellt. Diesem Verfahren der steuerbaren Pyramide ist Kapitel 6 gewidmet. In dem Kapitel werden desweiteren die Erzeugung eines Texturlayers und seine Einführung in die multispektrale Klassifikation erläutert. Die Ergebnisse der Klassifikationen werden präsentiert und beurteilt.

Introduction

In this thesis the use of textural information in multispectral aerial images is investigated. An overview of methods that extract textural information from images is given in chapter 1. Multiresolutional methods, often used in image processing, are the subject in chapter 2. The relation between spacial and frequency resolution is described. A general definition of multiresolution techniques is given and the worth of some multiresolution methods is evaluated. Chapter 3 focusses on the use of Gaborfilters for textural analysis. A multiresolutional technique providing orientational information is selected. Chapter 6 is dedicated to this technique called steerable pyramid. Further the generation of a layer containing textural information and its addition to the process of multispectral classification is described. The results are presented and reviewed.

Ich versichere, daß ich diese Diplomarbeit selbständig verfaßt und keine anderen
als die angegebenen Hilfsmittel verwendet habe.

Hannover, den 16. März 2000

Inhaltsverzeichnis

1	Textur	8
1.1	Textur - was ist das?	8
1.2	Texturanalyse	10
1.2.1	Statistische Methoden	11
1.2.1.1	Statistik erster Ordnung	12
1.2.1.2	Statistik zweiter Ordnung	13
1.2.1.3	Strukturelle Methoden	15
1.2.1.4	Stochastische Methoden	16
1.2.1.5	Fraktale Methoden	17
1.2.1.6	Neuronale Methoden	19
1.3	Datengrundlage	20
2	Multiresolutionsansätze	23
2.1	Die Fouriertransformation	25
2.1.1	Fourierreihendarstellung	26
2.1.2	Fouriertransformation	27
2.1.2.1	Heisenbergsche Unschärferelation	29

2.1.2.2	Kompakter Träger	29
2.2	Short Time Fouriertransformation	29
2.3	Der Burt-Adelsonsche Pyramidenalgorithmus	30
2.3.1	Gaußpyramide	31
2.3.2	Laplacepyramiden	32
2.3.2.1	DoG-Pyramide	32
2.3.2.2	LoG-Pyramide	32
2.4	Texturenergie nach Laws	33
2.5	Ortsfrequenzraum-Darstellungen	35
2.6	Die Wavelettransformation	36
2.6.1	Kontinuierliche Wavelettransformation	37
2.6.2	Die diskrete Wavelettransformation	38
2.7	Konstruktion einer Mehrfachauflösung	39
2.7.1	Vektorräume, Skalarprodukt, Basen und Orthogonalität	39
2.7.2	Definition einer Mehrfachauflösung	40
2.7.3	Konstruktion einer Mehrfachauflösung nach Mallat	42
2.7.3.1	Skalierungsfunktion	42
2.7.3.2	Das Detailsignal	44
2.7.4	Anwendung auf Bilder	46
2.8	Weitere Zerlegungen mit Wavelets	47
2.8.1	Aliasing und Verschiebungsinvarianz	49
3	Gaborfilterung in der Texturanalyse	52
3.1	Hilberttransformation und Quadraturfilterpaare	52

3.1.1	Energie eines Signales	52
3.1.2	Hilberttransformation	53
3.2	Gaborfilter	53
3.2.1	Definition der Gaborfilter	54
3.2.2	Bandbreiten	55
3.2.3	Gaborfilter und die Unschärferelation	57
3.3	Gaborfilter in der Texturanalyse	57
3.3.1	Auswahl der Filter	58
3.3.2	Lokale Energiefunktion	60
3.4	Untersuchungen mit Gaborfiltern	60
4	Erzeugung eines Layers mit Texturinformation	66
4.1	Auswahl von Testbereichen	67
4.2	Der Ansatz nach Laws	67
4.3	Gauß- und Laplacepyramiden	69
4.4	Bildpyramiden mit orthogonalen Basen	70
5	Klassifikation	72
5.1	Ablauf der Klassifizierung	72
5.1.1	Erzeugung von mehrkanaligen Bildern	74
5.2	Ergebnisse der multispektralen Klassifikation	75
6	Die steuerbare Pyramide	80
6.1	Funktionsweise der steuerbaren Pyramide	80
6.2	Die Pyramidentoolbox für MATLAB	85

6.3	Berechnungen mit der steuerbaren Pyramide	87
6.3.1	Funktionalitäten unter MATLAB	87
6.3.2	Probleme bei den Berechnungen unter MATLAB	87
6.4	Berechnungen mit den Pyramidentools	88
6.5	Nutzen der Orientierungsinformation	91
6.6	Erzeugung eines Texturlayers aus den Merkmalsbildern	92
6.7	Klassifikation mit Texturkanal	101
6.8	Nutzen von Texturinformationen in der Klassifikation	104
	Zusammenfassung	107
	A Contingency Matrizen der Klassifikationen	114
	B Daten der Feldbegehung	118

Abbildungsverzeichnis

1.1	Die Luftbildszene in Infrarotfarbdarstellung	22
2.1	Fouriertransformation von Bildern	28
2.2	Frequenzselektivität der eindimensionalen Lawsfiltermasken	35
2.3	Resultierende zweidimensionale Aufteilung der Frequenzebene . .	35
2.4	Zerlegung der Ortfrequenzebene durch die STFT	36
2.5	Ortsfrequenzebenenentteilung im Falle der Wavelettransformation . .	36
2.6	Zerlegung des Frequenzraumes durch Tiefpaß- und Hochpaßanteile	47
2.7	Anordnung der $D_{2^j}^k f$ und $A_{2^j}^d f$	47
2.8	Ausschnitt aus der Luftbildszene	48
2.9	Die Zerlegung des Bildes 2.8	48
2.10	Durch eine Energietransformation gewonnenes Merkmalsbild . . .	48
2.11	Durch Aliasing entstandene Differenzen	51
3.1	Impulsantworten von vier Gaborfiltern	56
3.2	Gaborfilter im Fourierraum	56
3.3	Gaborfilterbank nach [1], eingetragen sind Halbwertsbandbreiten .	58
3.4	tanh-Funktion	61
3.5	Aus verschiedenen Texturen zusammengesetztes Bild	62

3.6	Die Merkmalsbilder der Gaborfilterung mit acht Filtern	63
3.7	Merkmalsbild: <i>Merkmale, die in zwei senkrecht aufeinanderstehenden Orientierungen vorhanden sind</i>	64
3.8	Merkmalsbild: <i>Merkmale, die in nur einer von zwei senkrecht aufeinanderstehenden Orientierungen vorhanden sind</i>	64
4.1	Filterungen nach Laws	68
4.2	Gaußpyramide	69
4.3	Laplacepyramide	70
4.4	Texturklassen aus der Klassifikation einer orthogonalen Pyramidenzerlegung	71
5.1	Klassifikationsergebnis der multispektralen Klassifikation	75
5.2	Der signature editor von ERDAS, Legende für Abb. 5.1	76
6.1	Zerlegungsschema einer Stufe der steuerbaren Pyramide	82
6.2	Filter $L_0(\omega)$	82
6.3	Zerlegungsschema der steuerbaren Pyramide	83
6.4	Orientierungselektive Filter der steuerbaren Pyramide	84
6.5	Die steuerbaren Filter im Frequenzraum	84
6.6	Pyramidenstufe 1, Orientierung 3	90
6.7	Das zu Abb. 6.6 gehörige Merkmalsbild	90
6.8	Merkmalsbild der Pyramidenstufe 1, Orientierung 6	90
6.9	Merkmalsbild der Pyramidenstufe 2, Orientierung 6	90
6.10	Testbild	93

6.11 Kombination der Merkmalsbilder einer Orientierung aus den ersten drei Stufen der Pyramide	93
6.12 Die pyramidale Zerlegung des Testbildes 6.12	94
6.13 Farbiges Orientierungsbild	95
6.14 Ablauf zur Erzeugung des Texturlayers	96
6.15 Merkmalsraum für die Klasse Wald	97
6.16 Merkmalsraum für die Klasse Siedlung	97
6.17 Merkmalsraum für die Klasse Feld	97
6.18 Farbkodiertes Orientierungsmerkmalsbild aus drei Richtungen . .	99
6.19 Texturlayer, Ergebnis der Klassifikation aus dem Orientierungsbild	100
6.20 Texturlayer aus dem Orientierungsbild der verbleibenden drei Richtungen	100
6.21 Klassifikationsergebnis mit eingeführtem Texturlayer	102
6.22 Differenzen, die sich durch die Einführung des Texturkanales in die Klassifikation ergeben	104
6.23 Die Änderungen in der Klassenzuordnung in Abhängigkeit von der Klassenzuweisung im Texturlayer	104

Kapitel 1

Textur

1.1 Textur - was ist das?

Der Begriff Textur wird in vielen verschiedenen wissenschaftlichen Disziplinen verwendet. Es gibt keine eindeutige Definition.

Das Wort hat seinen Ursprung im lateinischen „textura“ und heißt wörtlich übersetzt „Gewebe“. In der Kristallographie versteht man darunter die Gesamtheit der Orientierungen der Kristalle in einem vielkristallinen Stück [2], in der Holztechnik wird damit die Zeichnung an längsgeschnittenen Holz- oder Furnierflächen, in Kunst und Design oft eine erfühlbare Qualität einer Oberfläche bezeichnet [3] [4].

Es besteht zwar kein einheitlicher Texturbegriff, aber es gibt qualitative Eigenschaften, die mit anschaulichen Attributen wie fein, grob, körnig, zufällig, gerichtet, regulär oder wellig bezeichnet werden können [5] [6].

Es wurden viele Versuche der Definition gemacht, die sich aber zumeist an einer Begründung des jeweils gewählten Ansatzes zur Texturanalyse orientieren und nicht an einer einheitlichen, verfahrensübergreifenden Definition des Begriffes ([7], S.10).

Weil es keine eindeutige Definition gibt, seien hier einige gängige Definitionen wiedergegeben. Einige Ansätze gehen davon aus, daß sich Textur aus Grund-

elementen aufbaut, die bei verschiedenen Autoren Texel (abgeleitet von *texture element*) [8], Textone, Primitive oder Grundmuster genannt werden. Texturen können darüber hinaus auch hierarchisch aus Elementen aufgebaut sein, die ihrerseits wiederum texturiert sind. So hat z.B. im Fernerkundungsbild ein Siedlungsbereich eine Textur, die sich bei lokaler Betrachtung noch in weitere Texturen von Häusern, Straßen, Bäumen und anderen Objekten unterteilen lässt.

Für die Bildverarbeitung findet man an Definitionen:

Textur ist ein Attribut, das die räumliche Anordnung der Grauwerte von Pixeln in einer Region darstellt [9].

Der Ausdruck Textur bezieht sich auf die Wiederholung von Basis-texturelementen, genannt Texel. Ein Texel enthält verschiedene Pixel, deren Anordnung periodisch, quasiperiodisch oder zufällig sein kann. Natürliche Texturen sind generell zufällig, wohingegen künstliche Texturen oft deterministisch oder periodisch sind.

Eine Region besitzt eine konstante Textur, wenn ein Satz lokaler Statistiken oder Eigenschaften der Bildfunktion konstant, sich langsam ändernd oder nahezu periodisch ist [10] (nach [11]).

Hier soll der Begriff nur für den Bereich der Bildverarbeitung gebraucht werden.

Deshalb bezeichnet Textur in dieser Arbeit eine charakteristische Eigenschaft eines Bildes oder einer Bildregion.

Da in der Fernerkundung die Erdoberfläche oder Objekte auf ihr abgebildet werden, resultiert die im Bild sichtbare Textur aus den Oberflächeneigenschaften. Die Analyse dieser Textur soll zur Erkennung und Unterscheidung von Objekten beitragen.

1.2 Texturanalyse

Ziel der Texturanalyse ist es, geeignete Parameter zu finden, die Texturen beschreiben und von anderen Texturen unterscheiden. Mit den in diesem Kapitel angeführten Methoden lassen sich typische Merkmale von Texturen finden. Die Merkmale, nach denen sich Texturen unterscheiden lassen, sollen der Abgrenzung von Gebieten mit unterschiedlicher texturaler Ausprägung dienen. Das untersuchte Bild muß daher in Bereiche mit gleichen Texturen eingeteilt werden. Diesen Vorgang nennt man Segmentierung des Bildes. Den einzelnen Bereichen, die dadurch gebildet werden, können dann Attribute zugewiesen werden. Diese Zuweisung der Attribute zu den Bereichen ist die Klassifizierung.

Die Bildanalyse hat diese Zuweisung zu den Objektklassen als Ziel. Bisher stehen für die Klassifikation aus multispektralen Bildern einige Verfahren zur Verfügung, die radiometrische Informationen aus verschiedenen Kanälen kombinieren und so eine Klassifikation ermöglichen (s. Kapitel 5). Diese spektralen Merkmale sollen nun um die Texturinformationen ergänzt werden, um das Klassifikationsergebnis zu verbessern und sicherer zu machen.

Das Gebiet der Texturanalyse wird schon seit einigen Jahren erforscht. Hier sollen einige erprobte Verfahren vorgestellt werden, wie sie in der Literatur zu finden sind. Es gibt viele Ansätze und die Verfahren lassen sich nicht immer eindeutig zuordnen. Häufig werden *statistische* und *strukturelle* Verfahren unterschieden. Hinzu kommen *syntaktische* und *stochastische* Verfahren. Es hat sich herausgestellt, daß außerdem die mehrfach auflösende Signalanalyse (*Multi Resolution Analysis*, MRA, s. Kapitel 2) eine der vielversprechendsten Herangehensweisen ist, weshalb ihr ein größerer Teil der Arbeit gewidmet wird.

Wie ein Erkennungssystem gestaltet werden muß, damit ähnlich gut und sicher erkannt wird wie mit dem menschlichen Sehapparat¹, ist die Fragestellung in der *computational vision*². Bisher ist die Leistung des Sehapparates bei der Unterscheidung von Texturen durch Nachbildung in Algorithmen nicht erreicht worden und seine Funktionsweise kann auch nicht vollständig beschrieben werden.

¹Mit Sehapparat ist das kombinierte System Auge–Gehirn gemeint.

²Wissenschaft, die sich mit der maschinellen Erkennung beschäftigt.

Denn das Erkennen ist ein komplizierter Prozeß, der auf verschiedenen Ebenen abläuft. In diesen gehen sowohl Erinnerung als auch Erlerntes ein, die auf hoher kognitiver Stufe stehen. Dieser Prozeß kann nicht in allen Details nachgebildet werden.

Es gibt allerdings Ansätze, die Erkenntnisse der *human vision*³ mathematisch umzusetzen. Dabei ergibt sich aber die Schwierigkeit, daß physiologische oder psychologische Sachverhalte meist nicht in Zahlen und algorithmischen Regeln ausgedrückt sind, sondern in Worten. Trotzdem gibt es vielversprechende Ansätze zur algorithmischen Darstellung menschlichen Sehens.

1.2.1 Statistische Methoden

Statistische Methoden bei der Merkmalextraktion beruhen auf einer Untersuchung der Grauwerte einzelner Pixel. Wird die Statistik für jeweils nur ein Pixel betrachtet, so spricht man von *Statistik 1. Ordnung*, betrachtet man die Statistik für zwei Pixel, also für Paare von Pixeln, die in bestimmter Richtung und Entfernung liegen, so handelt es sich um die *Statistik 2. Ordnung* (auch: *Verbundstatistik*). Es wurden auch Untersuchungen zur *Statistik 3. Ordnung* gemacht, die aber nicht zu signifikanten Ergebnissen geführt haben. Untersuchungen hierzu hat vor allem Julesz gemacht, die Muster generierte und deren Unterscheidbarkeit durch das Auge untersuchte. Sie stellte fest, daß der Mensch auf Änderungen in der Statistik 1. und 2. Ordnung der Grauwerteverteilung sehr empfindlich reagiert, daß bei Verbundstatistiken höherer Ordnung allerdings eine Unterscheidung durch das menschliche Auge nicht möglich ist [7].

Julesz unterscheidet zwei Arten der Wahrnehmung des Menschen: Die differenzierte Wahrnehmung lokaler Textureigenschaften und die globale Wahrnehmung unterschiedlicher Texturmuster. Die differenzierte Wahrnehmung beruht auf der Fähigkeit des visuellen Systems, statistische Eigenschaften zu erkennen und zu unterscheiden. Die globale Wahrnehmung hingegen bezieht sich auf die Kontextinformation

³Wissenschaft, die sich mit der menschlichen Wahrnehmung beschäftigt.

und ermöglicht es, Figur und Hintergrund voneinander zu trennen. [...] Lokale Wahrnehmung differenziert Muster aufgrund statistischer Merkmale. Unterschiede in der Statistik erster Ordnung werden als Dichteschwankung, Unterschiede in der zweiten Ordnung dagegen in Schwankungen der Körnigkeit wahrgenommen ([12], S.7).

Im folgenden werden die Methoden vorgestellt, mit denen Merkmale der Statistik 1. und 2. Ordnung erhalten werden können.

1.2.1.1 Statistik erster Ordnung

Die Statistik erster Ordnung läßt sich mathematisch durch den Mittelwert und die Varianz der Verteilung der Grauwerte beschreiben :

Sei N die Anzahl der möglichen Grauwerte eines Bildes und

$$p(s), \quad 1 \leq s \leq N$$

die relative Häufigkeit des Grauwertes s in dem betrachteten Bild mit

$$0 \leq p(s) \leq 1 \forall s, \quad \sum_{s=1}^N p(s) = 1, \quad (1.1)$$

so erhält man hierfür die Momente n -ter Ordnung μ_n entsprechend der Gleichung

$$\mu_n = \sum_{s=1}^N p(s) \cdot s^n, \quad n \in \mathbb{N}, \quad (1.2)$$

sowie die zentralen Momente n -ter Ordnung σ_n^2 zu

$$\sigma_n^2 = \sum_{s=1}^N (s - \mu_1)^n \cdot p(s). \quad (1.3)$$

Dabei sind das Moment erster Ordnung μ_1 (Mittelwert) und das zentrale Moment zweiter Ordnung σ_2^2 (Varianz) die gebräuchlichsten Momente in der Bildanalyse. Eine allgemeingültige Aussage über die

Brauchbarkeit der Statistik erster Ordnung ist kaum möglich, da sie zum einen für viele Texturen bereits gute Merkmale zur Unterscheidung liefert, auf der anderen Seite auch leicht unterscheidbare Texturen zu finden sind, die eine identische Statistik erster Ordnung besitzen ([13], S.12).

Die Momente höherer Ordnung beschreiben z.B. die Abweichung von der Symmetrie der Verteilung (Moment dritter Ordnung). Ist eine Verteilung symmetrisch, so verschwinden das dritte und alle höheren ungeraden Momente. ([14], S.76)

1.2.1.2 Statistik zweiter Ordnung

Zur Bestimmung von Merkmalen der Statistik zweiter Ordnung werden die Grauwertabhängigkeiten jeweils zweier im Abstand δ und Richtung θ liegender Grauwerte in eine Matrix eingetragen. Diese Matrix wird als **CO-OCCURENCE Matrix** (auch: *spatial gray-tone dependence matrices*, *Verbundverteilung*) bezeichnet. Aus ihr können mehrere Parameter abgeleitet werden. In die Matrix wird die Häufigkeit, mit der die Kombination *Grauwert*₁—*Grauwert*₂ auftritt, eingetragen. *Grauwert*₁ und *Grauwert*₂ haben zueinander eine durch δ und θ definierte Lage. Die Dimension der Matrix hängt von der Anzahl der möglichen Grauwerte ab. So ist die Größe für eine 8 Bit tiefe Grauwertdifferenzierung (256 Grauwerte) entsprechend 256×256 . Für jede relative Position der Grauwerte, die untersucht wird, erhält man eine Matrix. Aus der Verteilung und dem Wert der Matrixelemente lassen sich jetzt Rückschlüsse auf die Beschaffenheit der Grauwertverteilung ziehen. Eine Rekonstruktion des Ausgangsbildes ist mit diesen Matrizen nicht möglich. Eine geringe Änderung der Grauwerte führt auf eine stark besetzte Diagonale der Matrix. Sind die Elemente entfernt von der Diagonalen, weist das dagegen auf eine starke Grauwertschwankung hin. Die Matrizen sind symmetrisch, wie man sich durch Vertauschung der Zählindizes klarmachen kann. Es hat keinen Sinn, den Abstand der Grauwertpaare zu groß zu wählen, da die daraus abgeleiteten Merkmale nicht sehr aussagekräftig sind. Die menschliche Wahrnehmung analysiert nur die lokalen Zusammenhänge. Häufig wird zur Untersuchung der Abstand zweier angrenzender Pixel unter den Winkeln 0° , 45° ,

90°, 135° gewählt. In der Literatur findet man größtenteils Merkmale, die auf die grundlegenden Untersuchungen von *Haralick* zurückgehen, der 14 Merkmale aufstellte. Sie werden auch *Haralick-Momente* genannt.

Sie erfassen statistische und informationstheoretische Eigenschaften von Texturen und sind in folgende Merkmalklassen einteilbar:

1. Merkmale, die Grauwertabhängigkeiten von Texturen ausdrücken. Sie können visuelle Textureigenschaften quantifizieren. Zu diesen Maßen gehören: das zweite Winkelmoment, der Kontrast und die Korrelation.
2. Merkmale, die sich auf statistische Eigenschaften gründen. Zu diesen Maßen gehören: Trägheitsmoment, lokale Homogenität, Varianz, Summe der Durchschnitte, Summe der Varianzen und Differenz der Varianzen.
3. informationstheoretische Maße: Summe der Entropien, Entropie, Differenz der Entropien.
4. Korrelationsmaß: Korrelationskoeffizient([12], S.21).

Ich gebe hier die Formeln für vier oft benutzte Merkmale an, wie sie im Programmsystem `meastex` implementiert sind⁴:

$$\text{Energie} = \sum_{i,j} f(i,j)^2 \quad (1.4)$$

$$\text{Entropie} = - \sum_{i,j} f(i,j) \log(f(i,j)) \quad (1.5)$$

$$\text{Trägheit} = \sum_{i,j} (i-j)^2 f(i,j) \quad (1.6)$$

$$\text{Korrelation (Haralick)} = \frac{\sum_{i,j} (ij) f(i,j)^2 - \mu_x \mu_y}{\sigma_x \sigma_y} \quad (1.7)$$

⁴Erhältlich unter: <http://www.cssip.elec.uq.edu.au/~guy/meastex/meastex.html>, 1999

Unter **meastex** wurden zusätzlich zwei von Conners and Harlow vorgeschlagene Merkmale implementiert:

$$\text{Korrelation} = \sum_{i,j} \frac{(i - \mu_i)(j - \mu_j)f(i,j)}{\sigma_i\sigma_j} \quad (1.8)$$

$$\text{Inverses Differenz Moment} = \sum_{i,j} \frac{1}{1 + (i - j)^2} f(i,j) \quad (1.9)$$

Diese Merkmale wurden von Conners, Trivedi und Harlow um folgende zwei Merkmale erweitert, die die oben angegebenen ergänzen:

$$\text{Clusterschatten} = \sum_{i,j} ((i - \mu_i) + (j - \mu_j))^3 f(i,j) \quad (1.10)$$

$$\text{Clusterhervortreten} = \sum_{i,j} ((i - \mu_i) + (j - \mu_j))^4 f(i,j) \quad (1.11)$$

$$\text{worin } \mu_i = \sum_i i \sum_j f(i,j)$$

$$\mu_j = \sum_j j \sum_i f(i,j)$$

$$\sigma_i = \sum_i (i - \mu_i)^2 \sum_j f(i,j)$$

$$\sigma_j = \sum_j (j - \mu_j)^2 \sum_i f(i,j)$$

1.2.1.3 Strukturelle Methoden

Bei den strukturellen Methoden geht man davon aus, daß sich die Textur aus Primitive zusammensetzt. Diese Primitive, auch Textone genannt, stellen Grundmuster dar. Sind diese Primitive bekannt, kann im Falle einer Textursynthese eine Textur daraus zusammengesetzt oder eine vorhandene Textur auf die enthaltenen Primitive hin untersucht werden. Ist es also möglich, eine gegebene Textur aus Textonen zu synthetisieren, so erhält man als Texturmerkmale die Lage und Ausrichtung der benutzten Textone im Bild. Für diese Herangehensweise muß a priori eine gute Kenntnis der Textur vorhanden sein.

Als Texturelemente lassen sich angeben: Flächen gleichen Grauwertes, die durch Größe, Form, Ausrichtung und den mittleren Grauwert charakterisiert

sind. Andere Texturelemente können z.B. „Flecken“ sein oder lokale Extrema (*Min-Max-Merkmale*), die entlang einer Geraden ausgezählt und aufgrund der ersten Ableitung (Gradient, Steigung der Flanken) weiter differenziert werden. Die Auswertung in nur einer Richtung (entlang einer Geraden) ist hierbei nur für rotationsinvariante Texturen sinnvoll. Die räumliche Anordnung der Texturelemente wird in einer Grammatik oder einem Graphen kodiert. ([15], s.56ff)

Das Problem bei allen rein strukturellen Ansätzen ist die Tatsache, daß sie deterministisch sind und sich somit praktisch nicht auf natürlich auftretende Texturen anwenden lassen, da diese immer statistischen Strukturschwankungen unterworfen sind.([13], S.11)

Deshalb ist die Anwendung auf Luft- und Satellitenbilder nicht aussichtsreich. Sie wird hier nicht weiter verfolgt. Es ist nicht möglich, eine endliche Anzahl Primitive zu definieren, die in solchen Bildern vorkommen. Ein solcher Ansatz wäre immer unvollständig, da viele Texturen, z.B. die von Wald und felsigem Gebiet erzeugten, nicht regelmäßiger Natur sind.

1.2.1.4 Stochastische Methoden

Die Texturbeschreibungen, die bisher vorgestellt wurden, erfassen aber nicht alle Eigenschaften von Bilddaten, die in der Fernerkundung zu finden sind. Selbst wenn man extern bedingte Variationen durch Aufnahmegeometrie, Beleuchtung und Kameraparameter vernachlässigt, entspricht das abstrakte Modell den Texturen nur unvollständig. Abweichungen entstehen durch die fundamentale und wichtige Eigenschaft von natürlichen Texturen: ihrer Zufälligkeit.

Obwohl zufällige Abweichungen niemals exakt vorausgesagt werden können, ist es möglich, sie zu beschreiben, zu analysieren und ihr erwartetes Verhalten durch stochastische Modellierung vorherzusagen. In diesem Rahmen können komplexe zufällige Abweichungen mit wenigen Parametern erfaßt werden und die Modelle können durch optimale maximum-likelihood-Techniken angepaßt werden. (nach [16], S.42)

Stochastische Methoden basieren auf der Annahme, daß die Textur aus einem zweidimensionalen stochastischen Prozeß auf einem Gitter entstanden ist. Daher wird die Modellierung über Zufallsfeldmodelle vorgenommen, in denen das Bild durch eine multivariate Wahrscheinlichkeitsverteilung dargestellt wird. Der Grauwert eines einzelnen Pixels wird als Zufallsvariable behandelt und es werden die Verteilungsfunktionen parametrisiert. Anhand dieser Parameter, die Mittelwert, Varianz und Korrelationen erfassen, können Texturen beschrieben und klassifiziert sowie von anderen Texturen unterschieden werden, so daß eine Segmentierung des Bildes möglich ist. (nach [16], S.43; [17])

Ein Beispiel hierfür ist das autoregressive (AR) Signalmodell, bei dem man sich das Signal aus einem weißen Rauschprozeß und einer Linearkombination vergangener Signalwerte additiv zusammengesetzt denkt. [...] Das AR-Modell kann auch als Linearprädiktion des Grauwertes eines Bildpunktes aus seinen Nachbargrauwerten verstanden werden, wobei der Rauschprozeß den Prädiktionsfehler charakterisiert. ([13], S.14)

Es wurden verschiedene AR-Modelle entwickelt und ihre Eigenschaften analysiert, sie werden eingeteilt in kausale AR-, semikausale AR-, nicht kausale AR-, SAR- (simultaneous autoregressive) und bedingte Markov- Modelle, um einige wichtige zu nennen (nach [17], [13]).

Neben dem zweidimensionalen AR-Prozeß ist das Nachbarschafts-Gibbs-Feld ein wichtiges Modell. Pan untersucht diese Modelle und entwickelt Parameterschätzungsmethoden und Kriterien für die Texturklassifizierung und -segmentierung [17].

1.2.1.5 Fraktale Methoden

Seit der Einführung der fraktalen Geometrie steht eine mathematische Methode zur Verfügung, die es ermöglicht, für komplizierte Funktionen und Formen im euklidischen Raum einfache Bildungsgesetze zu finden. Dabei handelt es sich z.B. um Linien im zweidimensionalen Raum, die extrem unstetig sind und eine

über alle Grenzen wachsende Länge haben, so daß sie sich damit der euklidischen Beschreibung entziehen. Als grundlegendes Werk auf diesem Gebiet gilt das Werk von Mandelbrot [18]. Es hat sich gezeigt, daß sich mit Fraktalen natürliche Formen auf einfache Bildungsgesetze zurückführen lassen. So kann man für sehr unregelmäßige Gebilde wie Küstenlinien, Wolken, Gebirge, Pflanzenaufbau und Anderes Parameter ableiten, die diese Objekte charakterisieren. Mit Einführung der „fraktalen Dimension“, die alle reellen Werte auch zwischen $n = 1, 2, \dots$ annehmen kann, steht ein Parameter zur Verfügung. So kann eine fraktale Linie im zweidimensionalen Raum eine fraktale Dimension von 1,6 aufweisen.

Eine grundlegende Eigenschaft ist die Selbstähnlichkeit der Formen. Die Selbstähnlichkeit beschreibt die Ähnlichkeit der Figur mit sich selbst bei verschiedenen aufgelöster Betrachtung. Als Beispiel nehme man eine Gebirgsform, die in den Kleinformen (Mikro-Bereich) Ähnlichkeit zur Gesamtform (Makro-Bereich) des Gebirges aufweist. Die fraktale oder Hausdorff-Besicovitch-Dimension wird in der Texturanalyse häufig verwendet und durch die Ähnlichkeitsdimension geschätzt. Für selbstähnliche Figuren, Figuren die sich aus N_r nicht überlappenden und mit dem Faktor r skalierten Kopien ihrer selbst mittels einer Ähnlichkeitsabbildung zusammensetzen lassen, ist die Ähnlichkeitsdimension dann definiert durch

$$N_r r^D = 1, \quad D = \frac{\lg(N_r)}{\lg(\frac{1}{r})}$$

Sie genügt immer der Gleichung $0 \leq D \leq E$, worin E die euklidische Dimension ist [13]. Für weitergehende Informationen sei auf die Literatur verwiesen ([19], [20], [21]).

Inzwischen ist es nicht nur möglich, solche Fraktale zu erzeugen, sondern für gegebene Strukturen fraktale Beschreibungen zu finden. Dazu werden sogenannte iterierte Funktionensysteme (*iterated function systems*, IFS) benutzt, mit denen sich komplexe Muster in einem enorm kompakten Code unterbringen lassen. Es werden Kompressionsraten in der Größenordnung bis zu 1 : 10 000 erreicht. Damit ist diese Darstellung für die Texturanalyse und -diskriminierung bestens geeignet [14] [13]. Problematisch ist allerdings die Rechenzeit, die hunderte Stunden pro Bildberechnung erfordert. Mit steigender Rechenleistung ist es sicherlich ein Ansatz, der in der Zukunft relevante Ergebnisse liefert.

1.2.1.6 Neuronale Methoden

Ein Ansatz aus dem Bereich der künstlichen Intelligenz basiert auf neuronalen Netzen. Dabei muß man einschränkend sagen, daß hierunter meistens künstliche neuronale Netze verstanden werden (ANN, *artificial neural networks*). Sie unterscheiden sich von biologischen neuronalen Netzen, deren Aufbau weitaus komplexer ist als die mathematischen Modelle. Hierbei handelt es sich um ein Netzwerk aus vielen einfachen Prozessoren (*units*), die untereinander durch Kommunikationskanäle verbunden sind (*connections*). Diese Einheiten bearbeiten nur ihre lokalen Daten, die sie durch die Verbindungen erhalten. Einige neuronale Netze sind Modelle von biologischen neuronalen Netzen und wurden entwickelt, um künstliche Systeme mit „intelligenten“ Rechenregeln zu schaffen, die denen des Gehirns ähneln und deshalb auch Einblicke in die Vorgänge dort geben können. Haykin gibt die folgende Definition ([22], S. 2):

Ein neuronales Netzwerk ist ein stark parallelverteilter Prozessor, der eine natürliche Neigung zur Speicherung experimentellen Wissens hat und es zur Nutzung bereitstellt. Es gleicht dem Hirn in zwei Aspekten:

1. Wissen wird vom System durch einen Lernprozeß erworben.
2. Interneuronale Verbindungsstärken, bekannt als synaptische Gewichte, werden zur Wissensspeicherung genutzt.

Neuronale Netze sind sinnvoll für Probleme der Klassifikation und der Funktionsnäherung einzusetzen, die eine gewisse Unsicherheit tolerieren und eine große Menge Trainingsdaten zur Verfügung haben, auf die klare Regeln nicht einfach angewendet werden können. Es wurden und werden viele neuronale Netze entwickelt, die sich grob in zwei Kategorien einteilen lassen: Neuronale Netze mit überwachtem und unüberwachtem Lernprozeß. Bei überwachtem Lernen gibt es einen „Lehrer“, der in der Lernphase dem Netz „sagt“, wie gut die Leistung ist (*reinforcement learning*) oder wie das richtige Verhalten wäre (*fully supervised learning*). Beim unüberwachten Lernprozeß ist das System autonom: es hat nur die Daten, die ihm eingegeben werden, findet etwas über die Eigenschaften der

Daten heraus und lernt diese im Ergebnis darzustellen. Um welche Eigenschaften es sich dabei handelt, hängt in erster Linie vom verwendeten Netzwerkmodell und der Lernmethode ab⁵.

1.3 Datengrundlage

Als Datengrundlage stand ein Teil einer Szene des Kamerasystems DPA (*Digital Photogrammetry Assembly*) zur Verfügung. Das System ist ein flugzeuggestütztes digitales Aufnahmesystem, das für die Herstellung topographischer Karten im Maßstab 1:25000 und 1:50000 entwickelt wurde. Das System erlaubt die gleichzeitige Aufnahme hochauflösender panchromatischer Stereobilder sowie multispektraler Daten. Das System arbeitet nach dem 3-Zeilen-Prinzip, mit dem simultan Vor-, Rückwärts- und Senkrechtaufnahmen gemacht werden. Aus den nach vorne und hinten verschwenkten Daten (Stereoaufnahmeprinzip) können DTMs (*Digital Terrain Models*) mit Genauigkeiten von unter 3m abgeleitet werden. In dieser Arbeit wurden die multispektralen Daten verwendet. Der Sensor FAIRCHILD CCD 191 für die multispektralen Daten besteht aus 4 CCD-Zeilen (2 pro Linie) mit einer Pixelgröße von $10\mu\text{m} \times 10\mu\text{m}$ und einer Optik mit der Brennweite von 40mm. Es werden vier Kanäle folgender Wellenlängenbereiche erzeugt:

- Blau: 440-525nm
- Grün: 520-600nm
- Rot: 610-685nm
- nahes Infrarot (IR): 770-890nm

Die radiometrische Auflösung liegt bei 12 Bit, die auf 8 Bit reduziert wird. Die geometrische Auflösung ist von der Flughöhe abhängig und beträgt in der vorliegenden Szene 0,77m. Panchromatische Daten wurden hier nicht verwendet, sie haben eine doppelt so hohe geometrische Auflösung und arbeiten mit separaten CCD-Zeilen (Daten aus [23]).

⁵nach: FAQ Neuronale Netze, <ftp://ftp.sas.com/pub/neural/FAQ.html> (26.7.1999)

Die multispektralen Daten lagen geometrisch entzerrt und georektifiziert vor. Um einen Kanal für die Texturanalyse aus den vier multispektralen Kanälen zu erzeugen, wurde eine PCA (*Principal Components Analysis*, Hauptkomponentenanalyse) durchgeführt. Durch die PCA wird ein neuer Kanal erzeugt, die Redundanz zwischen den 4 multispektralen Kanälen wird reduziert.

Der vorliegende Ausschnitt enthält ein Gebiet in den Gemeinden Rosdorf und Friedland (Landkreis Göttingen, Regierungsbezirk Braunschweig). Zusätzlich lagen die Daten einer Feldbegehung zum Befliegungszeitpunkt vor, die in den DGK5 Obernjesa (Nr. 15 der DGK25 Nr. 4525) und Klein Schneen (Nr. 21 der DGK25 Nr. 4525) verzeichnet sind. Sie sind im Anhang B zu finden.

In der Abbildung 1.1 ist die Falschfarbendarstellung⁶ aus der Kombination dreier Kanäle zu finden. IR, Rot und Grün werden in den Kanälen Rot, Grün und Blau dargestellt.

⁶Der Begriff ist weit verbreitet, obwohl es keine „falschen“ Farben gibt. Eine bessere Bezeichnung ist Farbinfrarotdarstellung (*CIR-Color Infra Red*), da sie der Aufnahme mit Farbinfrarotfilmen entspricht.



Abbildung 1.1: Die Luftbildszene in Infrarotfarbdarstellung

Kapitel 2

Multiresolutionsansätze

Für die Analyse von Texturen stehen also verschiedene Ansätze zur Verfügung. In diesem Kapitel geht es um Transformationen, mit denen die Bildinformation vom Ortsraum in den Frequenzraum transformiert wird. Im Frequenzraum lassen sich dann charakteristische Merkmale der Texturen ableiten oder Filterungen vornehmen.

Multiresolutionstechniken versuchen, Bilder durch Transformationen in Teile zu zerlegen, in denen sowohl Orts- als auch Frequenzinformationen vorliegen. Dazu gibt es Ansätze wie die Zerlegung in Bildpyramiden oder die Texturenergieuntersuchungen nach Laws, die schon früh entwickelt wurden und daher zuerst beschrieben werden sollen. In den letzten Jahren wurde die Wavelettheorie entwickelt, die einen mathematischen Rahmen bildet und eine formal solide und einheitliche Herangehensweise an Multiresolutionsansätze gestattet. Sie bildet sozusagen den Überbau und gibt die Möglichkeit, eine an ein Problem angepasste Transformation zu entwerfen, ihre Eigenschaften zu beschreiben und vor allem einzuschätzen. In dieser Arbeit sollen die grundlegenden Aspekte dargestellt und ein Überblick über die in der Texturanalyse häufiger verwendeten Transformationen gegeben werden. Da die Forschungen auf diesem Gebiet mit enormer Geschwindigkeit vorangetrieben und erweitert werden, besteht kein Anspruch auf Vollständigkeit.

Für eine eingehende Einführung sei auf die zahlreich vorhandene Literatur

verwiesen. Die Waveletanalyse wird in sehr unterschiedlichen Disziplinen eingesetzt, generell gehört sie auf das Gebiet der Signalanalyse.

Multiresolutionsansätze stellen das Bild in verschiedenen Auflösungsstufen dar. Dabei zeigt jede Stufe andere Frequenzinformationen. Ziel der Analyse ist es, diese Frequenzinformationen als Merkmalsraum zu nutzen und damit eine über die räumliche Lage hinausgehende Information zu erhalten. Zum Verständnis des Verhältnisses Ortsraum - Frequenzraum wird in dieser Arbeit die Fouriertransformation in Kapitel 2.1 behandelt.

Motivationen für den Einsatz von mehrfachauflösenden Verfahren ergeben sich aus der Biologie: so belegen psychovisuelle Studien, daß das menschliche visuelle System eine Mehrfachauflösungsanalyse durchführt. Dieser biologische Sachverhalt kann durch mathematische Methoden nachvollzogen werden. Dabei wird das Bild als Signal betrachtet, das analysiert wird. Das Signal wird von Rezeptoren auf der Netzhaut (Zäpfchen) abgetastet und über Nerven an das Gehirn weitergeleitet. Die Auswertung erfolgt in Rezeptivfeldern, die aus Neuronen aufgebaut sind. Die Rezeptivfelder sprechen Untersuchungen zufolge auf bestimmte Eigenschaften des Bildes an, sie entsprechen den Filtern in der mathematischen Analyse. Diese Felder sind in der Lage, Frequenzen des Bildes mit bestimmter Orientierung zu orten. Hierbei liegt der Vergleich zu den Wavelets nahe, die Filter mit dieser speziellen Eigenschaft sind. Die neuronale Antwort, also die Reaktion der einzelnen Neuronen, entspricht dann der Filterantwort einzelner Filter (Wavelets). Dem Sehen entspricht eine kombinierte Orts-Frequenzanalyse. Dabei steht nicht die Komprimierung des Bildes im Vordergrund, sondern die Mustereerkennung. Deshalb werden für die Mustereerkennung und Texturanalyse mehrfachauflösende Verfahren konstruiert. Die Ergebnisse sind vielversprechend, denn auf diese Weise lassen sich Texturen analysieren und Bilder nach Texturgehalt segmentieren. Für die Texturanalyse bedeutet es, daß sich dominante Größen und Orientierungen von einzelnen Texturen unterscheiden lassen. In den bei der Filterung entstehenden Bildern werden meist einfache Statistiken der Grauwerte als Texturmerkmale benutzt, da sie Aussagen über den Energiegehalt zulassen (s. Kapitel 2.4 und 3.1.1).

Im weiteren Verlauf dieser Arbeit geht es um Methoden, die eine Transfor-

mation des Bildes vom Ortsraum in den Orts-Frequenzraum vornehmen. Die Grauwerte des Bildes $S(x, y)$ sind in der Form, in der das Bild vorliegt, ortsabhängig. x sowie y geben die Position innerhalb des Bildes an, an denen der Grauwert S auftritt. Es gibt nun Transformationen, die in den Frequenzraum transformieren (z.B. Fouriertransformation Kapitel 2.1), so daß die Darstellung nicht orts-, sondern frequenzabhängig ist. Das Bild wird dabei als Signal aufgefaßt, dessen spektrale Komponenten untersucht werden. In der Signalverarbeitung wird oft mit Zeitreihen gearbeitet, daher findet sich in der Literatur meist die Abhängigkeit des Signales von der Zeit t . Stellt man sich das Bild als zeitlich abhängig aufgezeichnet (wie das z.B. bei Scannerdaten der Fall ist) vor, so ergibt sich eine anschauliche Parallelität von der Ortsabhängigkeit der Bilder zu der Zeitabhängigkeit der Zeitreihen auf der einen Seite der Transformation. Mit diesen Transformationen ist es also möglich, Aussagen über auftretende Frequenzen (oder auch Wellenzahlen) im Bild zu machen. Die Grauwertschwankungen im Bild lassen sich dann im Frequenzraum analysieren. Häufig nutzt man den Frequenzraum, um dort Filterungen vorzunehmen. So läßt sich Rauschen, das hauptsächlich mit hoher Frequenz auftritt, lokalisieren und eliminieren. Werden diese Rauschfrequenzen zu Null gesetzt und mit den verbleibenden Parametern zurück in den Ortsraum transformiert, so ist das Ergebnis ein rauschfreies Bild.

Die Fouriertransformation wird als grundlegende Transformation zunächst behandelt.

2.1 Die Fouriertransformation

Die Fouriertransformation transformiert ein Signal vom Orts- (oder Zeit-) in den Frequenzraum. Man erhält eine Darstellung im Frequenzraum, in der zu allen Frequenzen die Amplituden und Phasen, daraus ableitbar die Energien, aufgetragen werden. Dabei geht die Ortsinformation verloren. Es kann zwar festgestellt werden, daß die Frequenzen auftreten, aber nicht mehr, an welchem Ort.

2.1.1 Fourierreihendarstellung

Der Fouriersche Satz besagt, daß jede periodische Funktion als Summe von Kosinus- und Sinustermen darstellbar ist. Eine Darstellung hierfür findet sich in der Fourierreihe, diese Entwicklung einer Funktion nennt man harmonische Analyse:

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} a_k \cos(kx) + b_k \sin(kx) = \sum_{k=-\infty}^{+\infty} c_k \exp(ikx). \quad (2.1)$$

Die a_k, b_k, c_k sind dann die Fourierkoeffizienten, die c_k sind die Koeffizienten in der Eulerschen Darstellung¹.

Allerdings ist die Entwicklung in eine Fourierreihe mit einigen Nachteilen verbunden:

- Es kann Punkte geben, z.B. an Unstetigkeitsstellen, an denen die Reihe nicht konvergiert.
- Die Reihe kann Signale, die nicht auf ein Zeitmaß synchronisiert sind, wie das bei Sprache oder Bildern der Fall ist, nicht effizient darstellen.
- Ein stetiges periodisches Signal mit einer von 2π verschiedenen Periode wird als nichtstetiges Signal über $[-\pi, +\pi]$ dargestellt. Daher konvergiert die Reihe schlecht.

Eine Lösung bietet die Fouriertransformation, in der die Frequenzen kontinuierlich variiert werden.

¹Eulersche Darstellung einer komplexen Zahl:

$$\exp(i\phi) = \cos(\phi) + i \sin(\phi)$$

2.1.2 Fouriertransformation

Die Fouriertransformation ist für den eindimensionalen Fall wie folgt definiert:

$$\hat{f}(k) = \int_{-\infty}^{+\infty} dx \quad f(x) \exp(-i k x) \quad (2.2)$$

$$\text{mit } k = \frac{2\pi}{\lambda} = \text{Wellenzahl}$$

Die dazugehörige inverse Fouriertransformation, die $\hat{f}(k)$ als Summe von Sinusfunktionen darstellt, lautet:

$$f(x) = \int_{-\infty}^{+\infty} dk \quad \hat{f}(k) \exp(ikx). \quad (2.3)$$

Die zweidimensionale Fouriertransformation ist analog folgendermaßen definiert und läßt sich aufgrund ihrer Eigenschaften auf den eindimensionalen Fall zurückführen.

$$\hat{f}(\vec{k}) = \frac{1}{(2\pi)^2} \int_{-\infty}^{+\infty} d^2 x \quad \exp(-i\vec{k}\vec{x}) f(\vec{x}) \quad (2.4)$$

$$\text{mit } |\vec{k}| = \frac{2\pi}{\lambda}.$$

Um die Berechnung zu vereinfachen, kann das Skalarprodukt in zwei Komponenten aufgeteilt werden:

$$-i\vec{k}\vec{x} = -i(k_1 x_1 + k_2 x_2), \quad (2.5)$$

und der Kern kann als Produkt geschrieben werden:

$$\exp(-i\vec{k}\vec{x}) = \exp(-ik_1 x_1) \exp(-ik_2 x_2). \quad (2.6)$$

Damit ist der Kern separabel in x_1 - und x_2 - Richtung, und die Transformation kann durch zwei aufeinanderfolgende eindimensionale Transformationen berechnet werden (nach [14], S. 289 ff.). Für separable Funktionen ist der Rechenaufwand

geringer. Die Funktionen $\hat{f}(\vec{k})$ sind in der Frequenz (Wellenlänge) gut bestimmt, da es sich um periodische Funktionen von unendlicher Länge handelt.

In Abbildung 2.1 sind die Fouriertransformierten zweier Bilder dargestellt. Im Frequenzraum sind die Frequenzen im komplexen Koordinatensystem aufgetragen. Polare Koordinaten geben in diesem Koordinatensystem die Orientierung als Winkel zur ω_x -Achse und die Frequenz als Abstand vom Ursprung an. Man erkennt, daß die Orientierung des Feldes sich im Fourierraum ausdrückt. Es liegt keine Information vor, an welcher Stelle die Frequenzen im Ortsraum zu finden sind. Typisch ist der hohe Energiegehalt der niedrigen Frequenzen.

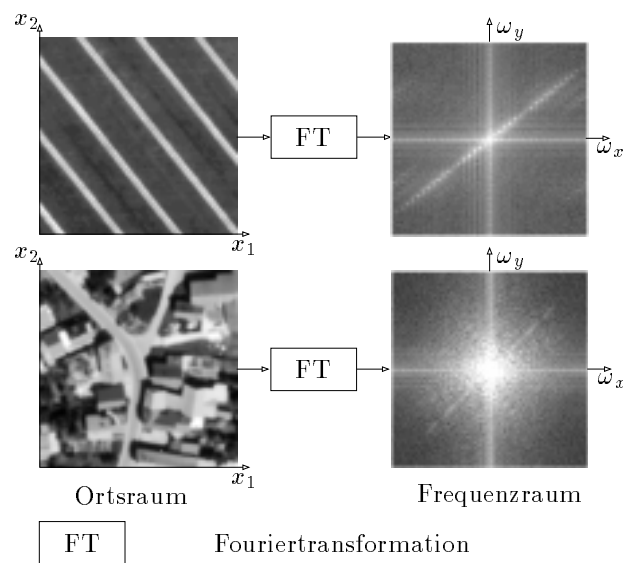


Abbildung 2.1: Fouriertransformation von Bildern

Um ein Signal orts- und frequenzlokalisiert darzustellen, braucht man Analysefunktionen, die sowohl im Ort als auch in der Frequenz beschränkt sind. Sie müssen in Orts- und Fourierraum beschränkte Träger besitzen.

Die Orts-Frequenz-Auflösung wird durch zwei Sachverhalte limitiert, durch die Heisenbergsche Unschärferelation und die Kompaktheit der Träger.

2.1.2.1 Heisenbergsche Unschärferelation

Wenn f aus L^2 , dann sind die Standardabweichung im Ortsraum σ_x und die Standardabweichung im Frequenzraum σ_ω definiert, und es gilt:

$$\sigma_x \sigma_\omega \geq \frac{1}{4} \quad (2.7)$$

Im Fourierraum ist σ_x unendlich und σ_ω null. Im Ortsraum ist σ_x null und σ_ω unendlich.

2.1.2.2 Kompakter Träger

Eine Funktion besitzt einen kompakten Träger, wenn ihre Ausdehnung begrenzt ist, d.h. ab einer unteren Grenze bis $-\infty$ und ab einer oberen Grenze bis $+\infty$ gleich 0 ist. Für diskretisierte Funktionen gilt, daß sie eine endliche Anzahl von Werten aufweisen, die nicht Null sind.

Ist eine Funktion mit kompaktem Träger nicht Null, dann kann ihre Fouriertransformierte auf einem Intervall nicht Null sein. Dasselbe gilt umgekehrt. Deshalb kann es keine Funktion in L^2 geben, die im Zeit- und Frequenzraum kompakt ist. Das bedeutet, daß es keine Frequenzanalyse für einen bestimmten Ort geben kann. Die Orts-Frequenz-Lokalisation kann nur im Sinne kleinster Quadrate erreicht werden. Die σ in Gleichung 2.1.2.1 sind Parameter einer Verteilung nach kleinsten Quadraten. Die Darstellung der Orts-Frequenzauflösung geschieht mit Hilfe von Heisenbergkästchen in Kapitel 2.5.

2.2 Short Time Fouriertransformation

Das Problem der Fouriertransformation ist, daß nur die Frequenzanteile eines Signales erhalten werden. Eine Ortslokalisierung der spektralen Komponenten ist nicht mehr möglich. Um sie einzubeziehen, braucht man eine Transformation, die eine Orts-Frequenz-Darstellung des Signales liefert. Dazu gibt es verschiedene Transformationen (Wigner-Distributionen, STFT - *Short Time Fourier Transformation*, Wavelettransformation), von denen hier die STFT vorgestellt wird.

Danach werden einige Aspekte der Orts-Frequenz-Darstellung geklärt (Kapitel 2.5), um dann zur Definition und zu den Eigenschaften der Wavelettransformation in Kapitel 2.6 überzugehen.

Die STFT wurde aus der FT entwickelt, um eine Orts-Frequenz-Lokalisation zu erreichen. Dazu wird das Signal durch Multiplikation mit einer Fensterfunktion in verschiedene ortslokalisierte Teile zerlegt, die dann jeweils einer Fouriertransformation unterworfen werden. Von diesen Teilen wird angenommen, daß sie stationär² sind. Damit hat man also die Möglichkeit, für Ausschnitte des Signales den Frequenzgehalt zu bestimmen. Nach der Heisenbergschen Unschärferelation nimmt die Frequenzauflösung mit höherer Ortsauflösung ab, was der Wahl eines schmaleren Fensters entspricht, so daß der Wahl der Fensterfunktion im Vorfeld besondere Bedeutung zukommt. Die Wahl sollte anhand von a-priori-Wissen dem zu untersuchenden Signal angepaßt werden. Als grundlegende Beziehung gilt: Ein schmales Fenster liefert gute Orts-, aber schlechte Frequenzauflösung, ein weites Fenster eine gute Frequenz-, aber schlechte Ortsauflösung.

Es bleibt das Problem, daß die Stationaritätsbedingung mit größer gewähltem Fenster eventuell nicht eingehalten wird, und daß die Fensterbreite für den gesamten Prozeß gleich bleibt [24].

2.3 Der Burt-Adelsonsche Pyramidenalgorithmus

Zur Mustererkennung in der Bildverarbeitung wurde die Multiskalenanalyse von Burt und Adelson vorgeschlagen. Sie prägten die Begriffe der „Gaußschen“ und „Laplaceschen“ Bildpyramide. Zerlegungen eines Bildes sind die Grundlagen der Multiresolutionsanalyse, die in den folgenden Kapiteln ausführlich behandelt werden.

²Stationäre Signale sind periodisch und ändern ihre Frequenzanteile nicht über den Ort.

2.3.1 Gaußpyramide

Die Gaußpyramide zerlegt ein Bild in verschiedene Glättungsstufen, ohne den Berechnungsaufwand exponentiell ansteigen zu lassen. Hier wird das Abtasttheorem nach Nyquist konsequent ausgenutzt: filtert man die Hälfte des Frequenzspektrums eines Bildes mit einem Tiefpaß heraus, so läßt sich das entstehende Bild ohne Verlust an Information mit der Hälfte der Stützstellen (Bildpunkte) darstellen. Das Abtasttheorem stellt hierbei sicher, daß das Bild mit höherer Auflösung aus dem mit niedrigerer Auflösung wiederhergestellt werden kann (s.a. Gleichung 2.12, S. 41). Der von Burt und Adelson vorgeschlagene Algorithmus verwendet eine symmetrische, separable 5×5 - Glättungsmaske. Dazu eignen sich Binominalfilter, die eine Gaußfunktion annähern und aus den Binominalkoeffizienten gewonnen werden können (z.B. $\frac{1}{16}[1 \ 4 \ 6 \ 4 \ 1]$)³. Die Gaußfunktion hat optimale kombinierte Auflösung in Orts- und Frequenzraum. Das Bild wird durch die Filterung geglättet und die hohen Frequenzen werden unterdrückt. Es wird ein Resampling durchgeführt, das die Stützstellen auf die Hälfte reduziert. Das zweidimensionale Bild reduziert seine Größe auf ein Viertel, da der Filterkernel in beide Achsrichtungen angewendet wird. Das wird bei separierbaren Filtern direkt sichtbar. Damit entsteht eine neue Stufe, die nur die Hälfte der Frequenzinformation enthält und nur ein Viertel von der Größe der Ausgangsstufe besitzt. Auf die so entstandene Stufe kann der Algorithmus iterativ angewendet werden. Er erzeugt nacheinander die verschiedenen Auflösungsstufen. Der Name „Pyramide“ folgt aus der Tatsache, daß die einzelnen, immer kleiner werdenden Stufen übereinandergeschichtet die Form einer Pyramide ergeben (nach [14], S. 142 ff.).

³Jähne merkt an, daß die Größe 5×5 zu klein ist, um die Reduktionsoperation fehlerfrei durchführen zu können ([14], S. 164). Außerdem kann die Einführung eines Bias den Einfluß der vernachlässigten Werte der Gaußfunktion außerhalb der Maske korrigieren.

2.3.2 Laplacepyramiden

2.3.2.1 DoG-Pyramide

Subtrahiert man je zwei aufeinanderfolgende Stufen der Gaußpyramide, wobei ein Resampling nötig ist, so erhält man bandpaßgefilterte Ergebnisbilder. Ein Bandpaßfilter kann durch die Differenzbildung der Filterantworten von zwei Glättungsmasken mit unterschiedlicher Standardabweichung erzeugt werden. Wählt man hierfür die Gaußschen Filtermasken, so spricht man von einer DoG-Operation (englisch *Difference of Gaussian*). Der DoG –Operator stellt eine Näherung für die LoG–Funktion (*Laplacian of Gaussian*) dar, die nachfolgend beschrieben wird. Oft wird der Name „Laplace-Pyramide“ für die DoG-Pyramide verwendet, obwohl sie eigentlich ihren Namen dem im folgenden beschriebenen LoG–Operator verdankt (nach [25], S. 51 ff.).

2.3.2.2 LoG-Pyramide

Eine fortgesetzte Bandpaßfilterung wird erreicht, indem man eine Bildglättung mittels eines Gaußfilters unterschiedlicher Varianz vornimmt und anschließend die zweite Ableitung bildet. Mit dem Laplaceoperator benötigt man zur Bildung der zweiten Ableitung nur eine einzige Bildfilterung pro Frequenzbereich. Statt einer sequentiellen Berechnung der Gauß- und der Laplacefunktion kann eine LoG-Filtermaske verwendet werden, die man auch Mexikanerhut nennt. Dabei handelt es sich um eine zweidimensionale Laplacefunktion einer Gaußkurve (nach [25], S. 46 ff.).

Der Vorteil der Pyramidendarstellung ist, daß in den verschiedenen Stufen, die ja nur einen begrenzten Frequenzbereich beinhalten, Bildoperatoren gezielt angewendet werden können. Die Größe der Kernel bleibt hierbei klein, ihre Ausdehnung im Originalbild wäre das 2^n -fache (mit $n = \text{Stufe}$).

Für die Texturanalyse bedeutet diese Vorgehensweise, daß sich die Frequenzanteile einer Textur in den verschiedenen Stufen ausprägen und daß sie zur Unterscheidung verschiedener Texturen herangezogen werden können. Die Untersu-

chungen mit Gauß- und Laplacepyramiden werden in Kapitel 4.3 beschrieben.

2.4 Texturenergie nach Laws

Die von Laws zur Texturanalyse vorgeschlagene Methode gehört nach Vollmerhaus zu den räumlich-statistischen Verfahren, bei denen lokale Bildeigenschaften durch die statistische Verteilung ihrer Grauwerte beschrieben werden ([25], S. 69 ff.). Dazu werden verschiedene Faltungsmasken vorgeschlagen, die lokale Merkmale wie Spots, kleine Wellen, Kanten und Ecken verschiedener Ausdehnung und Orientierung sowie Regionen bestimmter Form extrahieren. Die Filterantwort ist groß, wenn die untersuchten lokalen Grauwerte die gleiche Verteilung wie die der Maske aufweisen. Sind nun also Kanten, Spots oder andere lokale Merkmale, die extrahiert werden sollen, in der Faltungsmaske modelliert, so ist diese in der Lage, entsprechende Elemente im Bild zu lokalisieren, indem die Antwort als Ergebnis der Konvolution dort maximal wird.

Die Faltungsmasken sind nicht sehr groß, es werden Größen von 3x3 bis 7x7 Pixel vorgeschlagen. Sie werden Mikrofenster genannt. Nach der Konvolution dieser Faltungsmasken mit dem Bild ergibt sich das Ergebnisbild, das mit einem Makrofenster (vorgeschlagene Größe: 15x15 Pixel) auf die Häufigkeit der lokalen Textur-eigenschaft hin untersucht wird. Diesen Schritt nennt Laws eine Texturenergie-transformation. Basierend auf der Verteilung der extrahierten Eigenschaften werden Texturenergie-Merkmale gebildet. Dazu schlägt er statistische Momente vor, wie sie in Kapitel 1.2.1.1 zu finden sind. Die Standardabweichung, errechnet durch das Quadrat der Mikrofilterergebnisse, wird berechtigterweise Texturenergie genannt (s. auch Energie eines Signales, Kapitel 3.1.1).

Das Verfahren benutzt also Masken in zwei verschiedenen Größen (Mikro- und Makrofenster), um den Bildinhalt zu untersuchen. Dabei ist das Ergebnis von den absoluten Größen dieser Masken und von dem Größenverhältnis der Masken untereinander abhängig. Wählt man feste Größen und Größenverhältnisse, so ist die Anpassung an verschiedene Texturen sicherlich nicht optimal. Es wäre also sinnvoll, die Maskengrößen an die tatsächlich auftretenden Texturen und an ihre

Ausdehnung anzupassen. Diesen Schritt kann man mit Bildpyramiden vollziehen. Dazu gibt es zwei gangbare Wege:

- Man erzeugt die Bildpyramide und berechnet die Konvolution der entstehenden Stufen mit den Mikrofiltern. Es lassen sich dann Texturenergiepyramiden zusammenstellen, die jeweils durch einen Mikrofilter entstanden sind.
- Man erzeugt die mikrogefilterten Bilder aus dem Ausgangsbild und anschließend die Bildpyramiden. Mit den Mikrofiltern läßt sich die Texturrauhigkeit erfassen, deshalb wird das Verfahren um den ersten Schritt der Generierung der Tiefpaßpyramide gekürzt. Man variiert dadurch lediglich die Größe des Makrofensters ([25], S. 80).

Die Mikrofilter teilt man nach ihrer Größe und ihrer Filtercharakteristik ein. Dazu werden ein Buchstabe für die Charakteristik und eine Zahl (meist ungerade, damit bessere Symmetrie entsteht) für die Größe zusammengestellt. Die Buchstaben sind abgekürzte Begriffe aus dem Englischen, L steht für Level, E für Edge, S für Spot, W für Wave, R für Ripple, O für Oscillation. Wie diese Filtermasken aussehen, zeigen folgende Beispiele:

$$\begin{aligned}
 \text{L7} &= [1 \ 6 \ 15 \ 20 \ 15 \ 6 \ 1] \\
 \text{E7} &= [-1 \ -4 \ -5 \ 0 \ 5 \ 4 \ 1] \\
 \text{S7} &= [-1 \ -2 \ 1 \ 4 \ 1 \ -2 \ -1] \\
 \text{W7} &= [-1 \ 0 \ 3 \ 0 \ -3 \ 0 \ 1] \\
 \text{R7} &= [1 \ -2 \ -1 \ 4 \ -1 \ -2 \ 1] \\
 \text{O7} &= [-1 \ 6 \ -15 \ 20 \ -15 \ 6 \ -1]
 \end{aligned}$$

Randen [26] weist auf die Frequenzselektivität dieser Filter hin, die in Abb. 2.2 für 5 der eindimensionalen Filter⁴ und in Abb. 2.3 für die zweidimensionale separable Kombination je zweier Filter aufgetragen ist. Die Untersuchungen nach diesem Ansatz sind in Kapitel 4.2 zu finden.

⁴Im einzelnen: L,E,W,R,O

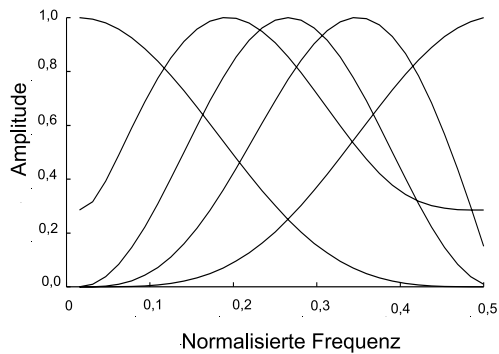


Abbildung 2.2: Frequenzselektivität der eindimensionalen Lawsfiltermasken

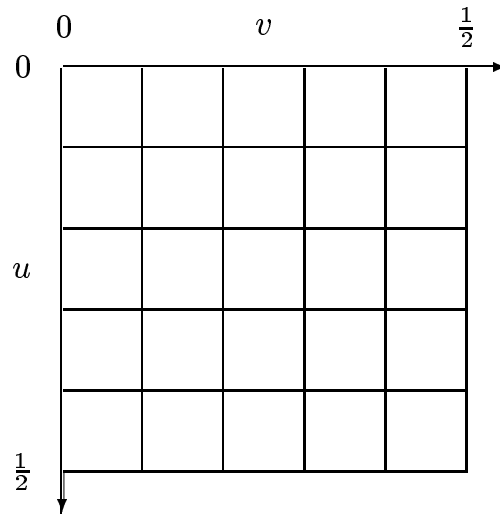


Abbildung 2.3: Resultierende zweidimensionale Aufteilung der Frequenzebene

2.5 Ortsfrequenzraum-Darstellungen

Die Darstellung von Transformationen in der Ortsfrequenzebene macht den Zusammenhang zwischen Orts- und Frequenzauflösung klar. Trägt man die Frequenz über dem Ort ab, so erhält man für jeden Koeffizienten der Transformation einen Bereich in dieser Ebene. Die entstehenden Kästchen werden Heisenbergkästchen genannt, und man kann an ihnen die kombinierte Orts- und Frequenzauflösung der für die Transformation verwendeten Basisfunktionen ablesen. Eine kleine Ausdehnung in Ortsrichtung bedeutet eine gute Ortsauflösung, eine kleine Ausdehnung in Frequenzrichtung entsprechend eine gute Frequenzauflösung. Nach Heisenberg (s. Kapitel 2.1.2.1) ist eine beliebige Steigerung beider Auflösungen nicht möglich, in der Darstellung drückt sich die kombinierte Auflösung als Fläche der Kästchen aus. Da die kombinierte Auflösung nur im Sinne kleinster Quadrate erfolgen kann, wird die Kantenlänge zu σ_x und σ_f festgelegt. Die Fläche ist nach unten hin durch die Ungleichung 2.7 begrenzt. Die Form der Kästchen variiert mit der verwendeten Transformation: so nehmen die Kästchen im Falle der Fouriertransformation in Ortsrichtung unendliche Länge an, entsprechend der Ortsauflösung (keine Ortsinformation), dafür sind sie in der Frequenz exakt lokalisiert

durch die Frequenz der zur Transformation verwendeten Sinusfunktionen. Im Falle der STFT (Kap. 2.2) ist die Form durch die Wahl des Fensters festgelegt und bestimmt die Breite in Ortsrichtung, und die Ebene wird von gleichgroßen und gleichgeformten Rechtecken überdeckt (Abb. 2.4). Bei der im folgenden vorgestellten Wavelettransformation ändert sich die Form entlang der Frequenzachse, wie in Abb. 2.5 dargestellt.

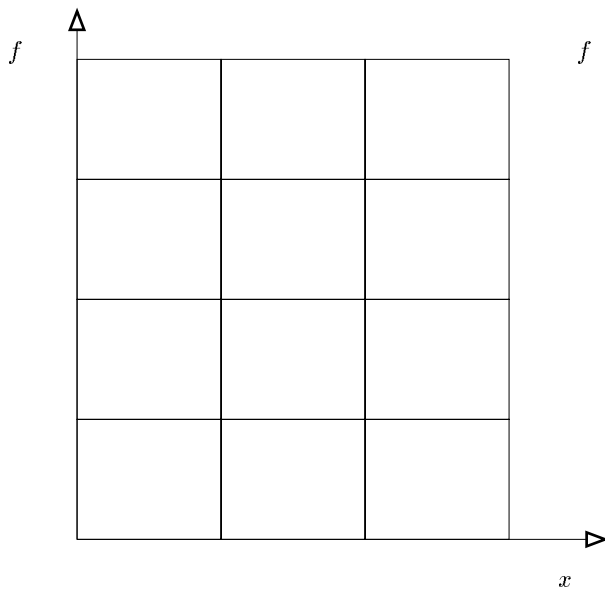


Abbildung 2.4: Zerlegung der Ortsfrequenzebene durch die STFT

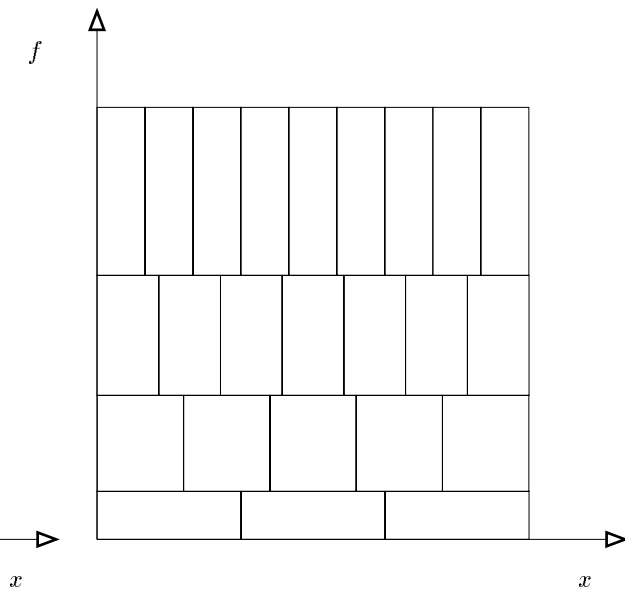


Abbildung 2.5: Ortsfrequenzebenenenteilung im Falle der Wavelettransformation

2.6 Die Wavelettransformation

Die Wavelettransformation schafft einen allgemeinen Zugang zu Ortsfrequenzuntersuchungen. Obwohl die Orts- und Frequenzauflösungsproblematik physikalische Ursachen hat (Heisenberg) und unabhängig von der Transformation existiert, kann ein Signal mit diesem alternativen Ansatz untersucht werden. Die Wavelettransformation analysiert das Signal bei verschiedenen Frequenzen in unterschiedlicher Auflösung. Bei diesem Konzept erhält man für hohe Frequenzen gute Zeitauflösung bei schlechter Frequenzauflösung und für niedrige Frequenzen

gute Frequenzauflösung bei schlechter Zeitauflösung. Da Signale in der Praxis tiefe Frequenzen über längere Zeiträume und höhere Frequenzen in kurzen Zeitabschnitten aufweisen, ist dieser Ansatz sinnvoll [24].

2.6.1 Kontinuierliche Wavelettransformation

Um ein Signal zu analysieren, wird bei der kontinuierlichen Wavelettransformation das Signal mit einer Familie von Funktionen multipliziert, die aus einer fest gewählten Funktion durch Translation und Dilatation hervorgehen. Dieser „Prototyp“ wird Mutterwavelet genannt, aus dem sich alle zur Analyse verwendeten Wavelets ableiten. Die Funktion wird im allgemeinen mit $\psi(t)$ bezeichnet und die Wavelets leiten sich folgendermaßen daraus ab:

$$\psi_{\tau}^s(t) = \psi\left(\frac{t - \tau}{s}\right) \quad (2.8)$$

Die Variable τ ist die Translationsvariable, s steht für Skale. Der Begriff Wavelet bedeutet soviel wie „kleine Welle“, resultierend aus der endlichen Länge der Funktion. Damit hat auch das Fenster endliche Länge. Die Besonderheit ist, daß Wavelets mit der Skalierung s sowohl die Ausdehnung des Fensters als auch ihre Frequenz ändern, weil die Anzahl der Oszillationen gleich bleibt. Hier ist der Unterschied zur STFT zu sehen, bei der die Fenstergröße fest bleibt und die Frequenz gewählt wird. Der Skalierungsparameter ist reziprok zur Frequenz definiert. Kleine Skalen erlauben eine globale Sicht auf das Signal (tiefe Frequenzen), große Skalen liefern dagegen eine Sicht auf die Details (hohe Frequenzen) des Signals.

Mit kontinuierlicher Veränderung der Skale und des Translationparameter erhält man die kontinuierliche Wavelettransformation als Skalarprodukte der Wavelets mit dem Signal. Die Berechnung eines Skalarproduktes (auch: inneres Produkt) von zwei Funktionen ist definiert durch:

$$\langle f(t), g(t) \rangle = \int_{-\infty}^{+\infty} g(t)f(t)dt. \quad (2.9)$$

Die kontinuierliche Wavelettransformation (CWT - *Continuous Wavelet Trans-*

form) ist definiert durch:

$$CWT_x^\psi(\tau, s) = \frac{1}{\sqrt{|s|}} \int x(t) \psi\left(\frac{t - \tau}{s}\right) dt. \quad (2.10)$$

Bei der kontinuierlichen Wavelettransformation wird das Signal also mit den durch Skalierung entstehenden Wavelets untersucht, wobei jedes Wavelet mit der Translation über das gesamte Signal verschoben wird. Das Skalarprodukt entspricht dabei einer Korrelationsrechnung, d.h. die Größe des Ergebnisses ist maximal an den Stellen des Signals, die wie das Wavelet geformt sind. Der Faktor $\frac{1}{\sqrt{|s|}}$ normalisiert das Ergebnis des Skalarproduktes, so daß unter verschiedenen Skalen Energiegleichheit herrscht.

Die kontinuierliche Wavelettransformation liefert ein unendlich redundantes Ergebnis mit unendlich vielen Koeffizienten, daher ist das Signal überabgetastet. Kontinuierliche Darstellungen sind translationsinvariant, und da sich die Koeffizienten bei kleinen Verschiebungen nicht ändern, spielt es bei der kontinuierlichen Wavelettransformation keine Rolle, an welcher Stelle des Signals man mit der Kodierung beginnt ([27], S.55). Nicht redundante Dekompositionen erreicht man mit Orthogonaltransformationen, wie sie in Kapitel 2.7 beschrieben werden. In der Praxis müssen nicht alle Koeffizienten berechnet werden.

2.6.2 Die diskrete Wavelettransformation

Die diskrete Wavelettransformation (DWT) wird zur Berechnung der Wavelettransformation benutzt. Die Signale liegen in diskreter Form vor. Mit der diskreten Wavelettransformation wird ein Berechnungsschema gefunden, das den Rechenaufwand, der zur Berechnung einer nötig ist, reduziert. Dafür wird eine redundanzfreie Darstellung des Signales gesucht und die Formeln für diskrete Signale und Filter bereitgestellt. Es gibt eine Parallelität zu den Mehrfachauflösungsverfahren der Pyramidenzerlegungen und der in der Signalanalyse verwendeten Teilbandzerlegung (*subband coding*). Diese Techniken wurden parallel entwickelt und weisen Ähnlichkeit in der Verfahrensweise auf.

Die Formeln zur Herleitung der dyadischen DWT finden sich in Kapitel 2.7.3. Die Umsetzung dieser Formeln auf Reihen machen die DWT über Faltungen

berechenbar (siehe z.B. [24]).

2.7 Konstruktion einer Mehrfachauflösung

Ziel der Mehrfachauflösung ist es, das Bild in verschiedene Auflösungsstufen zu zerlegen, die jeweils bestimmte Frequenzinformationen enthalten. Dazu wird das Bild in zwei Anteile aufgespalten: Ein Tiefpaßfilter erzeugt ein Ergebnisbild, in dem die niedrigen Frequenzen enthalten sind, in einem zweiten Ergebnisbild werden die hochfrequenten Anteile gespeichert, die den Details im Bild entsprechen. Durch diese Operationen entsteht eine neue Stufe, auf die die Zerlegung wiederum angewendet wird, so daß eine Folge von Bildern entsteht, die das Originalbild approximieren. Die Eigenschaften des zerlegenden Operators bestimmen den Inhalt der entstehenden Stufe. Der Operator, der die neue Stufe erzeugt, ist ein Tiefpaßfilter und wird Skalierungsfiler genannt. Verschiedene Skalierungsfiler erzeugen unterschiedliche Mehrfachauflösungen, wie bei den Bildpyramiden in Kapitel 6.2 sichtbar wird. Es besteht ein enger Zusammenhang zwischen dem Skalierungsfiler und den Wavelets, mit denen die Dekomposition durchgeführt wird. Dieser Zusammenhang wird in Kapitel 2.7.2 geklärt.

Zuerst werden einige Grundlagen der Funktionentheorie erläutert, die für das weitere Verständnis nötig sind.

2.7.1 Vektorräume, Skalarprodukt, Basen und Orthogonalität

Allgemein können Funktionen als Punkte in einem unendlich-dimensionalen Raum definiert werden. Sie sind dann über ihre Funktionswerte in einem Funktionenraum lokalisiert. Jeder Punkt eines n -dimensionalen Raumes kann durch seine Projektion auf n Achsen eindeutig festgelegt werden. Aus der Analysis ist die Zusammenstellung der einzelnen Achsprojektionen zu einem Vektor bekannt. Jeder Punkt hat dann über seinen Ortsvektor eine eindeutige Zuordnung, wenn die n Achsen einen n -dimensionalen Raum aufspannen, oder anders ausgedrückt, wenn

die Basisfunktionen linear unabhängig sind. Eine Basisfunktion legt eine Achsrichtung fest, mehrere Basisfunktionen bilden eine Basis des Raumes. Eine Basis heißt orthogonal, wenn ihre Basisfunktionen paarweise senkrecht aufeinander stehen. Orthonormal wird sie genannt, wenn alle Basisfunktionen darüber hinaus die Länge 1 haben. In orthogonalen Basen lassen sich die Entwicklungskoeffizienten durch die Skalarprodukte mit je einer Basisfunktion berechnen. Die Information, die in einem Koeffizienten kodiert wird, wird nicht an anderer Stelle noch einmal kodiert, wie das bei nichtorthogonalen Basen der Fall ist.

Nimmt man obige Definitionen und betrachtet das Signal als Funktion, dann kann diese Funktion durch Skalarprodukte nach einer Basis zerlegt werden. Die Entwicklungskoeffizienten sind für verschiedene Basen unterschiedlich. Die Basisfunktionen werden auch Analysefunktionen genannt und können je nach Ausprägung orthogonal oder nichtorthogonal sein. In der Waveletanalyse sind die Wavelets die Analysefunktionen, nach denen zerlegt wird. Lassen sich nun orthogonale Wavelets finden, erhält man eine nicht redundante Zerlegung des Signals. Andere Beispiele für Basen sind im Falle der Fourieranalyse von $-\infty$ bis $+\infty$ oszillierende Sinus- und Kosinusfunktionen, bei der gefensterten Fourieranalyse (STFT) ist es das Produkt einer zeitlich begrenzten Welle mit trigonometrischen Funktionen.

2.7.2 Definition einer Mehrfachauflösung

Es kann eine allgemeine Definition von Mehrfachauflösungen gegeben werden: Existiert eine Folge von Auflösungen eines Bildes $(V_j)_{j \in \mathbb{Z}}$, dann sind die Details einer Stufe der Auflösung V_j definiert als die Differenz der Information zwischen der Approximation der Stufe V_j und der folgenden niedriger aufgelösten Stufe V_{j-1} . Mit einer Waveletdekomposition ist es möglich, diese Details, die die Differenz zweier aufeinander folgender Stufen sind, zu erhalten. Dazu muß das Bild mit einer orthogonalen Waveletbasis zerlegt werden. Die Berechnung einer Waveletdekomposition kann über einen Pyramidenalgorithmus, basierend auf Konvolutionen mit Quadraturfiltern, erfolgen. Im folgenden werden Mehrfachauflösungen betrachtet, die auf einem dyadischen Schema beruhen. Dyadisch bedeutet, daß

der Auflösungsfaktor, mit dem gearbeitet wird, zwei ist. Die Auflösung wird also von Stufe zu Stufe um den Faktor zwei reduziert. Diese Beschränkung erlaubt eine einfachere Berechnung.

Es existieren sechs Bedingungen, die eine Mehrfachauflösung erfüllen muß:

1. Der Raum V_{j+1} wird aus V_j durch Reskalierung mit dem Faktor 2 gewonnen:

$$\forall (j, k) \in \mathbb{Z}^2, \quad f(t) \in V_j \Leftrightarrow f(t - 2^j k) \in V. \quad (2.11)$$

2. Für alle j ist V_{j+1} ein Unterraum von V_j

$$\forall j \in \mathbb{Z}, \quad f(t) \in V_{j+1} \subset V_j. \quad (2.12)$$

3. V_j ist 2^j translationsinvariant

$$\forall j \in \mathbb{Z}, \quad f(t) \in V_j \Leftrightarrow f\left(\frac{t}{2}\right) \in V_{j+1}. \quad (2.13)$$

4. Die Schnittmenge der V_j ist 0 in L^2

$$\lim_{j \rightarrow +\infty} V_j = \bigcap_{j=-\infty}^{+\infty} V_j = \{0\}. \quad (2.14)$$

Interpretation: Bei Auflösung 0 ist das bleibende Energiesignal 0.

5. Die Vereinigung der V_j ist abgeschlossen in L^2

$$\lim_{j \rightarrow +\infty} V_j = L^2(\mathbb{R}).$$

6. Es existiert eine Funktion, so daß die ganzzahligen Translationen von Θ eine Riesz-Basis von V_Θ bilden.

Interpretation: Die Auflösung V_j wird von einer Basis erzeugt, die durch 2^j Translationen einer mit 2^j skalierten Funktion Θ entsteht. Eine Riesz-Basis bildet einen Rahmen⁵ unabhängiger Vektoren (nach [28]).

⁵Rahmen sind eine stabile, evtl. redundante Darstellung von Signalen. Sie sind eine Verallgemeinerung des Basiskonzeptes im linearen Raum, bestehend aus Vektoren, die jedes endliche Energiesignal über Skalarprodukte darstellen können.

2.7.3 Konstruktion einer Mehrfachauflösung nach Mallat

Die Konstruktion einer Mehrfachauflösung wird zunächst für den eindimensionalen Fall dargestellt, der sich danach auf zwei Dimensionen erweitern lässt, wie sie in der Bildverarbeitung existieren. Die folgende Herleitung findet sich in [29], [30].

2.7.3.1 Skalierungsfunktion

Bei Mallat [29] findet man ein Theorem, das eine Mehrfachauflösung mit der Skalierungsfunktion verbindet:

Theorem 2.7.1

Ist $(V_{2^j})_{j \in \mathbb{Z}}$ eine mehrfachauflösende Approximation von $L^2(\mathbb{R})$, dann existiert genau eine Funktion $\Theta(x) \in L^2(\mathbb{R})$, genannt Skalierungsfunktion in der Art, daß mit $\Theta_{2^j}(x) = 2^j \Theta(2^j x)$ für $j \in \mathbb{Z}$ (Dilatation von $\Theta(x)$ um 2^j) gesetzt,

$$\left(\sqrt{2^{-j}} \Theta_{2^j}(x - 2^{-j}n) \right)_{n \in \mathbb{Z}}$$

eine orthonormale Basis von V_{2^j} bildet.

Die orthogonale Projektion auf V_{2^j} kann nun berechnet werden als Dekomposition des Signales $f(x)$ auf die orthonormale Basis von Theorem 2.7.1:

$$\forall f(x) \in L^2(\mathbb{R}), \quad A_{2^j} f(x) = 2^{-j} \sum_{n=-\infty}^{+\infty} \langle f(u), \Theta_{2^j}(u - 2^{-j}n) \rangle \Theta_{2^j}(x - 2^{-j}n). \quad (2.15)$$

Damit ist $A_{2^j} f(x)$ die Approximation des Signales $f(x)$ in der Auflösungsstufe 2^j . Sie wird die diskrete Approximation von $f(x)$ genannt. $A_{2^j} f(x)$ ist durch ein inneres Produkt gekennzeichnet und wird im folgenden bezeichnet durch:

$$A_{2^j} f = \left(\langle f(u), \Theta_{2^j}(u - 2^{-j}n) \rangle \right)_{n \in \mathbb{Z}}. \quad (2.16)$$

Jedes innere Produkt kann auch als Faltungsprodukt interpretiert werden, das am Punkt $2^{-j}n$ berechnet wird:

$$A_{2^j} f = (f(u) * \Theta_{2^j}(-u)) (2^{-j}n)_{n \in \mathbb{Z}}. \quad (2.17)$$

Da $\Theta(x)$ ein Tiefpaßfilter ist, kann dieses diskrete Signal als Tiefpaßfilterung von $f(x)$ mit anschließendem Sampling mit Grad 2^j interpretiert werden. In der Approximation werden durch Entfernen der Details von $f(x)$, die kleiner als 2^{-j} sind, die hohen Frequenzen dieser Funktion unterdrückt.

Die Implementierung in der Praxis geschieht mit Hilfe von diskreten Filtern. Die Skalierungsfunktion $\Theta_{2^j}(x - 2^{-j}n)$ ist Teil von V_{2^j} , der wiederum Teil von $V_{2^{j+1}}$ ist. Die Berechnung kann folgendermaßen erfolgen:

$$\begin{aligned} \Theta_{2^j}(x - 2^{-j}n) = \\ 2^{-j-1} \sum_{k=-\infty}^{+\infty} \langle \Theta_{2^j}(u - 2^{-j}n), \Theta_{2^j}(u - 2^{-j-1}k) \rangle \cdot \Theta_{2^{j+1}}(x - 2^{-j-1}k). \end{aligned} \quad (2.18)$$

Es sei H der diskrete Filter mit der Impulsantwort $h(n)$

$$\forall n \in \mathbb{Z}, h(n) = \langle \Theta_{2^{-1}}(u), \Theta(u - n) \rangle. \quad (2.19)$$

und \tilde{H} der Spiegelfilter mit der Impulsantwort $\tilde{h}(n) = h(-n)$, so erhält man:

$$\langle f(u), \Theta_{2^j}(u - 2^{-j}n) \rangle = \sum_{k=-\infty}^{+\infty} \tilde{h}(2n - k) \langle f(u), \Theta_{2^{j+1}}(u - 2^{-j-1}k) \rangle. \quad (2.20)$$

Der Gleichung 2.20 läßt sich entnehmen, daß $A_{2^{j+1}}^d f$ über eine Faltung von $A_{2^j}^d f$ mit \tilde{H} und Erhaltung jeder zweiten Stützstelle des Ergebnisses berechnet werden kann.

Alle folgenden Approximationen können durch Wiederholung dieses Prozesses aus $A_{2^j}^d f$ gewonnen werden. Damit ist die Pyramidentransformation definiert. Die Behandlung der Ränder des Bildes wird meistens über eine symmetrische oder periodische Erweiterung des Bildes vorgenommen.

Die Skalierungsfunktion kann also aus dem diskreten Filter H erhalten werden. Zusätzlich wird noch eine Regularitätsbedingung für die Skalierungsfunktion gefordert: Die Funktion $\Theta(x)$ soll stetig differenzierbar sein und der asymptotische Abfall von $\Theta(x)$ und $\Theta'(x)$ soll

$$|\Theta(x)| = O(x^{-2}) \quad \text{und} \quad |\Theta'(x)| = O(x^{-2})$$

genügen.

Theorem 2.7.2

Es sei $H(\omega)$ die Fourierreihe

$$H(\omega) = \sum_{n=-\infty}^{+\infty} h(n)e^{-in\omega} \quad (2.21)$$

und $\Theta(x)$ eine Skalierungsfunktion, sowie H definiert nach 2.19. $H(\omega)$, erfüllt die zwei Bedingungen⁶:

$$|H(0)| = 1 \text{ und } h(n) = O(-n^{-2}) \text{ im Unendlichen.} \quad (2.22)$$

$$|H(\omega)|^2 + |H(\omega + \pi)|^2 = 1 \quad (2.23)$$

Werden diese beiden Bedingungen so erfüllt, daß

$$|H(\omega)| \neq 0 \quad \text{für } \omega \in [0, \pi/2], \quad (2.24)$$

dann ist die Funktion

$$\hat{\Theta}(\omega) = \prod_{p=1}^{+\infty} H(2^{-p}\omega) \quad (2.25)$$

die Fouriertransformierte einer Skalierungsfunktion.

Filter, die die Bedingung 2.23 erfüllen, werden konjugierte Filter genannt.

2.7.3.2 Das Detailsignal

Das Detailsignal der Stufe 2^j enthält die Differenz der Informationen zwischen zwei Auflösungsstufen 2^{j+1} und 2^j . Es kann gezeigt werden, daß das Detailsignal durch orthogonale Projektion des Signales auf das orthogonale Komplement von V_{2^j} in $V_{2^{j+1}}$ erhalten werden kann. Mit der Bezeichnung O_{2^j} gilt:

O_{2^j} ist orthogonal zu V_{2^j} ,

$$O_{2^j} \oplus V_{2^j} = V_{2^{j+1}}.$$

⁶siehe auch Kapitel 3.1

Um diese Projektion zu berechnen, benötigt man eine orthogonale Basis von O_{2^j} , die von Wavelets erzeugt werden kann. Die Fouriertransformierte der Funktion $\psi(x)$ ist gegeben durch:

$$\hat{\psi}(x) = G\left(\frac{\omega}{2}\right) \hat{\theta}\left(\frac{\omega}{2}\right) \quad (2.26)$$

mit $G(\omega) = e^{(-i\omega)} \overline{H(\omega + \pi)}$. Es sei $\psi_{2^j}(x) = 2^j \psi(2^j x)$ die Dilatation von $\psi(x)$ um 2^j . Dann bildet

$$\left(\sqrt{2^{-j}} \psi_{2^j}(x) (x - 2^{-j} n) \right)_{n \in \mathbb{Z}}$$

eine orthonormale Basis von O_{2^j} und

$$\left(\sqrt{2^{-j}} \psi_{2^j}(x) (x - 2^{-j} n) \right)_{(n,j) \in \mathbb{Z}^2}$$

eine orthogonale Basis von $L^2(\mathbb{R})$. $\psi(x)$ wird orthogonales Wavelet genannt.

Eine orthonormale Basis von O_{2^j} kann durch Skalierung des Wavelets $\psi(x)$ mit einem Koeffizienten 2^j und durch Translation auf ein Gitter, dessen Intervall proportional zu 2^{-j} ist, erhalten werden.

Die Eigenschaften der Skalierungsfunktion und des Wavelets hängen von der Wahl der Funktion H ab. So kann eine gute Lokalisation in Frequenz- und Ortsraum gegeben sein. Es gibt sehr verschiedene Wavelets, die sich vor allem in ihrer Regularität⁷ und Kompaktheit unterscheiden. Eine häufig gestellte Forderung ist, daß das Wavelet glatt ist und eine Mindestanzahl verschwindender Momente⁸ aufweist.

Das Detailsignal läßt sich berechnen nach

$$D_{2^j} f = \left(\langle f(u), \psi_{2^j}(u - 2^{-j} n) \rangle \right)_{n \in \mathbb{Z}} \quad (2.27)$$

und enthält die Differenz der Information zwischen $A_{2^j}^d f$ und $A_{2^{j+1}}^d f$. Das innere Produkt kann über eine Faltung des Signales $f(x)$ mit $\psi_{2^j}(-x)$ berechnet werden:

$$\langle f(u), \psi_{2^j}(u - 2^{-j} n) \rangle = (f(u) * \psi_{2^j}(-u))(2^{-j} n). \quad (2.28)$$

⁷Der Regularitätsgrad ist die Anzahl der stetigen Ableitungen eines Wavelets.

⁸Das k -te Moment ist definiert als $m_k = \int_{-\infty}^{+\infty} f(x) x^k dx$ und bestimmt das Verhalten des Wavelets gegenüber Funktionen der Ordnung k .

Damit ist das diskrete Detailsignal über ein Sampling von $f(u) * \psi_{2^j}(-u)(x)$ mit dem Grad 2^j zu berechnen:

$$D_{2^j} f = (f(u) * \psi_{2^j}(-u))_{(2^{-j}n)_{n \in \mathbb{Z}}}. \quad (2.29)$$

Das Wavelet kann als Bandpaßfilter angesehen werden, dessen Frequenzbänder nahezu $[-2\pi, \pi] \cup [\pi, 2\pi]$ sind. Damit wird das Signal approximiert durch

$$\left(A_{2^{-j}}^d f, (D_{2^j} f)_{-j \leq j \leq -1} \right). \quad (2.30)$$

2.7.4 Anwendung auf Bilder

Der zweidimensionale und mehrdimensionale Fall lassen sich aus dem eindimensionalen analog entwickeln. Von Vorteil ist hier eine separable Skalierungsfunktion $\Phi(x, y) = \Theta(x)\Theta(y)$. Die Approximation der Stufe 2^j lautet dann:

$$A_{2^j} f = \left(\langle f(x, y), \psi_{2^{-j}}(x - 2^j n) \psi_{2^{-j}}(y - 2^j m) \rangle \right)_{(n, m) \in \mathbb{Z}^2}. \quad (2.31)$$

Mit der Aufteilung des Zerlegungsproblems auf Zeilen und Spalten ergeben sich drei Kombinationen der eindimensionalen Funktion Θ mit dem eindimensionalen ihr zugehörigen Wavelet ψ :

$$\begin{aligned} \Phi^1(x, y) &= \Theta(x)\psi(y), \\ \Phi^2(x, y) &= \psi(x)\Theta(y), \end{aligned} \quad (2.32)$$

$$\Phi^3(x, y) = \psi(x)\psi(y)$$

Mit diesen wird die Basis analog zum eindimensionalen Fall gebildet (mit $e = 1, 2, 3$):

$$\text{für } O_{2^j} : \quad (2^{-j} \Phi_{2^j}^e(x - 2^{-j} n, y - 2^{-j} m))_{(n, m) \in \mathbb{Z}^2} \quad (2.33)$$

$$\text{für } L^2(\mathbb{R}^2) : \quad (2^{-j} \Phi_{2^j}^e(x - 2^{-j} n, y - 2^{-j} m))_{(n, m) \in \mathbb{Z}^3}. \quad (2.34)$$

Damit ergeben sich drei Hochpaßanteile $D_{2^j}^e f$:

$$D_{2^j}^e f = \left(\langle f(x, y), \Phi_{2^j}^e(x - 2^{-j}n)\psi_{2^j}(y - 2^{-j}m) \rangle \right)_{(n,m) \in \mathbb{Z}^2} . \quad (2.35)$$

Diese bilden zusammen mit dem Tiefpaßanteil A_{2^j} die gesamte Information des Bildes ab. Die Zerlegung der Frequenzebene durch die $D_{2^j}^e f$ und A_{2^j} wird in Abb. 2.6 dargestellt, das Schema der Anordnung der Bilder in Abb.2.7. Die Zerlegung des Bildes 2.8 nach diesem Schema ist in Abb. 2.9 dargestellt. Die Koeffizienten mußten auf den Grauwertebereich [0 255] skaliert werden. Aus der Zerlegung läßt sich ein Merkmalsbild gewinnen, indem eine Energiefunktion verwendet wird (s. Kap. 3.3.2). Das Merkmalsbild ist in Abb. 2.10 dargestellt.

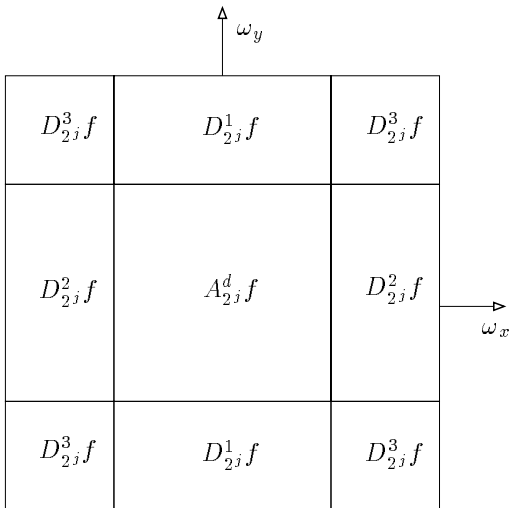


Abbildung 2.6: Zerlegung des Frequenzraumes durch Tiefpaß- und Hochpaßanteile

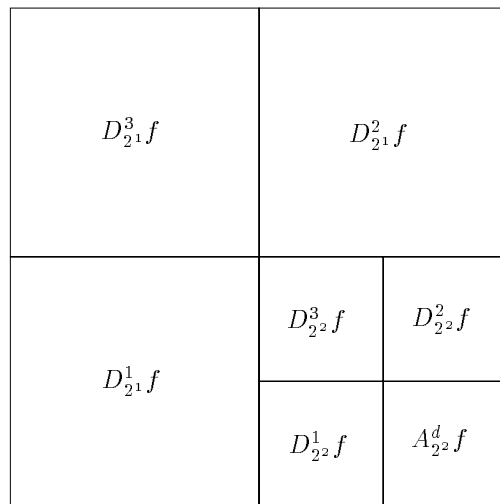


Abbildung 2.7: Anordnung der $D_{2^j}^k f$ und $A_{2^j}^d f$

2.8 Weitere Zerlegungen mit Wavelets

Die Basen, mit denen die in Kapitel 2.7.2 vorgestellte Dekomposition durchgeführt wird, sind aufgrund ihrer Orthogonalität festen Bedingungen unterworfen. So kann eine Skalierungsfunktion mit kompaktem Träger nicht symmetrisch oder antisymmetrisch und stetig sein. Die Symmetrie ist für die Analyse endlicher Signale nützlich [28]. Die Nutzung symmetrischer oder antisymmetrischer Filter

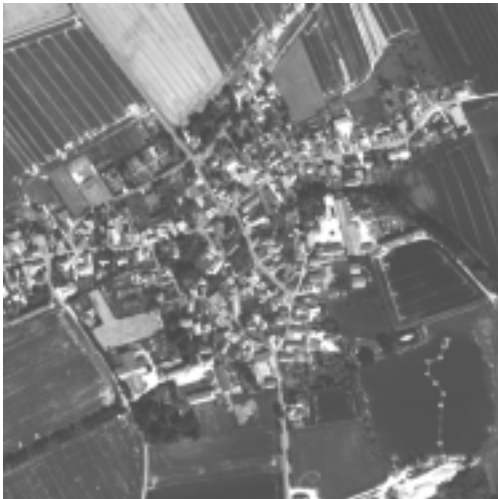


Abbildung 2.8: Ausschnitt aus der Luftbildszene

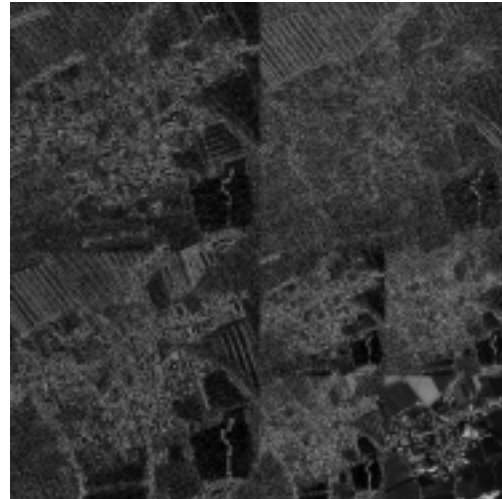


Abbildung 2.9: Die Zerlegung des Bildes 2.8

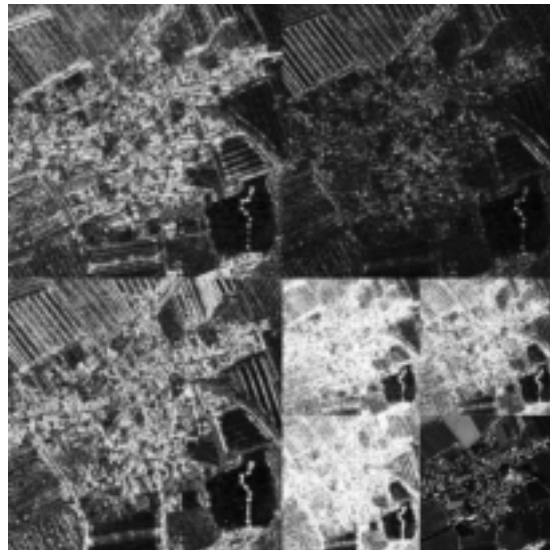


Abbildung 2.10: Durch eine Energietransformation gewonnenes Merkmalsbild

erhöhen die Genauigkeit der Texturkantenerkennung und vereinfachen die Randbehandlung [31]. Mit der Einführung biorthogonaler Mehrfachaufösungen können diese Zwänge gelockert werden [32].

In der Texturanalyse werden auch andere Waveletmethoden eingesetzt. Die Wavelet Packet Transformation zerlegt nicht nur die Tiefpaßanteile sukzessiv,

auch die Hochpaßanteile werden weiter zerlegt. Mit Entscheidungsregeln kann über eine weitere Zerlegung eines Anteils entschieden werden und z.B. die entstehende Baumstruktur der Zerlegung als Texturmerkmal dienen. Chitre und Dhawan untersuchen die Nutzung von M-Band Wavelets, mit denen sich hochfrequente Signale mit geringer Bandbreite untersuchen lassen [33].

2.8.1 Aliasing und Verschiebungsinvarianz

Je nach Wahl der Basisfunktionen einer Transformation entstehen in den Ergebnissen Aliasingeffekte durch das Subsampling. Hohe Frequenzen (höher als die Samplingfrequenz) drücken sich dann als niedrigere (niedriger als die Samplingfrequenz) aus. Für die aufeinanderfolgenden Stufen einer Pyramide bedeutet dies, daß es den Stufen Frequenzen gibt, die nicht im Bild enthalten sind. Die Rekonstruktion eines Bildes aus den Pyramidenstufen ist davon nicht betroffen, da sich diese Effekte genau aufheben. Für die Auswertung und Weiterverarbeitung einzelner Stufen können solche Effekte allerdings eine Rolle spielen. Bei der Texturanalyse müssen also Aliasingeffekte bedacht werden. Die Tatsache, daß sich einzelne Frequenzen über Stufen verteilt finden, kann als Korrelation der Stufen untereinander gedeutet werden. Mallat [29] schreibt dazu, daß zumindest für die Gauß- und Laplacepyramiden noch kein Modell zur Beschreibung der Korrelationen zwischen den Stufen besteht. Es ist also schwierig, das Verhalten von Aliasingeffekten zu modellieren oder vorherzusagen. Nach Simoncelli [34] tritt Aliasing in den Wavelet- und QMF⁹-Pyramiden verstärkt auf. Es besteht ein Zusammenhang zwischen kritisch abgetasteten (*critically sampled*) Dekompositionen¹⁰ und ihrem Aliasingverhalten: Die kritische Abtastung eines Signales mit Wavelets oder QMF's führt zwangsläufig zu Dekompositen mit starkem Aliasing, da hier das Nyquisttheorem verletzt werden kann, ohne Information zu verlieren. Diese Eigenschaft läßt zwar eine exakte Rekonstruktion zu, bei der Bearbeitung und Interpretation der einzelnen Stufen müssen diese Effekte aber bedacht werden, da sie ohne Modellierung nicht auf andere Probleme übertragbar sind. Ein weiteres Problem der orthogonalen Zerlegungen ist ihre Eigenschaft,

⁹Quadrature Mirror Filters, s. Kapitel 3.1.2.

¹⁰kritische Abtastung führt auf nichtredundante (orthogonale) Zerlegungen, s. Kap. 2.7.

nicht translationsinvariant zu sein. Eine Verschiebung des Signales ändert die Koeffizienten der Zerlegung erheblich. Sowohl innerhalb einer Zerlegungsstufe als auch zwischen verschiedenen Stufen treten Änderungen in den Energien auf. Die Translationsinvarianz ist für die Übertragbarkeit der Ergebnisse auf andere Fälle aber wünschenswert. Denn das Ergebnis der Analyse ist sonst abhängig von der Lage des Signales gegenüber dem Gitter, auf dem es diskretisiert ist. Da die Energieanteile aus der Transformation in den Stufen als Merkmale der Textur dienen sollen, wird klar, daß eine Verschiebung der Anteile innerhalb oder zwischen den einzelnen Stufen die Merkmale verändert und somit eine Ableitung von wiederverwendbaren Merkmalen erschwert bis unmöglich macht. Die Nutzung von orthogonalen und damit nichtredundanten und den Rechenaufwand verkleinernden Dekompositionen muß also für die Texturanalyse unter dem Vorbehalt der Nichtreproduzierbarkeit und Nichtübertragbarkeit auf ähnliche Probleme gesehen werden.

Aliasing soll am Beispiel der Laplacepyramide demonstriert werden (s. Abb. 2.11). Ein Weg, die Aliasingeffekte sichtbar zu machen, ist die Erzeugung von Variationen durch Verschiebung des Signals. Zu diesem Zweck wird das Bild um einige Pixel verschoben und geglättet (Tiefpaß), dann das Subsampling und danach das Upsampling durchgeführt. Das nicht verschobene Bild wird ebenfalls geglättet und dem Subsampling mit anschließendem Upsampling unterworfen und die Verschiebung zuletzt durchgeführt. Das Ergebnis ist dasselbe, wenn es kein Aliasing gibt. Die Glättungs- und Verschiebungsoperation verhalten sich kommutativ zueinander, und die Reihenfolge der Operationen sollte keine Änderung des Ergebnisses hervorrufen (nach [35]). Das Bild zeigt die entstandene Differenz.

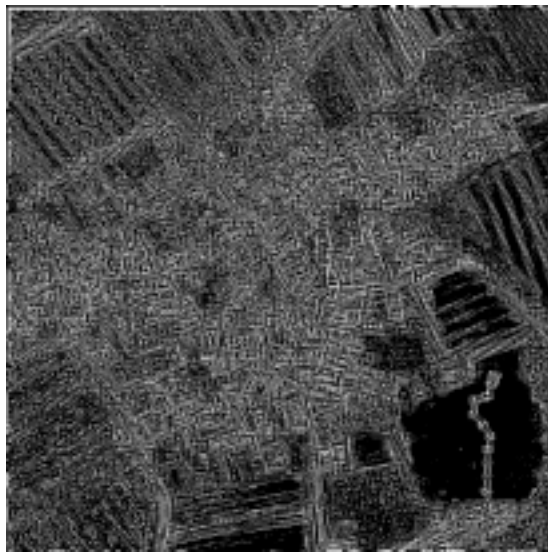


Abbildung 2.11: Durch Aliasing entstandene Differenzen

Kapitel 3

Gaborfilterung in der Texturanalyse

3.1 Hilberttransformation und Quadraturfilterpaare

Ein entscheidendes Merkmal eines Signales ist die Energie. Hier soll darauf eingegangen werden, wie die Energieanteile eines Signales berechnet werden können. Dazu wird die Hilberttransformation eingeführt, die die Konstruktion von Quadraturfilterpaaren ermöglicht, wie sie in der Bildverarbeitung häufig eingesetzt werden. Diese Quadraturfilterpaare werden im Englischen *Quadrature Mirror Filter* genannt und mit QMF abgekürzt.

3.1.1 Energie eines Signales

Die Energie einer Welle oder eines Signales ist physikalisch proportional zum Quadrat der Amplitude. Mit Hilfe des Additionstheorems $(\sin \phi)^2 + (\cos \phi)^2 = 1$ läßt sich eine Energiefunktion konstruieren. Nimmt man zwei Funktionen mit dem Phasenunterschied von $\pi/2$ und addiert ihre Quadrate, dann erhält man eine zur Energie proportionale Funktion, die phasenunabhängig ist. Ohne den phasenverschobenen Anteil würde die Energiefunktion mit doppelter Wellenzahl variieren, wäre demnach also nicht phasenunabhängig (nach [14], S. 125).

3.1.2 Hilberttransformation

Die Hilberttransformation kann ein Signal um $\frac{\pi}{2}$ verschieben, ohne die Amplitude zu verändern. Dabei entsteht aus einer geraden Funktion eine ungerade und umgekehrt, die sin- und cos-Funktionen sind ein Beispiel. Der Betrag dieser Funktion ist eins, zwei ausgezeichnete Punkte ausgenommen, deshalb wird die Amplitude nicht verändert. Eine Funktion und ihre Hilberttransformierte zusammen nennt man Quadraturpaar. Die Hilberttransformation bildet die Grundlage für eine ganze Familie von Filtern, die Quadraturfilterpaare heißen. Sie haben die gleichen Eigenschaften wie eine Funktion und ihre Hilberttransformierte: beide Teilfilter haben die gleiche Amplitudenfunktion und unterscheiden sich in der Phasenlage. Der gerade Teilfilter ist rein reell, erzeugt also keine Phasenverschiebung. Der ungerade Teilfilter ist rein imaginär; er erzeugt daher eine Phasenverschiebung um $\frac{\pi}{2}$. Jeder Teilfilter geht durch die Hilberttransformation in in den anderen über ([14],S. 125).

3.2 Gaborfilter

Gaborfilter sind Filter mit einer komplex gaußförmig harmonischen Filterantwort. Ihr Name geht auf D. Gabor zurück, der sie als erster zur kombinierten Zeit-Frequenz-Analyse vorschlug. In der Bildverarbeitung haben sie besondere Bedeutung erlangt, da die Stoßantworten der Filter eine biologische Entsprechung im visuellen Kortex besitzen. So wurde durch Messungen gezeigt, daß rezeptive Felder mit einer Phasenverschiebung von 90° bei sonst gleichen Eigenschaften auftreten, was eine phasenunabhängige Extraktion von Energiemerkmale ermöglicht. In der Bildverarbeitung werden daher komplexwertige Gaborfunktionen verwendet, deren Real- und Imaginärteil ein Quadraturfilterpaar bilden. Durch physiologische Messungen zeigt sich, daß die Parameter des Gabor- Modells von rezeptiven Feldern einfacher Zellen zum einen begrenzt und zum anderen untereinander korreliert sind und nicht beliebige Werte annehmen. So liegt zum Beispiel die radiale Bandbreite durchschnittlich bei 1,3 Oktaven, sowie die azimutale Bandbreite bei 35° . Außerdem sind die Bandbreiten stark miteinander korreliert, was zu einem

Öffnungsverhältnis der gaußförmigen Einhüllenden in den Grenzen von $\frac{1}{4}$ bis 1 führt. Es läßt sich auch eine hohe Korrelation zwischen der Orientierung der Einhüllenden eines rezeptiven Feldes und der Orientierung seiner modulierenden Schwingungen beobachten, und zwar in der Art, daß sie näherungsweise gleich sind, was die Einführung von sogenannten Daisy-Petal-Gaborfiltern nahelegt (s. Gleichung (3.4)) ([13], S.22).

Mit Gaborfiltern lassen sich keine orthogonalen Transformationen, wie sie in Kapitel 2.7 beschrieben werden, erzeugen.

3.2.1 Definition der Gaborfilter

Die Faltungsmasken der 2D-Gaborfunktion haben die generelle Form ($i = \sqrt{-1}$):

$$h(x, y) = g(x', y') \cdot \exp [2\pi i(Ux + Vy)] . \quad (3.1)$$

worin $(x', y') = (x \cos \phi + y \sin \phi, -x \sin \phi + y \cos \phi)$ rotierte Koordinaten sind, und mit

$$g(x, y) = \left(\frac{1}{2\pi\lambda\sigma^2} \right) \cdot \exp \left(-\frac{(x/\lambda)^2 + y^2}{2\sigma^2} \right) . \quad (3.2)$$

Damit ist $h(x, y)$ eine komplexe sinusoidale Funktion, die moduliert wird von einer 2D-Gaußfunktion mit dem Öffnungsverhältnis λ , dem Skalierungsparameter σ und der Hauptachsenrichtung ϕ von der x - Achse. Ist $\lambda = 1$, braucht ϕ nicht spezifiziert werden, da $g(x, y)$ kreisförmig ist. Die Orts-Frequenzantwort der Gaborfunktion ist:

$$H(u, v) = \exp \{ -2\pi^2\sigma^2 [(u' - U')^2\lambda^2 + (v' - V')^2] \} , \quad (3.3)$$

worin $(u', v') = (u \cos \phi + v \sin \phi, -u \sin \phi + v \cos \phi)$ und (U', V') die entsprechende Rotation der Mittenfrequenz (U, V) . Damit ist $H(u, v)$ eine Bandpass-Gaußfunktion mit der Orientierung der kleinen Hauptachse im Winkel ϕ von der u -Achse, Öffnungsverhältnis $1/\lambda$, radialer Mittenfrequenz $F = \sqrt{U^2 + V^2}$ (gemessen in Perioden/Bild) und Orientierung $\Phi = \arctan(V/U)$ (Grad oder Radiant, gemessen von der u -Achse).

Gaborfilter mit willkürlichen Orientierungen der modulierenden Gaußfunktion auszuschließen bedeutet eine Einschränkung der Möglichkeiten zur Variation von Parametern. Trotzdem ist es bequemer jene auszuwählen, deren modulierende Gaußfunktion dieselbe Orientierung hat wie die komplexe Sinusfunktion ($\phi = \Phi$). In diesem Fall werden die Filter Daisy-Petal-Gaborfilter genannt und haben die einfachere Form [13] [1]:

$$h(x, y) = g(x', y') \cdot \exp(2\pi i F x') \quad (3.4)$$

und

$$H(u, v) = \exp \left\{ -2\pi^2 \sigma^2 [(u' - F)^2 \lambda^2 + (v')^2] \right\} . \quad (3.5)$$

Abbildung 3.1 zeigt vier Impulsantworten von Gaborfiltern mit den Orientierungen $0^\circ, 45^\circ, 90^\circ$ und 135° . In Abbildung 3.2 sind die Fouriertransformierten der Filter zu finden. Hier erkennt man die Frequenz- und Richtungsselektivität.

3.2.2 Bandbreiten

Gaborfilter können nach Bandbreiten eingeteilt werden. Sie machen eine Aussage über den Frequenzbereich, der von den Funktionen abgedeckt wird. Die Bandbreiten in der allgemeinen Definition nach Gleichung 3.1 haben keine fest definierte Bedeutung, da die radiale Breite mit der Orientierung der Filter variiert. Für die Daisy-Petal-Konfiguration lassen sie sich aber einfach angeben und sind dort auch richtungsunabhängig.

Es werden zwei Maße für die Bandbreiten unterschieden: die *radiale* und *azimutale* (auch Frequenz- und Orientierungs-) Bandbreite. Es ist üblich, Halbwertsbandbreiten anzugeben, sie sind in der Literatur unter Bandbreite meist gemeint. Sie werden über Frequenzen definiert, für die der Betrag der Übertragungsfunktion die Hälfte des Maximalwertes annimmt. Die radiale Halbwertsbandbreite ist definiert als das logarithmische Verhältnis der beiden Halbwertsfrequenzen in radialer Richtung vom Ursprung:

$$B = \log_2 \left[\frac{\pi F \lambda \sigma + \alpha}{\pi F \lambda \sigma - \alpha} \right] . \quad (3.6)$$

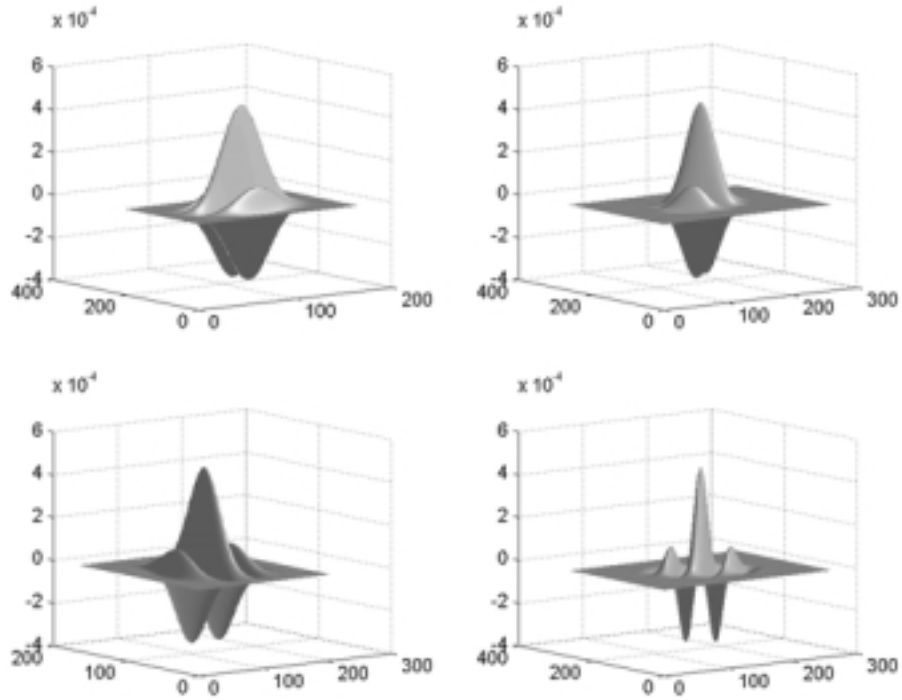


Abbildung 3.1: Impulsantworten von vier Gaborfiltern

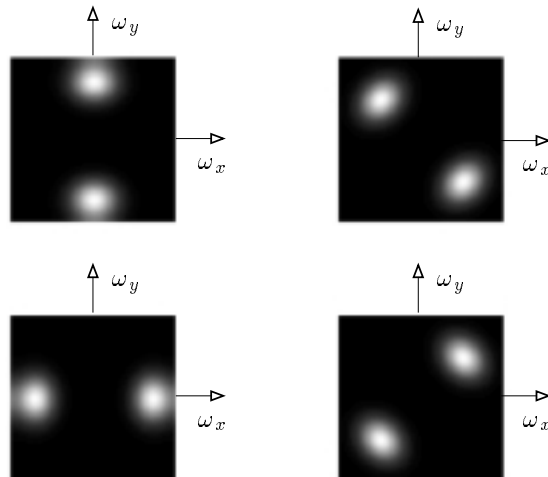


Abbildung 3.2: Gaborfilter im Fourierraum

Diese radiale Halbwertsbandbreite wird in Oktaven angegeben. Die azimutale Halbwertsbandbreite bezeichnet den Winkel, der durch die Halbwertsfrequenzen an der im Ortsfrequenzbereich in v -Richtung breitesten Stelle der Gaußglocke aufgespannt wird. Sie wird in Radiant oder Grad angegeben ([13], S. 26 f.; [1]):

$$\Omega = 2 \arctan \left(\frac{\alpha}{\pi F \sigma} \right) \quad (3.7)$$

$$\text{mit } \alpha = \sqrt{\frac{\ln 2}{2}} .$$

Die Bandbreiten sind in Abbildung 3.3 eingezeichnet.

3.2.3 Gaborfilter und die Unschärferelation

Gaborfilter erfüllen die Unschärferelation mit dem Gleichheitszeichen. Sie haben also maximales kombiniertes Auflösungsvermögen im Orts- und Ortsfrequenzbereich. Die Unschärferelation lautet für die kombinierte Ausdehnung von 2D-Signalen im Ortsfrequenzbereich:

$$\Delta x \Delta y \Delta u \Delta v \geq \frac{1}{4} \quad (3.8)$$

Das Gleichheitszeichen gilt, wenn die Koordinatenachsen des Signals sowohl im Orts- als auch im Ortsfrequenzbereich mit deren Koordinatenachsen zusammenfallen ([13], S. 25). Es läßt sich zeigen, daß die Relation unabhängig von der Wahl der Parameter mit dem Gleichheitszeichen erfüllt wird, Es konnte sogar für die 2D-Gaborfunktionen gezeigt werden, daß es die einzigen Funktionen mit dieser Eigenschaft sind [27].

3.3 Gaborfilter in der Texturanalyse

Es gibt viele Untersuchungen zur Texturanalyse mit Gaborfiltern ([13], [1], [36], [37], [38]). Die Verfahren unterscheiden sich in der Auswahl der Filter und der Merkmalsbildung. Für die Klassifikation mit den jeweils entstehenden Merkmalsbildern werden jeweils andere Klassifikationsalgorithmen verwendet. Hier soll eine kleine Auswahl der häufiger verwendeten Wege gegeben werden.

3.3.1 Auswahl der Filter

Die Auswahl der Filter läßt sich auf verschiedene Weise vornehmen. Entweder werden feste Filterbänke benutzt, in denen ein Satz Gaborfilter mit festgelegten Orientierungen und Bandbreiten verwendet wird, oder die Auswahl der Filter ist adaptiv an die Bildszene angepaßt. Feste Filterbänke werden meist so konstruiert, daß die Ortsfrequenzebene nahezu vollständig überdeckt wird.

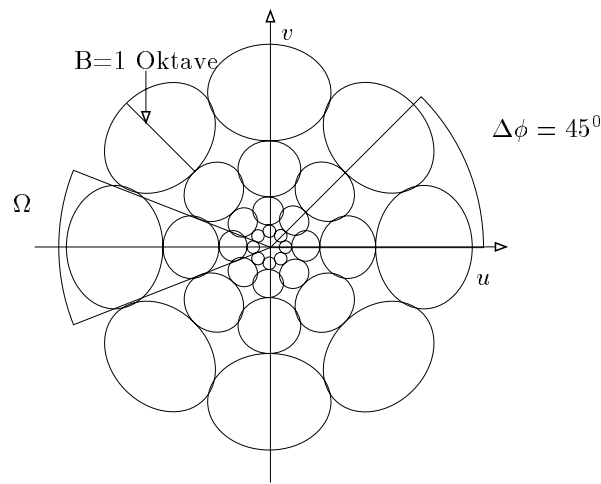


Abbildung 3.3: Gaborfilterbank nach [1], eingetragen sind Halbwertsbandbreiten

Dazu schlagen Jain und Farokhnia [1] eine Filterbank vor, bei der die Orientierungen 0° , 45° , 90° und 135° betragen. Die radialen Frequenzen der Filter werden in Abhängigkeit von der Größe des Bildes errechnet. Für ein Bild mit der Breite $N_c = 2^n$, $n \in \mathbb{Z}$ ergeben sich die radialen Frequenzen u_0 zu:

$$1\sqrt{2}, 2\sqrt{2}, 4\sqrt{2}, \dots \text{ in Perioden/Bild.}$$

Damit ergeben sich für ein Bild der Breite 256 insgesamt 28 Filter. Es werden die geraden Gaborfilter benutzt, die dem Realteil entsprechen, der mit dem Kosinus errechnet wird. Abbildung 3.1 zeigt vier Gaborfilter mit oben genannten Orientierungen.

Auf die Filter mit niedriger Frequenz kann man eventuell verzichten, da sie Variationen erfassen, die zu groß sind, um Texturvariationen auszudrücken. Aus der Anzahl der Filter ergibt sich der hohe Rechenaufwand, der mit Filterbänken verbunden ist. Wenn die Hauptfrequenzen der Texturen nicht mit denen der Filter zusammenfallen, ist die feste Frequenz- und Orientierungsselektivität bei der Nutzung von Filterbänken nicht optimal an die im Bild auftretenden Texturen angepaßt. Daher wurden Ansätze entwickelt, die die Auswahl weniger Filter erlauben. Dunn und Higgins [39] schlagen ein Filterauswahlverfahren vor, das einen Filter zur Unterscheidung von zwei Texturen auswählt. Damit können aber nur Bilder, in denen zwei Texturen auftreten, behandelt werden. In der Fernerkundung hat man es aber generell mit Bildern zu tun, die mehrere Texturen aufweisen. Daher wird dieser Ansatz nicht weiter verfolgt.

Einen Ansatz zur Auswahl mehrerer Filter machen Bovik und Clark [36]. In der Fouriertransformierten $T(u, v)$ des Bildes $t(x, y)$ werden die Spitzen aufgesucht (als Maxima von $|T(u, v)|^2$ in einem Bereich der Halbebene $u_0 > 0$) und in polarer Form (F, Θ) über $[-\pi/2, +\pi/2]$ aufgelistet. F und Θ sind dann mit der Bandbreite, die zu 0.7, 1.0 oder 1.3 gewählt wird, die Parameter für die Gaborfilter. Das Öffnungsverhältnis λ wird zu 1 gewählt. Eine vorhergehende Filterung der Ortsfrequenzebene mit einem Gaußfilter verbessert das Auffinden der Maxima. Die Suche der Maxima ist dabei von Hand vorgenommen worden. Die Annahme bei diesem Ansatz ist, daß sich das Bild aus sogenannten Schmalbandtexturen zusammensetzt, Texturen also, die im wesentlichen nur Frequenzanteile in der direkten Umgebung einer Hauptfrequenz besitzen, die als Maximum gefunden wird. Texturen, die sich durch andere, weniger energiereiche Frequenzen unterscheiden, kann dieser Algorithmus nicht unterscheiden. Außerdem bleiben Texturen unberücksichtigt, deren Flächen- und damit Energieanteil klein ist ([13], S. 31).

Hier wird der Ansatz von Jain und Farokhnia untersucht, indem ein Satz Gaborfilter mit festen Parametern auf das Bild angewendet wird. Die Filter haben die oben genannten vier Orientierungen und es werden drei Frequenzen verwendet. Damit erhält man eine Filterbank mit 12 Filtern. Mehr Filter erhöhen den Rechenaufwand erheblich.

3.3.2 Lokale Energiefunktion

Aus den gefilterten Bildern müssen Merkmalsbilder erzeugt werden, die die lokale Energie widerspiegeln und zu einheitlichen Grauwertebereichen führen. Da bei dem Ansatz nur die geraden Teilfilter verwendet werden, schwanken die Grauwerte mit der Frequenz der Filter. Es gibt mehrere Funktionen, die die gefilterten Bilder zu Merkmalsbildern transformieren können. Diese Transformationen sind nichtlinear. Möglich ist die Berechnung des Absolutwertes oder des Quadrates der Grauwerte der gefilterten Bilder. In [1] wird die sogenannte tanh- Nichtlinearität verwendet, die folgendermaßen definiert ist:

$$T(t) = \tanh(\alpha t) = \frac{1 - e^{-2\alpha t}}{1 + e^{-2\alpha t}}. \quad (3.9)$$

Experimentell ermittelten sie einen optimalen Wert für α von 0.25, der auch verwendet wird. Das Merkmalsbild wird folgendermaßen aus dem gefilterten Bild $r_k(a, b)$ errechnet:

$$e(x, y) = \frac{1}{M^2} \sum_{(a,b) \in W_{xy}} |T(r_k(a, b))|. \quad (3.10)$$

Hierin sind $T(\cdot)$ die Funktion 3.9 und W_{xy} ein $M \times M$ großes Fenster um das Pixel mit den Koordinaten (x, y) . Den Graph der tanh-Funktion zeigt Abbildung 3.4 Die Fenstergröße ist ein wichtiger Faktor, große Fenster führen auf ein gutes Texturmaß, kleine Fenster lokalisieren die Grenzen besser. Hier wurde ein Wert von 7 Pixeln für das Fenster gewählt. In [1] wird die Wahl eines gaußgewichteten Fensters in Abhängigkeit von der Filtergröße vorgeschlagen.

3.4 Untersuchungen mit Gaborfiltern

Um die Tauglichkeit der Gaborfilter zu untersuchen, wurde ein Bild aus 128×128 Pixeln großen Teilen der Luftbildszene zusammengesetzt, die zu unterscheidende Texturen enthalten. Das zusammengesetzte Bild ist in Abb. 3.5 dargestellt. Es enthält eine Zeile mit Ausschnitten von Feldbereichen unterschiedlicher Orientierung, eine Zeile mit Ausschnitten aus besiedeltem Gebiet. In der dritten

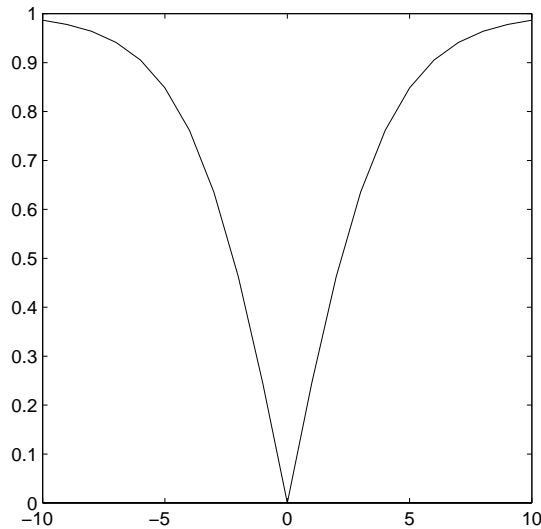


Abbildung 3.4: tanh-Funktion

Zeile sind Ausschnitte aus Waldbereichen und eine Wasserfläche zu finden, in der vierten Zeile werden Ausschnitte aus bebauten Gebieten, die inhomogene Größen aufweisen (Industrieflächen), sowie ein Feld, das von einer mit Bäumen bestandenen Straße geteilt wird, und ein weiterer Feldausschnitt zusammengefaßt. Die Filterung des Bildes wurde mit einem Satz Gaborfilter vorgenommen, der 4 Orientierungen aufweist. In Abb. 3.6 sind die Merkmalsbilder der vier Orientierungen nebeneinander für zwei Mittenfrequenzen dargestellt. Deutlich erkennt man das orientierungsselektive Verhalten der Filter: Das Feld oben links findet sich im Merkmalsbild der entsprechenden Orientierung wieder, in den anderen nicht. Hier erkennt man die Eigenschaft von Siedlungsbereichen, in allen Orientierungen deutliche Merkmale zu erzeugen. Ebenso hat Wald die Eigenschaft, in allen Orientierungen Merkmale zu bilden, allerdings nicht so ausgeprägt wie die Siedlungsbereiche. Die Begründung dafür ist der relativ geringe Kontrast der Waldbereiche im Bild, der dann das Filterergebnis mit geringerem Energieanteil ausfallen läßt. Die Abgrenzung zu Feldbereichen mit „Flecken“ ist nicht auf einfache Weise möglich, da diese ähnliche Merkmale erzeugen.

Was kann nun aus den Merkmalen an Regeln zur Unterscheidung der verschiedenen Bereiche abgeleitet werden?



Abbildung 3.5: Aus verschiedenen Texturen zusammengesetztes Bild

Eine Abgrenzung von Wald-, Feld- und Siedlungsbereichen wäre ein wünschenswertes Ziel. Ihre Eigenschaften lassen sich wie folgt zusammenfassen:

- Siedlung: Merkmale treten in allen Orientierungen auf, die Merkmale sind maximal (besonders hell im Merkmalsbild)
- Wald: Merkmale treten in allen Orientierungen auf
- Felder: Felder mit parallelen Bearbeitungsspuren erzeugen in einer Orientierung deutliche Merkmale, in den anderen nicht

Mit dem im folgenden vorgestellten Ansatz wird der Versuch gemacht, diese Eigenschaften durch Bildoperationen zu extrahieren. Eine andere Möglichkeit, die Textureigenschaften zu extrahieren, besteht in einer Klassifikation mit den gebildeten Merkmalen. Dieser Ansatz wird später in Kapitel 6 verfolgt. Merkmale aus Gaborfilterungen werden oft mit einer Learning Vector Quantisation klassifiziert, wenn eine unüberwachte Klassifizierung stattfinden soll (s. z.B. [40]). Hier sollen die Eigenschaften der Texturen, wie sie oben beschrieben sind, zu einer Segmentierung des Bildes führen.

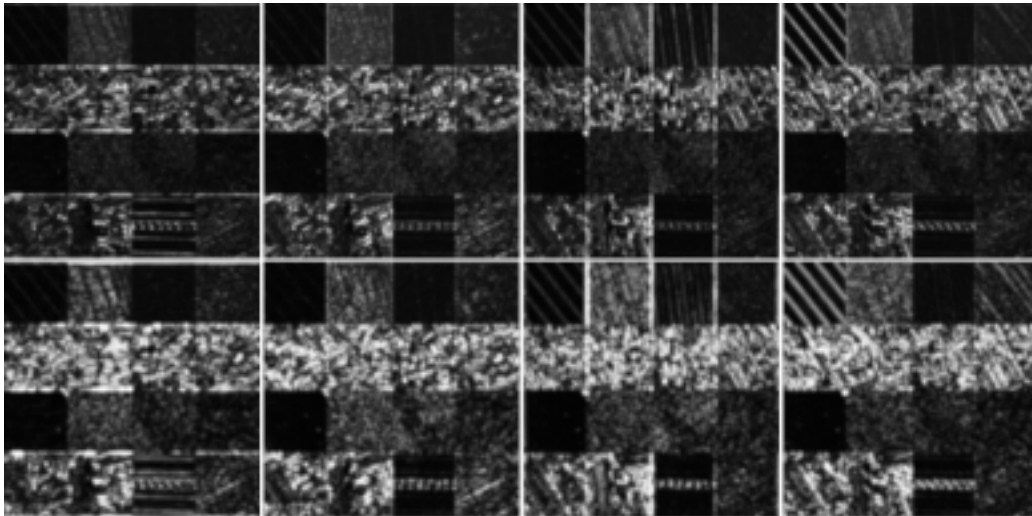


Abbildung 3.6: Die Merkmalsbilder der Gaborfilterung mit acht Filtern. Vier Orientierungen in den Spalten, zwei Mittenfrequenzen in den Zeilen dargestellt

Um die Information aus den Merkmalsbildern zusammenzufassen, wurde folgendermaßen vorgegangen:

Zur Unterscheidung der Eigenschaften *eine Orientierung vorherrschend* und *alle Orientierungen vorhanden* wurden jeweils die Merkmalsbilder zweier senkrecht aufeinander stehender Orientierungen miteinander verrechnet. Geht man davon aus, daß Merkmale für orientierte Bereiche (Felder) nur in einem der beiden Merkmalsbilder vorhanden sind, kann man aus den beiden Merkmalsbildern ein weiteres erzeugen, das aus den Minimalwerten der beiden besteht. Damit werden Merkmale, die nur in einem der beiden eingehenden Bilder vorhanden sind, eliminiert oder zumindest auf den im anderen Bild vorhandenen Wert reduziert. Für

Merkmale, die in beiden Bildern vorhanden sind, ergibt sich keine große Änderung. Damit bedeuten verbleibende Merkmale im entstehenden Bild *in zwei senkrecht stehenden Richtungen Merkmale vorhanden*. Ein Beispiel findet sich in Abb. 3.7, in dem die in Abb. 3.6 links angeordneten vier Merkmalsbilder und die vier rechts angeordneten verrechnet wurden. Das entstehende Bild wurde noch einer Schwellwertoperation unterworfen. Die Siedlungsbereiche treten deutlich hervor, die Merkmale für das Feld oben links (Merkmale in nur je einem der verrechneten Bilder) fallen erwartungsgemäß heraus. Mit der Minimumoperation kann also zumindest eine Merkmalsbildung für Siedlungsbereiche stattfinden. Wald hat keine signifikanten Merkmale erhalten. Dafür sind die Merkmale, wie in Abb. 3.6 zu sehen ist, nicht ausgeprägt genug und zu inhomogen.

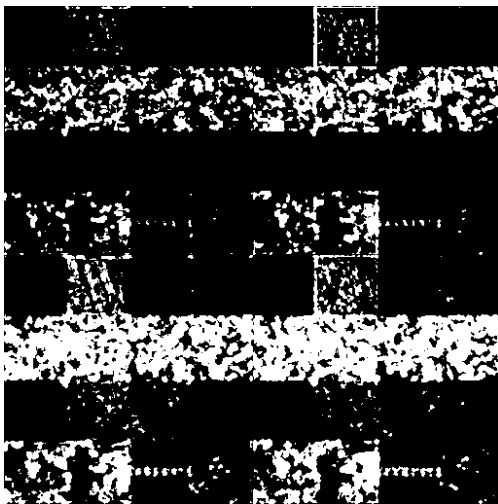


Abbildung 3.7: Merkmalsbild: *Merkmale, die in zwei senkrecht aufeinanderstehenden Orientierungen vorhanden sind*

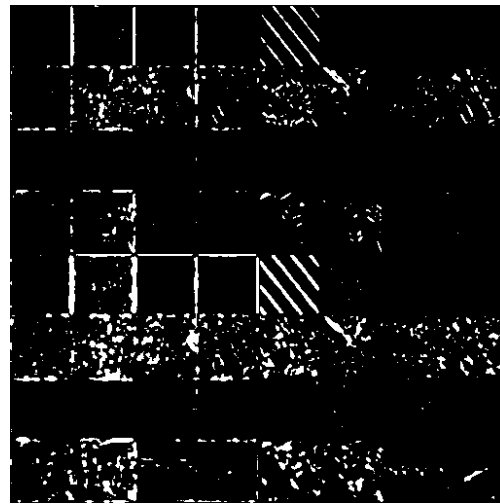


Abbildung 3.8: Merkmalsbild: *Merkmale, die in nur einer von zwei senkrecht aufeinanderstehenden Orientierungen vorhanden sind*

Die Extraktion der Eigenschaft *nur in einer Richtung Merkmale vorhanden* wurde über eine ähnlich einfache Verrechnung zweier Merkmalsbilder senkrecht aufeinander stehender Orientierung versucht. Dazu wurden die Bilder voneinander subtrahiert, damit Merkmale, die in beiden Bildern auftauchen, herausfallen und Merkmale, die nur in einem Bild enthalten sind, erhalten bleiben. Dabei stellt sich das Problem, daß zwei Subtraktionen möglich sind ($Bild_1 - Bild_2$,

$Bild_2 - Bild_1$). Die so entstehenden zwei Bilder wurden addiert und einer Schwellwertoperation unterworfen, das Ergebnis ist in Abb. 3.8 zu sehen. Erwartungsgemäß sind Merkmale für das Feld oben links vorhanden, aber auch in anderen Bereichen treten diese Merkmale auf. Damit ist eine Segmentierung nach dieser Orientierungseigenschaft nicht möglich. Diese Merkmalsbildung ist demnach nicht optimal. Als Fazit läßt sich festhalten, daß eine Merkmalsbildung wie oben beschrieben nur für die Extraktion von bebauten Gebieten zu akzeptablen Ergebnissen führt. Die Berechnung der Merkmalsbilder mit Gaborfiltern unterlag außerdem Beschränkungen in der Rechenkapazität, wie sie auch in Kapitel 6.3.2 beschrieben werden. Große Filtermasken führen zu erheblichen Rechenzeiten und Speicherplatzbedarf unter `MATLAB`. Deshalb wurde zur steuerbaren Pyramide übergegangen, wie in Kapitel 6 beschrieben.

Kapitel 4

Erzeugung eines Layers mit Texturinformation

Das Ziel dieser Arbeit ist, die multispektrale Klassifikation¹ durch Einführung eines Layers mit Texturinformation zu verbessern. Problematisch ist die Klassifikation aufgrund der spektralen Rückstrahlung vor allem in Bereichen mit stark wechselnder Helligkeit, die oft durch Schattenwurf verursacht wird. In diesen Bereichen kommt es dann häufiger zu Fehlzuweisungen in den Klassen. So hat das Verfahren Schwächen in der Klassifikation von Waldgebieten, und besonders die besiedelten Gebiete mit ihrer stark wechselnden Farbigekeit und Helligkeit führen zu stark inhomogenen Ergebnissen. Ein Texturkanal hat die Aufgabe, solche Bereiche starker Helligkeitsvariationen zu einem homogenen Farb- oder Grauwert umzuwandeln und damit die Homogenität des Klassifikationsergebnisses in diesem Bereich zu verbessern. Ziel ist also die Bildung von Merkmalen, die innerhalb der Bereiche homogen sind und damit eine einheitliche Textur anzeigen. Auf der anderen Seite müssen die Merkmale zwischen den Texturen unterschiedlich ausfallen, um diese abgrenzbar zu machen.

Die Erzeugung eines Layers mit Texturinformationen wurde mit verschiedenen Ansätzen durchgeführt und der erfolgversprechendste Ansatz ausgewählt (Ergebnisse der steuerbaren Pyramide, Kapitel 6.6) und in die Klassifikation eingeführt.

¹s. Kapitel 5.2.

4.1 Auswahl von Testbereichen

Um die Verfahren auf ihre Nutzbarkeit für die Texturanalyse beurteilen zu können, wurden sie auf Teile der Luftbildszene (s. Kapitel 1.3) angewendet, in denen sich verschiedene Texturen befinden. Um die Verfahren in Hinsicht auf ihre Diskriminanzleistung beurteilen zu können, kamen sowohl Ausschnitte des Bildes sowie aus Texturen zusammengesetzte Bilder zum Einsatz. Bei der Einteilung von Luftbildern anhand der Textur interessieren in erster Linie folgende Klassen:

- Siedlung
- Wald
- Felder

Für die zusammengestellten Testbilder wurden 128×128 Pixel große Teilbilder aus den aufgezählten Bereichen verwendet. Dabei zeigte sich schon die große Vielfalt an Texturen, die sich in einer Luftbildszene finden lassen. So gibt es z.B. für Felder stark variierende Ausschnitte: einige Felder weisen (meist parallele) Bearbeitungsspuren auf, die je nach Anbauverfahren einen bestimmten Abstand haben. An den Rändern der Felder finden sich Wendespuren der Fahrzeuge. Die Homogenität der Textur ist an diesen Stellen gering. In den Siedlungsbereichen variiert die Größe und der Abstand der Bebauung, Straßen und Freiplätze haben wechselnde Größe und Form. Auch in den Waldbereichen finden sich Inhomogenitäten in den Baumkronendurchmessern. Ein Texturmaß muß sich deshalb von Anfang an an den globaleren Merkmalen orientieren, ansonsten ist die Anzahl der Texturklassen so groß, daß sie sich nicht mehr eindeutig den interessierenden Klassen zuordnen lassen.

4.2 Der Ansatz nach Laws

Um einen Texturkanal zu erhalten wurde zunächst der von Laws vorgeschlagene Ansatz (Kapitel 2.4) verfolgt. Dazu wurden ausgewählte Testbilder mit Lawskerneln gefaltet. Die Kombination der 6 Kernel und die Filterung mit nur je einem

Kernel in einer Richtung ergibt 49 gefilterte Bilder. Diese vielen Kombinationen in verschiedenen Stufen einer Pyramide zu berechnen erhöht den Rechenaufwand zusätzlich. Durch die separabel durchführbare Berechnung mit den Lawskernen ergibt sich eine feste Richtungsselektivität in horizontaler und vertikaler Richtung. Die Ergebnisse sind nur für rotationssymmetrische Texturen übertragbar. Die Abbildung 4.1 zeigt einige Ergebnisse der Filterungen mit Lawskernen. Die Filterung erfolgte mit den Kombinationen von fünf Kernen, sie sind in der Reihenfolge L5, E5, S5, W5, R5 horizontal und vertikal miteinander kombiniert worden und die Ergebnisse zu dem Bild zusammengestellt.

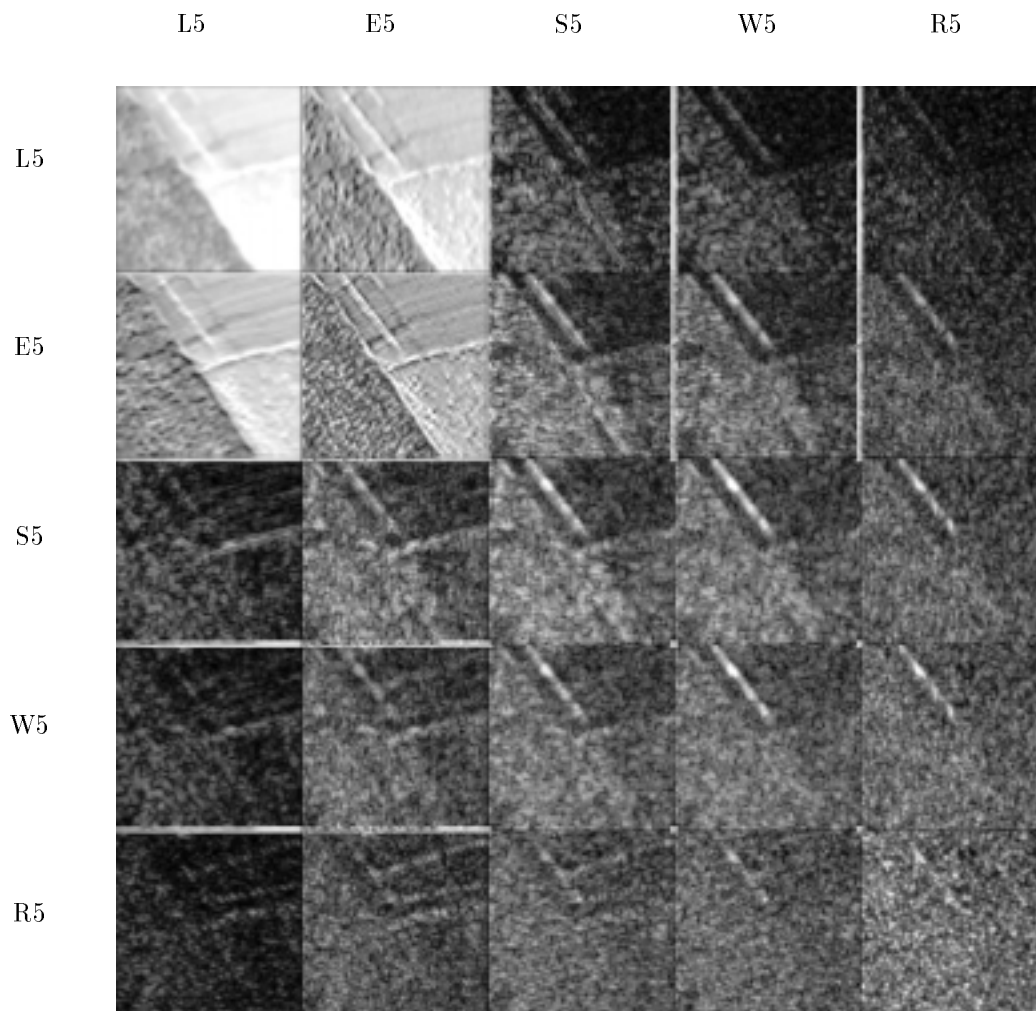


Abbildung 4.1: Filterungen nach Laws

Insgesamt läßt sich feststellen, daß die Suche nach passenden Kombinationen der Kernel sowie der Merkmalsbildung rechenintensiv und zeitintensiv ist. Eine gute Diskriminanzleistung für die hier interessierenden Texturen wird in den untersuchten Fällen nicht erreicht. Es bildeten sich in keinem der untersuchten Fälle homogene Merkmalsbereiche aus, die den Texturen zugeordnet werden können.

4.3 Gauß- und Laplacepyramiden

Die weit verbreitete Erzeugung von Gauß- und Laplacepyramiden wurde auch in dieser Arbeit untersucht. Eine Beschreibung der Funktionsweise findet sich in Kapitel 2.3. Die Gaußpyramide in Abbildung 4.2 für das zusammengesetzte Bild aus Abbildung 3.5 (S. 62) läßt gut die Wirkungsweise erkennen: Von Stufe zu Stufe findet eine Glättung statt, die hohen Frequenzen werden jeweils ausgefiltert. Die fein strukturierten Felder bilden trotzdem erst ab der 4. Stufe einigermaßen homogene Bereiche, die sich dort aber gut gegen die stark schwankenden Grauwerte der Siedlungsbereiche abheben. Eine Varianzberechnung kann diese Bereiche also voneinander trennen. Die Abgrenzung der Feldbereiche gegenüber den Waldbereichen ist hingegen schon schwieriger, da die Glättung für beide Bereiche ähnlich verläuft und nicht für die Felder in einer Stufe zu einem deutlich homogeneren Ergebnis führen, wie aufgrund der fehlenden tiefen Frequenzen zu erwarten wäre. Das Feld oben links mit seiner tiefen Frequenz (helle Streifen großen Abstands) tritt nicht wie zu erwarten in einer höheren Stufe deutlich hervor. Auffällig ist in der Pyramide das Verhalten der besiedelten Bereiche, die sich in allen Stufen durch stark wechselnde Grauwerte bemerkbar machen.

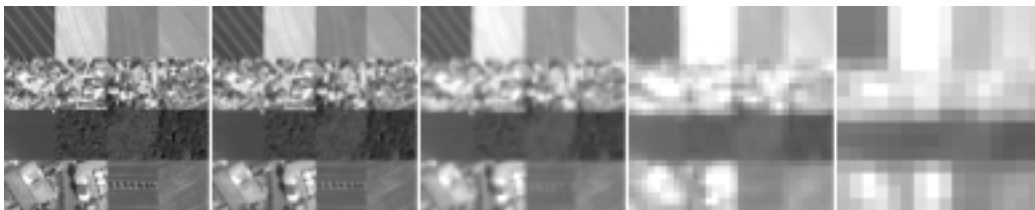


Abbildung 4.2: Gaußpyramide

Die Laplacepyramide desselben Bildes (Abbildung 4.3) zeigt deutlich die Fähigkeit zur Kantenbetonung (Gradientenwirkung des Kernels). Hier fallen ebenfalls die Siedlungsbereiche mit stark schwankenden Grauwerten in allen Stufen auf. Die Waldbereiche haben über die ersten drei Stufen verteilt eine gleichmäßige Struktur, sind jedoch insgesamt nicht sehr ausgeprägt, besonders im Vergleich mit den Feldbereichen.

Die Ableitung von robusten Merkmalen, die sich den oben genannten Klassen zuordnen lassen, ist aus diesen Pyramiden schwierig. Daher wurde versucht, mit der Einführung von Richtungsinformation Merkmale abzuleiten (Kapitel 6).



Abbildung 4.3: Laplacepyramide

4.4 Bildpyramiden mit orthogonalen Basen

Die in Kapitel 2.7 beschriebene Zerlegung liefert das Merkmalsbild in Abbildung 2.10 (S. 48). Andere orthogonale Zerlegungen, die mit verschiedenen Wavelets durchgeführt wurden, führten auf ähnliche Ergebnisse. Die Merkmalsbilder aus den Stufen wurden übereinandergelegt und klassifiziert. Ein Ergebnis ist in Abbildung 4.4 zu sehen. Weitere Klassifikationen, die für andere Kombinationen der Merkmalsbilder vorgenommen wurden, indem Auswahlen der Merkmalsbilder getroffen wurden, lieferten vergleichbare Ergebnisse. Die unüberwachte Klassifikation mit einem ISODATA-Klassifikationsalgorithmus erfolgte mit unterschiedlicher Anzahl vorgegebener Klassen. Das Bild zeigt das Klassifikationsergebnis für fünf Klassen. Eine eindeutige Zuordnung der Klassen zu den Objektklassen kann nicht vorgenommen werden, daher ist das Ergebnis nicht als Textur-layer brauchbar. Eine überwachte Klassifikation erbrachte auch keine Ergebnisse,

die eine Trennung der interessierenden Texturklassen zuließ. Ursache dafür ist, daß die Merkmale in keiner Zerlegungsstufe die gesuchten Texturklassen eindeutig differenzieren, wie sich schon im Merkmalsbild erkennen läßt.

Die einzige Klasse, die eine Zuordnung sinnvoll erscheinen läßt, ist die im Bild weiß dargestellte, die man im Siedlungsbereich findet. Die Überdeckung ist aber nicht kontinuierlich, was die Nutzung als Maske für diesen Bereich ausschließt. Maskierung ist eine Vorgehensweise, bei der das zu klassifizierende Bild nur in einem durch die Maske definierten Bereich klassifiziert wird. Deckt die Maske z.B. den Siedlungsbereich ab, so werden Fehlklassifikationen der Klasse Siedlung in den anderen Bereichen des Bildes ausgeschlossen.

Die Erzeugung des Texturlayers, der in die Klassifikation eingeführt wird, ist in Kapitel 6.6 beschrieben.

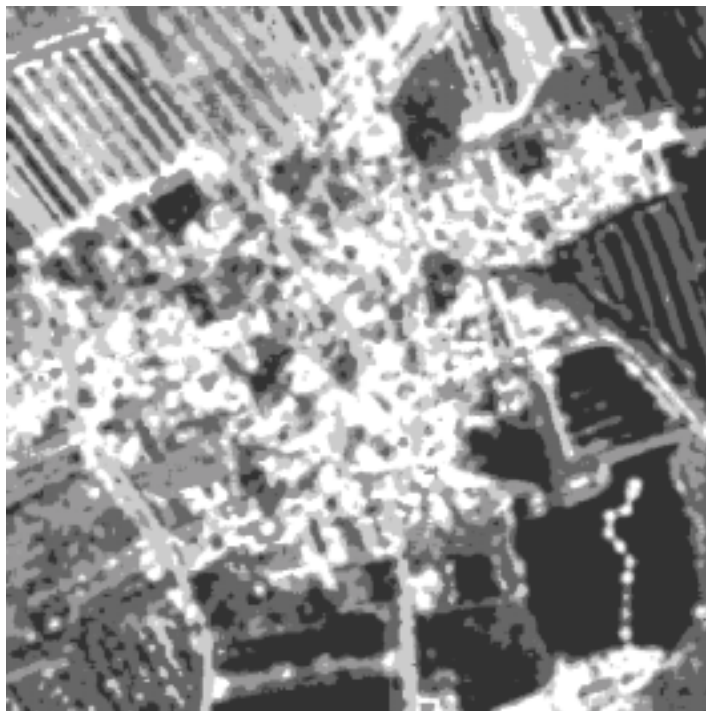


Abbildung 4.4: Texturklassen aus der Klassifikation einer orthogonalen Pyramidenzerlegung

Kapitel 5

Klassifikation

Die Klassifikation der Bildszene wurde unter ERDAS auf einer Silicon Graphics Workstation vorgenommen. Zunächst wird der generelle Ablauf einer Klassifizierung erläutert, dann werden die Untersuchungen im einzelnen beschrieben.

5.1 Ablauf der Klassifizierung

Das Programm ERDAS stellt alle benötigten Hilfsmittel für die Klassifikation von mehrkanaligen Bildern zur Verfügung. Die Schritte zur Durchführung einer Klassifikation sind im einzelnen:

1. Man öffnet das zu klassifizierende Bild in einem Viewer. Es muß ein von ERDAS unterstütztes Bildformat vorliegen.
2. In dem Viewer werden nun sogenannte AOIs (*Area Of Interest*) definiert. Die AOIs definieren die Trainingsgebiete der Klassifikation und sollten repräsentativ eine Klasse enthalten. Die AOIs können dann mit der Endung .aoi abgespeichert werden, um sie für weitere Klassifikationen zu verwenden.
3. Mit dem Tool *signature editor*¹ werden die Klassenspezifikationen aus den

¹Eine Abbildung der Oberfläche des signature editors findet man auf Seite 76 in Abbildung 5.2.

AOIs berechnet und die Klassen gebildet. Diese Klasseneinteilung kann ebenfalls abgespeichert werden (Endung .sig). Es besteht die Möglichkeit, Klassen zusammenzufassen und eine statistische Auswertung der gebildeten Klassen vorzunehmen. Dazu werden zwei Matrizen errechnet:

- Separability: Der Separability Matrix kann entnommen werden, in welchen Kanälen des Bildes die beste Unterscheidung der Klassen möglich ist. Mit ihrer Hilfe läßt sich der Beitrag der Kanäle zur Trennung der Klassen abschätzen. Ist die spektrale Signatur zweier Klassen in einem Band ähnlich, ergibt sich ein kleiner Wert für die Separability. Es ergab sich, daß alle Bänder verwendet werden sollten.
 - Contingency²: Der Contingency Matrix kann die Zuordnung von Pixeln der Trainingsgebiete zu den Klassen entnommen werden. Dazu werden die Pixel der Trainingsgebiete, die eine Klasse definieren, klassifiziert, und in der Matrix die prozentuale Verteilung der Pixel auf die Klassen eingetragen. Es lassen sich Hinweise auf schlecht trennbare Klassen und problematische Festlegung der Trainingsgebiete finden, wenn eine große Anzahl der Pixel eines Trainingsgebietes nicht der zugehörigen Klasse zugeschlagen werden.
4. Aus dem signature editor heraus wird die Klassifikation gestartet. Es gibt die Möglichkeit der unüberwachten Klassifikation (ohne Trainingsgebiete) und der überwachten Klassifikation mit den zuvor gebildeten Klasseneinteilungen. Das gesamte Bild wird klassifiziert. Jedes Pixel wird einer Klasse zugeordnet. Dabei kann die Klassenbegrenzung zu *maximum likelihood*, *Mahalanobis distance* und *minimum distance* gewählt werden. Dieser Parameter bestimmt die Form der Näherung für die Cluster im mehrdimensionalen Merkmalsraum. Ein Cluster umfaßt eine Häufung der kombiniert auftretenden Merkmale³. Diese Häufung wird als Charakteristik der Klasse genutzt, um in der Klassifikation die Pixel gleicher Charakteristik derselben Klasse zuzuordnen. Es wird die statistische Verteilung errechnet, deren Schwerpunkt und Standardabweichung zur Klassengrenzenbildung genutzt

²Contingency Matrizen befinden sich in Anhang A.

³Beispiele für Cluster im Merkmalsraum finden sich in Abb. 6.15 ff. auf S.97.

wird. Die Oberfläche der Klassenbegrenzung bildet eine Entscheidungsgrenze im Merkmalsraum. Sie umschließt den Entscheidungsraum. Der Klassifikationsalgorithmus findet die Position eines Pixels im Merkmalsraum. Fällt die Position in den Entscheidungsraum, wird es der Klasse zugeordnet (nach [41], S. 283 f.). Es wurde für alle Klassifikationen *maximum likelihood* gewählt. Der Entscheidungsraum nimmt dann die Form eines (Hyper-) Ellipsoides an.

5. Das Ergebnisbild der Klassifikation enthält farbige Pixel, die der Klassenzuordnung entsprechen. Die Zuordnung der Farbe zu den Klassen kann im signature editor festgelegt werden.

Damit ist die Klassifikation abgeschlossen. Die Wahl der Trainingsgebiete hat entscheidenden Einfluß auf das Klassifikationsergebnis, deshalb ist bei der Definition auf Einheitlichkeit und Repräsentanz zu achten. Für die untersuchte Szene lagen die Daten einer Feldbegehung zum Befliegungszeitpunkt vor (s. Anhang B), damit konnten gemischte Gebiete (z.B. Weizen, Rest Roggen) als Trainingsgebiete ausgeschlossen werden.

5.1.1 Erzeugung von mehrkanaligen Bildern

Um ein mehrkanaliges Bild aus mehreren Bildern zu erzeugen, müssen im layer stack tool die entsprechenden Layer aus den Bildern zusammengefügt werden. Diese Funktion wurde für das Hinzufügen des Texturkanales zu der Szene verwendet. Um Kanäle zusammenzufügen, müssen sie georektifiziert sein, also in demselben Koordinatensystem vorliegen. Da die Szene bereits auf Gauß-Krüger-Koordinaten rektifiziert ist, mußten die Texturkanäle, die nur in Pixelkoordinaten vorlagen, auf die Koordinaten der Szene angepaßt werden. Die Referenzierung kann mit Hilfe der Funktion *Layer-Info* aufgerufen und unter dem Menüpunkt *Edit* bearbeitet werden.

5.2 Ergebnisse der multispektralen Klassifikation

Die vorliegende Luftbildszene (Abbildung 1.1) wurde überwacht klassifiziert. Dazu wurden Trainingsgebiete für zehn Klassen definiert, die sich der Abbildung 5.2 auf Seite 76 entnehmen lassen. Das Klassifikationsergebnis ist in Abbildung 5.1 zu sehen. Der signature editor in Abbildung 5.2 ist eine Legende für die Zuordnung der Farben zu den Klassen.



Abbildung 5.1: Klassifikationsergebnis der multispektralen Klassifikation

Class #	Signature Name	Color	Red	Green	Blue	Value	Order	Count	Prob.	P	I	A	T
1	Wald1	Dark Green	0.200	0.332	0.000	3	3	20474	0.000	X	X	X	
2	Wald2	Light Green	0.226	0.431	0.039	4	6	55634	0.000	X	X	X	
3	Weizen	Yellow	1.200	0.843	0.000	9	13	47198	0.000	X	X	X	
4	Rüben	Red	1.047	0.106	0.165	10	14	23241	0.000	X	X	X	
5	Wiese	Bright Green	0.200	1.000	0.000	5	17	23737	0.000	X	X	X	
6	Gerste	Tan	0.924	0.776	0.549	7	21	26939	0.000	X	X	X	
7	Raps	Yellow	1.200	1.000	0.000	6	27	62247	0.000	X	X	X	
8	Wasser	Blue	0.000	0.000	1.000	8	27	22012	0.000	X	X	X	
9	Brache	White	1.200	1.000	1.000	1	20	13635	0.000	X	X	X	
10	Siedlung	Black	0.200	0.000	0.000	3	32	37039	0.000	X	X	X	

Abbildung 5.2: Der signature editor von ERDAS, Legende für Abb. 5.1

Zunächst kann mit Hilfe der Contingency Matrix⁴ die innere Genauigkeit der Klassifikation abgeschätzt werden. Da sich die Werte der Matrix nur auf die Trainingsgebiete beziehen, macht sie keine Aussage über die Richtigkeit der Klassifikation im restlichen Bild. Es kann aber die Homogenität der Trainingsgebiete überprüft werden. Werden eine große Anzahl der Pixel in einem Trainingsgebiet einer anderen Klasse zugewiesen, kann das zwei Gründe haben:

- Die Trainingsgebiete sind inhomogen, was auf eine große Standardabweichung in der Statistik führt. Der Cluster für diese Klasse fällt zum einen entsprechend groß aus und überdeckt Cluster anderer Klassen. Andererseits enthält ein inhomogenes Trainingsgebiet Pixel anderer Klassen, die in anderen Trainingsgebieten definiert werden. Ist die Definition der anderen Klassen enger, d.h. mit kleinerem Cluster, werden diese Pixel richtigerweise diesen Klassen zugeschlagen und zeigen die Inhomogenität auf.

Dieser Fall ist in der multispektralen Klassifikation in der Klasse Siedlung aufgetreten. Die Definition durch die Trainingsgebiete umfasste eine große Bandbreite multispektraler Signaturen, da von Straßen, Plätzen und Dächern unterschiedlicher Farbe große Farbvariationen eingebracht wurden.

⁴Die Contingency Matrix befindet sich in Anhang A auf Seite 116, große Werte abseits der Hauptdiagonalen sind unterstrichen.

Die Klasse faßt alle diese Objekte zusammen. Hier kann die Einführung separater Klassen für diese Objekte die Klassendefinition verbessern.

Ebenso weist die Klasse Wiese eine solche Inhomogenität auf, da sie ihrer Natur nach ein Mischgebiet mit unterschiedlichen Vegetationstypen (Gräser und Kräuter) ist. Die Klasse Wald1 umfaßt auch einen relativ großen Farbbereich, da hier verschiedene Baumarten zu finden sind. Bei der Klasse Wald2 handelt es sich um Aufforstungen, die sich nur durch ihr Alter von der Klasse Wald1 unterscheiden. Die größte Fehlzuweisung in der Klasse Wald1 findet sich dementsprechend aus der Klasse Wald2.

- Die Definition der Klassen ist nicht sinnvoll gewählt. Werden zwei Klassen angelegt, die ihrer Natur nach keine sinnvolle Unterscheidung bringen, ist mit Fehlklassifikationen zu rechnen. Eine Zusammenfassung dieser Klassen sollte überdacht werden.

Diesen Fall gab es bei der Aufteilung in die Klassen Wiese und Brache. Ein brachliegendes Feld hat ähnliche Vegetation wie eine Wiese. Es finden sich dort ebenfalls mehrere Vegetationsarten, da keine Bestellung stattfindet. Die Pflanzengesellschaft einer Wiese und einer Brache sind ähnlich, da beide aus holzfreien Vegetationsarten bestehen.

Die Interpretation der Klassifikation wird visuell vorgenommen. Das geschieht aus mehreren Gründen, die ich im folgenden skizziere. Eine Definition von Trainingsgebieten, die aussagekräftig sind, erfordert Erfahrung. Diese Erfahrung ist bei erstmaliger Durchführung einer Klassifikation nicht vorhanden. So ist z.B. die Einführung der beiden Klassen Wiese und Brache nicht sinnvoll gewesen und zeigt diese Unzulänglichkeit auf. Die Definition der Trainingsgebiete hat entscheidenden Einfluß auf die Klassifikationsgüte. Eine statistische Auswertung zeigt demnach in erster Linie die Güte in der Definition der Klassen durch die Trainingsgebiete.

Für eine statistische Auswertung, die aussagekräftig ist, müssen unabhängige Testgebiete definiert werden. Die Auswahl dieser unabhängigen Testgebiete hat wiederum Einfluß auf das statistische Ergebnis. Eine aussagekräftige Statistik erfordert die Einteilung der gesamten Szene, denn wenn nur Teile als Testgebiete

definiert werden, die eindeutig zu einer Klasse gehören, erhält man keine Information über die Grenzfälle. Eine statistische Aussage der Art *x% dieser Klasse wurden richtig klassifiziert* erfordert die Einbeziehung *aller* Flächen dieser Klasse. Die Abgrenzung der Flächen stellt dabei ein Problem dar, da diese Abgrenzung aus unabhängigem Wissen eingeführt werden muß. Wird diese Definition der Testgebiete nicht sorgfältig vorgenommen, so ist das statistische Ergebnis von geringem Wert. Eine quantitative Auswertung erfordert demnach wiederum viel Erfahrung in der Auswertung von Infrarotbildern, die nicht vorhanden war. Hinzu kommt, daß quantitative Untersuchungen für größere Szenen gemacht werden sollten, da einige Klassen nicht stark genug vertreten sind, um statistisch sichere Aussagen zu machen (z.B. Raps).

Mit einer visuellen Interpretation lassen sich qualitative Aussagen machen. Das Prinzip läßt sich erfassen. Das Augenmerk kann auf spezielle Fälle gerichtet werden, in denen die Klassifikation fehlschlägt. Da es letztendlich um den Vergleich der reinen multispektralen Klassifikation mit einer Klassifikation unter Zuhilfenahme von Texturinformation geht, ist die Ableitung von qualitativen Aussagen über eine Verbesserung oder Verschlechterung des Klassifikationsergebnis visuell möglich. Wichtig ist letztendlich, in welcher Weise die Texturinformation auf das Ergebnis wirkt. Diese qualitativen Aussagen sollen als Ergebnis getroffen werden.

Eine erste Überprüfung zeigt, daß die Klassenzuordnung in den Waldbereichen richtig ist. Wald1 und Wald2 werden, soweit aus dem Luftbild erkennbar ist, richtig getrennt. Vereinzelt werden dort allerdings Pixel der Klasse Weizen zugeschlagen. Die Klassen Weizen und Wald1 werden nicht immer richtig zugeordnet, Fehlklassifizierungen treten besonders in Weizenfeldern auf. Auf nahezu allen Weizenfeldern werden Pixel der Klasse Wald1 zugeordnet. Die Fehlklassifizierung ist vom Umfang her besonders ausgeprägt.

Eine andere Fehlklassifizierung tritt mit der Klasse Siedlung auf: Die Siedlungsbereiche werden zwar richtig zugewiesen, die einzelnen Straßen und Häuser sind gut von den umgebenden Vegetationsbereichen getrennt. Die Zuordnung der Bäume und Wiesen im besiedelten Bereich ist weitgehend richtig, vereinzelt finden sich Weizen zugeschlagene Pixel für Bäume. Allerdings kommt es zu

Fehlklassifizierungen in Feldbereichen, dort wird Siedlung klassifiziert.

Die angesprochene schlechte Trennung der Klassen Wiese und Brache zeigt sich im benachbarten Auftreten dieser Klassen.

Die Wasserflächen des Baggersees im unteren Teil des Bildes wurden nicht korrekt klassifiziert, der südliche See wird der Klasse Siedlung zugeordnet. Die spektrale Signatur gleicht der von schattierten Bereichen in Siedlungsbereichen.

Kapitel 6

Die steuerbare Pyramide

Mit den in Kapitel 2.3.1 beschriebenen Gauß- und Laplacepyramiden erhält man in den einzelnen Stufen keine Orientierungsinformationen. Das Konzept der steuerbaren Pyramide (*steerable pyramid*) erweitert die Zerlegung um diese Information. Die steuerbare Pyramide ist eine translationsinvariante mehrfachauflösende Dekomposition, die auch Orientierungen darstellt. Die einzelnen Stufen werden mit richtungsabhängigen Filtern zerlegt, die in Orts- und Frequenzraum lokalisiert sind. Man erhält k orientierte Bänder für jede Stufe, damit ist die Dekomposition mit dem Faktor $\frac{4k}{3}$ überabgetastet. Die Dekomposition hat außerdem den Vorteil, Aliasing zu vermeiden. Eine Rekonstruktion des Ausgangsbildes kann mit denselben Filtern durchgeführt werden, die man zur Dekomposition benutzt. Man spricht in diesem Fall von einer Transformation mit „engem Rahmen“ (*tight frame*) [35]. Allerdings hat die Rekonstruktion aus der Sicht der Texturanalyse keine große Bedeutung, wie in Kapitel 2.8.1 deutlich wird. Die Berechnung der Pyramide erfolgt mit den Pyramidentools unter MATLAB (s. Kapitel 6.2).

6.1 Funktionsweise der steuerbaren Pyramide

Simoncelli [34] untersucht den Zusammenhang von Verschiebungen des Signales (Bildes) in Position, Skale und Orientierung mit Änderungen des Transformationsergebnisses. Dafür definiert er Bedingungen, die Basisfunktionen jeweils erfül-

len müssen, um sich invariant gegenüber den Verschiebungen zu verhalten. Dabei bedeuten im einzelnen:

- **Translationsinvarianz:** Die Energien in den Bändern dürfen sich bei Verschiebungen des Signales nicht ändern.
- **Rotationsinvarianz:** Die Energie in den Bändern soll sich bei Rotation des Signales für den Rotationspunkt nicht ändern.
- **Skaleninvarianz:** Die Energien sollen sich bei Änderung der Skale nicht ändern.

Da alle drei Eigenschaften nicht gleichzeitig erreicht werden können, muß mindestens eine Bedingung zur Erhaltung der anderen gelockert werden. Eine Möglichkeit ist, in einer dieser Eigenschaften die gesamte Auflösung zu erhalten (keine sukzessive Auflösung).

Bei der steuerbaren Pyramide handelt es sich um eine zweidimensionale Transformation, deren Basisfunktionen durch Translation, Rotation und Dilatation eines einzigen Kernels erzeugt werden. Sie ist als rekursive Pyramide angelegt. Anders als die orthogonalen Pyramiden erzeugt diese Pyramide eine redundante Darstellung.

In der steuerbaren Pyramide soll das Signal in Tiefpaß- und Bandpaßanteile zerlegt werden. Die Konstruktion unterscheidet sich etwas von den klassischen Pyramidenkonstruktionen (s. Kapitel 2.3.1 und 2.7.2). Das Signal wird mit einem Tiefpaß $L_1(\omega)$ und einem Bandpaß $B(\omega)$ gefaltet. Um Aliasing zu vermeiden wird das bandpaßgefilterte Ergebnis keinem Subsampling unterworfen. Das tiefpaßgefilterte Ergebnis wird einem Subsampling mit dem Faktor 2 unterworfen und durchläuft einen zweiten Tiefpaß $L_0(\omega)$. Die Antwort des gesamten Systems läßt sich dann schreiben als:

$$S(\omega) = |B(\omega)|^2 + |L_1(\omega)|^2 |L_0(\omega)|^2 \quad (6.1)$$

Die bisherige Zerlegung kann graphisch wie in Abbildung 6.1 dargestellt werden. Die Impulsantwort des gesamten Systems (Abb. 6.1) gleicht der Impulsant-

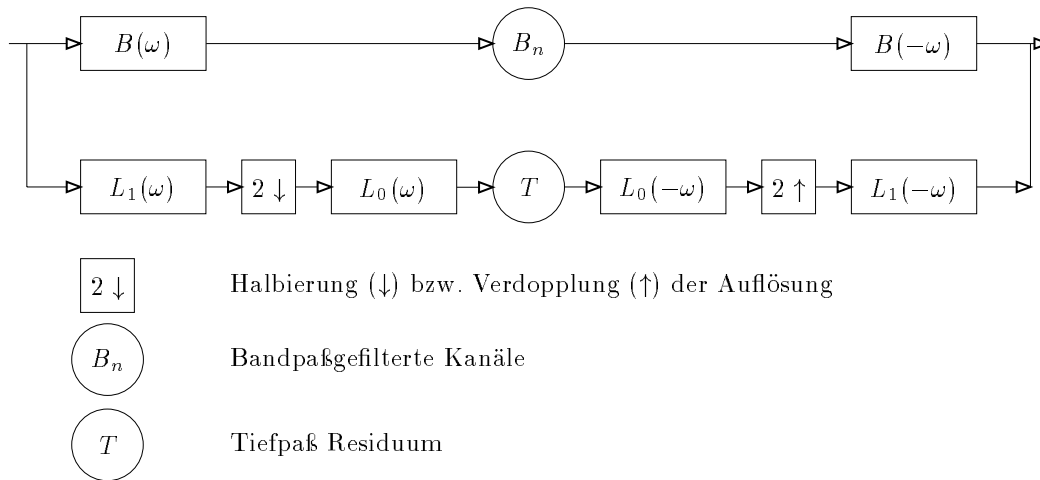


Abbildung 6.1: Zerlegungsschema einer Stufe der steuerbaren Pyramide

Abbildung 6.2: Filter $L_0(\omega)$

wort des Filters $L_0(\omega)$ (Abbildung 6.2, s. auch Gleichung 6.2). Daher kann das System rekursiv angewendet werden, wie in Abbildung 6.3 graphisch aufgetragen ist. Der Filter $L_1(\omega)$ soll als Tiefpaßfilter die Forderung nach der Antwort 0 auf dem Intervall $\omega = [\frac{\pi}{2}, \pi]$ erfüllen. Diese Forderung muß nicht streng eingehalten werden, da die zweite Tiefpaßfilterung die meisten verbleibenden Aliaskomponenten herausfiltert. Es wird ein Binominalfilter der Länge 7 benutzt ($\frac{1}{16}[16\ 15\ 20\ 15\ 6\ 1]$). Der Filter $L_0(\omega)$ bzw. die Gesamtantwort $S(\omega)$ wird optimiert auf den Wert 1 im Intervall $\omega = [0, \frac{\pi}{2}]$ und den Wert 0 bei π . Ein Filter der Länge 13 wurde gewählt. Der Bandpaßfilter ist nun durch folgende Beziehung festgelegt.

$$|L_0(\omega)|^2 = |B(\omega)|^2 + |L_1(\omega)|^2 |L_0(2\omega)|^2 \quad (6.2)$$

Um richtungsselektive Filter zu erhalten, wird das im Frequenzraum ringförmige Bandpaßspektrum in Orientierungsbänder aufgeteilt. Die inverse Fouriertransformation führt dann auf die Basisfilter im Ortsraum. Nach der Anwendung

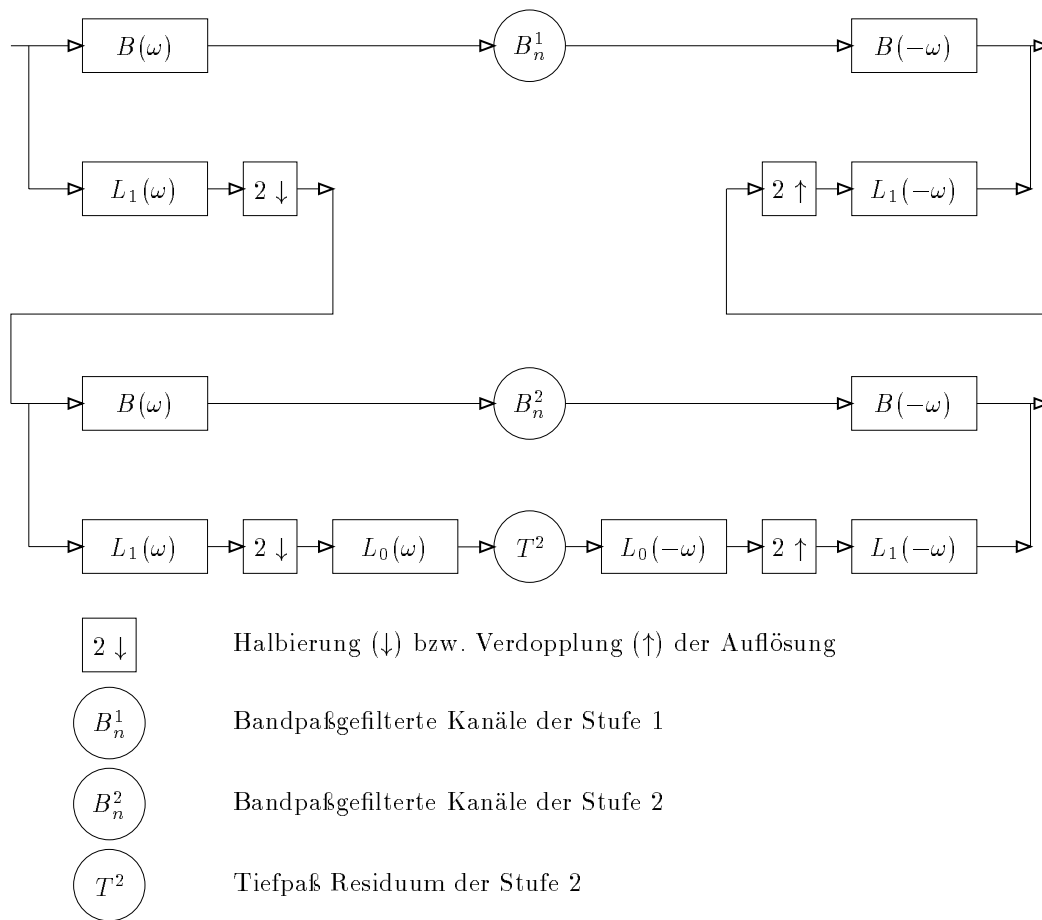


Abbildung 6.3: Zerlegungsschema der steuerbaren Pyramide

der orientierten Bandpaßfilter wird kein Downsampling durchgeführt. Zusammen mit dem aliasingfreien Entwurf der Tiefpässe ergibt sich für die Teilbänder räumliche Verschiebungsinvarianz. Der Entwurf der radialen und der Winkelkomponenten im Frequenzraum stellt sicher, daß die Summe der Quadrate der Basisfunktionen konstant ist in Hinblick auf die Orientierung und relevante Skale (nach [34]).

Zur Veranschaulichung werden in Abbildung 6.4 die Impulsantworten von 4 orientierten Bandpaßfiltern und in Abbildung 6.5 die Filter im Frequenzraum dargestellt:

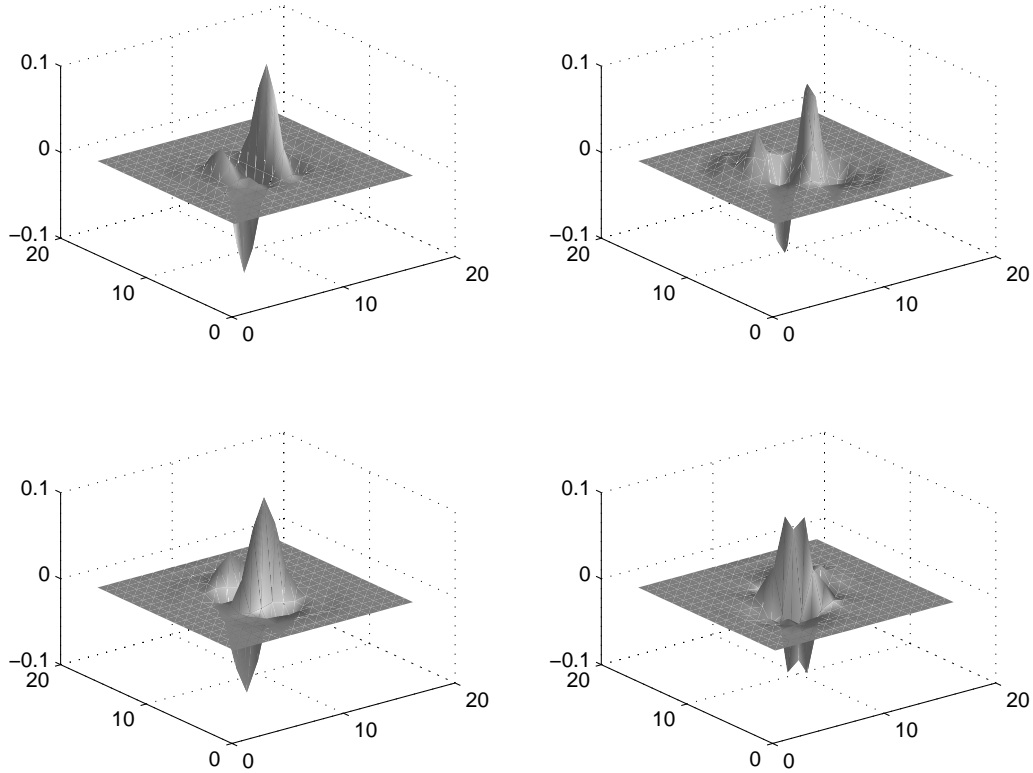


Abbildung 6.4: Orientierungselektive Filter der steuerbaren Pyramide

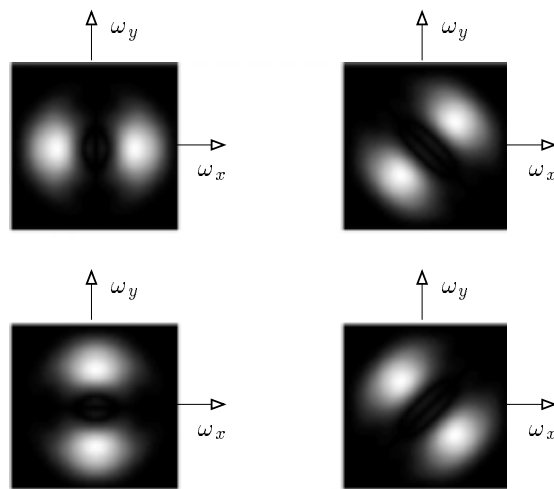


Abbildung 6.5: Die steuerbaren Filter im Frequenzraum

6.2 Die Pyramidentoolbox für MATLAB

Die Pyramidentoolbox für MATLAB ist frei über das Internet erhältlich [35]. Mit der Toolbox wird die Erzeugung verschiedenster Pyramiden unterstützt. Es sind Funktionen enthalten, die eine Zerlegung eines Bildes in eine Pyramide sowie seine Rekonstruktion erlauben. Die Pyramiden, die erzeugt werden können, sind im einzelnen:

- Gauß
- Laplace
- QMF¹
- Wavelet
- Steuerbare

Es stehen eine ganze Reihe Filter zur Verfügung, die für die Dekomposition benutzt werden können. Die Nutzung eigener Filter wird unterstützt. Die Funktionen sind modular aufgebaut und können in die folgenden Klassen eingeteilt werden, in denen eine kurze Beschreibung der wichtigsten Funktionen vorgenommen wird:

1. Funktionen zum Erzeugen einer Stufe oder einer kompletten Pyramide
 - `corrDn`: Berechnet die Korrelation der Bildmatrix mit einem Filter und führt ein Downsampling durch. Mit dieser Funktion wird eine Pyramidenstufe erzeugt. Es läßt sich die Randbehandlung angeben.
 - `buildGpyr`, `buildLpyr`, `buildWpyr`, `buildSpyr`, `buildSFpyr`: Funktionen, die eine Pyramide erzeugen. G steht für Gauß, L für Laplace, W für Wavelet oder QMF, S und SF für steuerbare Pyramiden (S – im Ortsraum berechnet, SF – im Frequenzraum berechnet). Die Pyramide wird in einem Vektor gespeichert, die Information über die Größe und Anzahl der Stufen in einer Index-Matrix.

¹Quadrature Mirror Filters, s. auch Kapitel 3.1.2, S. 53.

2. Funktionen zur Rekonstruktion einer Stufe oder des gesamten Bildes aus der Pyramide. Die zur Rekonstruktion verwendeten Stufen können explizit angegeben werden.
 - **upConv**: Mit dieser Funktion wird eine Bildmatrix durch Resampling auf die Größe der nächstgrößeren Stufe gebracht und mit einem angegebenen Filter gefaltet. Sie ist das Gegenstück zu **corrDn** und erzeugt aus einer höheren Stufe der Pyramide die nächstniedrigere. Hier läßt sich ebenfalls die Randbehandlung angeben.
 - **reconLpyr**, **reconWpyr**, **reconSpyr**, **reconSFpyr**: Funktionen zur Rekonstruktion des Bildes aus der Pyramide. Die **recon**- Funktionen bilden das Gegenstück zu den **build**- Funktionen. Es können explizit die Stufen angegeben werden, die zur Rekonstruktion verwendet werden sollen.
3. Funktionen, mit denen die einzelnen Pyramidenstufen angesprochen werden können.
 - **pyrBand**: erlaubt den Zugriff auf einzelne Teile einer Pyramide.
 - **wpyrBand**, **spyrBand**: wie **pyrBand**, es kann die Stufe und das Band innerhalb einer Stufe angegeben werden.
4. Funktionen, mit denen eine graphische Ausgabe erzeugt werden kann.
 - **showLpyr**, **showSpyr**, **showWpyr**: Stellt die angegebene Pyramide mit allen Stufen dar. Es wird eine automatische Skalierung der Grauwerte auf den Grauwertebereich [0 255] vorgenommen. Die einzelnen Stufen werden nach einem festen Schema angeordnet.
 - **showIm**: stellt eine Matrix dar, nachdem in den Grauwertebereich [0 255] skaliert wurde und gibt den Minimum- und Maximumwert der Matrix sowie ihre Dimension aus.
 - **imStats**: Erzeugt eine Statistik eines Bildes (Minimum, Maximum, Mittelwert, Standardabweichung, Schiefe der Verteilung) oder vergleicht zwei Bilder durch eine Statistik der Differenz miteinander.

6.3 Berechnungen mit der steuerbaren Pyramide

6.3.1 Funktionalitäten unter MATLAB

Die Berechnung von Merkmalsbildern geschah unter MATLAB, einem auf Matrizenrechnung basierenden Programm. Im Programm sind standardmäßig Ein- und Ausgabeoperationen für Bilder implementiert. Es liegt ferner die Image Processing Toolbox vor, mit der weitere Funktionalitäten zur Verfügung gestellt werden. Ein Bild wird als Matrix eingelesen und es können über Matrizenoperationen Faltungen und andere Matrizenoperationen durchgeführt werden. Die Matrizen in Matlab sind einfach zu handhaben, ihre Elemente sind je nach Erforderlichkeit reell oder komplex und können jederzeit auf Festplatte gesichert und von dort wieder eingelesen werden.

6.3.2 Probleme bei den Berechnungen unter MATLAB

Leider ist die Darstellung als Matrize sehr speicherintensiv, so daß Berechnungen mit großen Bildern nicht möglich waren, da sie mit einer Fehlermeldung (out of memory) abgebrochen wurden. Die Speicherverwaltung von Matlab ist nicht optimal, da der Speicher fragmentieren kann, wenn viele Variablen angelegt werden. Aus diesem Grund wurde die Berechnung der Pyramiden auf Ausschnitten der Szene vorgenommen, die nach der Erzeugung der Merkmalsbilder wieder zusammengefügt wurden. Da Matlab eine Scriptsprache zur Verfügung stellt, die recht einfach zu erlernen ist, kann eine Programmierung ohne allzugroßen Einarbeitungsaufwand vorgenommen werden. Allerdings ist die Arbeitsgeschwindigkeit der Scripte langsam, wenn größere Schleifen durchlaufen werden, wie das z.B. für die Merkmalerzeugung mit der tanh- Nichtlinearität² der Fall war, so daß erhebliche Rechenzeiten in Kauf zu nehmen waren. Da MATLAB die Möglichkeit bietet, C-Source mit einem plattformspezifischen Compiler zu kompilieren und die entstehenden Programme (oder DLLs³ unter Windows) als Funktionen in MATLAB zur Verfügung zu stellen, können solche rechenintensiven Schritte vorteil-

²s. Kapitel 3.3.2.

³Dynamic Link Library.

haft in C/C++ programmiert werden. Die Pyramidentoolbox enthält auch einige solcher Sourcen, allerdings gestaltete sich die Einbindung problematisch, da sie offensichtlich auf einer anderen Plattform entwickelt wurden und unter Windows zu von mir nicht behebbaren Fehlermeldungen beim kompilieren und linken führten oder die DLLs bei der Ausführung Laufzeitfehler erzeugten. Die Berechnung fand auf einem PC statt, der mit 128MB RAM und einem Prozessor mit 400MHz ausgestattet ist. Im Laufe der Berechnungen wurde zur Erweiterung des Speichers teilweise eine sehr große Auslagerungsdatei angelegt (bis zu 700MB), ein Hinweis auf den enormen Speicherplatzbedarf von **MATLAB**. Um den Speicher zu verwalten, gibt es den Befehl **pack**, der alle aktuellen Variablen auf der Festplatte speichert und sie dann wieder einliest, um das Fragmentierungsproblem des Speichers zu beheben. Allerdings funktioniert dieser Befehl nicht innerhalb von Schleifen, so daß teilweise eine umständliche schleifenlose Programmierung vonnöten war und die Speicherverwaltung mitprogrammiert werden mußte. Die Variable, in der die Pyramide eines 1024×1024 Pixel großes Bild gespeichert wurde, hatte eine Größe von 75MB und mußte regelmäßig mit dem **clear**-Befehl aus dem Speicher entfernt und vor der Benutzung wieder eingelesen werden.

Als Fazit dieser doch recht hinderlichen Beschränkungen bleibt die Notwendigkeit einer besseren Implementierung, wenn die Berechnungen auf die sehr großen Bilddatenmengen, wie sie bei Luft- und Satellitenbildern vorliegen, angewendet werden sollen.

6.4 Berechnungen mit den Pyramidentools

Um Merkmalsbilder aus dem Graustufenbild zu erhalten, wurde folgendermaßen vorgegangen:

- Aufteilung der Szene in 1024×1024 Pixel große Teilbilder.
- Berechnung der Pyramiden dieser Teilbilder: Es wurden 5 Stufen erzeugt, die jeweils in 6 Orientierungsbänder zerlegt wurden.
- Berechnung der Energie in den einzelnen Teilen der Pyramide: Hier erwies

sich die tanh-Nichtlinearität als am besten geeignet, da andere Energietransformationen⁴ (Absolutwert, Quadrierung, Varianz) nicht zu einheitlichen Grauwertebereichen führten. Die Fenstergröße zur Berechnung der Nichtlinearität wurde für die erste Stufe zu 7×7 Pixel, für die zweite Stufe zu 5×5 Pixel und für die dritte Stufe zu 3×3 Pixel gewählt.

- Die Energiebilder bilden den Merkmalsraum für die Klassifikation und wurden deshalb auf Originalgröße skaliert (1024×1024 Pixel) und als TIFF⁵-Bilder gespeichert. Dafür mußte eine Skalierung auf den Grauwertebereich $[0 \ 255]$ vorgenommen werden, da die tanh-Nichtlinearität Werte auf dem Intervall $[0 \ 1]$ liefert (s. Kap. 3.3.2, Abb. 3.4).
- Die so entstandenen Bilder wurden pro Stufe und Orientierung zusammengesetzt, so daß für die Szene Merkmalsbilder in 5 Stufen und 6 Orientierungen vorlagen.

Die durch Filterung entstandenen Pyramidenstufen (s. Kapitel 6.1) enthalten Werte, die um 0 verteilt sind. Zur Darstellung als Bild muß eine Skalierung auf den Grauwertebereich $[0 \ 255]$ erfolgen. Dem Wert 0 wird also ein Grauwert zugewiesen. Ein Beispiel dafür findet sich in Abb. 6.6. Es handelt sich um die erste Stufe der Pyramide mit Orientierung 3 (Ausschnitt der Größe 1024×1024 Pixel aus der Szene, Abb. 3.5). Das zugehörige Merkmalsbild wird in Abb. 6.7 dargestellt. Die anderen Orientierungen werden gegen den Uhrzeigersinn gezählt und teilen den Bereich $[0 \ \pi]$ gleichmäßig. Es zeigte sich, daß die Merkmalsbilder bis einschließlich der dritten Stufe brauchbar waren, in den höheren Stufen ist die Ortsauflösung zu gering, um Texturmerkmale darstellen zu können (s. auch Kapitel 3.2). Die Ortsauflösung wird durch die Wahl der Maske bestimmt, die für die Berechnung der tanh-Nichtlinearität verwendet wurde. Sie variiert über die Stufen. In der ersten Stufe ergibt sich eine Ortsauflösung von $7px \times 0.77m = 5.39m$, für die zweite Stufe $5px \times 1.54m = 7.7m$ und für die dritte Stufe $3px \times 3.08m = 9.24m$. In der ersten Stufe sind die hochfrequenten Anteile enthalten, wie am Ortsbereich in der Mitte des Bildes deutlich wird.

⁴s. auch Kapitel 3.3.2.

⁵Tagged Image File Format, ein weit verbreitetes Bildformat.

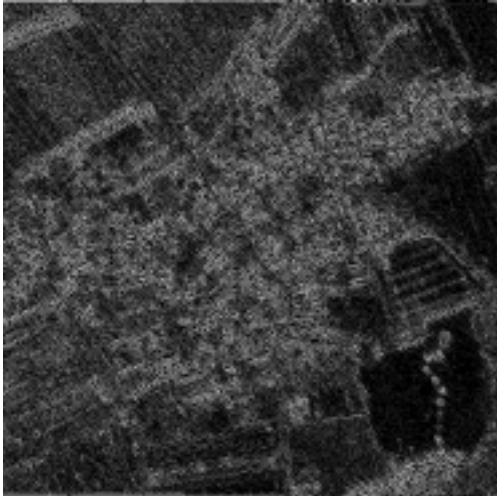


Abbildung 6.6: Pyramidenstufe 1, Orientierung 3



Abbildung 6.7: Das zu Abb. 6.6 gehörige Merkmalsbild



Abbildung 6.8: Merkmalsbild der Pyramidenstufe 1, Orientierung 6

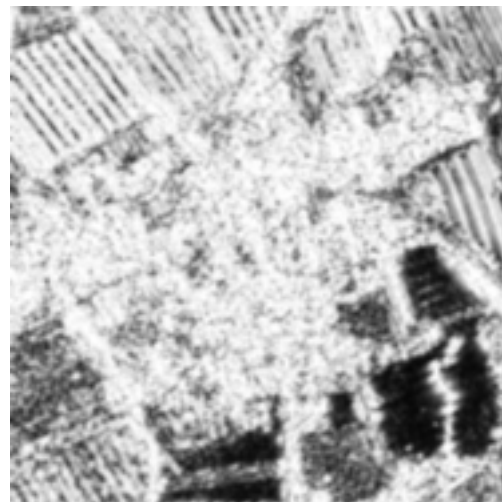


Abbildung 6.9: Merkmalsbild der Pyramidenstufe 2, Orientierung 6

In den Abbildungen 6.8 und 6.9 sind die Merkmalsbilder der ersten und zweiten Stufe mit der Orientierung 6 dargestellt. Durch Vergleich der Abbildungen 6.7 und 6.8 erkennt man, wie sich die Orientierung in den Merkmalsbildern ausprägt. Im Merkmalsbild der zweiten Stufe (ebenfalls mit Orientierung 6, Abb. 6.9) werden die tieferen Frequenzen detektiert, wie an den Feldern gut nachzuvollziehen ist: Die Felder mit größerem Abstand der Bearbeitungsspuren haben

hier fast kontinuierliche Merkmalsausprägung. Diese Beispiele zeigen, daß mit der vorgestellten Vorgehensweise Orientierungsmerkmale erhalten werden können. Wie sich die Orientierungsinformation in Luftbildern nutzen läßt, soll in den folgenden Betrachtungen geklärt werden.

6.5 Nutzen der Orientierungsinformation

Wie schon festgestellt, ist in Luftbildern eine große Bandbreite an Texturen zu finden. Aus dem vorgestellten Ansatz werden Orientierungsinformationen erhalten. In welcher Weise bilden diese Informationen einen brauchbaren Merkmalsraum für die Texturen in Luftbildern?

Da die Texturinformation die multispektrale Information ergänzen soll, ist es sinnvoll, zunächst die Ergebnisse der multispektralen Klassifikation zu interpretieren, um Problemfälle eingrenzen zu können. Für Objektarten, die sich multispektral hinreichend unterscheiden, brauchen keine weiteren Merkmale gefunden werden. Das ist der Fall für alle Klassen, die in Feldbereichen zu finden sind (s. Kapitel 5.2). Die Einführung von zu vielen Merkmalen hat sogar einen negativen Effekt auf die Klassifikationsgüte, der unter dem Namen *curse of dimensionality* bekannt ist [42]. Mit steigender Anzahl der Merkmale kommt es zur Überdeckung signifikanter Merkmale durch weniger signifikante. Anzustreben ist daher die Ableitung von wenigen, aussagekräftigen und robusten Merkmalen. Der Effekt zeigte sich bei den Untersuchungen in der Tatsache, daß die Klassifikation aus allen sechs Orientierungslayern zu wenig homogenen und aussagekräftigen Klassen im Texturlayer führte. Die Reduktion auf drei Orientierungslayer brachte zum einen ein homogeneres Klassifikationsergebnis für die Texturklassen, zum anderen ergab sich die Möglichkeit, durch eine zweite, unabhängige Klassifikation mit den verbleibenden drei Richtungen die erstere zu überprüfen und die Robustheit des Verfahrens zu prüfen. Sollte sich herausstellen, daß die unabhängigen Klassifikationen auf verschiedene Ergebnisse führen, so muß an der Robustheit des Verfahrens gezweifelt werden. Denkbar ist zum Beispiel, daß die zufällig vorgenommene Wahl der Richtungen das Ergebnis der Klassifikation entscheidend beeinflusst, wenn in der vorliegenden Szene bestimmte Richtungen vorherrschend sind, die mit den

gewählten Richtungen zusammenfallen. Das unabhängige Klassifikationsergebnis zeigt, daß nur wenige Änderungen in der Klassenzuordnung auftreten (s. Abbildungen 6.19 und 6.20, S. 100 und 100). Damit kann die Vorgehensweise als robust gegenüber Rotationen der gesamten Szene bezeichnet werden.

Die multispektrale Klassifikation wird in Kapitel 5.2 ausführlich behandelt. Ihre Ergebnisse haben Einfluß auf die Ableitung der Texturmerkmale.

Nach den bisherigen Ergebnissen (Gaborfilterung Kap. 3.2, Lawssche Filtermasken, Gauß- und Laplacepyramiden Kap. 4) ist nur die Trennung von Siedlungsbereichen möglich, da hier die größten Helligkeitsvariationen auftauchen, die bei den meisten Filterungen zu einem großen Energiewert führen. Andere Textureigenschaften lassen sich nicht in den Merkmalen trennen.

Das Merkmal der Orientierung hat in Luftbildern besondere Bedeutung, da sich die Klassen Siedlung, Wald und Felder genau darin unterscheiden: Siedlung und Wald haben Frequenzen in jeder Richtung, Felder aufgrund ihrer Bearbeitungsspuren in einer bevorzugten Richtung. Siedlung und Wald lassen sich zum einen spektral gut voneinander unterscheiden, zum anderen weist Wald eine geringere Amplitude der Frequenzen auf. Wie die folgenden Ergebnisse zeigen, lassen sich aufgrund dieser einfachen Einordnung drei Texturklassen bilden.

6.6 Erzeugung eines Texturlayers aus den Merkmalsbildern

Anhand eines Testbildes soll die Vorgehensweise zur Erzeugung des Texturlayers erklärt werden: Das Testbild in Abbildung 6.10 besteht aus einer radial vom Mittelpunkt verlaufenden Sinusfunktion, deren Frequenz nach außen hin abnimmt. Dieses Bild wurde der Pyramidenzerlegung mit anschließender Merkmalsbildung unterworfen. Am Ergebnis kann man sowohl die Frequenz- als auch die Orientierungsselektivität der Zerlegung erkennen. Die Zerlegung wurde auch hier in sechs Orientierungen vorgenommen, von denen drei in Abbildung 6.12 auf Seite 94 dargestellt sind.

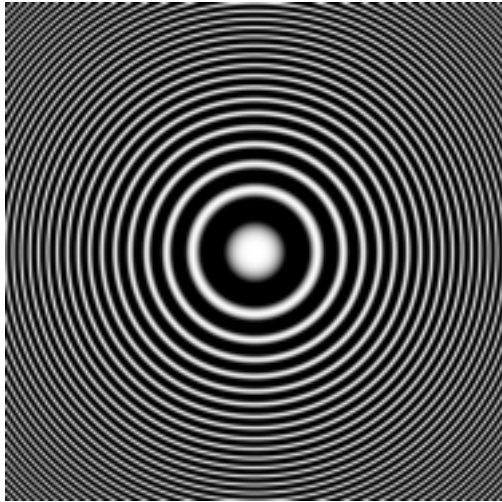


Abbildung 6.10: Testbild

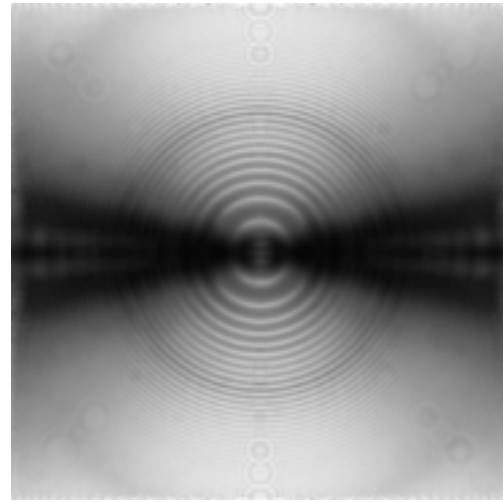


Abbildung 6.11: Kombination der Merkmalsbilder einer Orientierung aus den ersten drei Stufen der Pyramide

Aus den dargestellten neun Merkmalsbildern werden jeweils die drei gleicher Orientierung zusammengefaßt. Das geschieht über eine Hauptachsentransformation (PCA, s. a. Kapitel 1.3). Das Ergebnis sind drei Merkmalsbilder, die jeweils die Orientierungsinformation aus drei Stufen in sich vereinigen. Eines der Bilder ist in Abbildung 6.11 dargestellt.

Die drei Orientierungen können nun zu einem Farbbild zusammengesetzt werden, indem jeder von ihnen eine der Farben Rot, Grün und Blau zugewiesen werden. Im resultierenden Bild ist die Orientierung dann als Farbe kodiert, wie Abbildung 6.13 zeigt.

Mit einer überwachten Klassifizierung (Funktionsweise s. Kapitel 5) werden dann aus diesem Orientierungsbild drei Klassen gebildet (Wald, Siedlung, Felder). Das Ergebnis der Klassifikation ist ein Texturlayer, der das Bild aufgrund seiner Textureigenschaften segmentiert.

Den gesamten Ablauf zeigt Abbildung 6.14

Als Trainingsgebiete kamen die zuvor für die multispektrale Klassifikation verwendeten zum Einsatz. Diese Vorgehensweise ist von Vorteil, da zum einen der

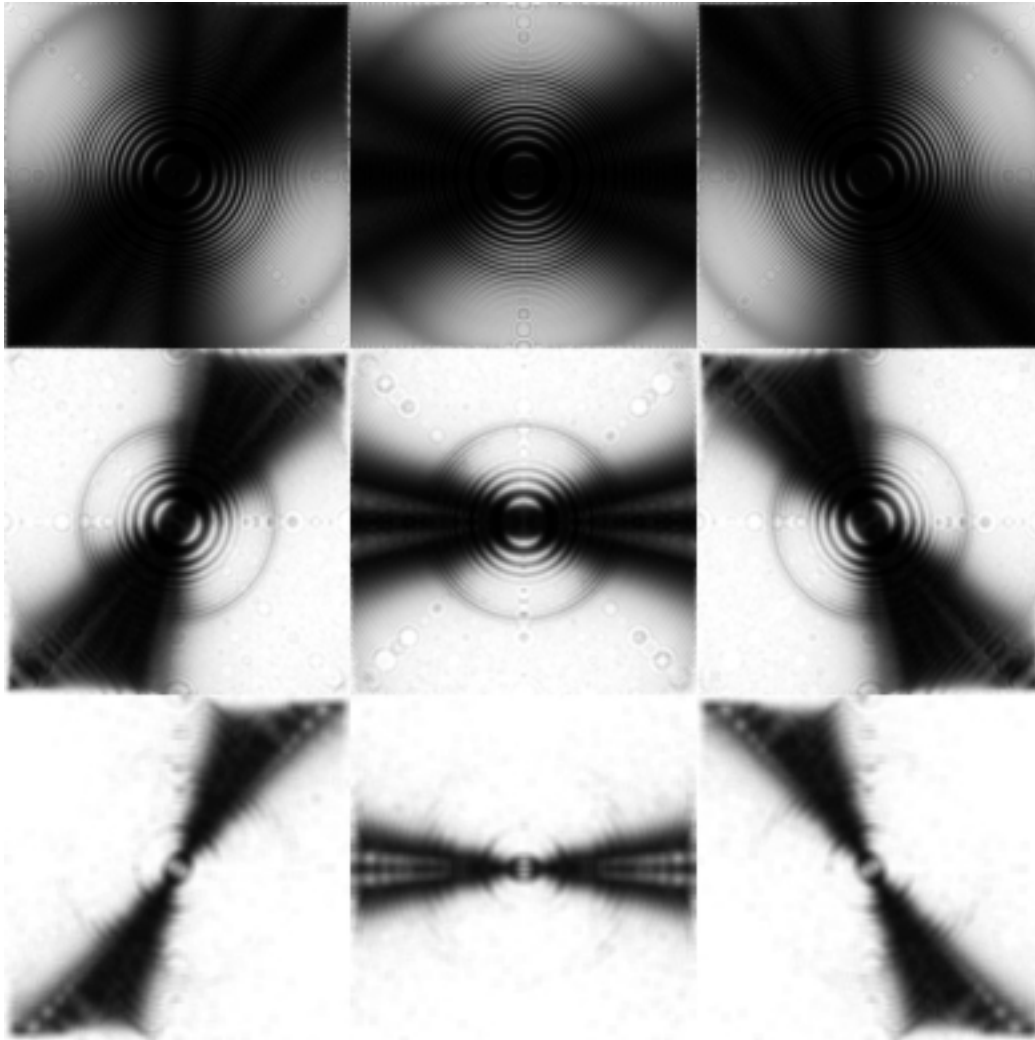


Abbildung 6.12: Die pyramidale Zerlegung des Testbildes 6.12

Aufwand einer erneuten Definition von Trainingsgebieten wegfällt, zum anderen kann keine Verfälschung des Ergebnisses durch die Wahl neuer Trainingsgebiete auftreten. Für die überwachte Klassifizierung wurde aus allen Waldgebieten eine Signatur errechnet, ebenso wurden die Siedlungsbereiche zusammengefaßt und die verbleibenden Trainingsgebiete bildeten die dritte Signatur. Diese dritte Klasse hat besondere Eigenschaften: In ihr wird das Merkmal *eine Richtung bevorzugt* zusammengefaßt, das sich im Farbbild durch *eine bevorzugte Farbe* ausdrückt. Der genaue Farbwert ist nicht von Bedeutung, da nicht die spezielle Orientierung interessiert, sondern nur die Eigenschaft, überhaupt eine bevorzug-

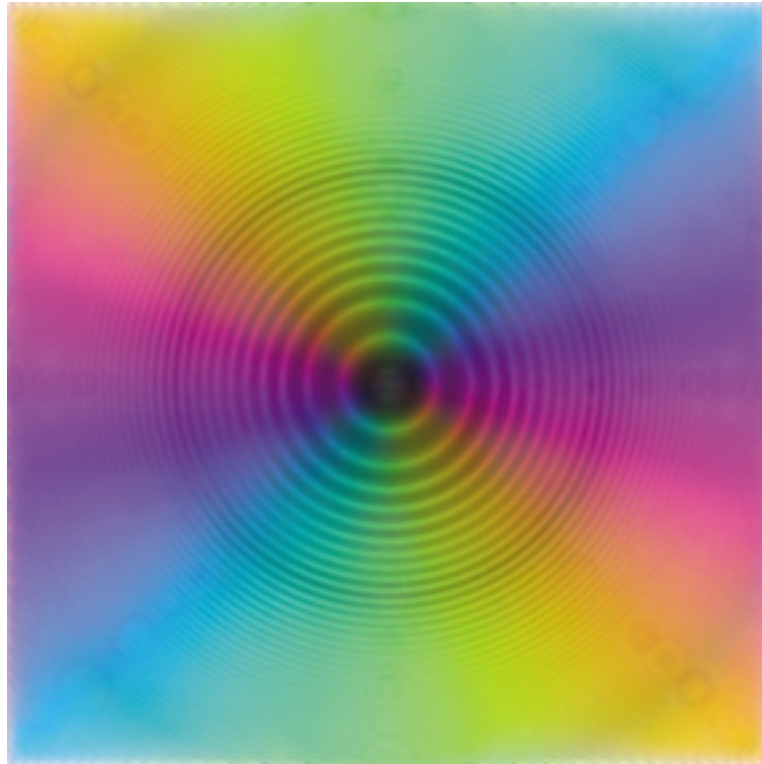


Abbildung 6.13: Farbiges Orientierungsbild

te Orientierung aufzuweisen. Farbigkeit entspricht im Merkmalsraum der Lage abseits der Unbuntgeraden. Damit sind hier die Klassengrenzen sehr weit gefaßt, sie umfassen den gesamten Merkmalsraum außer einer Umgebung der Unbuntgeraden. Die beiden anderen Klassen Wald und Siedlung liegen dann in einer Umgebung der Unbuntgeraden, da hier in allen Orientierungen zu in etwa gleichen Teilen Merkmale vorliegen. In Farben ausgerückt bedeutet diese Lage, daß alle drei Grundfarben zu gleichen Teilen vorhanden sind und in der Addition im Farbbild zu einem Grauwert führen. Im Farbbild erkennt man sofort, daß sich die Farbwerte im Bereich der Siedlungsgebiete zu Weiß addieren.

Tatsächlich läßt sich diese Eigenschaft durch die Betrachtung des Merkmalsraumes bestätigen. Die Merkmalsräume für die Trainingsgebiete Wald, Siedlung und Feld sind in den Abbildungen 6.15, 6.16 und 6.17 dargestellt. Da der Merkmalsraum dreidimensional ist (entsprechend den drei Grundfarben Rot, Grün, Blau), werden die zweidimensionalen Projektionen auf die Achsebenen darge-

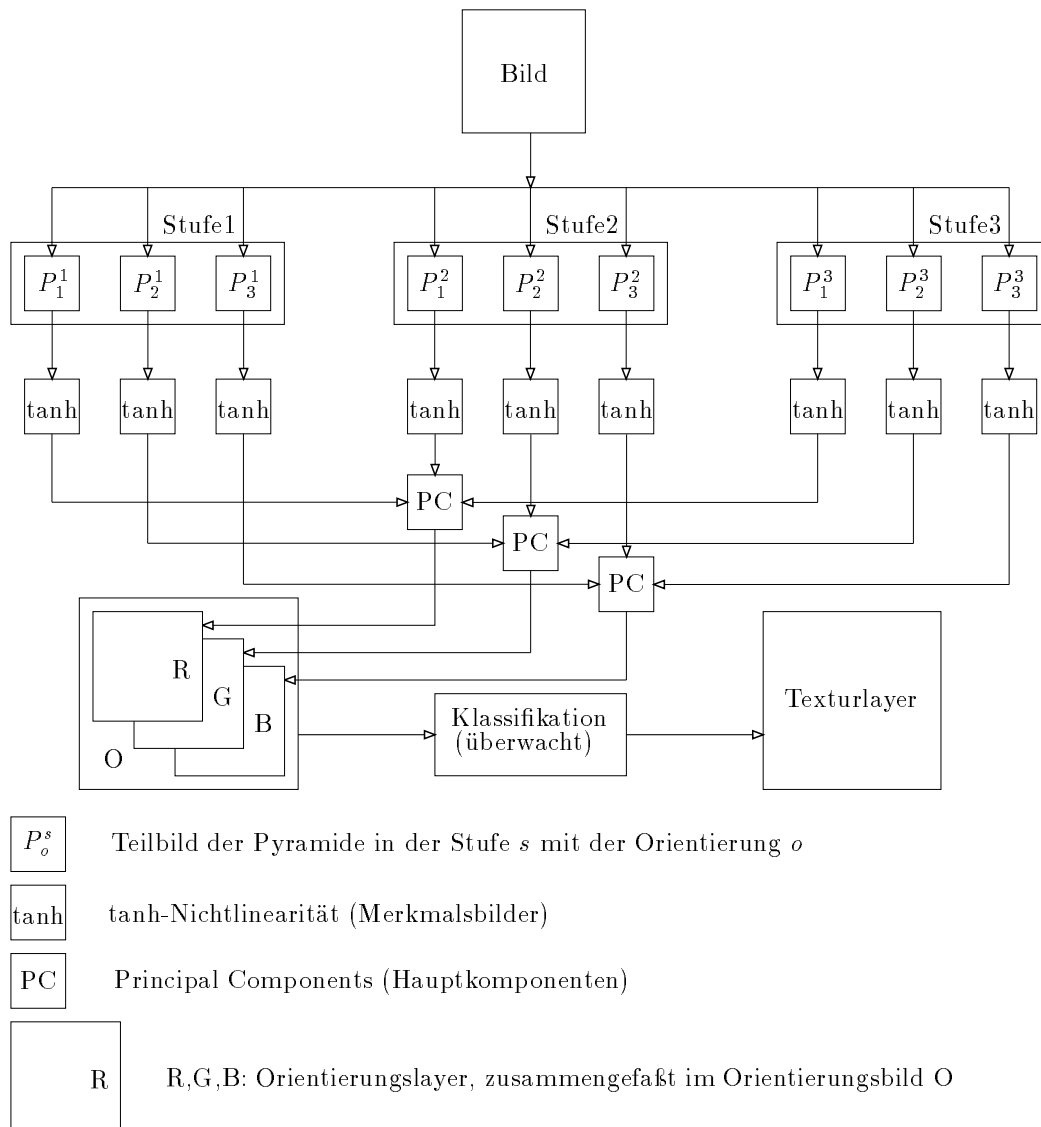


Abbildung 6.14: Ablauf zur Erzeugung des Texturlayers

stellt. Bei der Darstellung handelt es sich um zweidimensionale Histogramme, in denen die Häufigkeit des kombinierten Auftretens von Grauwerten in je zwei Farbkanälen durch die Helligkeit kodiert wird.

Wie aus den vorangegangenen Überlegungen zu erwarten, zeigt sich für die Siedlungsbereiche eine deutliche Clusterbildung⁶ im ursprungsfernen Bereich der

⁶Als Cluster werden die bevorzugten Bereiche im Merkmalsraum bezeichnet, s.a. Kapitel 5.

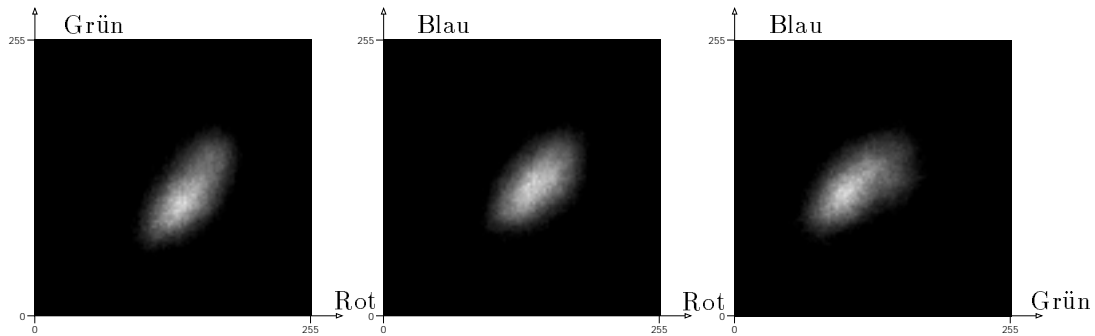


Abbildung 6.15: Merkmalsraum für die Klasse Wald

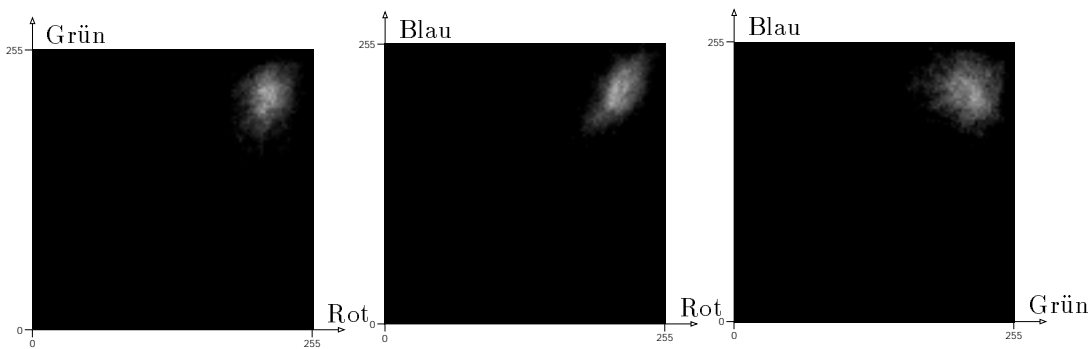


Abbildung 6.16: Merkmalsraum für die Klasse Siedlung

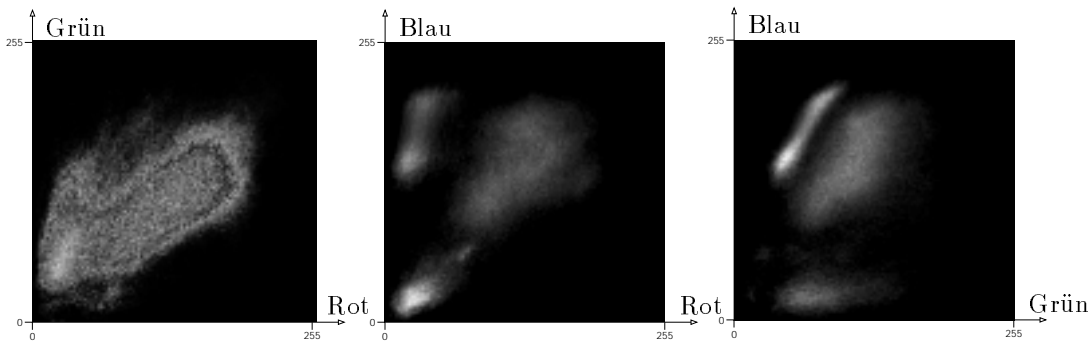


Abbildung 6.17: Merkmalsraum für die Klasse Feld

Unbuntgeraden (entspricht Weiß), die jeweils diagonal vom Ursprung durch das Bild verläuft. Ähnliches ist für die Klasse Wald festzustellen, allerdings ist die Ausdehnung des Clusters etwas größer und seine Lage ist ursprungnäher (ent-

spricht Grau). Wie im Farbbild 6.18 (S. 99) zu sehen ist, bildet sich nicht so deutlich ein einzelner Grauwert für diese Klasse aus, vielmehr ist das Gebiet „fleckig“ und damit inhomogen. Es hat zusätzlich eine gewisse Farbigkeit (blau bis türkis). Die Erklärung für diese Orientierung anzeigende Farbigkeit liegt meiner Meinung nach im Schattenwurf, der immer in einer Richtung erfolgt und im Bild zu langgezogenen Strukturen quer zur Einfallrichtung führt, deren Orientierung diese Farbigkeit hervorrufen. Gestützt wird diese These durch den direkten Vergleich der durch die Farbe angezeigten Orientierung mit der Einfallrichtung des Lichts in der Szene: Blau entspricht der Richtung senkrecht zum Lichteinfall, der in der Luftbildszene aus Südost erfolgte. Bei genauerer Betrachtung des Merkmalsraumes fällt auch die Lage des Clusters etwas neben die Unbuntgerade. Für die Klassifikation weist der Cluster aber ausreichende Kompaktheit auf. Die dritte Klasse, in der Orientiertheit zusammengefaßt wird, hat im Merkmalsraum wie erwartet eine große Ausdehnung, die Häufung in bestimmten Bereichen ergibt sich aus der Tatsache, daß die Trainingsgebiete diese Klasse gebildet haben. Sie wiesen Bearbeitungsspuren bestimmter Orientierungen auf, die nicht gleichmäßig verteilt sind. Größere Trainingsgebiete bringen ihre Orientierung mit größerem Gewicht ein. Die Bildung dieser Klasse kann aufgrund der oben angestellten Überlegungen besser deterministisch erfolgen, indem die Klasse aus dem gesamten Merkmalsraum abzüglich der von Wald und Siedlung belegten Cluster gebildet wird. Für eine weitergehende Implementierung ist diese Vorgehensweise zu empfehlen.

Das Klassifikationsergebnis ist in Abbildung 6.19 dargestellt. Zum Vergleich ist das Klassifikationsergebnis aus den verbleibenden drei Richtungen in Abbildung 6.20 dargestellt. Man kann erkennen, daß die Segmentierung aufgrund der Texturinformation die Klassen Wald, Siedlung und Feld trennen kann. Helle und fleckige Felder werden nicht der Klasse Feld zugewiesen. Die Texturmerkmale dieser Felder sind denen von Wald und Siedlung durch die Art der Merkmalsbildung ähnlich. Diese Felder werden deshalb nicht der Klasse Feld zugewiesen. Ihre Orientierung ist nicht ausgeprägt. In der Farbdarstellung (Abb. 6.19) sieht man, daß sie zum einen inhomogen sind und zum anderen eine ähnliche Farbigkeit wie die Waldbereiche aufweisen. Hinzu kommt, daß keine Felder als Trainingsgebiete



Abbildung 6.18: Farbkodiertes Orientierungsmerkmalsbild aus drei Richtungen

eingingen, die im farbigen Orientierungsbild die Farbe Blau erhalten haben, so daß die Zuweisung zu der Klasse Wald erfolgt. Im Merkmalsraum drückt sich dieser Sachverhalt in den hellen Bereichen aus, die nur auf einer Seite der Unbuntgeraden liegen (Abb. 6.17). Er zeigt sich auch in der Contingency Matrix (Anhang A, S. 117), in der abseits der Hauptdiagonalen große Werte für diese beiden Klassen stehen.

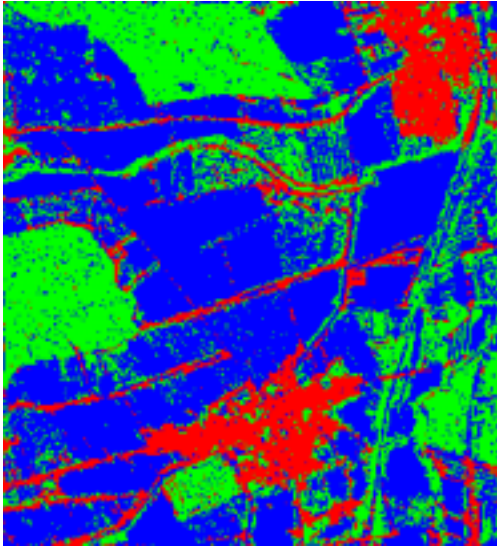


Abbildung 6.19: Texturlayer, Ergebnis der Klassifikation aus dem Orientierungsbild

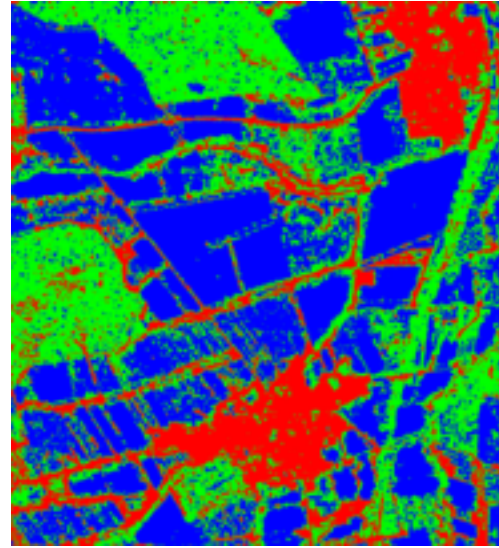


Abbildung 6.20: Texturlayer aus dem Orientierungsbild der verbleibenden drei Richtungen

Der Vergleich der Klassifikationen aus je drei Richtungen ergibt, daß es im zweiten Fall (Abb. 6.20) zu mehr Fehlzugeweisungen der Klassen Wald und Siedlung kommt. Das zweite Merkmalsbild fiel insgesamt heller aus. Es kommen damit schon in der Trainingsphase mehr helle Bereiche in den Trainingsgebieten vor, die Inhomogenität ist damit größer und die Trennbarkeit herabgesetzt. Eine bessere Definition der Cluster kann hier noch Verbesserungen bringen. Die Clusterverteilung für die Feldklasse ist nicht optimal (s. Abbildung 6.17), da sie auf nur einer Seite der Unbuntgeraden einen Schwerpunkt hat. Es müßte zumindest eine weitere Klasse auf der anderen Seite der Unbuntgeraden geben, damit der Schwerpunkt nicht in den Waldcluster fällt.

Es kann festgehalten werden, daß eine Trennung der drei Klassen Wald, Siedlung und Feld aufgrund von Textur möglich ist. Die Einführung des Texturkanales in die multispektrale Klassifikation kann nun erfolgen.

Es wurde auch eine Klassifikation mit vier Klassen durchgeführt, indem Wasserflächen als eigene Klasse eingeführt wurden. Das Ergebnis der Klassifikation zeigte, daß das Merkmal *in keinem der Orientierungsbilder Merkmale vorhan-*

den, das Wasserflächen kennzeichnet, auch in Feldbereichen besteht. Der südliche Baggersee wurde nicht dieser Klasse zugeordnet. Nach der Einführung des Texturkanales in die Klassifikation war die Fehlklassifikation verstärkt.

6.7 Klassifikation mit Texturkanal

Der Texturkanal wurde den vier Kanälen des multispektralen Bildes hinzugefügt⁷. Die Klassifikation erfolgte wiederum überwacht mit denselben Trainingsgebieten. Das Ergebnis der Klassifikation ist in Abbildung 6.21 zu sehen.

Zunächst fällt auf, daß eine Homogenisierung der Feldbereiche stattfindet. Die Fehlzugeweisungen von Siedlung und Wald zu Feldbereichen reduziert sich. Außerhalb der Feldbereiche mit falscher Zuweisung im Texturkanal ist eine Verbesserung der Klassifikation festzustellen.

Die Texturinformation wirkt sich entscheidend auf die Unterscheidung der Klassen Wald1 und Weizen aus. Die Entscheidung des Klassifikators fällt auf Weizen, wenn die Texturinformation Feld beinhaltet. Auf der anderen Seite fällt die Entscheidung auf Wald, wenn die Texturinformation Wald bedeutet. Das führt zu neuen Fehlklassifikationen in den Feldbereichen, in denen die Textur nicht die gewünschte Zuordnung liefert.

Aus der Texturinformation wird in Grenzfällen entschieden. Vereinzelt auftretende fehlerklassifizierte Pixel in der rein multispektralen Klassifikation deuten auf eine Lage nahe einer Entscheidungsgrenze hin.

Ungünstig wirkt sich die Fehlklassifikation im Texturlayer für das Trainingsgebiet Wiese aus. Das Trainingsgebiet wird im Texturkanal der Klasse Wald zugeordnet. Damit wird eine falsche Signatur trainiert. Eine der Auswirkungen ist, daß das Wegenetz jetzt eher der Klasse Wiese zugeordnet wird als vorher. So fällt die Entscheidung auf Wiese statt auf Weizen, wenn im Texturlayer die Information Wald steht.

Die Klasse Weizen hat also besondere Sensibilität in Hinblick auf die Textur-

⁷Vorgehensweise unter ERDAS s. Kapitel 5.1.1.

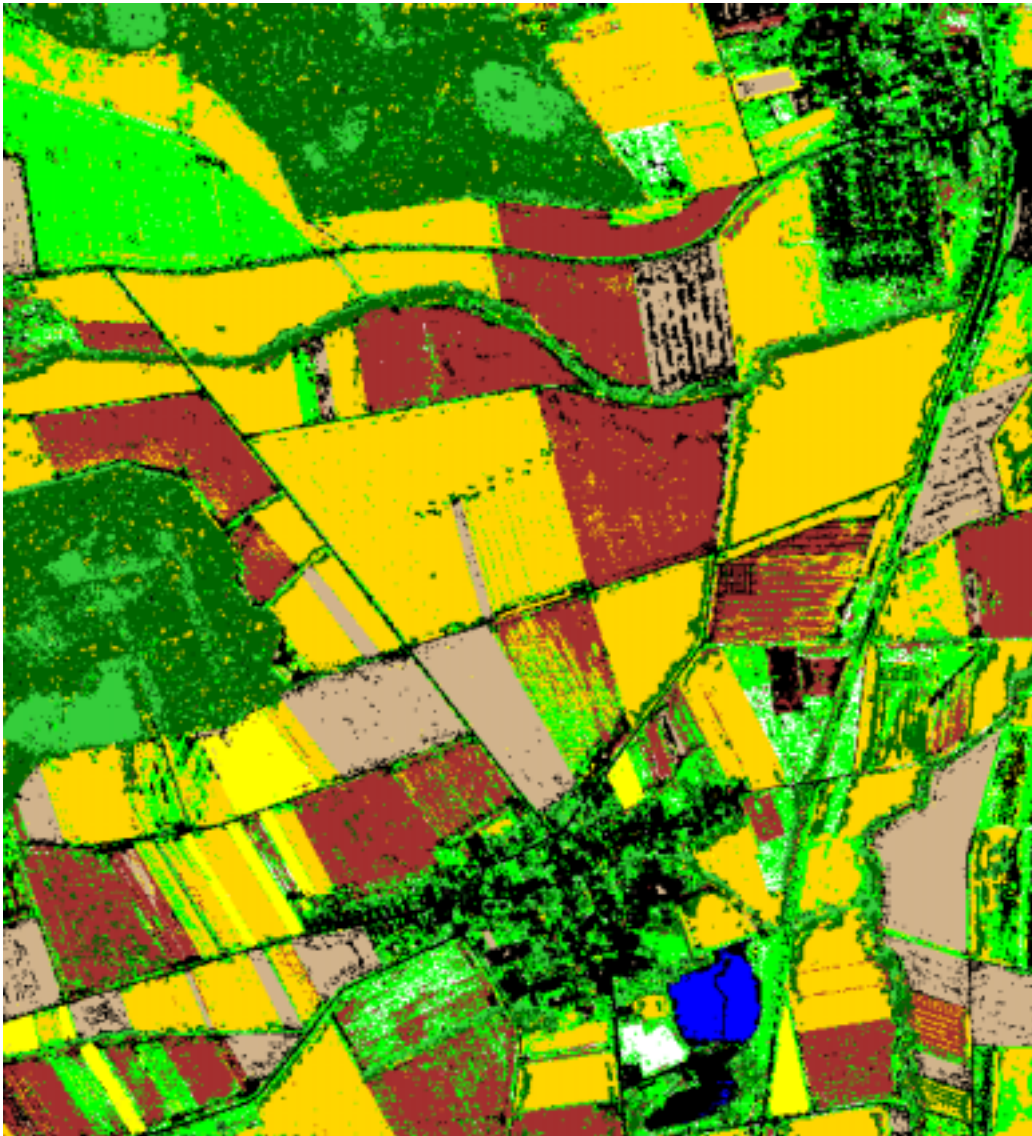


Abbildung 6.21: Klassifikationsergebnis mit eingeführtem Texturlayer

information.

Es ergeben sich viele Änderungen in der Klassifikation von Gebieten, die im Texturkanal fälschlicherweise der Klasse Wald angehören.

Um die Veränderungen zu visualisieren, habe ich die Differenz der Klassifikationen gebildet und ihnen Farben zugewiesen. Es sind zwei Differenzbildungen möglich, die jeweils etwas anderes bedeuten. In dem einen Fall bedeutet sie z.B.,

daß Waldfehlklassifikationen auf Weizenfeldern abnehmen, im anderen Fall bedeutet die Differenz, daß sie hinzukommen. Daher wurde für diese zwei Fälle eine unterschiedliche Farbe gewählt. Dann wurden die zwei gebildeten Differenzen addiert und zeigen die Stellen an, an denen die Klassifikation sich geändert hat. Schwarz bedeutet, daß es keine Änderung gab. Die entstehenden Farben können durch Vergleich der beiden Klassifikationsbilder einer speziellen Änderung zugeordnet werden. Das Bild ist in Abbildung 6.22 zu sehen. Am auffälligsten sind folgende Farben:

- Helles Blau: Statt Wald wird Weizen klassifiziert.
- Gelb: Statt Weizen wird Wald klassifiziert.
- Rot: Statt Weizen wird Wiese klassifiziert.
- Dunkles Grün: Statt Gerste wird Siedlung klassifiziert.
- Grau: Statt Siedlung wird Gerste klassifiziert.
- Violett: Statt Siedlung wird Wiese klassifiziert.

Durch den Vergleich mit dem Originalbild ergeben sich die Bereiche, in denen die Klassenzuweisung durch Klassifikation verbessert wird, ebenso lassen sich hinzugekommene Fehlklassifikationen feststellen. Hellblaue Bereiche in Feldbereichen bedeuten eine Verbesserung der Klassenzuordnung, da Weizen klassifiziert wird. In Waldbereichen bedeutet eine hellblaue Farbe Fehlklassifikation. Die Farbe Gelb in Feldbereichen zeigt eine Änderung des Klassifikationsergebnisses, die zusätzliche Fehlklassifikation zu der Klasse Wald bedeutet. In Waldbereichen wird dadurch eine richtige Zuweisung angezeigt. Um den Einfluß des Texturkanals auf diese Änderungen der Klassifikation zu zeigen, wurde der Texturkanal mit dem Differenzenbild verrechnet, indem eine Mittelwertbildung der Farbwerte vorgenommen wurde. Damit erhalten Pixel, die keine Änderung erfahren, einen Farbwert von 128 ($= \frac{255+0}{2}$) in dem der Klassenfarbe entsprechenden Kanal (Rot, Grün, Blau). Ist eine Veränderung eingetreten, so ist dieser Farbwert höher bzw. es entstehen Werte in den zwei anderen Farbkanälen, was eine Änderung der Farbe hervorruft. Das Resultat ist eine hellere Farbe bzw. eine Farbverschiebung.

Damit wird die Wirkung des Texturkanales deutlich. Das Bild ist in Abbildung 6.23 zu finden.

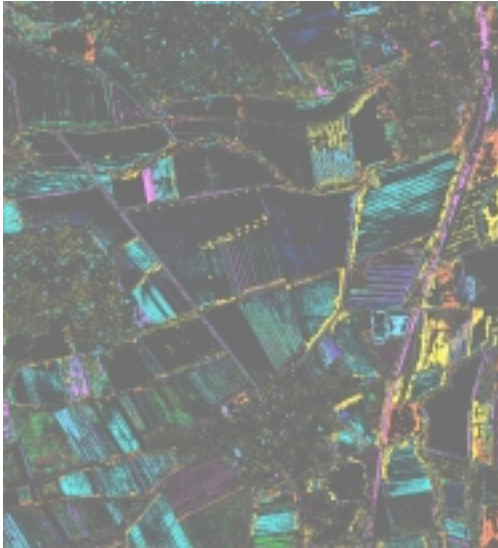


Abbildung 6.22: Differenzen, die sich durch die Einführung des Texturkanales in die Klassifikation ergeben

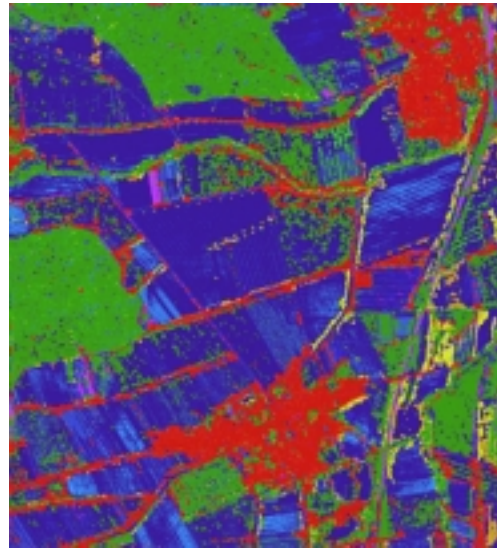


Abbildung 6.23: Die Änderungen in der Klassenzuordnung in Abhängigkeit von der Klassenzuweisung im Texturlayer

6.8 Nutzen von Texturinformationen in der Klassifikation

Die Ergebnisse zeigen, daß diese Informationen das Ergebnis nicht nur verbessern, sondern auch verschlechtern können. Es stellt sich die Frage, ob Texturinformationen in den Klassifikationsprozeß eingeführt werden sollten. Um diese Frage zu beantworten, werden im folgenden einige grundsätzliche Überlegungen angestellt.

Zunächst wird festgehalten, daß die Textur, die man in Luftbildern findet, komplex ist. Die Ableitung von Merkmalen, die eindeutig den gesuchten Objektklassen zugeordnet werden können, ist nicht immer möglich, da für einzelne Objektklassen Variationen auftreten und die Merkmale sich überdecken. Die Auflösung des Bildes hat besondere Bedeutung für den Texturgehalt. Hier wurden hochauflösende Daten verwendet. Je höher die Auflösung ist, desto mehr Textur-

information ist vorhanden und die Komplexität steigt. Damit ergibt sich die Notwendigkeit, die Texturen in Hinblick auf die jeweilige Auflösung zu untersuchen. Von der Auflösung unabhängige Merkmale lassen sich durch die mehrfachauflösenden Verfahren finden. In diesen Verfahren wird das Bild in mehreren Auflösungsstufen untersucht. Damit lassen sich durch die Wahl der Auflösungsstufe Vergleiche zwischen unterschiedlich aufgelösten Bildern anstellen und Merkmale finden, die sich entsprechen. Es stellt sich die Frage nach der Anzahl der Texturen, die in einem Luftbild unterschieden werden sollten. In der Antwort auf diese Frage wird ein grundsätzliches Problem deutlich:

Werden mehr Texturen durch Bildung von spezielleren Merkmalen unterschieden, erhält man mehr Texturklassen. Die Zuordnung der Texturklassen zu den Objekten wird gleichzeitig erschwert, da die Textur der Objekte nicht immer einheitlich ist. Dadurch ist die Gefahr einer Fehlzusammenordnung erhöht. Die Segmentierung eines Bildes in viele Texturklassen bringt also die Notwendigkeit mit sich, die Textureigenschaften der Objekte genau zu kennen und festzulegen. Diese Definition ist mit einem hohen Aufwand verbunden und kann nur durch eine genaue Kenntnis der Objekte vorgenommen werden. Denkbar sind hier wissensbasierte Ansätze, die zusätzliche Informationen z.B. über die Nachbarschaftsbeziehungen einbringen. Fehlen solche Informationen, kommt es zu Fehlzusammenordnungen, wenn die Merkmale für unterschiedliche Objekte sich ähneln.

Globalere Merkmale, wie sie in dieser Arbeit verwendet wurden, führen auf wenige Texturklassen und sind robuster. Sie fassen mehrere Objektklassen in einer Klasse zusammen, eine Unterscheidung aufgrund der Textur ist dann in mehr Fällen nicht mehr möglich. Diese Unterscheidung muß aufgrund anderer Merkmale der Objekte getroffen werden, in diesem Fall aufgrund der spektralen Eigenschaften.

Die Ableitung von Texturinformationen sollte sich immer an nachvollziehbaren Eigenschaften der Objekte orientieren. Die Feststellung, daß verschiedene Texturen auftreten und getrennt werden können, ist ohne Interpretierbarkeit ihrer Bedeutung wertlos.

Hinzu kommt, daß für die Untersuchung von Texturen viele Verfahren zur

Verfügung stehen, die je nach vorliegender Textur besser oder schlechter geeignet sind. Die Auswahl des Verfahrens muß anhand der Eigenschaften der vorliegenden Texturen erfolgen, damit die Merkmale eine Unterscheidbarkeit ermöglichen, die für die Segmentierungsaufgabe Voraussetzung ist.

In dieser Arbeit wurden Texturmerkmale für die gesamte Szene abgeleitet. In den Fällen, in denen aufgrund des gewählten Verfahrens keine richtige Unterscheidung erfolgte, wie z.B. bei Feldern und Wiesen ohne eindeutig bevorzugte Orientierung, ist die Texturinformation kontraproduktiv und verschlechtert das Klassifikationsergebnis.

Einen Ausweg aus dem Problem sehe ich in der Nutzung von Texturinformationen nur an den Stellen, an denen sie eine sinnvolle Unterscheidung liefern kann. Dazu muß ein wissensbasiertes System entworfen werden, das über die Einführung dieser Informationen entscheidet. Dann wird die Information nur an den Stellen verwendet, an denen man sicher sein kann, daß die Information differenzierende Wirkung hat. Mit der Untersuchung nur in lokalen Bereichen kann eine Eingrenzung auf wenige Texturmerkmale erfolgen und auch das passende Verfahren zur Ableitung dieser Merkmale ausgewählt werden.

Zusammenfassung

In dieser Arbeit wurde untersucht, wie sich Texturinformationen in Luftbildern nutzen lassen, um die rein multispektrale Klassifikation zu unterstützen. Dazu erfolgte die Untersuchung mehrerer Verfahren, die Texturmerkmale erfassen können. Es wurde ein mehrfachauflösender Ansatz ausgewählt, der Orientierungsinformationen bereitstellt. Anhand von Eigenschaften der zu trennenden Texturen wurde die Erzeugung eines Texturlayers vorgenommen, der das Bild in die Klassen Wald, Siedlung und Feld mit einer überwachten Klassifikation segmentiert. Dieser Layer wurde als zusätzlicher Kanal in die multispektrale Klassifikation eingeführt und die Auswirkung auf das Klassifikationsergebnis beurteilt. Es zeigte sich, daß die Texturinformation das Klassifikationsergebnis für schlecht multispektral trennbare Klassen verbessert oder verschlechtert. Die Klassenzuordnung wird für diese Klassen aufgrund der Texturinformation getroffen, so daß eine richtige Zuordnung der Klasse im Texturlayer zu einer Verbesserung führt, eine falsche Zuordnung dagegen Fehlklassifikationen hervorruft.

Conclusion

In this thesis it was investigated how textural information can be used to improve the results of multispectral classification. Several techniques that recognize textural features were analysed. A multiresolutional technique that provides orientational information was chosen. The properties of the textures to be recognized lead to a texture layer that segments the picture with a supervised classification into the classes wood, settlements and field. This layer was added to the multispectral layers and the changes in classification were evaluated. It was shown that textural information can improve or worsen the classification results of poorly separable classes. The assignment to one of the classes depends on the basis of the textural information. A correct assignment in the textural layer leads to an improvement, an incorrect assignment leads to misclassification.

Danksagung

Mein Dank gilt allen Menschen, die mich in der Erstellung dieser Arbeit unterstützt haben. Besonders danke ich den Mitarbeitern des Institutes für Photogrammetrie, die immer offen für Fragen waren und Anregungen gegeben haben, sowie meinen Eltern, die mir den Weg bis zu dieser Arbeit ermöglicht haben.

Literaturverzeichnis

- [1] A. K. Jain; F. Farrokhnia: *Unsupervised texture segmentation using Gabor filters*. Pattern Recognition, 24(12):1167–1186, 1991.
- [2] Kurt Helming: *Texturapproximation durch Modellkomponenten*. Cuvillier, Göttingen, 1996.
- [3] A. Ravishankar Rao: *A taxonomy for texture description and identification*. Springer series in perception engineering. Springer, New York [u.a.], 1990.
- [4] Rainer Müller: *Untersuchung von Verfahren der Texturanalyse zur Differenzierung und Extraktion verschiedener Objektarten aus Fernerkundungsdaten*. Diplomarbeit, Universität Hannover, Institut für Photogrammetrie und Ingenieurvermessungen, 1996.
- [5] Rafael C. Gonzalez: *Digital image processing*. Addison-Wesley, Reading, Mass. [u.a.], 1994.
- [6] R. M. Haralick (Herausgeber): *Statistical and structural approaches to Texture*, Nummer 67 in *Proceedings of the IEEE*. IEEE, 1979.
- [7] Diederich Wermser: *Automatische Texturauswertung auf der Basis einer Analyse des Visuellen Systems*. VDI-Verl., Düsseldorf, 1985.
- [8] Sehran Tatari: *Auswahl von Textur-Analyseverfahren zur Automatisierung industrieller Sichtprüfungen*, Band 67 der Reihe *Fortschritt-Berichte VDI. Reihe 10*. VDI Verl., Düsseldorf, 1987.
- [9] IEEE Standard 610.4-1990: *IEEE Standard glossary of image processing and pattern recognition terminology*. IEEE Press, New York, 1990.

-
- [10] Sklansky: *Image segmentation and feature extraction*. IEEE Trans Syst. Man cybern., 8:237–247, 1978.
- [11] Gert Van de Wouwer: *Wavelets for texture analysis*. <http://wcc.ruca.ua.ac.be/~visielab/wta/textu.html>, August 1999.
- [12] Hans Jürgen Grotepaß: *Ein Verfahren zur dreidimensionalen Texturanalyse in Grauwertvolumina*. VDI-Verl., Düsseldorf, 1992.
- [13] Olaf Pichler: *Unüberwachte Bild- und Bildfolgenanalyse mit Mehrkanalfiltermethoden*. VDI-Verl., Düsseldorf, 1998.
- [14] Bernd Jähne: *Digitale Bildverarbeitung*. Springer, Berlin, 3. Auflage, 1993.
- [15] Wolfgang Jourdan: *Einbeziehung der Farbinformation zur Texturparameterberechnung für die Klassifikation von Luftbildaufnahmen*. Fortschritt-Berichte der VDI-Zeitschriften : Reihe 10, Angewandte Informatik; 28. VDI Verl., Düsseldorf, 1983.
- [16] Christopher Philip Winder: *Markovian analysis of texture: serial and parallel paradigms in low-level vision*. Programming Research Group, Oxford Univ. Computing Laboratory, Oxford, England, 1992.
- [17] Yanjun Pan: *Autoregressive Modelle für die Texturanalyse in digitalen Bildern*. Verl. der Bayer. Akad. der Wiss. [u.a.], München, 1994.
- [18] Benoit B. Mandelbrot: *The fractal geometry of nature*. New York, New York, 1983.
- [19] Herbert Zeitler und Wolfgang Neidhard: *Fraktale und Chaos: Eine Einführung*. Wissenschaftliche Buchgesellschaft, Darmstadt, 2 Auflage, 1994.
- [20] Kenneth J. Falconer: *Fraktale Geometrie : Mathematische Grundlagen und Anwendungen*. Spektrum, Akad. Verl., Heidelberg [u.a.], 1993.
- [21] Michael F. Barnsley: *Fraktale : Theorie und Praxis der deterministischen Geometrie*. Spektrum Akad. Verl., Heidelberg [u.a.], 1995.

- [22] S. Haykin: *Neural Networks: A comprehensive foundation*. Macmillan, New York, 1994.
- [23] Jessica Schnellert: *Klassifizierung von DPA Bildern unter Einbeziehung der Methoden der Neuronalen Netze und Fuzzy Sets*. Diplomarbeit, Universität Hannover, Institut für Photogrammetrie und Ingenieurvermessungen, Dezember 1999.
- [24] Robi Polikar: *The wavelet tutorial*.
<http://www.public.iastate.edu/~rpolikar/WAVELETS/WTacknowledment.html>, 1999.
- [25] Dieter Vollmerhaus: *Texturanalyse auf der Basis hierarchischer Bildpyramiden*. Hanser, Muenchen [u.a.], 1992.
- [26] Trygve Randen: *Filter and filter bank design for image texture recognition*. Doktorarbeit, Norwegian University of Science and Technology, Stavanger, Norway, 1997.
- [27] Barbara Burke Hubbard: *Wavelets: Die Mathematik der kleinen Wellen*. Birkhaeuser, Basel, 1997.
- [28] Stephane Mallat: *A wavelet tour of signal processing*. Academic Press, San Diego, Calif. [u.a.], 2. Auflage, 1999.
- [29] Stephane G. Mallat: *A theory for multiresolution signal decomposition: The wavelet representation*. IEEE Transactions on pattern analysis and machine intelligence, 11(7):674–693, Juli 1989.
- [30] Stephane Mallat: *Multiresolution representations and wavelets*. Doktorarbeit, University of Pennsylvania, Philadelphia, August 1988.
- [31] Andrew Laine; Jian Fan: *An adaptive approach for texture segmentation by multichannel wavelet frames*. Technischer Bericht TR-93-025, Center for Computer Vision and Visualization, University of Florida, August 1993.
<ftp://ftp.cis.ufl.edu/cis/tech-reports/tr93/tr93-025.ps>.

-
- [32] Vidya Manian; Ramon Vasques: *Scaled and rotated texture classification using a class of basis functions*. Pattern recognition, 31(12):1937–1948, 1998.
- [33] Yateen Chitre; Atam P. Dhawan: *M-band wavelet discrimination of natural textures*. Pattern recognition, 32:773–789, 1999.
- [34] Eero P. Simoncelli; William T. Freeman; Edward H. Adelson; David J. Heeger: *Shiftable multi scale transforms*. IEEE Trans. Information Theory, 38(2):587–607, März 1992.
- [35] Eero Simoncelli: *Image pyramid tutorial der Pyramidentools (pyramids.m)*, 1999. Mit den Pyramidentools distributiert:
ftp.cns.nyu.edu/pub/eero/matlabPyrTools.tar.gz.
- [36] A.C. Bovik; M.Clark; W. Geisler: *Multichannel texture analysis using spatially localized spatial filters*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 12(1):55–73, Januar 1990.
- [37] J. R. Smith: *Integrated spatial and feature image systems: Retrieval, compression and analysis*. Diplomarbeit, Graduate School of Arts and Sciences, Februar 1997.
- [38] A. K. Jain; N. K. Ratha und S. Lakshmanan: *Object detection using Gabor filters*. Pattern Recognition, 30(2):295–309, 1997.
- [39] D. F. Dunn und W. E. Higgins: *Optimal Gabor filters for texture segmentation*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 4(7):947–964, July 1995.
- [40] Trygve Randen; John Håkon Husøy: *Filtering for texture classification: A comparative study*. IEEE Transactions on pattern analysis and machine intelligence, 21(4):291–310, April 1999.
- [41] Floyd F. Sabins: *Remote sensing, principles and interpretation*. W. H. Freeman Company, New York, 1997.
- [42] P. Scheunders; S. Livens; G. Van de Wouwer; P. Vautrot; D. Van Dyck: *Wavelet-based texture analysis*. International Journal Computer Science

and Information Management, Dezember 1997. Vorgelegt zum Erschei-
nen in angegebener Ausgabe, erhältlich als elektronisches Dokument:
<http://visiehp4.ruca.ua.ac.be/papers/wouwer/ipa.ps.gz>.

Anhang A

Contingency Matrizen der Klassifikationen

Classified Data	Wald1	Wald2	Weizen	Rueben	Wiese	Gerste	Raps	Wasser	Brache	Siedlung	Row Total
Wald1	<u>89.11</u>	<u>2.58</u>	0.28	0.00	0.98	0.00	0.00	0.00	0.00	0.48	265729
Wald2	<u>5.86</u>	<u>95.64</u>	0.02	0.00	0.37	0.00	0.00	0.00	0.00	<u>2.12</u>	71369
Weizen	<u>2.31</u>	0.05	<u>98.18</u>	<u>2.63</u>	0.19	0.00	0.00	0.00	0.00	0.00	452304
Rüben	0.13	0.03	0.49	<u>94.86</u>	<u>3.31</u>	0.00	0.44	0.00	0.16	0.29	224224
Wiese	<u>2.27</u>	0.86	0.90	1.24	<u>62.60</u>	0.00	0.15	0.00	<u>9.75</u>	<u>2.84</u>	31240
Gerste	0.00	0.00	0.00	0.01	0.00	<u>98.52</u>	0.42	0.00	0.01	0.49	265895
Raps	0.00	0.00	0.03	0.10	0.67	0.11	<u>98.62</u>	0.00	0.00	0.19	62287
Wasser	0.00	0.00	0.00	0.00	0.00	0.00	0.00	<u>99.50</u>	0.00	0.00	21903
Brache	0.03	0.04	0.03	0.17	<u>30.18</u>	0.00	0.00	0.00	<u>89.73</u>	0.27	20088
Siedlung	0.29	0.80	0.06	0.99	1.71	1.37	0.36	0.50	0.35	<u>93.32</u>	37295
Column Total	294743	55694	447456	232397	23706	269437	62247	22013	13609	31032	1452334

Tabelle A.1: Contingency Matrix der multispektralen
Klassifikation mit hinzugefügtem Texturlayer, Angaben
in %

Classified Data	Brache	Wasser	Wald1	Wald2	Gerste	Raps	Wiese	Siedlung	Weizen	Rueben	Row Total
Brache	<u>90.47</u>	0.00	0.04	0.04	0.00	0.00	<u>31.62</u>	0.32	0.00	0.17	20341
Wasser	0.00	<u>99.50</u>	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	21903
Wald1	0.00	0.00	<u>88.50</u>	<u>2.50</u>	0.00	0.00	0.61	0.38	0.96	0.00	157381
Wald2	0.00	0.00	<u>6.51</u>	<u>95.58</u>	0.00	0.00	0.32	<u>2.21</u>	0.01	0.00	65189
Gerste	0.00	0.00	0.00	0.00	<u>99.33</u>	0.36	0.00	0.42	0.00	0.00	268006
Raps	0.00	0.00	0.00	0.00	0.12	<u>98.75</u>	0.94	0.23	0.02	0.07	62345
Wiese	<u>9.00</u>	0.00	<u>1.84</u>	0.78	0.00	0.14	<u>61.55</u>	<u>2.94</u>	<u>1.24</u>	<u>1.29</u>	28955
Siedlung	0.38	0.50	0.32	0.82	0.55	0.32	<u>1.32</u>	<u>92.70</u>	0.86	<u>1.07</u>	38267
Weizen	0.01	0.00	<u>2.73</u>	0.25	0.00	0.01	0.11	0.49	<u>95.25</u>	0.30	431868
Rueben	0.14	0.00	0.06	0.02	0.00	0.42	<u>3.52</u>	0.31	<u>1.65</u>	<u>97.10</u>	234391
Column Total	13540	22013	171124	55694	269437	62247	23706	31032	447456	232397	1328646

Tabelle A.2: Contingency Matrix der multispektralen
Klassifikation, Angaben in %

Classified Data	Wald	Siedlung	Rest	Row Total
Wald	<u>94.54</u>	1.34	<u>23.80</u>	586644
Siedlung	0.32	<u>97.83</u>	1.64	49049
Rest	<u>5.14</u>	0.82	<u>74.56</u>	816726
Column Total	350446	31032	1070941	1452419

Tabelle A.3: Contingency Matrix der Klassifikation des Orientierungsbildes, Angaben in %

Anhang B

Daten der Feldbegehung

Auf den folgenden Seiten befinden sich die Daten der Feldbegehung, die in Ausschnitten der DGK5 Obernjesa (Nr. 15 der DGK25 Nr. 4525) und Klein Schneen (Nr. 21 der DGK25 Nr. 4525) verzeichnet sind.