

# Objektextraktion von Bäumen aus Luftbildern

---

Vergleich und Steuerung von Segmentierungsverfahren  
zur Vorbereitung eines Expertensystems

## **Diplomarbeit**

am Institut für Geographie der Universität Hamburg

---

### **Im Dezember 2007 vorgelegt von:**

Benjamin Bechtel | bennib@gmx.net

### **Erstgutachter:**

Prof. Dr. Jürgen Böhner

### **Zweitgutachter:**

Prof. Dr. Dietmar P. F. Möller



# Inhaltsverzeichnis

<b>Inhaltsverzeichnis</b> .....	<b>i</b>
<b>Abbildungsverzeichnis</b> .....	<b>iv</b>
<b>Tabellenverzeichnis</b> .....	<b>vii</b>
<b>Abkürzungsverzeichnis</b> .....	<b>vii</b>
<b>1 Einleitung</b> .....	<b>1</b>
1.1 Motivation und Fragestellung.....	1
1.2 Aufbau der Arbeit.....	3
<b>2 Grundlagen</b> .....	<b>5</b>
2.1 Fernerkundung.....	5
2.1.1 Struktur und Reflexionsspektrum grüner Vegetation.....	7
2.1.2 LIDAR.....	10
2.2 Mustererkennung.....	11
2.2.1 Methodologische Struktur der Mustererkennung.....	12
2.2.2 Bayes'sche Entscheidungstheorie.....	15
2.2.3 Parametrische Verfahren.....	17
2.2.4 Nicht parametrische Verfahren.....	20
2.2.5 Unüberwachtes Lernen und Clusteranalyse.....	21
2.3 Bildverarbeitung.....	22
2.3.1 Bildrepräsentation.....	23
2.3.2 Operationen und ihre Eigenschaften.....	25
2.3.3 Homogene Punkttransformationen.....	27
2.3.4 Faltung und lineare Filter.....	28
2.3.5 Standardfilter und ihre Effekte.....	30
2.3.6 Fouriertransformation.....	34
2.3.7 Filter im Frequenzraum.....	38
2.3.8 Mathematische Morphologie.....	40
2.3.9 Multiskalenrepräsentationen.....	43
2.4 Segmentierung.....	45
2.4.1 Pixel- und kantenbasierte Verfahren.....	46
2.4.2 Regionenbasierte Segmentierung.....	46
2.4.3 Split-and-merge, regionsplitting und pyramid-linking.....	47
2.4.4 Region Growing.....	48
2.4.5 Morphologische Segmentierung: Wasserscheidentransformation.....	48
2.4.6 Clustering und Segmentierung.....	49
<b>3 Objektextraktion von Bäumen aus Fernerkundungsdaten</b> .....	<b>50</b>
3.1 Grundlagen der Objektextraktion.....	51
3.1.1 Steuerung.....	51

3.1.2	Modell.....	53
3.1.3	Maßstab.....	54
3.1.4	Bewertung der Ergebnisse .....	54
3.2	<i>Template-matching nach POLLOCK, LARSEN &amp; OLOFSSON</i> .....	55
3.3	<i>Blob-detection: PINZ, KÄTSCH, DRALLE &amp; RUDEMO</i> .....	57
3.4	<i>Valley-following approach nach GOUGEON</i> .....	58
3.5	<i>Fouriertechniken nach KRUMMHEUER</i> .....	59
3.6	<i>Morphologischer Ansatz nach ANDERSEN et al.</i> .....	61
3.7	<i>WST im Maßstabsraum nach STRAUB</i> .....	62
3.8	<i>NATSCAN: DIEDERSHAGEN et al.</i> .....	67
3.9	<i>HIGHSCAN: HYYPPÄ et al.</i> .....	70
3.10	<i>Hill-Climbing nach PERSSON</i> .....	72
3.11	<i>Bandpassfilter nach KÖRBER, MÖLLER &amp; KÄTSCH</i> .....	75
3.12	<i>Region-Growing nach ERIKSON</i> .....	77
3.13	<i>Vergleichende Ansätze</i> .....	80
<b>4</b>	<b>Ein Expertensystem zur Baumextraktion.....</b>	<b>82</b>
4.1	<i>Aufbau des Expertensystems</i> .....	82
4.1.1	Kategorisierung der Problemstellungen.....	84
4.1.2	Wissensrepräsentation und Inferenzmechanismus.....	85
4.2	<i>Steuerung</i> .....	85
4.3	<i>Steuerungsparameter</i> .....	86
4.4	<i>Modell</i> .....	87
4.4.1	Annahmen.....	88
4.4.2	Merkmale.....	90
4.4.3	Klassen und Wahrscheinlichkeitsdichten.....	91
4.4.4	Kontext und Komplexität.....	92
4.4.5	Modell und Maßstab .....	92
4.5	<i>Verfahren</i> .....	92
<b>5</b>	<b>Segmentierungsverfahren.....</b>	<b>96</b>
5.1	<i>Wasserscheidentransformation im Maßstabsraum (WST)</i> .....	96
5.2	<i>Saatgut</i> .....	98
5.2.1	Saatpunkte durch regionale Maxima: (REGMAX).....	98
5.2.2	Saatpunkte durch Distanz-Transformation (DISTRANS) .....	98
5.3	<i>Basic Region Growing (BRG)</i> .....	99
5.4	<i>Brownsche Bewegung (BROWNMO)</i> .....	101
5.5	<i>Übersicht</i> .....	107

<b>6</b>	<b>Test der Segmentierungsverfahren und Diskussion .....</b>	<b>109</b>
6.1	<i>Datensätze und Untersuchungsgebiete.....</i>	<i>109</i>
6.1.1	La Motte Plantation, Western Cape Province, Südafrika .....	109
6.1.2	Göttinger Wald .....	111
6.1.3	Hans Kayinga Forest, Kavango, Namibia .....	114
6.1.4	Übersicht.....	116
6.2	<i>Wasserscheidentransformation im Maßstabsraum (WST) .....</i>	<i>117</i>
6.3	<i>Saatpunkte.....</i>	<i>123</i>
6.4	<i>Basic Region Growing (BRG) .....</i>	<i>127</i>
6.5	<i>Brownsche Bewegung (BROWNMO) .....</i>	<i>131</i>
6.6	<i>Ein einfaches Modell zur Hypothesenauswahl.....</i>	<i>140</i>
6.7	<i>Genauigkeit .....</i>	<i>142</i>
<b>7</b>	<b>Zusammenfassung und Ausblick.....</b>	<b>145</b>
	<b>Literaturverzeichnis .....</b>	<b>I</b>
	<b>Danksagung.....</b>	<b>VIII</b>
	<b>Anhang A: Beispielcode .....</b>	<b>IX</b>
	<b>Anhang B: Programmausgaben .....</b>	<b>XI</b>
	<b>Anhang C: Erklärung .....</b>	<b>XII</b>

## Abbildungsverzeichnis

Abb. 2.1:	Aufbau eines Laubblattes (Wikipedia) .....	7
Abb. 2.2:	Absorptionsspektren von Chlorophyll a und b (Wikipedia) .....	8
Abb. 2.3:	Reflexionsspektrum grüner Vegetation. Links: bestimmende Bereiche (CAMPBELL 2002: 463); Mitte: Spektren verschiedener Pflanzen (CAMPBELL 2002: 462); Rechts: Spektren verschiedener Baumarten (STRAUB 2003A: 47) .....	9
Abb. 2.4:	LIDAR Aufnahme: Messpunkte und Interpolation (CAMPBELL 2002: 238).....	10
Abb. 2.5:	LIDAR Aufnahme: first pulse und last pulse. (CAMPBELL 2002: 238).....	11
Abb. 2.6:	Merkmalsextraktion: Länge (links) und Helligkeit (rechts) von zwei Fischarten (DUDA <i>et al.</i> 2001) .....	13
Abb. 2.7:	Merkmalsraum (DUDA <i>et al.</i> 2001).....	13
Abb. 2.8:	Entscheidungsgrenzen (DUDA <i>et al.</i> 2001) .....	14
Abb. 2.9:	Ablauf in der Mustererkennung (Veränderte Darstellung nach Wikipedia) .....	14
Abb. 2.10:	Hypothetische klassenabhängige PDF: wie wahrscheinlich ist die Beobachtung von $x$ bei Zustand $\omega_i$ ? (DUDA <i>et al.</i> 2001: 21).....	16
Abb. 2.11:	A posteriori Wahrscheinlichkeiten für $P(\omega_1)=2/3$ und $P(\omega_2)=1/3$ : Wie wahrscheinlich ist der Zustand $\omega_1$ bei Beobachtung von $x$ ? (DUDA <i>et al.</i> 2001: 23) .....	17
Abb. 2.12:	Moiré-Effekte beim sampling: Bild mit verschiedenen Ortsfrequenzen, resampelt auf 0.8, 0.6, 0.4 und 0.2 der Originalgröße .....	24
Abb. 2.13:	Negativtransformation: Originalbild, Kennlinie, transformiertes Bild .....	27
Abb. 2.14:	Verschiedene Filtermasken als Bild und in 3D: Rechteck, Gauß, 3D-Gauß, Laplace, DoG, 3D-DoG .....	31
Abb. 2.15:	Wirkungsweise von Schärfungs- und Glättungsfiltern .....	32
Abb. 2.16:	Diskretes Signal mit zwei Kanten und leichtem Rauschen, erste und zweite Ableitung. ....	33
Abb. 2.17:	Links: die ersten sechs Basisvektoren der DFT für $N=16$ , Realteil und Imaginärteil. Rechts: die komplette Basis als Bild, Real- und Imaginärteil .....	36
Abb. 2.18:	Fouriertransformierte des Bildes in Abb. 2.13. Real- und Imaginärteil (Histogramm beschnitten), Spektrum (gamma-korrigiert), Phase .....	38
Abb. 2.19:	Rechteckfunktion, Spektrum der sinc-Transferfunktion als Plot und Bild .....	40
Abb. 2.20:	Abstandsmaße: Einheitskreis in verschiedenen Normen: L1 (cityblock), L2 (euklidisch), L3, L6 und $L^\infty$ (chessboard).....	43
Abb. 2.21:	Links: 1D-Signal (Bildzeile) im Maßstabsraum für $t = 0, 2, 8, 32$ und $128$ . Rechts: dasselbe Signal im ‚kontinuierlichen‘ Maßstabsraum (logarithmisch skaliert) .....	44
Abb. 3.1:	Modellgetriebener Ansatz zur Baumextraktion nach STRAUB (2003A, S. 28) .....	52
Abb. 3.2:	Datengetriebener Ansatz zur Baumextraktion nach STRAUB (2003A, S. 28) .....	52
Abb. 3.3:	Vorgehensweisen bei der Bildanalyse nach HINZ (2003: 12).....	53
Abb. 3.4:	Visualisierung der Begriffe False Negatives, True Positives und False Positives .....	55
Abb. 3.5:	Zugehörigkeitsfunktionen für Größe und Kreisförmigkeit einer Baumkrone nach STRAUB (2003A: 43f) .....	63
Abb. 3.6:	Schnitt durch idealisierte Pollockbäume und Ableitungen. Dunkelgrau: Höhe; schwarz: Betrag des Gradienten; hellgrau: $\Delta$ -Operator nach STRAUB (2003A: 45) .....	64
Abb. 3.7:	Zugehörigkeitsfunktionen für die Merkmale Konvexität und Vitalität nach STRAUB (2003a: 44, 46).....	64
Abb. 3.8:	Verfahren nach STRAUB (2003A: 55) .....	65

Abb. 3.9:	Verfahren nach SCHARDT <i>et al.</i> (2002: 305) .....	71
Abb. 4.1:	Schematischer Aufbau des Expertensystems.....	83
Abb. 4.2:	Ablauf des Verfahrens bei der Baumextraktion, Einzelschritte (Module).....	93
Abb. 4.3:	Alternatives Verfahren bei mehreren Baumarten .....	95
Abb. 5.1:	Alle Steuerungsparameter werden als PREF gespeichert.....	96
Abb. 5.2:	Gewichtsfunktionen $w_p$ für $p = 0.7, 1, 1.5$ und $2$ .....	103
Abb. 5.3:	Idealisiertes Experiment. Links: Position des Teilchens nach 1000 Schritten, Verteilung bei 10000 Versuchen. Rechts: Besuchshäufigkeit.....	104
Abb. 5.4:	Relative Aufenthaltshäufigkeiten für verschiedene Parameter. Oben: 10 Durchläufe; Unten: 3000. Links: 100 Schritte; Mitte: 1000 Schritte; Rechts: 1000 Schritte bei konstanter Gewichtsfunktion von $0.5$ .....	105
Abb. 5.5:	Ablauf des Moduls Segmentierung und implementierte Algorithmen .....	107
Abb. 6.1:	Szene 1: Kiefern-Plantage in der Nähe von Franschoek, Südafrika .....	110
Abb. 6.2:	Ausschnitte aus Szene 1: Bilder lamotte1 (links) und lamotte2 (rechts) .....	111
Abb. 6.3:	Szene 2: Luftaufnahme aus dem Göttinger Wald, Niedersachsen.....	112
Abb. 6.4:	Verwendete Ausschnitte aus Szene 2. Oben: goett{1}, goett{2} und goett{3}; Unten: goett{4}, goett{5} und goett{6}.....	113
Abb. 6.5:	Ausschnitt aus goett{6} (links); bilinear-resampelt auf halbe Größe (Mitte) und auf Viertel (rechts).....	114
Abb. 6.6:	Plot 155 aus der Namibia-Quickbird-Szene (namb1). Oben: panchromatisch, blau und grün; Unten: rot, NIR, NDVI.....	115
Abb. 6.7:	Geometrische Beschaffenheit verschiedener Bäume im Hans Kanyinga Forest (a-c); Team bei der Forstinventur im April 2007 (d).....	115
Abb. 6.8:	Ergebnisse von WST auf lamotte1. Oben: Steuerungsparameter, Bild, gewählter Kanal 2G-R; Unten: Maßstabebene, Segmentierungsfunktion, Segmentierungsergebnis.....	118
Abb. 6.9:	Links: Maßstabebene des Kanals in Abb. 6.8 (rechts unten) als 3D- Grauwertfunktion; Rechts: Histogramm des Kanals .....	119
Abb. 6.10:	Ergebnis der Hypothesenauswahl mit einfachem Modell, unterschiedliche Darstellungen .....	119
Abb. 6.11:	Verschiedene Kanäle (PREF.WST.COLOUR_CHANNEL) von lamotte1: G, DOA, und 2G-R.....	120
Abb. 6.12:	Test von Segmentierungsfunktionen (PREF.WST.SEG_FUNCTION) auf lamotte1 .....	121
Abb. 6.13:	Schnitt durch eine Bildzeile aus lamotte1 (oben) und entsprechende Segmentierungsfunktionen Betrag des Gradienten (Mitte) und invertiertes Quadrat der Summe der zweiten Ableitungen (unten) .....	121
Abb. 6.14:	Auswirkung des Parameters WST.SIGMA bei lamotte2: $\sigma = 3, 5, 7$ , Segmentierungsfunktionen (oben) und Segmentierungsergebnis (unten) .....	122
Abb. 6.15:	Verschiedene Filtergrößen bei goett{3}, $\sigma = 26$ , WST.FSIZE: 201(links) und 271 (Mitte), rechts: Gültiger Bereich für FSIZE: 201(rot) und 271(blau).....	122
Abb. 6.16:	Auswirkungen der Steuerungsparameter bei WST.....	123
Abb. 6.17:	Saatpunkte durch REGMAX auf goett{4}, SEED.SIGMA = 18, Maßstabebene als Grauwertfunktion .....	124
Abb. 6.18:	Ablauf des DISTRANS-Verfahrens auf goett{2} in Einzelschritten .....	125
Abb. 6.19:	Ergebnisse für $\sigma 7, 10, 13, 16$ für REGMAX (oben), und DISTRANS (unten)....	126
Abb. 6.20:	Auswirkungen des Parameters DISTRANS_THRESH auf goett{2} : Schwellenwert bei $0.15, 0.25, 0.35$ und $0.45$ .....	126

Abb. 6.21:	BRG auf lamotte1: Parameter, Bild, Saatpunkte, erzeugte Regionen (Segmente und Maske) .....	129
Abb. 6.22:	BRG auf lamotte2: Saatpunkte und erzeugte Regionen (2 mal) .....	130
Abb. 6.23:	BRG auf goett{6}: Saatpunkte mit REGMAX, und erzeugte Regionen (2 mal)....	130
Abb. 6.24:	BRG auf goett{6} für verschiedene Baumgrößen. MOD.MAX_RAD = 40, 60, 80 .....	131
Abb. 6.25:	Auswirkungen des Parameters RAD_MAX bei BRG .....	131
Abb. 6.26:	NumPass-Ergebnisse von BROWNM0 auf goett{4} von 3 Saatpunkten. Oben: numPass-Werte (BM.WEIGHT = 'RGB_mean', WEIGHT_POT = 1.5, NUM_RUNS = 10, NUM_STEPS = 100000); Unten: Korrespondierende Ausschnitte im Bild..	133
Abb. 6.27:	Ergebnisse von BROWNM0 auf goett{4}; Oben: Parameter, Bild, Saatpunkte; Unten: numPass, Segmentierungsergebnis (Regionen und Maske).....	134
Abb. 6.28:	Ergebnisse von BROWNM0 auf goett{1}; Oben: Parameter, Bild, Saatpunkte; Unten: numPass (gamma-korrigiert), Segmentierungsergebnis (Regionen und Maske) .....	135
Abb. 6.29:	Ergebnisse von BROWNM0 auf goett{1}, vergrößerter Ausschnitt .....	135
Abb. 6.30:	Ergebnisse von BROWNM0 auf namb1: Saatpunkte, numPass, Regionen, maskierter panchromatischer Kanal.....	135
Abb. 6.31:	NumPass-Ergebnisse von BROWNM0 auf lamotte1. Oben: numPass-Werte (BM.WEIGHT = 'RGB_mean', WEIGHT_POT = 1.5, NUM_RUNS = 100, NUM_STEPS = 1000); Unten: Korrespondierende Ausschnitte im Bild.....	136
Abb. 6.32:	Ergebnisse von BROWNM0 auf lamotte1; Oben: Parameter, Bild, Saatpunkte; Unten: numPass, Segmentierungsergebnis (Regionen und Maske).....	137
Abb. 6.33:	Auswirkungen des Parameters BM.WEIGHT_POT auf goett{4}, WEIGHT_POT = 0.8, 1.1, 1.5 und 2.0; Oben: numPass; Unten: Segmentierungsergebnis .....	137
Abb. 6.34:	Auswirkungen des Parameters BM.WEIGHT_POT auf goett{2}: Bild, Segmentierungsergebnisse für WEIGHT_POT = 0.8, 1.1, 1.5 und 2.0, .....	138
Abb. 6.35:	Veränderung der radiometrischen Charakteristik aller extrahierten Regionen bei Variation von WEIGHT_POT auf goett{2} .....	138
Abb. 6.36:	Radiometrische und geometrische Charakteristik extrahierter Regionen bei verschiedenen Gewichtsfunktionen: Intensität (links), Standardabweichung (Mitte) und Gesamtgröße (rechts) .....	139
Abb. 6.37:	Auswirkungen der Steuerungsparameter bei BROWNM0 .....	140
Abb. 6.38:	Unterschiedlich klassifizierte Bereiche bei mehreren Durchgängen .....	140
Abb. 6.39:	Merkmale aller Regionen in Abb. 6.21 .....	141
Abb. 6.40:	Hypothesenauswahl mit verschiedenen Modellen auf lamotte1 .....	142
Abb. 6.41:	lamotte2, extrahierte Regionen und Hypothesenauswahl .....	142
Abb. 6.42:	Referenzpolygone für lamotte1 und lamotte2, Regionen und Maske .....	143
Abb. 6.43:	Matching-Ergebnisse auf lamotte1 (links) und lamotte2(rechts).....	143
Abb. 6.44:	Genauigkeit der extrahierten Bäume bei BRG auf lamotte1 .....	144
Abb. 6.45:	Genauigkeit der extrahierten Bäume bei BROWNM0 auf lamotte2, rgb_mean, 1.5 .....	144
Abb. 7.1:	Erweiterter Ablauf für Region Growing .....	149
Abb. 7.2:	Alternativer Ablauf: Segmentierung im Maßstabsraum .....	149

## Tabellenverzeichnis

Tab. 1:	Variablen zur Kategorisierung von Problemstellungen.....	84
Tab. 2:	Voraussetzungen und benötigte Modell-Parameter für ausgewählte Segmentierungsverfahren.....	108
Tab. 3:	Erfüllte Voraussetzungen der Datensätze lamotte1, lamotte2, goett und namb1 ...	116
Tab. 4:	Eignung der verwendeten Datensätze für WST.....	117
Tab. 5:	Eignung der verwendeten Datensätze für die Saatgut-Verfahren REGMAX und DISTRANS .....	123
Tab. 6:	Eignung der verwendeten Datensätze für BRG.....	128
Tab. 7:	Eignung der verwendeten Datensätze für BROWNMOTION.....	132
Tab. 8:	Genauigkeit der extrahierten Ergebnisse für verschiedene Verfahren und Datensätze .....	145

## Abkürzungsverzeichnis

1D, 2D, 3D	eindimensional, zweidimensional, dreidimensional
4N, 8N	4 und 8-Nachbarschaft
Abb.	Abbildung
AN	Annahme
BMP	Windows Bitmap
BRDF	Bidirectional Reflectance Distribution Function
BRG	Basic Region Growing (Verfahren)
BROWNMOTION	Brownian Motion (Verfahren)
BV	Bildverarbeitung
bzw.	beziehungsweise
CIR	colour infrared (Foto)
d.h.	das heißt
DFT	Diskrete Fouriertransformation
DGM	Digitales Geländemodell
DIP	Digital Image Processing
DISTRANS	Distanz Transformation (Verfahren)
DN	Digital Numbers
DOA	degree of artificiality
DOM	Digitales Oberflächenmodell
<i>et al.</i>	<i>et alii</i> (lat. und andere)
EU	Europäische Union
f	folgende
FE	Fernerkundung
ff	fortlaufend folgende
FIR	finite impulse response
FT	Fouriertransformation
GIF	Graphics Interchange Format

GIS	Geographisches Informationssystem
goett	Göttinger Wald (Datensätze)
GPS	global positioning system
HPT	Homogene Punktransformation
IDFT	Inverse diskrete Foriertransformation
IMU	inertial measurements units
INS	inertial navigation system
IPT	Image Processing Toolbox
ISODATA	Iterative Self-Organizing Data Analysis Technique
ITC	individual tree crown
KI	Künstliche Intelligenz
kNN	k Nearest Neighbour
L.	Taxonomie nach Carl von Linné (bei Pflanzenamen)
lamotte	La Motte Plantation (Datensätze)
LIDAR	light detection and ranging
LMR	linearer Maßstabsraum
LOG	Laplacian of Gaussian
LSI	linear shift-invariant
LUT	Look-Up-Table
ME	Mustererkennung
MED	minimum Euclidean distance
ML	maximum likelihood
namb	Namibia (Datensatz)
nDOM	normalisiertes Digitales Oberflächenmodell
NDVI	normalized difference vegetation index
NIR	near infrared
PDF	probability density function
PSF	point spread function
RADAR	radiowave detection and ranging
REGMAX	Regionale Maxima (Verfahren)
RG	Region Growing
RGB	RotGrünBlau (Farbbild)
SE	Strukturelement
SSE	sum of squared error
Tab.	Tabelle
TF	Transferfunktion
TIFF	Tagged Image File Format
usw.	und so weiter
UV	ultraviolet
v.a.	vor allem
vgl.	vergleiche
WST	Wasserscheidentransformation (Verfahren)
z.B.	zum Beispiel

# 1 Einleitung

## 1.1 Motivation und Fragestellung

Der Bedeutungsgewinn der Geoinformatik an der interdisziplinären Schnittstelle zwischen Geowissenschaften und Informatik hat auch und vor allem in der Geographie den Bedarf an räumlich hochauflösenden und qualitativ hochwertigen Geo-Daten enorm gesteigert. Eine elegante Antwort auf die Frage nach der Herkunft solcher Daten ist die automatische Objektextraktion aus Fernerkundungsdaten, bei der die Position bestimmter Objekte auf der Erdoberfläche durch Bildanalyseverfahren gewonnen wird. Dafür werden zwei wissensbasierte Konzepte genutzt: *„Auf der einen Seite die computergerechte Formalisierung des Wissens (Modellbildung) und auf der anderen Seite die Nutzung des Wissens für die Extraktion (Strategie)“* (HINZ 2003: 6).

Die Extraktion von Bäumen stellt dabei aufgrund des uneinheitlichen Erscheinungsbilds und Maßstabs eine besondere Herausforderung dar, der sich seit nunmehr 20 Jahren Wissenschaftler aus verschiedenen Forschungsrichtungen widmen. Die meisten Impulse stammen dabei aus der Forstwirtschaft, daneben sind aber auch technische Hochschulen sowie Institute aus Geowissenschaften, Photogrammetrie und Informatik und vertreten.

In der Forstwirtschaft liegt das Hauptaugenmerk auf der Bewirtschaftung von Wäldern zur Produktion von Holz, dem als nachwachsendem Rohstoff im Hinblick auf einen nachhaltigen Ressourcenverbrauch eine wichtige Rolle zukommt. Zum Management von Wäldern sind aufgrund der langen Wachstumszyklen regelmäßige Erhebungen der Bestände erforderlich. In Deutschland wird eine solche Forstinventur alle zehn Jahre durchgeführt und besteht aus drei Schritten (vgl. DIEDERSHAGEN *et al.* 2003B: 157): Zunächst wird der gegenwärtige Status wichtiger Kenngrößen wie Artenverteilung, Gesundheitszustand, Holzvolumen und Zuwachs erhoben. In den Schritten zwei und drei werden dann die Planerfüllung für die vergangene Dekade betrachtet und die Ziele für die nächsten zehn Jahre gesetzt. Dabei werden gleichmäßig über den Wald verteilte Stichprobenflächen zugrunde gelegt, die am Boden erhoben und auf die Fläche hochgerechnet werden (vgl. KÖRBER *et al.* 2005). Unterstützend werden Luftbilddauswertung, photogrammetrische Messungen und Wachstumsmodelle eingesetzt (HYYPÄ *et al.* 2001: 58).

Die Notwendigkeit einer automatisierten Forstinventur wird dabei meistens mit den hohen Kosten der konventionellen Inventur begründet (KÄTSCH & KUNNEKE 2006, KÖRBER *et al.* 2005, HYYPÄ *et al.* 2001, HYYPÄ *et al.* 2005), die vor allem vom hohen Arbeitsaufwand der Erhebung im Feld und bei der manuellen Auswertung verursacht werden. So liegen die Kosten der konventionellen Forstinventur in Europa bei 15-20€/ha (HYYPÄ *et al.* 2001: 58, HYYPÄ & HYYPÄ 1999: 5). Der automatisierten Inventur wird typischerweise die Hypothese zu Grunde gelegt, dass durch Messungen von: *„major individual tree characteristics, such as height of the tree, tree species, and crown diameter, it is possible to derive other valuable tree characteristics, such as stem diameter, stem volume, and age for the same tree and to use these information to calculate standwise forest information“* (HYYPÄ *et al.* 2005: 144). Daraus resultiert für forstwirtschaftliche Anwendungen die zusätzliche Herausforderung, dass sekundäre Parameter abgeleitet werden müssen, die gar nicht direkt im Bild zu sehen sind.

Neben der ökonomischen Bewirtschaftung erfüllen Wälder aber auch eine Vielzahl weiterer wichtiger Funktionen, die ebenfalls erheblich von der Entwicklung und Verbesserung von automatisierten

Verfahren zur Baumextraktion profitieren könnten. Besonders hervorzuheben sind dabei die biologische Diversität und die Schutzfunktion von Wäldern sowie ihre Rolle in geosystemaren Stoffkreisläufen, so dass auch nationale und internationale Umweltpolitik und die damit verbundenen Programme und Vereinbarungen (Kyoto-Protokoll, FFH-Richtlinien, Biodiversitäts-Konvention) einen steigenden Bedarf an genauen und aktuellen Informationen über Wälder induzieren. Wesentliche Anwendungen liegen daher auch im Umwelt-Monitoring, in der Bewertung ökologischer Risiken, der Verhinderung von Bränden sowie im Klimaschutz allgemein, die natürlich ihrerseits Bestandteil einer zeitgemäßen Forstwirtschaft sind. An dieser Stelle sei darauf hingewiesen, dass mit der Entwicklung solcher Technologien die Notwendigkeit eines verantwortungsvollen Einsatzes einhergeht, da sie natürlich auch missbräuchlich, etwa zum gezielten Raubbau an wertvollen Spezies, eingesetzt werden können. Daneben gibt es eine breite Palette von nicht Wald-bezogenen Anwendungen, angefangen beim Schutz von Hochspannungsleitungen (vgl. DIEDERSHAGEN *et al.* 2003A: 259) bis hin zu 3D-Stadtmodellen (vgl. STRAUB & HEIPKE 2001).

Die Entwicklung von Verfahren zur automatischen Extraktion vom Bäumen aus Fernerkundungsdaten ist eng mit der Entwicklung der Hardware, also der Sensoren und der Rechenleistung der verarbeitenden Computer, verbunden. Während frühe Arbeiten auf dem Gebiet durchweg gescannte Luftbilder als Datengrundlage verwendeten (PINZ 1989, GOUGEON & MOORE 1988, GOUGEON 1995, POLLOCK 1996, BRANDBERG & WALTER 1998, LARSEN & RUDEMO 1997), kamen in jüngerer Zeit vorwiegend mit Hilfe von Laserbefliegungen (LIDAR) erhobene Digitale Oberflächenmodelle zum Einsatz (HYYPÄ *et al.* 2001, PERSSON 2001, PERSSON *et al.* 2002, HEURICH *et al.* 2004, Morsdorf *et al.* 2004). Neben den Oberflächenmodellen gibt es außerdem erste Ansätze zur direkten Extraktion aus den Vektor-Rohdaten (vgl. MORSDORF *et al.* 2004: 260). Technische Voraussetzung für die ausgereiften Systeme waren dabei die Verfügbarkeit von Differential-GPS sowie von *inertial measurements units* (CAMPBELL 2002: 237, NÆSSET *et al.* 2004: 484), die eine genaue Lagebestimmung der Sensorplattform ermöglichen. Mittlerweile existiert eine Vielzahl von LIDAR-basierten Verfahren (eine sehr gute Übersicht findet sich in NÆSSET *et al.* 2004), die insbesondere bei hochauflösenden Datensätzen und einfachen Waldsituationen bereits gute Ergebnisse liefern.

Andererseits bleiben die hohen Kosten für Laserbefliegungen (1-5 €/ha nach HYYPÄ *et al.* 2005: 145) und die damit verbundene geringe Verfügbarkeit insbesondere hochauflösender Daten entscheidende Hindernisse für eine breite Anwendung. Auch die Erzeugung von Höhenmodellen durch photogrammetrische Verfahren scheint vorerst keine praktikable Alternative, da die Objektregistrierung im Wald nur sehr eingeschränkt funktioniert (KÄTSCH 2002: 211).

Aus diesem Grund sollten Ansätze, die auf optischen Bildern (im Folgenden auch radiometrische Daten) arbeiten, nicht aus dem Auge verloren werden und bilden daher auch den Schwerpunkt dieser Arbeit. Insbesondere können durch die Verbindung beider Formen von Daten gleich bleibend gute Ergebnisse zu niedrigeren Kosten erzeugt werden, wie HYYPÄ *et al.* (2005) zeigen. Mittelfristig ebenfalls interessant werden könnten Bilder aus dem Weltall, da kommerzielle Satelliten mit Auflösungen von unter einem Meter in den für die Objektextraktion interessanten Bereich vorstoßen.

Zusammenfassend kann der Stand der Forschung dahingehend bewertet werden, dass mittlerweile eine Vielzahl von Verfahren existiert, die bei einzelnen Problemen gute Ergebnisse liefern. Gleichzeitig sind die untersuchten Szenen aber sehr unterschiedlich in Bezug auf ihren Inhalt (STRAUB & HEIPKE 2004: 62). Daher sollte das Hauptaugenmerk zukünftiger Arbeiten nicht mehr darauf gerichtet werden,

weitere Einzelverfahren zu entwickeln, sondern darauf, die bestehenden Verfahren zusammen zu führen und damit für eine größere Menge von Problemstellungen nutzbar zu machen. Dabei scheint es unwahrscheinlich, dass ein einzelnes Verfahren die Gesamtlösung sein wird: *"It is most likely that none of today's methods are alone able to handle all types of forests. Comparative studies of different segmentation methods with different types of forests are therefore of importance in order to clarify how much a method is reliable with respect to a certain type of forest. This knowledge can, for instance, be used to build up an expert system which is supposed to be able to extract individual tree crowns in any type of forests. Comparative studies among the existing methods are rare so far"* (ERIKSON *et al.* 2006: 1).

Die zentrale Anforderung an ein solches Expertensystem wäre also, dass es möglichst universell im Sinne von „vielseitig einsetzbar“ sein sollte. Damit rückt die Frage nach der „Übertragbarkeit“ von Verfahren auf andere Situationen in den Mittelpunkt des Interesses, zu der es bislang nur sehr wenige Untersuchungen gibt. Dazu soll die vorliegende Arbeit einen Beitrag leisten. Es wird der Fragestellung nachgegangen, wie ein künftiges Expertensystem zur Baumextraktion aufgebaut sein könnte, was die einzelnen Komponenten leisten müssen und wie ein konzeptioneller Zugriff auf die Frage der Eignung von Verfahren für eine gegebene Problemstellung gewonnen werden kann.

## 1.2 Aufbau der Arbeit

In Kapitel 2 wird zunächst in benötigte Grundlagen aus den Bereichen Fernerkundung, Mustererkennung und Bildverarbeitung eingeführt. Da es sich bei der bearbeiteten Thematik um ein sehr interdisziplinäres Feld handelt, das aus Sicht eines Geographen (und auch Geoinformatikers) Konzepte erfordert, die eher am Rande der fachspezifischen Methodik liegen oder bereits über diese hinaus gehen, werden die Grundlagen besonders ausführlich behandelt. Dies gilt insbesondere für die Mustererkennung und die Bildverarbeitung, bei der Fernerkundung wird sich dagegen auf die Vorstellung der unmittelbar relevanten Konzepte beschränkt. Da sie eine ausgezeichnete Stellung sowohl im Rahmen der Objektextraktion als auch in der vorliegenden Arbeit genießen, wird den Segmentierungsverfahren ein eigenständiger Abschnitt gewidmet, auch wenn diese als Teil der Bildverarbeitung zu verstehen sind.

In Kapitel 3 werden bekannte Verfahren zur Extraktion von Bäumen aus Fernerkundungsdaten vorgestellt und im Hinblick auf ihre Verwendbarkeit in einem künftigen Expertensystem untersucht und bewertet. Eine solch umfassende Bestandsaufnahme und systematisch vergleichende Evaluation wird als vorbereitender Schritt für ein Expertensystem als unabdingbar betrachtet. Da aber nicht alle Verfahren in die weitergehende Betrachtung einfließen, sei zum Verständnis der folgenden Kapitel empfohlen, zumindest die Abschnitte 3.1, 3.3, 3.7 sowie 3.12 zu lesen. Insbesondere soll untersucht werden, welche expliziten und impliziten Voraussetzungen die Verfahren an eingehenden Datensätze stellen, um so erste Anhaltspunkte für einen konzeptionellen Zugriff auf die Frage der Übertragbarkeit zu gewinnen.

In Kapitel 4 werden die Anforderungen an ein Expertensystem zur automatischen Extraktion von Bäumen aus Fernerkundungsdaten spezifiziert und ein Vorschlag für dessen Aufbau gemacht. Dabei werden die einzelnen Komponenten sowie ihre Relationen zueinander eingehender vorgestellt und diskutiert. Insbesondere wird der Modellbildung und der Frage geeigneter Merkmale zur Klassifikation von Bäumen besondere Aufmerksamkeit gewidmet. Da das System die doppelte Aufgabe hat,

sowohl ein geeignetes Verfahren auszuwählen als auch dieses zu steuern, wird auch der Frage nach geeigneten Steuerungsparametern nachgegangen und ein schematischer Ablauf für die Extraktion entwickelt, um die teilweise komplexen Verfahren in kleinere und individuell kombinierbare Module zu zerlegen.

Von zentraler Relevanz für die Fragestellung ist, wie sich die Verfahren steuern lassen und ob durch Einführung neuer Parameter der Funktionsumfang vergrößert und somit die Universalität erhöht werden kann. Dafür wird in Kapitel 5 exemplarisch eine Auswahl von Segmentierungsverfahren aus der Literatur durch neue Parameter erweitert. Bei den implementierten Verfahren wird weiterhin detailliert analysiert, welche Voraussetzungen in die Verfahren eingehen. Dafür wird auf ein Set von Modell-Annahmen zurückgegriffen, das in Kapitel 4 entwickelt wurde.

In Kapitel 6 werden die implementierten Verfahren dann an unterschiedlichen Bildern getestet, die vorab kurz vorgestellt werden. Das Ziel ist dabei ausdrücklich nicht, besonders genaue Ergebnisse für einzelne Datensätze zu erzielen. Stattdessen wird zum einen untersucht, wie die eingeführten Parameter das Ergebnis beeinflussen, ob der Funktionsumfang durch sie erhöht werden kann, und ob sie zur Steuerung des Verfahrens geeignet sind. Zum anderen wird geprüft, ob das Set von Modell-Annahmen geeignet ist, die Voraussetzungen von Verfahren zu charakterisieren und damit deren Eignung für einzelne Datensätze zu beschreiben.

Anschließend werden die erzielten Ergebnisse in Kapitel 7 diskutiert und bewertet. Die gewonnenen Erkenntnisse werden genutzt, um den vorgeschlagenen Ablauf zu erweitern, und es wird ein Ausblick gegeben, welche ergänzenden Studien als Bausteine auf dem Weg zu einem allgemeinen Expertensystem zur Extraktion von Bäumen aus Fernerkundungsdaten in Betracht gezogen werden können.

## 2 Grundlagen

### 2.1 Fernerkundung

Eine übliche Definition für Fernerkundung (*remote sensing*, FE) ist die Akquisition von Daten ohne in physischem Kontakt mit dem Objekt zu sein (z.B. ELACHI & VAN ZYL 2006: 1, SABBINS 1987: 1). In dieser allgemeinen Form wird also auch mit einem Computertomograph oder dem *Very Large Telescope* in Chile Fernerkundung betrieben. Im engeren Sinne wird unter Fernerkundung die Beobachtung der Erdoberfläche und der Atmosphäre durch elektromagnetische Strahlung verschiedener Wellenlängen verstanden. Die verwendbaren Frequenzbänder sind dabei stark von der Absorption und Transmission der Atmosphäre bestimmt. Während diese im Bereich des sichtbaren Licht am durchlässigsten ist (darum hat sich das menschliche Sehsystem auf dieses Frequenzband spezialisiert), ist sie beispielsweise für Röntgen- oder Mikrowellen weitgehend opak (eine genaue Übersicht findet sich z.B. in SABBINS 1987: 5).

Die erste Luftaufnahme wurde vom Pariser Fotografen Gaspard Felix Tournachon im Jahre 1858 von einem Ballon aus aufgenommen (LILLESAND & KIEFER 1979: 36). Heute werden als Trägerplattformen vor allem Flugzeuge (*airborne remote sensing*) und Satelliten (*spaceborne remote sensing*) eingesetzt. Während bei Flugzeugen bis heute mit Luftbildkameras auch noch konventionelle Fotografie zum Einsatz kommt, hat gerade die Notwendigkeit der nichtmateriellen Datenübertragung vom Satelliten die Entwicklung digitaler Sensoren sehr befördert.

Die digitalen Sensoren lassen sich in verschiedene Kategorien einteilen. Zunächst wird zwischen passiven und aktiven Sensoren unterschieden: Während erstere von der Erde reflektiertes Sonnenlicht oder emittierte thermische Strahlung registrieren, senden letztere die zu messende elektromagnetische Strahlung selbst aus. Weiterhin stellt jeder Sensor einen Kompromiss zwischen den konkurrierenden Zielen hoher räumlicher, spektraler, radiometrischer und temporaler Auflösung dar. Für die hier geschilderten Zwecke bieten sich Sensoren an, die vordefinierte spektrale Bänder (im Folgenden auch Kanäle) in hoher räumlicher Auflösung aufzeichnen und *multispectral mappers* genannt werden (eine komplette Kategorisierung findet sich in ELACHI & VAN ZYL 2006: 2). Demgegenüber wird beispielsweise bei der hyperspektralen Fernerkundung eine geringere räumliche Auflösung zu Gunsten einer sehr hohen spektralen Auflösung in Kauf genommen.

Da diese Arbeit nicht den Raum für eine umfassende Übersicht über alle Gebiete der FE bietet, werden nur exemplarisch einige relevante Bereiche vorgestellt. Um aber wenigstens die Vielfältigkeit anzudeuten, soll zum Einstieg die Frage aufgeworfen werden, was ein Fernerkundungssensor überhaupt misst. Wird ein Foto aufgenommen, beispielsweise mit einer Digitalkamera aus einem Flugzeug, beeinflussen eine große Anzahl von Prozessen das Ergebnis.

Die Kamera misst die Strahlung in einem bestimmten Wellenlängenbereich, die das Objektiv erreicht. Diese gelangt aber auf sehr unterschiedlichen Wegen dort hin. Sei vereinfachend angenommen, dass sämtliche Strahlung ursprünglich von der Sonne kommt (was zumindest tagsüber akzeptabel erscheint) und auf die Erklärung verzichtet, wie die Strahlung in der Sonne erzeugt wird, dann ist der relevante Teil der einfallenden Strahlung von der Erdoberfläche reflektiertes Sonnenlicht. Die vom Sensor gemessene physikalische Größe ist der ankommende Strahlungsfluss  $\Phi$  (Energie pro Zeit) bzw.

die Strahldichte  $L$  (*radiance*), also der Strahlungsfluss pro projizierte abstrahlende Fläche pro Raumwinkel:

$$(2.1) \quad L = \frac{d\Phi}{d\Omega dA \cos \theta}$$

Der Winkel  $\theta$  liegt dabei zwischen der Oberflächennormalen der abstrahlenden Fläche und der Beobachtungsrichtung.

Um zu bestimmen, was aufgenommen wurde, müssen Entfernung und Orientierung zwischen Kamera und Erdoberfläche bekannt sein. Verfahren zur Georeferenzierung und Projektion von Daten werden oft unter dem Begriff „geometrische Vorverarbeitung“ zusammengefasst. Für die Messung ist es weiterhin entscheidend zu wissen, wie die Kamera auf bestimmte Intensitäten nach Frequenzbereichen reagiert. Damit der Zusammenhang zwischen den Grauwerten im Bild (in der Fernerkundung oft auch *Digital Numbers*, DN, genannt) und der gemessenen Strahldichte bekannt ist, muss die Kamera kalibriert werden.

Die Strahldichte selbst ist noch keine Eigenschaft des beobachteten Objektes, sondern hängt insbesondere von der Beleuchtung und der Orientierung der Fläche ab. Verfahren, die das ausgleichen (etwa die Streifen im Bild, die durch die Befliegungsstreifen im Flugzeug entstehen), zählen zu den radiometrischen Korrekturverfahren.

Die erste echte Materialeigenschaft, die sich aber bei bekannten äußeren Größen ableiten lässt, ist die Albedo, also der Anteil des eingehenden Strahlungsflusses (pro Fläche = Bestrahlstärke  $E$ , *irradiance*) im sichtbaren Bereich, der diffus reflektiert wird. Der wellenlängenabhängige Reflexionsgrad (Reflektanz, *reflectance*) wird als spektrale Signatur eines Materials bezeichnet. Neben der Wellenlänge ist die Streuung aber meistens auch richtungsabhängig, und zwar sowohl von der Einfallrichtung als auch der Ausfallrichtung. Daraus ergibt sich die *Bidirectional Reflectance Distribution Function* (BRDF) des Materials. Materialien, bei denen diese konstant ist (die also aus allen Richtungen gleich aussehen), werden lambertsche Strahler genannt, in guter Näherung wird diese Bedingung von einem weißen Blatt Papier erfüllt. Es ist also notwendig, die Interaktion zwischen Materie und Strahlung genauer zu verstehen. Daneben besteht die Möglichkeit, die spektralen Signaturen und die BRDF empirisch zu ermitteln, was der Spektroskopie zugeordnet wird.

Leider fällt neben dieser erwünschten Strahlung auch weitere beim Sensor ein. So wird beispielsweise ein Teil des einfallenden Sonnenlichtes durch die Atmosphäre absorbiert und gestreut (Rayleigh-Streuung). Ebenso wird auf dem Weg vom Objekt zum Sensor ein Teil der Strahlung weg und Strahlung von anderen Objekten zum Sensor hin gestreut. Das Verhalten der Atmosphäre ist wiederum abhängig von der Wellenlänge sowie verschiedenen anderen Parametern, etwa der Luftfeuchtigkeit und anderen Stoffkonzentrationen. Um diese Fehler radiometrisch korrigieren zu können, ist es notwendig, die Wechselwirkung von Strahlung und Atmosphäre zu verstehen, was zu sehr komplexen Modellen führt. Die entsprechenden Methoden werden als atmosphärische Korrekturverfahren zusammengefasst.

### 2.1.1 Struktur und Reflexionsspektrum grüner Vegetation

Um die Interaktion von Strahlung mit grüner Vegetation zu verstehen, soll zunächst die physiologische Struktur eines Laubblattes betrachtet werden (siehe dazu CAMPBELL 2002: 459ff), die in Abb. 2.1 dargestellt ist.

#### Physiologische Struktur eines Laubblattes

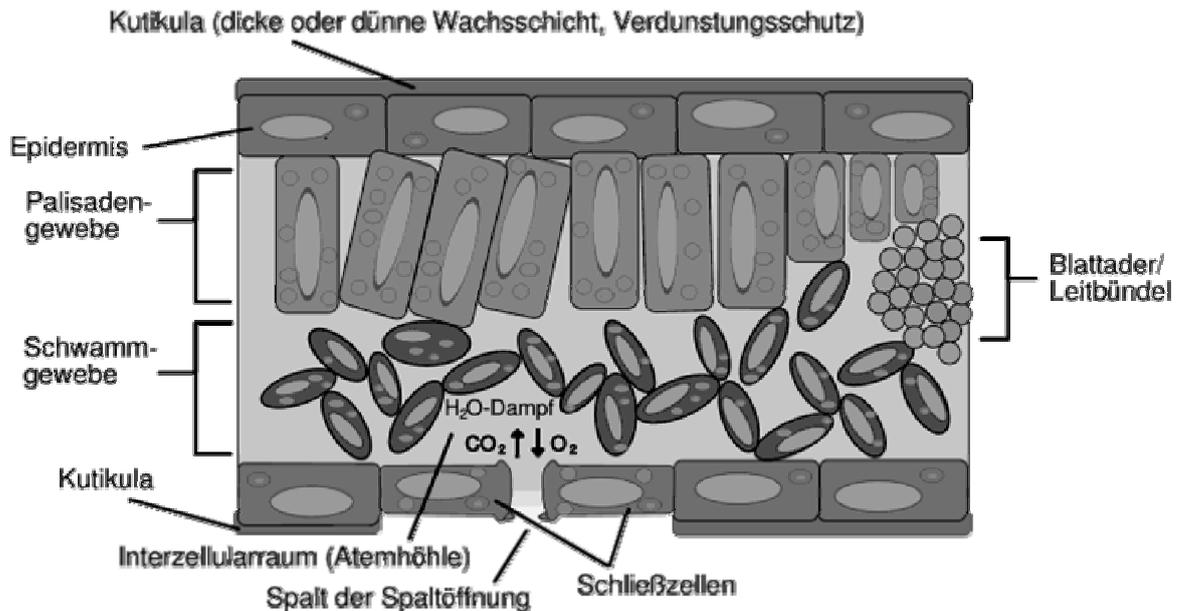


Abb. 2.1: Aufbau eines Laubblattes (Wikipedia)

Ein Laubblatt setzt sich aus mehreren Schichten zusammen. Die oberste Schicht wird als obere Epidermis bezeichnet und besteht aus lückenlos angeordneten Spezialzellen, die in der Regel keine Chloroplasten enthalten, also auch keine Photosynthese betreiben. Die Epidermis wird nach außen durch eine dünne durchsichtige Wachsschicht namens Kutikula abgeschlossen, die den Verlust von Wasser verhindert. Auf der Unterseite wird das Blatt durch die untere Epidermis abgeschlossen, die ähnlich aufgebaut ist wie die obere. Zusätzlich hat sie aber Spaltöffnungen (Stomata), die den Gasaustausch und damit den Wasserverlust der Pflanze regulieren. Die Stomata bestehen meistens aus zwei Schließzellen und ihre Position auf der Unterseite maximiert die Lichttransmission durch die obere Epidermis und minimiert den Flüssigkeitsverlust (CAMPBELL 2002: 460). Bei Schwimmpflanzen sind sie abweichend auf der Oberseite zu finden, bei Gräsern auf beiden Seiten.

Zwischen den Epidermisschichten befindet sich das Mesophyll, das sich wiederum in zwei Schichten unterteilt, das Palisadengewebe und das Schwammgewebe. Das Palisadengewebe besteht aus bis zu drei Schichten von langgestreckten, senkrecht angeordneten Zellen, die den Großteil der Chloroplasten enthalten und die meiste Photosynthese betreiben. Das darunter liegende Schwammgewebe besteht aus unregelmäßig geformten Zellen mit großer Oberfläche sowie zahlreichen Hohlräumen. Seine Hauptaufgabe ist es, den Austausch von  $\text{CO}_2$  und  $\text{O}_2$  zu gewährleisten. Nicht zum Mesophyll gehören die Leitbündel, die Wasser und Mineralien ins Blatt transportieren.

## Reflexionsspektrum grüner Vegetation

Das Spektrum des Blattes wird in den verschiedenen Wellenlängenbereichen von unterschiedlichen Schichten im Blatt dominiert. Generell gilt, dass die einfallende Strahlung teils reflektiert, teils absorbiert und teils transmittiert wird. Der Anteil der reflektierten Strahlung wird als Reflektanz oder Reflexionsgrad bezeichnet. Die Spektren von Reflexion und Transmission sind sehr ähnlich, da es sich ja letztendlich nur um nach oben respektive nach unten gestreute Strahlung handelt (darum hat ein Blatt von oben und von unten die gleiche Farbe).

Die Farbe im sichtbaren Bereich wird hauptsächlich vom Chlorophyll bestimmt. Im gesamten sichtbaren Bereich wird der Großteil der Strahlung zur Energiegewinnung durch Photosynthese absorbiert. Es gibt mehrere Arten von Chlorophyll, von denen Chlorophyll a und Chlorophyll b die wichtigsten sind (CURRAN 1985: 24). Deren Absorptionsspektren haben Maxima bei Wellenlängen von 430nm und 662nm bzw. 454nm und 643nm. Insgesamt wird also blaues und rotes Licht besonders stark absorbiert, grünes dagegen weniger. Deshalb erscheinen uns Pflanzen grün.

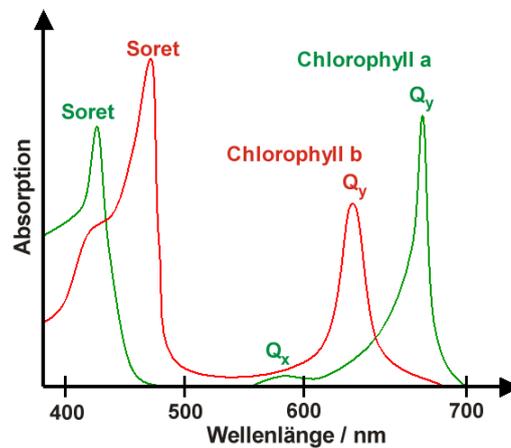
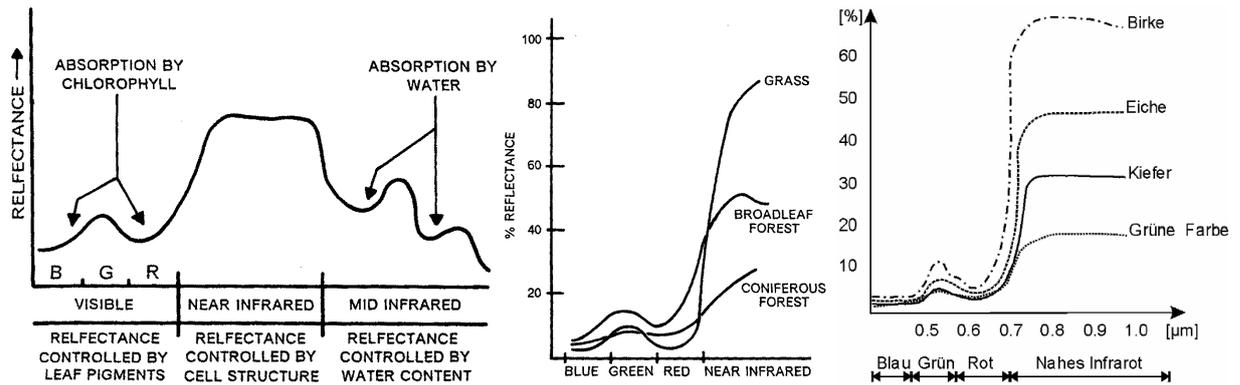


Abb. 2.2: Absorptionsspektren von Chlorophyll a und b (Wikipedia)

Im Nahen Infrarot wird die Wechselwirkung nicht von den Pigmenten sondern vom lückenhaften Schwammgewebe bestimmt, da Kutikula und Epidermis für diese Strahlung weitgehend durchsichtig sind. Im Bereich des nahen Infrarots wird der Großteil der Strahlung reflektiert bzw. transmittiert, daher steigt der Reflexionsgrad bei rund 700nm stark an. Wäre dieser Spektralbereich für das menschliche Auge sichtbar, würden uns Pflanzen also nicht grün, sondern (infra-)rot erscheinen. Der Anstieg wird auch als *red edge* bezeichnet und ist sowohl für einzelne Pflanzenarten als auch für ihren Zustand charakteristisch. So wurde von COLLINS (1978) gezeigt, dass bei älteren Pflanzen ein „red-shift“ auftritt, d.h. das *red edge* verschiebt sich zu längeren Wellenlängen hin. Dieser Frequenzbereich eignet sich somit besonders für Vegetationsstudien.

Im weiteren nahen und mittleren Infrarot wird das Spektrum schließlich von der Absorption durch Wasser dominiert. Das H<sub>2</sub>O-Molekül hat drei Hauptabsorptionslinien bei 1,4µm, 1,9µm und 2,7µm (CURRAN 1985:25), die bestimmten Schwingungsübergängen entsprechen. Die Absorption nimmt mit der Stärke des Blattes und dem Wassergehalt zu. Hohe Reflexion in diesem Bereich deutet daher oft auf einen schlechten Gesundheitszustand hin (BUSCHMANN 1993: 446). Es sei darauf hingewiesen, dass gerade diese Absorptionslinien nur schlecht per Fernerkundung zu messen sind, da das Wasser in der Atmosphäre natürlich ebenso Strahlung dieser Wellenlänge absorbiert.

Zusammengenommen ergibt sich für grüne Pflanzen das Reflexionsspektrum in Abb. 2.3.



**Abb. 2.3: Reflexionsspektrum grüner Vegetation. Links: bestimmende Bereiche (CAMPBELL 2002: 463); Mitte: Spektren verschiedener Pflanzen (CAMPBELL 2002: 462); Rechts: Spektren verschiedener Baumarten (STRAUB 2003A: 47)**

Das Reflexionsspektrum eines einzelnen Blattes erklärt natürlich noch nicht vollständig das Reflexionsverhalten einer ganzen Baumkrone. Da Baumkronen aus vielen Schichten von Blättern bestehen, die sich gegenseitig beschatten, haben sie insgesamt einen geringeren Reflexionsgrad (CAMPBELL 2002: 463). Der relative Rückgang ist im sichtbaren Bereich stärker als im NIR, da die unteren Schichten infrarote Strahlung teilweise wieder nach oben reflektieren, sichtbare dagegen eher absorbieren. Ein sehr komplexes Modell für den Strahlungstransport in Laubkronen findet sich bei ASRAR (1989). Weiterhin kann der Boden im Hintergrund Auswirkungen auf das Spektrum der einer Vegetationsdecke haben (siehe CURRAN 1985: 26f).

### Vegetationsindizes

Die charakteristischen Eigenschaften von Vegetationsspektren sind geeignet, Vegetation auf Luftbildern zu erkennen. Dafür werden standardmäßig verschiedene Vegetationsindizes verwendet, die spektrale Reflektanzen in verschiedenen Wellenlängenbereichen zueinander in Beziehung setzen. Sofern diese Bänder zur Verfügung stehen, eignen sich dafür am Besten Rot (R) und Nahes Infrarot (NIR), da das *red edge* besonders charakteristisch ist.

Der Ratio-Vegetationsindex (*ratio vegetation index*, RVI) ist:

$$RVI = \frac{NIR}{R}$$

Er wird groß für Vegetation und geht gegen Null für Wasserflächen. Noch besser ist der *normalized difference vegetation index* (NDVI), der einen einheitlichen Wertebereich von -1 bis 1 annimmt:

$$(2.2) \quad NDVI = \frac{NIR - R}{NIR + R}$$

Steht kein infraroter Kanal zur Verfügung, wie bei farbigen Luftbildern, so wird stattdessen der grüne Kanal (G) verwendet. Entsprechend zum NDVI wird der *Degree of Artificiality* definiert:

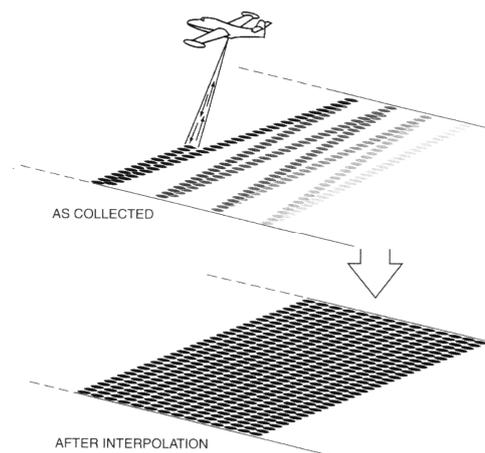
$$DoA = \frac{G - R}{G + R}$$

Daneben gibt es eine große Zahl weiterer Vegetationsindizes, etwa den bodenangepassten SAVI (*soil adjusted vegetation index*), der einen Korrekturparameter für die Vegetationsdichte einführt. Für eine umfangreiche Übersicht siehe BUSHMANN (1993: 445). In der Praxis macht es aber nicht so einen großen Unterschied, welcher Index verwendet wird (CAMPBELL 2002: 467).

### 2.1.2 LIDAR

LIDAR (*light detection and ranging*) ist ein aktives Verfahren zur Messung von Entfernungen. Das Verfahren ist mit dem RADAR (*radiowave detection and ranging*) verwandt, nur dass anstelle von Mikrowellen ein Laser im ultravioletten (UV), sichtbaren oder nahen infraroten (NIR) Spektrum verwendet wird. LIDAR sendet in regelmäßigen Abständen kohärente Lichtpulse aus und misst die Zeit, bis das reflektierte Signal eintrifft (CAMPBELL 2002: 237).

LIDAR wird als bildgebendes Verfahren noch nicht sehr lange eingesetzt, da eine genaue Positions- und Lagebestimmung erst durch *inertial measurements units* (auch *inertial navigation systems*, INS) und GPS (*global positioning system*) möglich wurden (CAMPBELL 2002: 237). Ein schnell oszillierender Spiegel bewegt dabei den Laser-Strahl quer zur Flugrichtung, wodurch viele tausend Punkte pro Sekunde erfasst werden können. Die Fläche, die der Sensor auf einmal beleuchtet, wird als Fußabdruck (*footprint*) bezeichnet. Je kleiner der Fußabdruck und je größer die Puls-Frequenz sind, desto höher auflösende Datensätze können erzeugt werden. Moderne Systeme arbeiten mit über 100kHz (NÆSSET *et al.* 2004: 483) und können dabei je nach Flughöhe 30 und mehr Punkte pro m<sup>2</sup> aufnehmen. Die Punktwolken in Form von Vektordaten können dann in Raster unterschiedlicher Größe überführt werden. Dabei muss in Zwischenräumen interpoliert werden, um die ungleiche Verteilung der Messpunkte durch das beim Scannen erzeugte charakteristische Muster auszugleichen (siehe Abb. 2.4). Gemeinsam mit der Orientierung und der Flugposition kann so aus den Pulsen ein Digitales Geländemodell (DGM, *digital ground model* oder *digital terrain model*) erzeugt werden.



**Abb. 2.4: LIDAR Aufnahme: Messpunkte und Interpolation (CAMPBELL 2002: 238)**

Neben der Erdoberfläche werden die Pulse natürlich auch von Objekten auf ihr (wie etwa Bäumen) gestreut – falls diese teilweise transparent sind, sogar mehrfach (siehe Abb. 2.5). Daher wird meistens neben dem ersten auch der letzte Puls (*first pulse*, *last pulse*) aufgenommen, bei modernen Systemen teilweise sogar alle Puls-Antworten.

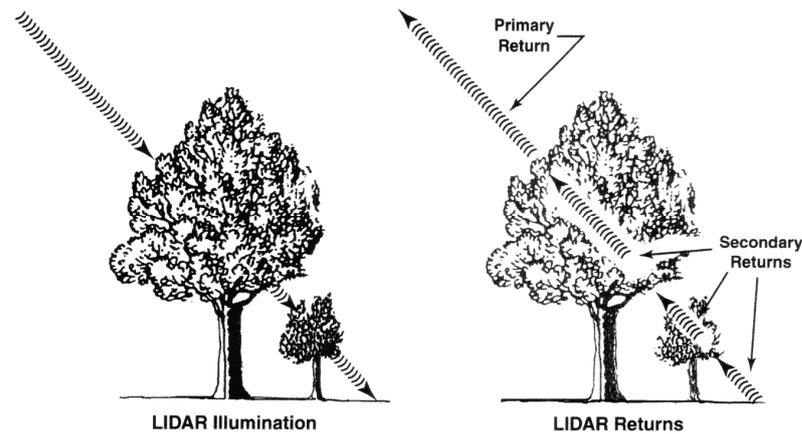


Abb. 2.5: LIDAR Aufnahme: *first pulse* und *last pulse*. (CAMPBELL 2002: 238)

Oftmals wird angenommen, dass der jeweils letzte Puls von der Erdoberfläche kommt und daraus das DGM abgeleitet. Diese Annahme ist aber unzutreffend, wenn undurchsichtige Objekte wie Häuser Pulse reflektieren. Es existiert eine große Bandbreite von unterschiedlichen Filtertechniken um Punkte von solchen Objekten zu eliminieren (VÖGTLE & STEINLE 2004: 316), wobei die mathematische Morphologie (siehe 2.3.8) einen besonders geeigneten Ansatz bietet.

Aus den ersten Pulsen wird dann ein Digitales Oberflächenmodell (DOM, *digital surface model* oder *digital elevation model*) erstellt, das die Objekte auf der Oberfläche mit einschließt. Aus der Differenz wird das für die Objektextraktion nützlichere normalisierte Oberflächenmodell ( $nDOM = DOM - DGM$ ) abgeleitet, das nur die Höhe der Objekte ohne die darunter liegende Geländehöhe enthält. Auf diesem Modell, das im Fall von Baumkronen im Folgenden Digitale Kronenmodell (DKM) genannt wird, können dann Objekte segmentiert und klassifiziert werden (TOVARI & VÖGTLE 2004). In Kapitel 3 werden solche Verfahren vorgestellt, für die Techniken aus der Bildverarbeitung (BV) und der Mustererkennung (ME) benötigt werden.

## 2.2 Mustererkennung

Die Mustererkennung (*pattern recognition*) ist ein Teilgebiet der Künstlichen Intelligenz (KI) und beschäftigt sich nach PAULUS & HORNEGGER (2003) mit der Klassifikation von Objekten anhand von typischen Merkmalen, den beobachteten Mustern. Dabei ist die Methodik unabhängig von der Signalquelle, und die Anwendungen der Mustererkennung sind folglich sehr vielfältig. Beispiele hierfür sind die Erkennung von Buchstaben, Sprache, Gesichtern oder Fingerabdrücken – aber auch von Spam, Aktienkursbewegungen, Krankheitssymptomen oder DNA-Sequenzen.

Zur Begriffsklärung unterteilt NIEMANN (1983: 4) in die Bereiche Klassifikation und Analyse und unterscheidet zwischen einfachen und komplexen Mustern. Einfache Muster sind demnach solche, die als Ganzes betrachtet und einer Klasse zugeordnet werden können, beispielsweise das Bild eines Buchstabens oder ein gesprochenes Wort. Komplexe Muster sind solche, bei denen das nicht möglich ist, also etwa ein Bild eines Schaltplans, das Foto einer Geburtstagsfeier oder ein gesprochener Satz. Komplexe Muster lassen sich jedoch in einfachere Bestandteile zerlegen (NIEMANN 1983: 22).

Die Klassifikation ist die Zuordnung eines einfachen Musters zu einer von endlich vielen Klassen, die Analyse ist die Zuordnung einer symbolischen Beschreibung zu einem komplexen Muster. Manche Autoren führen daneben die Klassifikation komplexer Muster als eigenes Teilgebiet auf (PAULUS &

HORNEGGER 2003). Heute wird unter Mustererkennung hauptsächlich die statistische Klassifikation sensorischer Muster verstanden, und die Analyse komplexer Muster anderen Teilgebieten der KI zugeordnet, etwa dem Bildverstehen oder dem Sprachverstehen. Manchmal werden sie aber auch unter wissensbasierte Mustererkennung oder Musterverstehen behandelt (PAULUS & HORNEGGER 2003: 41).

Im Folgenden sollen zunächst einige Grundbegriffe der Mustererkennung erläutert werden, die von NIEMANN (1983: 2ff) eingeführt und von vielen Autoren übernommen (PAULUS & HORNEGGER 2003, 6 ff, LEHMAN *et al.* 1997: 398 ff) wurden. Die Umwelt  $U$  (*environment*) ist im technischen Sinne die Gesamtheit aller physikalisch messbaren Größen, die durch Vektorfunktionen in verschiedenen Dimensionen dargestellt werden. Nichtmessbare Dimensionen werden nicht berücksichtigt. Da dies offensichtlich sehr umfassend ist, beschränkt sich die ME jeweils auf einen Problemkreis  $\Omega$  (*problem domain*). Jeder Problemkreis ist eine Untermenge der Umwelt  $\Omega \subset U$  und enthält nur die für ein bestimmtes Anwendungsgebiet relevanten Funktionen, die mit geeigneten Sensoren erfasst werden. Die Elemente des Problemkreises werden Muster genannt, haben alle die gleichen Dimensionen und können daher als Vektorfunktion zusammengefasst werden:

$$(2.3) \quad \mathbf{f}(\mathbf{x}) = \begin{pmatrix} f_1(x_1, \mathbf{K}, x_n) \\ f_2(x_1, \mathbf{K}, x_n) \\ \mathbf{M} \\ f_m(x_1, \mathbf{K}, x_n) \end{pmatrix}.$$

Im Fall eines Landsat-MSS-Bildes mit vier Kanälen ist beispielsweise  $n=2$  und  $m=4$ , der Ortsraum ist also zweidimensional und der Merkmalsraum vierdimensional. Beim Farbfernsehen ist  $n=3$  für zwei Raumdimensionen plus Zeit und  $m=3$  für die drei Farben rot, grün und blau.

Der Problemkreis wird disjunkt in Klassen oder Musterklassen  $\Omega_k$  zerlegt, und zwar im Allgemeinen so, dass sich die Muster innerhalb einer Klasse ähnlich sind. Es besteht die Möglichkeit, eine Rückweisungsklasse  $\Omega_0$  einzuführen, die nicht klassifizierbaren Mustern zugewiesen wird.

### 2.2.1 Methodologische Struktur der Mustererkennung

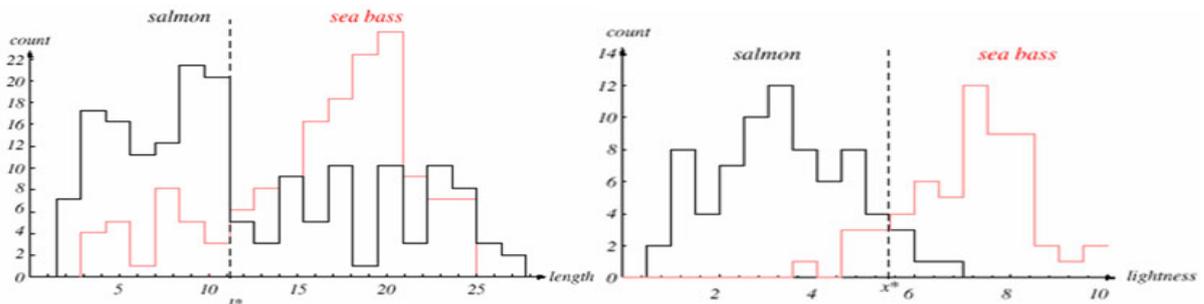
Das folgende Beispiel ist aus DUDA *et al.* (2001: 1 ff) entnommen und veranschaulicht sehr eindrücklich die Grundprobleme der Mustererkennung. Es handelt sich dabei um ein automatisches System zur Sortierung von Fischen in einer Fabrik. Das System soll dabei anhand eines Fotos auf dem Laufband die Arten Lachs (*salmon*) und Wolfsbarsch (*sea bass*) unterscheiden.

Dafür ist es erforderlich, zunächst Unterschiede zwischen den Arten festzustellen und somit unterschiedliche Modelle von ihnen zu entwerfen. Das Hauptziel der Mustererkennung ist nach DUDA *et al.* (2001: 2), Hypothesen von Modellen zu entwerfen, Rauschen, das nicht im Modell ist, zu entfernen und jedes Muster dem Modell zuzuweisen, zu dem es am besten passt.

Im vorliegenden Fall könnten beispielsweise Merkmale wie die Länge, die Farbe, die Anzahl der Flossen oder die Position der Augen zur Modellbildung in Betracht gezogen werden. Sei zunächst die Länge betrachtet: Um diese zu erhalten ist zunächst eine Vorverarbeitung (*preprocessing*) erforderlich, wie beispielsweise die Segmentierung des Bildes. Anschließend werden die gewünschten Merkmale extrahiert (*feature extraction*), wobei die Daten erheblich reduziert werden (im vorliegenden Fall von

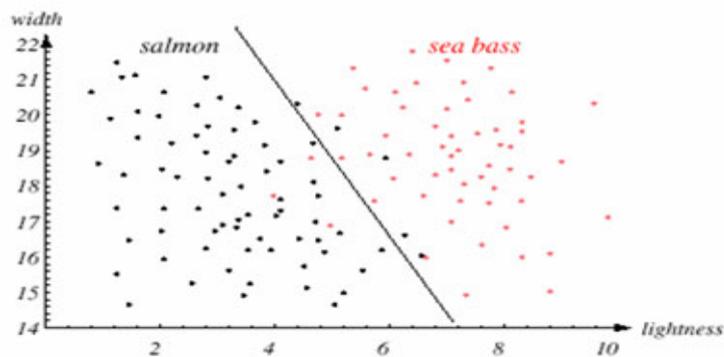
einem Bild zu einer Länge). Die gewonnen Merkmale (*features* oder *properties*) werden an den Klassifikator weitergegeben, der eine Entscheidung über die Art trifft.

Unter der Prämisse, dass Wolfbarsche normalerweise länger seien als Lachse, kann bereits ein einfaches Modell entwickelt werden, das besagt, dass jeder Wolfbarsch eine gewisse (also bestimmte und bestimmbare) Länge hat, die größer ist, als die Länge eines Lachses. Eine Möglichkeit zur näheren Bestimmung ist die Erhebung einer repräsentativen Stichprobe (siehe Abb. 2.6, links).



**Abb. 2.6: Merkmalsextraktion: Länge (links) und Helligkeit (rechts) von zwei Fischarten (DUDA et al. 2001)**

An der Verteilung ist unschwer erkennbar, dass die Länge alleine ein schlechtes Entscheidungskriterium ist. Auf der Suche nach einem besseren Kriterium, findet sich die Verteilung für die Helligkeit in Abb. 2.6 rechts. Die Helligkeit scheint also bereits merklich besser geeignet. Noch besser wird die Unterscheidbarkeit aber, wenn beide Merkmale zu einem Vektor zusammengefügt werden. In Abb. 2.7 sind die Elemente eines solchen zweidimensionalen Merkmalsraumes (*feature space*) dargestellt.



**Abb. 2.7: Merkmalsraum (DUDA et al. 2001)**

Die nächste Aufgabe ist nun, den Merkmalsraum in zwei Teile zu partitionieren. Die schwarze Linie in Abb. 2.7 könnte als Entscheidungsgrenze (*decision boundary*) für den Klassifikator dienen. Allgemeiner wird auch der Begriff Entscheidungsflächen benutzt, wobei diese im 2D Merkmalsraum Linien sind, bei 3 Dimensionen Flächen etc. Da jeweils verschiedene Grenzen möglich sind, stellt sich die Frage nach der optimalen Entscheidungsgrenze und wie sie gewonnen werden kann.

In Abb. 2.8 sind zwei weitere mögliche Entscheidungsflächen zu sehen. Die linke entspricht einem sehr komplexen Modell, würde aber bei der Klassifikation von neuen Mustern wahrscheinlich mehr Fehler produzieren, da sie zu sehr auf die Trainingsstichprobe zugeschnitten ist. Die rechte ist generalisierter und für neue Muster wahrscheinlich besser geeignet, in beiden Fällen bleibt jedoch ein gewisser Fehler.

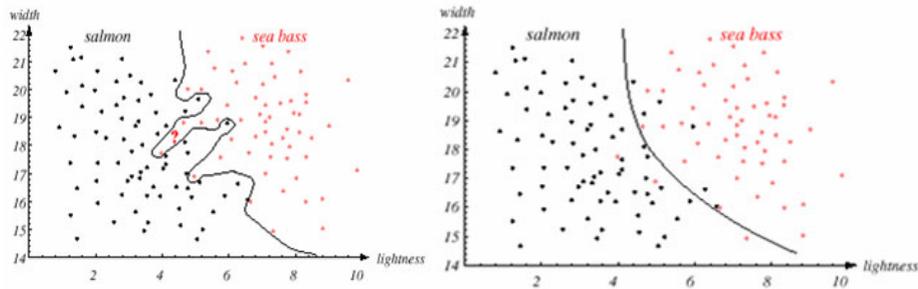


Abb. 2.8: Entscheidungsgrenzen (DUDA *et al.* 2001)

Die entscheidende Aufgabe in der ME ist es also, geeignete Charakteristika der gesuchten Klassen (hier Wolfsbarsche und Lachse) zu finden und zur Modellbildung heranzuziehen. Ein Weg hierzu wäre, den Umfang der Stichprobe drastisch zu erhöhen. Oftmals ist dies aber nur beschränkt möglich oder zumindest zeit- und kostenaufwendig. Alternativ kann versucht werden, zusätzlich *a priori* Wissen über die Klassen in die Modellbildung einfließen zu lassen. Da hier die Modellbildung auf einer höheren Repräsentationsebene stattfindet, könnte dies schon dem Bereich der Analyse oder des Verstehens zugeordnet werden.

Vorweggreifend sei darauf hingewiesen, dass es sich hierbei um zwei unterschiedliche Strategien des (Bild-)Verstehens handelt: von den Daten ausgehend ein Modell zu entwickeln (*bottom-up*) oder das Objektmodell abstrakt aus *a priori* Wissen zu gewinnen (*top-down*) (PINZ 1994: 117) genannt werden. Auf diese Strategien wird im Rahmen der vorgestellten Verfahren in Kapitel 3 und bei der Modellbildung in Kapitel 4.4 eingegangen.

Aus dem geschilderten Beispiel ergibt sich bereits der übliche Ablauf in der Mustererkennung, der in Abb. 2.9 noch einmal schematisch dargestellt ist. Im Folgenden werden die einzelnen Schritte noch einmal kurz diskutiert und abgegrenzt. Eine weiterführende Behandlung findet sich in PAULUS & HORNEGGER (2003: 36 ff) oder DUDA *et al.* (2001: 9 ff).

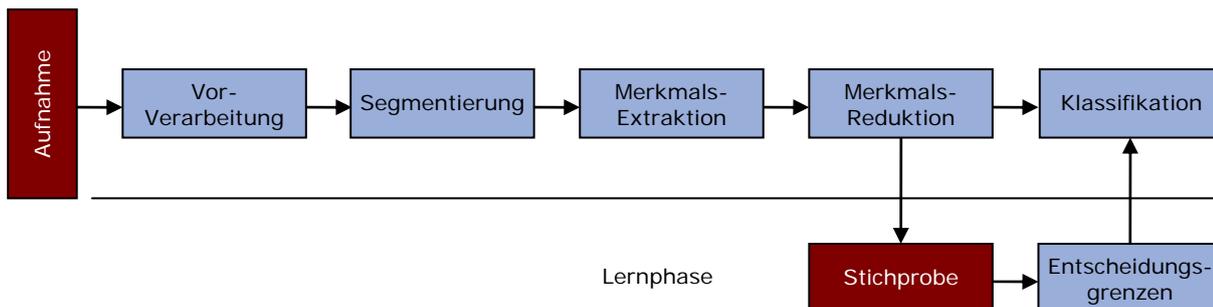


Abb. 2.9: Ablauf in der Mustererkennung (Veränderte Darstellung nach Wikipedia)

### Vorverarbeitung und Segmentierung

Die Vorverarbeitung dient dem Ziel, die weitere Verarbeitung zu vereinfachen. Beispielsweise werden in diesem Schritt unerwünschte Signalanteile wie Rauschen entfernt oder irrelevante durch Segmentierung eliminiert. Dabei bleibt die Art der Repräsentation aber gleich, aus dem Muster wird wieder ein Muster, ein Bild bleibt ein Bild und ein Sprachsignal ein Sprachsignal.

## Merkmalsgewinnung und Merkmalsreduktion

Bei der Merkmalsgewinnung werden die Muster in niedrigerdimensionale Merkmalsvektoren transformiert, die die Klasse des Musters charakterisieren sollen. Die methodische Herausforderung der ME ist es also Merkmale zu wählen, die zu möglichst großer Unterscheidbarkeit der Klassen im Merkmalsraum führen. Außerdem kommt es darauf an, sich auf möglichst wenige aussagekräftige Merkmale zu konzentrieren, da zu hochdimensionale Merkmalsräume die Qualität des Ergebnisses verschlechtern können. Dies widerspricht zunächst vielleicht der menschlichen Intuition (je mehr Merkmale, desto besser), was vermutlich darin begründet ist, dass hochdimensionalen Räume nicht den menschlichen Alltagserfahrungen entsprechen. PAULUS & HORNEGGER (2003: 37) führen zum Beispiel an, dass in einem 50-dimensionalen Hyperraum ein Würfel mit 91% der Kantenlänge des Einheitswürfels nur 1% von dessen (Hyper-)Volumen einnimmt. Daraus resultiert das Problem, dass eine sehr große Stichprobe benötigt wird, um einen hochdimensionalen Raum mit einer angemessenen Dichte zu füllen. Wird obiger Hyper-Würfel in ein Raster mit Kantenlänge 0.01 unterteilt, so besteht er aus  $10^{100}$  Rasterzellen, die mit Stichproben gefüllt werden müssen. Das sind deutlich mehr, als es Atome im Universum gibt, während in einem eindimensionalen Einheitswürfel 100 Proben reichen würden, um den Raum mit der gleichen Dichte zu füllen. Dieses Problem wird in der Literatur auch als der Fluch der Dimensionalität (*curse of dimensionality*, nach dem US Mathematiker Richard Bellman) bezeichnet. Daher wird die Merkmalsreduktion, also die Auswahl der besten Merkmale, gelegentlich als eigenständiger Schritt aufgeführt.

In der Fernerkundung sind die Merkmale meist durch die spektralen Bänder einer Aufnahme gegeben, und der Merkmalsraum wird daher auch (Multi-)Spektralraum genannt. Hier kommt es vor allem darauf an, die aussagekräftigsten Bänder auszuwählen, beispielsweise indem Bänder mit ähnlichen Informationen eliminiert werden (die Farben Rot, Grün und Blau sind beispielsweise meist hoch korreliert). Eine andere Möglichkeit ist durch eine Hauptachsentransformation (*principal components transformation*) das Koordinatensystem so zu wählen, dass der Großteil der Information in den ersten Koordinaten liegt (siehe dazu z.B. RICHARDS & JIA 2006: 127ff oder CAMPBELL 2002: 293f).

## Klassifikation

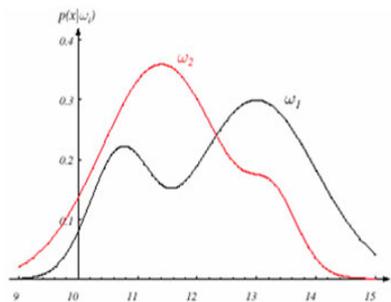
Das Ziel der Klassifikation besteht darin, jedem der Muster  $\mathbf{f}(\mathbf{x})$  anhand der extrahierten Merkmale eine der Klassen  $\Omega_k$  zuzuweisen. Als größte Einteilung wird zwischen numerischen und syntaktischen Klassifikatoren (LEHMANN *et al.* 1997: 397) unterschieden. Alle hier behandelten Verfahren gehören der ersten Kategorie an. Syntaktische Klassifikatoren transformieren die Muster dagegen in Symbolketten aus einem terminalen Alphabet, für die es im Gegensatz zu Merkmalsräumen nicht sinnvoll ist, Metriken zu definieren (NIEMANN 1983: 262). Die numerischen Klassifikatoren gliedern sich weiter in nicht-parametrische und parametrische Verfahren, sowie in verteilungsfreie und lernende Verfahren. Bei letzteren werden wiederum überwachte und nicht-überwachte Klassifikatoren differenziert (LEHMANN *et al.* 1997: 397f).

### 2.2.2 Bayes'sche Entscheidungstheorie

Die Bayes'sche Entscheidungstheorie ist ein statistisches Konzept, um eine bestmögliche Klassifikation zu finden. Zur Veranschaulichung noch einmal zurück zum Beispiel mit den Fischen aus DUDA *et al.* (2001): Betrachtet wird nun die Vorhersage der nächsten Fischart als Zufallsexperiment mit den

Ausgängen (oder Zuständen)  $\omega_1$  und  $\omega_2$ . Sei weiterhin angenommen, dass auf Grund von Wissen über die Beschaffenheit des Fanggebiets *a priori* die Wahrscheinlichkeit für Wolfsbarsch doppelt so groß ist wie die für Lachs ( $P(\omega_1)=2/3$  und  $P(\omega_2)=1/3$ ). Die Tatsache, dass sich beide zu eins summieren, beinhaltet die weitere Randannahme, dass es im Fanggebiet keine weiteren Fischarten gibt). Stehen neben diesen *a priori* Wahrscheinlichkeiten keine weiteren Informationen (Merkmale) zur Verfügung, so wird die beste Trefferquote erreicht, indem jedes Mal Wolfsbarsch vorhersagt wird.

Wird wieder ein Merkmal wie die Helligkeit hinzugenommen, so kann diese als kontinuierliche Zufallsvariabel betrachtet werden, die eine klassenbedingte Wahrscheinlichkeitsdichte  $p(\mathbf{x} | \omega_i)$  (*probability density function*, PDF) hat. Die PDF sind in Abb. 2.10 dargestellt, und geben die die Wahrscheinlichkeit (*likelihood*) für die Ausprägung  $\mathbf{x}$  der Zufallsvariablen unter der Bedingung an, dass die Umwelt im Zustand  $\omega_i$  ist. Anders ausgedrückt: Wie wahrscheinlich ist die Beobachtung des Merkmals  $\mathbf{x}$  bei einer bestimmten Fischart. Die PDF sind *a priori* natürlich nicht bekannt und müssen etwa durch Stichproben geschätzt werden.

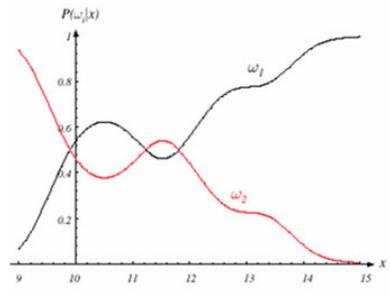


**Abb. 2.10: Hypothetische klassenabhängige PDF: wie wahrscheinlich ist die Beobachtung von  $\mathbf{x}$  bei Zustand  $\omega_i$ ? (DUDA *et al.* 2001: 21)**

Zur Klassifikation wird aber die *a posteriori* Wahrscheinlichkeiten  $P(\omega_i | \mathbf{x})$  benötigt, also die Wahrscheinlichkeit für den Zustand  $\omega_i$  unter der Bedingung, dass  $\mathbf{x}$  beobachtet wurde. Beide hängen über die Bayes'sche Formel zusammen:

$$(2.4) \quad P(\omega_i | \mathbf{x}) = \frac{p(\mathbf{x} | \omega_i)P(\omega_i)}{p(\mathbf{x})}, \quad p(\mathbf{x}) = \sum_{\forall i} p(\mathbf{x} | \omega_i)P(\omega_i).$$

Die *a posteriori* Wahrscheinlichkeit ist also das Produkt der klassenbedingten Wahrscheinlichkeit von  $\mathbf{x}$  und der *a priori* Wahrscheinlichkeit der Klasse gewichtet mit einem Faktor, der angibt wie wahrscheinlich die Beobachtung von  $\mathbf{x}$  überhaupt ist und damit die Summe aller posteriori Wahrscheinlichkeiten auf eins normiert. Es kommt damit sowohl auf die Wahrscheinlichkeit der Ausprägung  $\mathbf{x}$  des Merkmal für eine Klasse an, als auch auf die Wahrscheinlichkeit der Klasse selbst. Im Sonderfall von gleichen (oder unbekannt) *a priori* Wahrscheinlichkeiten  $P(\omega_1) = P(\omega_2)$  kommt es nur auf die PDF an. Im Sonderfall von  $p(\mathbf{x} | \omega_1) = p(\mathbf{x} | \omega_2)$  sind die Merkmale offensichtlich schlecht gewählt, da sie nicht zur Unterscheidung der Klassen beitragen.



**Abb. 2.11:** *A posteriori* Wahrscheinlichkeiten für  $P(\omega_1)=2/3$  und  $P(\omega_2)=1/3$ : Wie wahrscheinlich ist der Zustand  $\omega_1$  bei Beobachtung von  $x$ ? (DUDA *et al.* 2001: 23)

Die Bayes'sche Entscheidungsregel besagt, dass immer die Klasse gewählt wird, deren *a posteriori* Wahrscheinlichkeit am größten ist. Da  $p(\mathbf{x})$  dabei für alle Klassen gleich ist, ergibt sich unter Verwendung von (2.4):

$$(2.5) \quad \mathbf{x} \in \Omega_i, \text{ falls } p(\mathbf{x} | \omega_i)P(\omega_i) > p(\mathbf{x} | \omega_j)P(\omega_j) \text{ für alle } j \neq i$$

Die Entscheidungsgrenzen sind damit Hyperflächen im Merkmalsraum mit gleichen Wahrscheinlichkeiten von zwei Klassen:

$$p(\mathbf{x} | \omega_i)P(\omega_i) = p(\mathbf{x} | \omega_j)P(\omega_j).$$

Die Bayes'schen Entscheidungsregel minimiert also den Fehler, gleichzeitig lässt sich dieser aber auch noch genau angeben:

$$(2.6) \quad P(\text{error}) = \int_{-\infty}^{\infty} P(\text{error}, x) dx = \int_{-\infty}^{\infty} P(\text{error} | x) p(x) dx,$$

mit  $P(\text{error} | x) = \min_{\forall i} [P(\omega_i | x)]$

Die Fehlerfunktion kann zu einer Verlustfunktion erweitert werden, etwa wenn nicht alle Fehlklassifikationen gleich teuer oder gefährlich sind (siehe dazu DUDA *et al.* 2001: 24 ff), beispielsweise bei einem Klassifikator, der Diamanten von Steinen unterscheidet, oder der Blutspenden auf HIV untersucht.

Da der natürliche Logarithmus streng monoton wachsend ist, also die Rangordnung erhält, kann durch Definition von  $g_i(\mathbf{x}) = \ln(p(\mathbf{x} | \omega_i)P(\omega_i)) = \ln p(\mathbf{x} | \omega_i) + \ln P(\omega_i)$  die Multiplikation zur Addition vereinfacht werden. Die Entscheidungsregel vereinfacht sich damit zu:

$$(2.7) \quad \mathbf{x} \in \Omega_i, \text{ falls } g_i(\mathbf{x}) > g_j(\mathbf{x}) \text{ für alle } j \neq i,$$

wobei die  $g_i$  als Unterscheidungsfunktionen bezeichnet werden (RICHARDS & JIA, 2006: 198).

### 2.2.3 Parametrische Verfahren

Die zentrale Frage für alle Klassifikatoren bleibt nun, woher die klassenabhängigen Wahrscheinlichkeitsdichten und die *a priori* Wahrscheinlichkeiten zu erhalten sind. Parametrische Verfahren versuchen, die PDF durch Wahrscheinlichkeitsverteilungen mit gewissen Parametern anzunähern.

### Maximum Likelihood

Das wahrscheinlich bekannteste parametrische Verfahren ist die Maximum Likelihood (ML) Methode, bei der angenommen wird, dass die klassenspezifischen Wahrscheinlichkeitsdichten normalverteilt sind. Oftmals wird diese Voraussetzung aber nicht hinreichend verifiziert und das Verfahren vor allem deswegen eingesetzt, da es einfach zu handhaben ist. Die multivariate Form der Normalverteilung ist:

$$(2.8) \quad p(\mathbf{x} | \omega_i) = N(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i) = (2\pi)^{-N/2} |\boldsymbol{\Sigma}_i|^{-1/2} \exp\left[-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_i)' \boldsymbol{\Sigma}_i^{-1} (\mathbf{x} - \boldsymbol{\mu}_i)\right]$$

Der Term  $(\mathbf{x} - \boldsymbol{\mu}_i)'$  steht dabei für den transponierten Vektor  $(\mathbf{x} - \boldsymbol{\mu}_i)$ , der Mittelwert der  $i$ -ten Klasse  $\boldsymbol{\mu}_i$  und die Kovarianzmatrix  $\boldsymbol{\Sigma}_i$  sind die Parameter der Verteilung. Im univariaten Fall wird der Mittelwert zu  $\mu_i$  und die Kovarianz zur Varianz  $\sigma_i^2$ , im Fall unabhängiger Variablen wird sie zur Diagonalmatrix der Form:

$$\boldsymbol{\Sigma}_i = \begin{bmatrix} \sigma_1^2 & & 0 \\ & \ddots & \\ 0 & & \sigma_n^2 \end{bmatrix}.$$

Die Parameter werden aus den Trainingsdaten geschätzt:  $\boldsymbol{\mu} = E[\mathbf{x}]$ ,  $\boldsymbol{\Sigma}[(\mathbf{x} - \boldsymbol{\mu})(\mathbf{x} - \boldsymbol{\mu})']$ . Zu den damit verbundenen Schwierigkeiten siehe DUDA & HART (1973: 44ff).

Durch logarithmieren und Weglassen der konstanten Summanden ergibt sich als Unterscheidungsfunktion:

$$(2.9) \quad g_i(\mathbf{x}) = \ln P(\omega_i) - \frac{1}{2} \ln |\boldsymbol{\Sigma}_i| - \frac{1}{2} (\mathbf{x} - \boldsymbol{\mu}_i)' \boldsymbol{\Sigma}_i^{-1} (\mathbf{x} - \boldsymbol{\mu}_i).$$

Da die Unterscheidungsfunktion quadratisch in  $\mathbf{x}$  ist, sind die entstehenden Entscheidungsflächen  $g_i - g_j = 0$  Parabeln, Kreise und Ellipsen (RICHARDS & JIA 2006: 197).

Im Spezialfall gleicher *a priori* Wahrscheinlichkeiten und gleicher Kovarianzen  $\boldsymbol{\Sigma}_i = \boldsymbol{\Sigma}$  (z.B. wenn die Varianz tatsächlich nur von der statistischen Verteilung der Messfehler im Sensor herrührt) aller Klassen lässt sich die Unterscheidungsfunktion weiter vereinfachen zu:

$$(2.10) \quad g_i(\mathbf{x}) = -(\mathbf{x} - \boldsymbol{\mu}_i)' \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}_i) = -d(\mathbf{x}, \boldsymbol{\mu}_i)^2.$$

Das Inverse dieser Unterscheidungsfunktion ist das Quadrat des berühmten Mahalanobis-Abstandes. Der Klassifikator misst also quasi den Abstand zum Mittelwert der einzelnen Klassen mit einem besonderen Abstandsmaß, das die unterschiedlichen Varianzen in den verschiedenen Richtungen berücksichtigt.

### Minimum Distance

Wird der obige Sonderfall noch weiter auf gleiche Varianzen in alle Richtungen eingeschränkt, dann ist die Kovarianzmatrix diagonal und in allen Komponenten gleich:  $\Sigma = \sigma^2 \mathbf{I}$  und (2.10) wird zu:

$$(2.11) \quad g_i(\mathbf{x}) = -\frac{1}{\sigma^2}(\mathbf{x} - \boldsymbol{\mu}_i)^t(\mathbf{x} - \boldsymbol{\mu}_i).$$

Dies entspricht bis auf einen konstanten Faktor dem Quadrat der Euklidischen Distanz im  $\mathbb{R}^N$ :

$$(2.12) \quad \begin{aligned} d(\mathbf{x}, \boldsymbol{\mu}_i)^2 &= \left( \|\mathbf{x} - \boldsymbol{\mu}_i\|_2 \right)^2 = \langle (\mathbf{x} - \boldsymbol{\mu}_i) | (\mathbf{x} - \boldsymbol{\mu}_i) \rangle = \\ &(\mathbf{x} - \boldsymbol{\mu}_i)^t (\mathbf{x} - \boldsymbol{\mu}_i) = \mathbf{x}\mathbf{x} - 2\mathbf{x}\boldsymbol{\mu}_i + \boldsymbol{\mu}_i\boldsymbol{\mu}_i \end{aligned}$$

Entsprechend wird dieser Klassifikator *minimum euclidean distance* (MED) genannt. MED misst den Abstand mit der 2-Norm und ist daher (im Gegensatz zum Mahalanobis-Abstand) richtungsunabhängig.

Wird in (2.12) weiterhin der von der Klasse unabhängige Term  $\mathbf{x}\mathbf{x}$  entfernt und das Vorzeichen bei der Ungleichung berücksichtigt (kleinerer Abstand entspricht größerem Wert der Unterscheidungsfunktion wegen MINIMUM distance aber MAXIMUM likelihood), so ergibt sich die Unterscheidungsfunktion:

$$(2.13) \quad g_i(\mathbf{x}) = 2\mathbf{x}\boldsymbol{\mu}_i - \boldsymbol{\mu}_i\boldsymbol{\mu}_i.$$

Der Vorteil des MED-Klassifikators ist, dass pro Klasse nur noch ein Parameter, nämlich der Mittelwert, geschätzt werden muss. Dafür reicht bei gleicher Genauigkeit eine deutlich kleinere Stichprobe. Der Nachteil ist eine geringere Flexibilität durch das vereinfachte Klassenmodell. Da die Unterscheidungsfunktion linear in  $\mathbf{x}$  ist, sind auch die Entscheidungsflächen lineare Hyperflächen (RICHARDS & JIA 2006: 204). Klassen mit einer kleinen Varianz, die im Merkmalsraum scharf begrenzt sind, werden vom MED tendenziell überbewertet, solche mit einer großen Varianz dagegen unterbewertet. Daher ist es für alle parametrische Verfahren unerlässlich, den Klassifikatoren mit einer zweiten unabhängigen Stichprobe zu testen.

### Parallelepiped Klassifikator

Noch einfacher sind Parallelepiped- oder Box-Klassifikatoren, bei denen die Wahrscheinlichkeitsdichte als Rechteckfunktion modelliert wird. Die Grenzen des Rechtecks werden dabei direkt aus dem Histogramm (Häufigkeitsverteilung der Merkmale, siehe 2.3.3) durch die obere und untere Grenze der Merkmalsausprägung abgeleitet.

Der Vorteil ist, dass dies sehr schnell zu berechnen ist. Der Nachteil ist, dass die Partitionierung des Merkmalsraums weder überlappungsfrei noch vollständig ist. Einige Vektoren gehören also mehreren Klassen an, andere dagegen keiner (bzw. einer Rückweisungsklasse). Box-Klassifikatoren spielen zwar keine große Rolle mehr, sind aber dem Verständnis insbesondere bei *top-down* Ansätzen zuträglich.

### 2.2.4 Nicht parametrische Verfahren

Während bei den parametrischen Verfahren durch Annahme einer bestimmten statistischen Verteilung für die Wahrscheinlichkeitsdichte vorab eine Modellstruktur festgelegt wird, wird die Modellstruktur bei parameterfreien Verfahren aus den Daten selbst ermittelt, beispielsweise indem die Wahrscheinlichkeitsdichten direkt geschätzt werden. Parameterfrei bedeutet dabei nicht zwangsläufig, dass das Modell keine Parameter hat, sondern nur, dass deren Art und Anzahl nicht vorab festgelegt sind.

#### K-Nearest-Neighbour

Der *k-nearest-neighbour* Klassifikator (kNN) ist konzeptuell sehr einfach, aber aufwendig zu berechnen (RICHARDS & JIA 2006: 2007). Dem Verfahren liegt die Annahme zu Grunde, dass Elemente mit geringem Abstand im Merkmalsraum mit großer Wahrscheinlichkeit zur selben Klasse gehören. Daher werden für einen neuen Punkt die Klassenzugehörigkeiten der  $k$  nächsten Nachbarn betrachtet und im einfachsten Fall eine Mehrheitsentscheidung getroffen. Die Unterscheidungsfunktion für eine Klasse ist dann die Anzahl der Treffer bei den  $k$  Nachbarn. Der Lernprozess reduziert sich somit auf das Abspeichern der Stichprobe. Zur Berechnung muss aber jeweils der Abstand zu jedem einzelnen Element der Stichprobe berechnet werden (im Vergleich dazu wird beim MED nur der Abstand zu den Mittelwerten der Klassen benötigt). Dies ist insbesondere bei großen Stichproben und hoher Dimensionalität sehr aufwendig. Dabei können wieder verschiedene Abstandsmaße verwendet werden, von denen das euklidische am üblichsten ist.

Zur Verbesserung des Verfahrens können bei der Entscheidung die nächsten Nachbarn mit ihrer Distanz gewichtet werden. Im Fall von unverhältnismäßigen Klassenanteilen bei der Stichprobe ist es außerdem sinnvoll, die Unterscheidungsfunktion über die Bayes'sche Formel mit den *a priori* Wahrscheinlichkeiten zu gewichten. Dennoch bleibt das kNN Verfahren suboptimal, d.h. der Fehler liegt über dem Bayes'schen Fehler in (2.6). Es kann aber gezeigt werden, dass er unter dem doppelten Bayes'schen Fehler liegt (siehe DUDA & HART 1973: 98ff).

#### Dichteschätzung mit dem Parzen Fenster

Das kNN-Verfahren versucht, direkt die *a posteriori* Wahrscheinlichkeiten zu schätzen. Eine andere Möglichkeit ist, zunächst die klassenbedingten Wahrscheinlichkeitsdichten zu ermitteln. Dafür wird die Stichprobe als Zufallsexperiment betrachtet. Sei  $P$  die integrierte Wahrscheinlichkeitsdichte über eine kleine Region mit dem Volumen  $V$ . Dann sind die Wahrscheinlichkeiten, dass  $k$  von  $n$  Proben in die Region fallen, durch eine Binomialverteilung gegeben. Der Erwartungswert für  $k$  ist  $E[k] = nP$  und je größer die Stichprobe, desto genauer wird der Erwartungswert erreicht.

Umgekehrt kann auch die Wahrscheinlichkeitsdichte aus der Stichprobe geschätzt werden:

$$(2.14) \quad p(\mathbf{x}) \approx \frac{P}{V} \approx \frac{k/n}{V}.$$

Damit wurde zunächst natürlich nur die durchschnittliche Wahrscheinlichkeit in der Region geschätzt. Da aber die Wahrscheinlichkeitsdichte im Punkt  $\mathbf{x}$  gesucht ist, ist es zielführend, sich ihr durch eine Folge mit kleiner werdenden Volumina  $V_i$  anzunähern. Im hypothetischen Fall einer unendlich großen Stichprobe würde diese gegen den gewünschten Wert konvergieren.

In der Realität stehen natürlich immer nur endlich große Stichproben zu Verfügung, so dass die Folge fast überall gegen Null konvergieren würde. Daher müssen sowohl eine gewisse Varianz im Verhältnis  $k/n$  als auch eine gewisse räumliche Mittelung in Kauf genommen werden (DUDA & HART 1973: 87), wobei die eine jeweils nur auf Kosten der anderen verbessert werden kann.

Außerdem stellt sich die Frage, welche Folgen von Regionen sich eignen, diese Probleme zu minimieren. Dazu sollen nur einige Anmerkungen gemacht werden, für die genauen Konvergenzbedingungen sei auf DUDA & HART (1973: 89ff) verwiesen. Die Regionen werden durch Fensterfunktionen beschrieben, die auch als Parzen-Fenster bezeichnet werden. Diese müssen größer gleich Null sein und sich zu Eins aufsummieren damit  $p(\mathbf{x})$  zu einer gültigen Dichtefunktion konvergiert. Ein einfaches Fenster ist ein Hyperwürfel mit dem Wert Eins innerhalb des Würfels und Null außerhalb. Bei dieser Rechteckfunktion werden aber alle Samples unabhängig von ihrer Distanz gleich bewertet. Sehr beliebte Fensterfunktionen sind daher auch Normalverteilungen. Es soll noch darauf hingewiesen werden, dass die ermittelten Dichtefunktionen immer der gewählten Fensterfunktion ähneln (vergleichbar der Faltung mit einem Kern, siehe 2.3.4). Auch wenn alle Fenster gegen die gleiche Dichte konvergieren, wird bei endlichen Stichproben auch in diesem nicht-parametrischen Verfahren das Modell maßgeblich durch die Wahl der Fensterfunktion mitbestimmt. Generell sind für das Verfahren sehr große Stichproben notwendig. Die kNN Methode kann als eine ganz ähnliche Schätzung begriffen werden, bei der das Volumen  $V_n$  jeweils so gewählt wird, dass genau  $k$  Samples (aus allen Klassen) darin liegen.

### 2.2.5 Unüberwachtes Lernen und Clusteranalyse

Alle bislang beschriebenen Klassifikationsverfahren setzen voraus, dass die Musterklassen bekannt sind, und eine repräsentative Stichprobe mit Klassenzugehörigkeiten als Trainingsdatensatz zur Verfügung steht. Ist dies nicht der Fall, so muss auf unüberwachte Verfahren zurückgegriffen werden.

Bei der Clusteranalyse (auch Clusterung, *clustering*) werden quasi natürliche Klassen in den Daten gesucht. Zunächst muss dafür ein Maß für die Güte einer Klassifikation eingeführt werden, wofür häufig die Summe der quadratischen Fehler (*sum of squared error*, SSE) verwendet wird (RICHARDS & JIA 2006: 250):

$$(2.15) \quad SSE = \sum_{C_i} \sum_{x \in C_i} (\|\mathbf{x} - \boldsymbol{\mu}_i\|_2)^2.$$

Die  $C_i$  sind die zugewiesenen Klassen. Als Distanzmaß wurde wieder der euklidische Abstand gewählt, aber andere sind selbstverständlich auch möglich. Der Fehler nimmt mit der Anzahl der Klassen ab, bei einer Klasse pro Objekt verschwindet er.

Es gibt eine Vielzahl unterschiedlicher Cluster-Verfahren. Statt eine vollständige Übersicht zu geben, sollen hier nur exemplarisch zwei wichtige Verfahren vorgestellt werden, die hierarchische Clusterung und das k-means Verfahren (vergleiche MATLAB 1994-2007: Statistical Toolbox).

#### Hierarchische Clusterung

Bei der hierarchischen Clusterung werden die Objekte in verschiedenen Skalen gruppiert, indem ein Clusterbaum oder Dendrogramm erzeugt wird. Dabei handelt es sich um eine Hierarchie von Clustern, bei der Cluster eines Levels zu Clustern auf einem höheren Level zusammen gefügt werden. Am einen

Ende steht also jedes Muster als eigene Klasse, am anderen Ende sind alle zu einer Klasse zusammengefasst, wobei die Cluster mit dem geringsten Abstand zuerst vereinigt werden. Je nach dem, von welcher Seite ausgegangen wird, werden anhäufende und teilende Verfahren unterschieden. Für die Abstandsmessung zwischen den Clustern ist neben dem Abstandsmaß die *linkage*-Strategie entscheidend, also welcher Abstand zwischen den Clustern als Bemessungsgrundlage dient. Ein Standardverfahren ist das *centroid linkage*, bei dem der Abstand zwischen den Mittelwerten der Cluster benutzt wird (siehe auch 2.4.4).

Der Vorteil der hierarchischen Clusterung liegt in der objektivierten Identifikation einer geeigneten Anzahl von Klassen (üblicherweise derart, dass die nächste Vereinigung den Fehler unverhältnismäßig vergrößern würde). Da auf jeder Ebene sämtliche Abstände zwischen den Clustern berechnet werden müssen, ist das Verfahren jedoch aufwendig und wird für größere Datensätze zudem vergleichsweise unübersichtlich.

### **K-Means-Clustering und ISODATA**

Das *k-Means* Verfahren wird als partitionierendes Verfahren bezeichnet, da es die Daten in Klassen unterteilt, deren Anzahl  $k$  vorgegeben wird. Der Algorithmus versucht den SSE wie folgt zu minimieren: Zunächst wird er mit  $k$  (zufälligen oder vorgegebenen) Clusterzentren initialisiert. Dann wird jedes Objekt dem (nach gegebenem Abstandsmaß) nächsten Zentrum zugeordnet, anschließend werden die Zentren der Cluster neu berechnet. Dies wird so lange wiederholt, bis sie sich nicht mehr (oder nur noch wenig) ändern.

Der *k-Means* Algorithmus hat den Vorteil, dass er erheblichen Rechenaufwand spart (hier müssen nur die Abstände zu den  $k$  Clusterzentren berechnet werden) und meistens gute Ergebnisse liefert. Allerdings ist nicht garantiert, dass er konvergiert (z.B. können einzelne Cluster leer werden). Insbesondere muss er nicht zum besten Ergebnis konvergieren, sondern kann in Bezug auf das Qualitätsmaß in einem lokalen Minimum stecken bleiben. Um ein möglichst gutes Ergebnis sicherzustellen, wird der Algorithmus daher häufig mehrfach mit unterschiedlichen Initialwerten für die Clusterzentren ausgeführt. Ein weiterer Nachteil ist, dass die unterschiedlichen Varianzen verschiedener Klassen nicht berücksichtigt werden.

In DUDA & HART (1973: 201) wird dieser Algorithmus auch als *basic ISODATA (iterative self-organizing data analysis technique)* bezeichnet. Häufig wird in der Literatur unter ISODATA aber auch ein leicht abgewandelter Algorithmus verstanden, bei dem die Anzahl der Cluster automatisch angepasst wird. Dabei werden Cluster verworfen, wenn sie eine bestimmte Größe unterschreiten, vereinigt, wenn sie einen bestimmten Abstand unterschreiten und gespalten, wenn sie eine bestimmte interne Varianz überschreiten.

## **2.3 Bildverarbeitung**

Die Digitale Bildverarbeitung (BV, *digital image processing*) ist im engeren Sinne ein Teilgebiet der Signalverarbeitung. Etwas weiter gefasst ist sie jedoch ein sehr interdisziplinäres Feld zwischen Gebieten aus Physik, Informatik, Geoinformatik, Mathematik und Ingenieurwissenschaften mit Überschneidungen mit den Kognitionswissenschaften oder der Neurophysiologie (JÄHNE 2005). Da sie uneinheitlich gebraucht werden, sollen zunächst einige Begriffe abgegrenzt werden:

Die Bildverarbeitung (*image processing*) beschäftigt sich mit der maschinellen Verarbeitung von visuellen Informationen. Streng genommen wird darunter nur die Transformation von Bildern ohne Vorannahmen und ohne Interpretation verstanden, im weiteren Sinne schließt sie aber Erfassen, Analyse, Bearbeitung, Erkennen und Aufbereitung mit ein. In diesem Sinne ist die Überschrift dieses Kapitels zu verstehen. Die Bildbearbeitung beschäftigt sich dagegen ausschließlich mit der Aufbereitung und Verbesserung (*enhancement*) von Bildern zwecks optischer Darstellung und manueller Auswertung.

Bildanalyse (*image analysis*) bezeichnet die Extraktion von Informationen mit Bedeutung aus Bildern, z.B. das Erkennen von Objekten und deren Klassifikation mit Hilfe der Mustererkennung. Dabei geht Modellwissen über die analysierten Objekte ein. GONZALEZ & WOODS (1993: 571) gehen sogar noch einen Schritt weiter und bezeichnen *image analysis* als Prozess „*of discovering, identifying, and understanding patterns, that are relevant to the performance of an image based task*“. Dabei solle die Maschine sich menschlichen Fähigkeiten annähern.

Das Bildverstehen (*image understanding*) betrachtet einen Prozess. Ausgehend von einer Fragestellung wird das Bild in eine Szenenbeschreibung überführt. Der inverse Prozess wird in der Computergraphik verwendet, die ein Bild aus einer abstrakten Repräsentation oder Beschreibung erzeugt (PINZ 1994:5).

Die Art der Szenenbeschreibung kann dabei variieren. Eine übliche Methode ist aber, die Szene durch die enthaltenen Objekte und deren Eigenschaften (Ort, Form, Größe, Textur, etc.) zu beschreiben (PINZ 1994:5). Dazu wird neben dem Modellwissen über die einzelnen Objekte auch Wissen über den Kontext benötigt, also die Szene als Ganzes und die Beziehungen zwischen den Objekten (ein blaues Auto wartet an einem Zebrastreifen um einen Fußgänger durchzulassen).

Oft wird der Begriff Bildverstehen mit Computersehen (*computer vision*) gleich gesetzt. Computersehen konzentriert sich aber verstärkt auf 3D-Szenenmodelle, während das ebenfalls benachbarte Feld des Maschinellen Sehens (*maschine vision*) sich vor allem auf angewandte Bereiche wie industrielle Fertigungsprozesse bezieht (JÄHNE 2005:18). Das Bildverstehen ist eng mit der wissensbasierten Mustererkennung verflochten.

### 2.3.1 Bildrepräsentation

Eingangs wird kurz geklärt, was hier unter einem Bild verstanden werden soll. Da BV ein sehr interdisziplinäres Feld ist, existieren sehr unterschiedliche Konzepte, Terminologien und formale Beschreibungen von Bildern und Einheiten für dieselben Größen. In der Bildverarbeitung wird ein Bild meist als reelle Funktion von Helligkeiten (Grauwert, im Folgenden auch Intensität) im Ortsraum betrachtet, ein digitales Bild entsprechend als diskrete reelle (oder manchmal auch komplexe) Funktion. Physikalisch ist ein Bild eine Funktion der Intensität einer elektromagnetischen (Licht-)welle in einem gewissen Frequenz- und Zeitintervall. Aus Sicht der Signalverarbeitung ist es ein zweidimensionales Signal, wobei die ursprüngliche reelle Funktion durch ein optisches System in ein Ausgangssignal verwandelt und dabei gestört wird. Weitere Interpretationen sind das Bild als 2D-Muster in der ME, als Zufallsprozess in der Stochastik und als Menge in der Morphologie. Daraus resultieren sehr unterschiedliche formale Beschreibungen, wobei versucht wurde, eine möglichst konsistente Terminologie zu verwenden.

## Sampling und Quantisierung

Ausgehend vom Bild als reeller Funktion ist es sinnvoll, den Definitionsbereich auf  $D \subset \mathbb{R} \times \mathbb{R}$  zu beschränken, da jedes Bild nur einen endlichen Bereich darstellt. Da Licht eine Energieform ist, kann der Wertebereich auf positive Werte beschränkt werden, die Bildfunktion ist also:

$$(2.16) \quad f = f(x, y) = f(\mathbf{x}) : D \rightarrow \mathbb{R}^+$$

Dabei sind sowohl der Definitionsbereich als auch der Wertebereich zunächst kontinuierlich. Um Bilder in der digitalen Bildverarbeitung mit Computern verarbeiten zu können, müssen sie diskretisiert werden.

Der Definitionsbereich wird zu diesem Zweck in Rasterzellen, (auch Pixel von *picture elements*) eingeteilt. Am üblichsten sind dabei quadratische Raster, daneben sind aber auch regelmäßige Dreiecks und Hexagonalraster in Gebrauch. Der Vorteil des quadratischen Rasters ist, dass das Bild als  $M \times N$ -Matrix repräsentiert werden kann, wobei der erste Index die Spalte und der zweite die Zeile angibt. Ein großer Nachteil ist aber, dass die Nachbarschaft im quadratischen Raster uneindeutig ist. Daher werden unter 4-Nachbarschaft (4N) eines Pixels nur die mit gemeinsamer Kante angrenzenden 4 Pixel verstanden, unter 8-Nachbarschaft (8N) zusätzlich die diagonal angrenzenden (siehe auch (2.47)).

Der Übergang vom kontinuierlichen zum diskreten Signal wird digitalisieren oder abtasten (*sampling*) genannt, daneben ist auch der eingedeutschte Begriff sampeln gebräuchlich. Dabei gehen Informationen verloren: Veränderungen im Signal, deren Abstand kleiner ist als der zwischen zwei Pixeln, können nicht aufgelöst werden. Für solch hohe Ortsfrequenzen besagt das

Abtast-Theorem (auch Nyquist-Theorem):  $\Delta x < \frac{1}{2f_{\max}}$ .

Die Rasterweite muss also so gewählt werden, dass die größte Frequenz immer noch mit mindestens zwei Punkten pro Phase abgetastet wird, andernfalls kommt es zu Aliasing oder Moiré-Effekten (siehe Abb. 2.12).

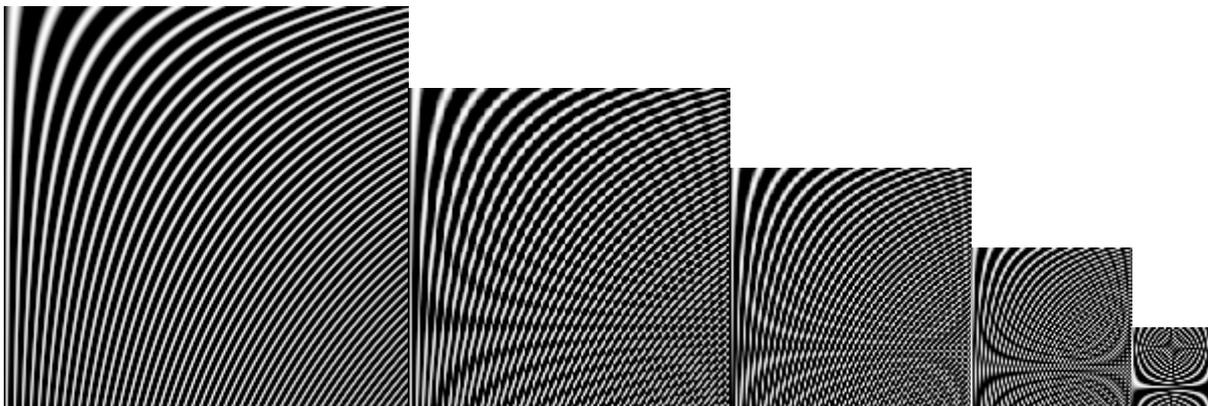


Abb. 2.12: Moiré-Effekte beim *sampling*: Bild mit verschiedenen Ortsfrequenzen, resampelt auf 0.8, 0.6, 0.4 und 0.2 der Originalgröße

Zur einfacheren Bearbeitung im Computer wird entweder direkt vom Sensor oder beim Scannen von Fotos meistens auch der Wertebereich diskretisiert (Quantisierung). Dabei wird er auf ganze Zahlen in

einem Intervall  $g \in G = \{0 \dots 2^n - 1\}$  abgebildet. Die angenommenen Werte  $g$  werden dann als Grauwerte bezeichnet, der Umfang der möglichen Werte ist der Dynamikbereich. Da bei der Quantisierung von einem unendlichen auf einen endlichen Bereich abgebildet wird, verringert sich dieser, daneben nimmt auch die Genauigkeit ab. Gemeinsam mit der Abtastung ergibt sich die diskrete Bildfunktion:

$f: M \times N \mapsto G$  mit  $M = \{0 \dots M_{max}\}$ ,  $N = \{0 \dots N_{max}\}$ . Diese Darstellung entspricht einer Abbildung zwischen Mengen, ebenso kann das Bild aber auch als Matrix geschrieben werden:  $\mathbf{B} = (b_{ij})$ ,  $b_{ij} \in G$ .

### Bildtypen

Die im letzten Abschnitt beschriebenen Bilder werden als Grauwertbilder bezeichnet. Im Spezialfall von  $n=0$  wird vom Binärbild gesprochen und die Werte 0 und 1 je nach Konvention als schwarz und weiß, falsch und wahr oder Hintergrund und Vordergrund bezeichnet.

Farbbilder werden für die Bildschirmdarstellung und Verarbeitung meistens als Kompositbild von drei Grauwertbildern in den Farben Rot, Grün und Blau codiert (RGB). Das entspricht der Darstellung der Farben durch additive Farbmischung in einem Computermonitor. Eine Farbcodierung von drei mal acht Bit heißt *True Colour*, obwohl der RGB Farbraum deutlich kleiner ist, als der vom menschlichen Auge wahrgenommene. Ein Farbbild kann somit als Vektor von drei Grauwertbildern verstanden werden:  $\mathbf{f}(x, y) = (f_R(x, y), f_G(x, y), f_B(x, y))$ .

Eine weitere Möglichkeit zur Codierung farbiger Bilder sind indexierte Bilder. Hier entspricht jeder Wert der Referenz auf eine Farbe, deren Mischungsvorschrift in einer gesonderten Tabelle verwaltet wird. Diese wird Index, Palette, *colourmap* oder *look-up-table* (LUT) genannt. Der Vorteil liegt in einer Ersparnis von Speicherplatz, für weitere Verarbeitung ist aber von Nachteil, dass die Werte nur noch nominales Skalenniveau haben.

Neben quantisierten Bildern sind in der digitalen BV auch kontinuierliche Wertebereiche gebräuchlich. Diese werden intern als Fließkommatypen (meist vom Typ *double*) realisiert und üblicherweise auf den Wertebereich  $[0, 1]$  normiert.

Für alle Bildcodierungen gibt es entsprechende Dateiformate. Die gängigen Formate TIFF (*Tagged Image File Format*) und BMP (*Windows Bitmap*) speichern Kodierungen mit 1 bis 24 Bit pro Farbkanal, das GIF (*Graphics Interchange Format*) speichert indexierte Bilder mit bis zu 256 Farben, das OpenEXR-Format Fließkommabilder.

### 2.3.2 Operationen und ihre Eigenschaften

Operationen in der Bildverarbeitung sind mathematische Abbildungen, die ein Eingabebild  $E(x, y)$  in ein Ausgabebild  $A(x, y)$  überführen (PINZ 1994:51).

$$(2.17) \quad f: E(x, y) \alpha A(x, y)$$

Die Abbildung kann in der Operatorschreibweise durch einen Operator repräsentiert werden:  $\mathcal{H}E = A$ . Stehen mehrere Operatoren hintereinander, so werden sie stets von rechts nach links angewendet:  $\mathcal{H}_2\mathcal{H}_1E = \mathcal{H}_2(\mathcal{H}_1E)$ , die  $n$ -te Potenz  $\mathcal{H}^n$  eines Operators ist seine  $n$ -fache Hintereinanderanwendung.

Operationen werden in der BV auch als Filter oder Transformationen bezeichnet. Spezieller wird der Begriff Transformation für solche Operationen verwendet, bei denen das Bild in einen anderen Raum (z.B. in Ortsfrequenzen bei der Fouriertransformation) abgebildet wird:  $t : E(x, y) \alpha T(u, v)$ .

Abhängig davon, welcher Teil des Eingabebildes sich bei der Abbildung auf ein Pixel des Ausgabebildes auswirkt, wird zwischen Punktoperationen, lokalen Operationen und globalen Operationen (PINZ, 1994:51) unterschieden. Punktoperationen bilden nur ein Pixel des Eingangsbildes auf jedes Pixel des Ausgangsbildes ab. Der Ausgabewert für ein Pixel wird also nur aus seinem eigenen Grau- oder Farbwert berechnet:  $A(x, y) = f_{xy}(E(x, y))$ . Der Index  $_{xy}$  weist darauf hin, dass die Abbildung für jeden Bildpunkt unterschiedlich sein kann. Wichtige Punktoperationen sind Kontrastveränderung durch Histogrammspreizung, Gammakorrektur, Invertierung eines Bildes (Negativtransformation) und Schwellenwertbildung.

Lokale Operationen oder Nachbarschaftsoperationen verknüpfen die Werte rund um einen Punkt und schreiben das Ergebnis zurück an diesen Punkt (JÄHNE 2005: 112). Ein diskreter Nachbarschaftsoperator bildet eine  $M \times N$  Matrix auf sich selbst ab:

$$(2.18) \quad A(x, y) = N(E(x - x', y - y'), \forall [x', y']^t \in M),$$

wobei die Menge M von Punkten als Maske, Fenster oder Strukturelement bezeichnet wird. Nachbarschaftsoperationen werden besonders häufig Filter genannt, wobei darunter teilweise auch nur die Filtermaske verstanden wird. Im weit verbreiteten Fall einer quadratischen Umgebung um das Pixel mit Fenstergröße  $2\delta + 1$  kann eine Fensterfunktion  $w_E(x, y, \delta)$  definiert werden, und (2.18) vereinfacht sich zu:

$$(2.19) \quad A(x, y) = f(w_E(x, y, \delta)).$$

Globale Operationen verwenden zur Transformation jedes einzelnen Punktes das ganze Bild. Dementsprechend sind sie sehr rechenaufwendig (Ordnung n hoch 4).

$$(2.20) \quad A(x, y) = f\{E\}$$

Eine weitere wichtige Eigenschaft von Operatoren ist die Linearität. Seien A und B Bilder und a, b Skalare aus  $\mathbb{R}$  oder  $\mathbb{C}$ , dann gilt: ein Operator  $\mathcal{H}$  ist genau dann linear, wenn gilt:  $\mathcal{H}(aA + bB) = a\mathcal{H}A + b\mathcal{H}B$ . Dies wird auch als Superpositionsprinzip bezeichnet (JÄHNE, 2005:119). Linearität bedeutet also, dass die Überlagerung von Signalen durch die Transformation nicht gestört wird.

Ein Operator ist translationsinvariant, verschiebungsinvariant oder homogen, wenn die Antwort des Operators nicht von der Position im Bild abhängig ist, wird das Bild verschoben, so verschiebt sich die Antwort des homogenen Operators ebenso. Ein Operator ist genau dann homogen, wenn er mit dem Verschiebungsoperator  $\mathcal{S}$  vertauscht werden kann:

$$\mathcal{H}^{mn}\mathcal{S} = {}^{mn}\mathcal{S}\mathcal{H}, \text{ mit } {}^{mn}\mathcal{S}I(x, y) = I(x - m, y - n).$$

Ein Operator der linear und homogen ist, wird als LSI-Operator (*linear shift-invariant*) bezeichnet. Ein Operator ist isotrop oder rotationsinvariant, wenn er in alle Richtungen gleich wirkt, also mit dem

Rotationsoperator vertauscht. Im diskreten Fall ist dies aber meist nur annäherungsweise möglich. Operatoren sind kommutativ, wenn sie vertauschen, d.h. die Reihenfolge in denen sie angewendet werden unerheblich ist:  $\mathcal{H}_1\mathcal{H}_2 = \mathcal{H}_2\mathcal{H}_1$ . Ein Operator ist assoziativ, wenn gilt:  $\mathcal{H}_1(\mathcal{H}_2\mathcal{H}_3) = (\mathcal{H}_1\mathcal{H}_2)\mathcal{H}_3$ . Das bedeutet, dass sich auch die sequentielle Anwendung mehrerer Operatoren durch einen einzelnen Operator ersetzen lässt. Schließlich sind Operatoren distributiv, wenn gilt:  $(\mathcal{H}_1 + \mathcal{H}_2)\mathcal{H}_3 = \mathcal{H}_1\mathcal{H}_3 + \mathcal{H}_2\mathcal{H}_3$ . LSI Operatoren sind kommutativ, assoziativ und distributiv. Ein Operator ist idempotent, wenn eine zweite Anwendung desselben Operators keine Veränderung mehr bewirkt (JÄHNE 2005: 535):  $\mathcal{H}\mathcal{H}I = \mathcal{H}I$ .

### 2.3.3 Homogene Punkttransformationen

In diesem Abschnitt werden einige gebräuchliche homogene Punkttransformationen (HPT, auch Histogrammoperationen) vorgestellt werden. Am häufigsten werden diese neben der Bildbearbeitung in der Stufe der Bildvorverarbeitung eingesetzt.

Da die Operationen homogen, also von der Position im Bild unabhängig sind, vereinfacht sich (2.17) zu:

$$(2.21) \quad A_{xy} = f(E_{xy}), \quad f: \{0..G\} \rightarrow \{0..G\}$$

$G$  sei dabei der maximale Grauwert, bei 8-bit Digitalisierung 255. Die HPT bilden also die Grauwerte auf sich selbst ab. Da sie im Allgemeinen mehrere Eingabewerte auf denselben Ausgabewert abbilden, also nicht injektiv sind, sind sie meistens nicht umkehrbar und nicht linear.

Die HPT können jeweils anhand ihrer Transformationskennlinie (auch Grauwertübertragungsfunktion oder Gradiationskurve) charakterisiert werden, die kennzeichnet, welcher Grauwert welchem zugewiesen wird. Im diskreten Fall wird sie als LUT realisiert.

Als Histogramm eines Bildes wird die statistische Häufigkeitsverteilung seiner Grauwerte bezeichnet. Im Falle kontinuierlicher Grauwerte wird der Wertebereich in gleiche Intervalle geteilt und damit quasi eine Quantisierung herbeigeführt (JÄHNE 2005: 88).

Eine einfache homogene Punkttransformation ist die Negativtransformation  $P_N$  oder Invertierung eines Bildes (JÄHNE 2005: 270), die in der Morphologie auch als Komplementierung bezeichnet wird:

$$(2.22) \quad P_N(g) = G - 1 - g.$$

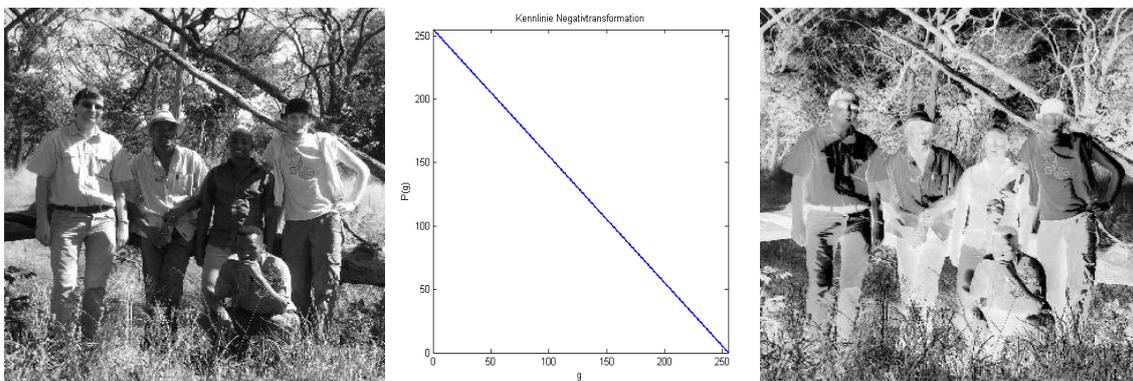


Abb. 2.13: Negativtransformation: Originalbild, Kennlinie, transformiertes Bild

Eine andere übliche Standardtransformation ist die Gammatransformation:

$$(2.23) \quad P_\gamma(g) = \frac{G}{G^\gamma} g^\gamma.$$

Die Gammatransformation bewirkt für  $\gamma < 1$  eine Spreizung der Grauwerte in den dunklen Bereichen und eine Stauchung in den hellen Bereichen. Das kommt der Dynamik des menschlichen Auges entgegen, das in einem weiten Helligkeitsbereich Grauwerte differenzieren kann, wenn sie einen relativen Helligkeitsunterschied von etwa 2% haben (JÄHNE 2005: 49). Die Charakteristik des menschlichen Sehsystems ist also eher logarithmisch und seine Dynamik ist wesentlich größer als die maschineller Systeme, die typischerweise nur 256 Grauwerte auflösen. Dabei ist die relative Auflösung in den hellen Bildbereichen zwar sehr hoch, in den dunklen dafür aber viel niedriger (bei einem Grauwert von 10 beispielsweise nur 10%). Daher erscheinen in Bildern mit hohem Kontrast oftmals die dunklen Bereiche unterlaufen, oder die hellen Bereiche überlaufen. An Stelle der der Gammatransformation wird auch eine Logarithmustransformation eingesetzt.

Bei der Histogrammspreizung oder Tonwertspreizung werden die Grauwerte eines bestimmten Intervalls üblicherweise linear auf den gesamten Grauwertbereich abgebildet, die Grauwerte außerhalb des Intervalls werden auf den minimalen bzw. maximalen Wert abgebildet. Die Histogrammspreizung wird vor allem zur Erhöhung des Kontrasts verwendet.

Bei der Histogrammäqualisation soll das Histogramm in ein uniformes Histogramm überführt werden, bei dem alle Grauwerte gleich häufig vorkommen. Da dies meistens nicht möglich ist, wird es stattdessen durch ein quasi-uniformes angenähert. Dabei werden häufige Grauwertbereiche gespreizt und seltene Grauwerte gestaucht (RICHARDS & JIA 2006: 90 ff). Die Übergangsfunktion wird mit Hilfe des kumulativen Histogramms  $C(g)$  berechnet:

$$(2.24) \quad C(g) = \sum_{n=0}^g H(n), \quad P_{EQU}(g) = \frac{G}{N} C(g)$$

Dabei ist  $H(g)$  das Histogramm des Ausgangsbildes und  $N$  die Gesamtzahl der Pixel.

Eine besondere HPT stellt der globale Schwellenwert dar, der allen Grauwerten nur die Werte 1 und 0 zuordnet, also ein binäres Bild erzeugt. Dieser wird in 2.4.1 näher vorgestellt.

### 2.3.4 Faltung und lineare Filter

Nachbarschaftsoperationen sind einfacher zu verstehen, wenn zunächst der eindimensionale Fall der linearen Systemtheorie betrachtet wird. Gegeben sei ein Signal  $x(t)$  in der Zeit. Geht dieses durch ein System (etwa einen Verstärker oder einen Filter), so wird es zum Ausgangssignal  $y(t)$  transformiert. Die Eigenschaften des Systems sind dabei vollständig durch seine Impulsantwort  $h(t)$  beschrieben (RICHARDS & JIA 2006: 110).

Die Impulsantwort  $h(t)$  ist das Antwortsignal des Systems auf die Impulsfunktion und wird auch Diracsche Deltafunktion genannt. Sie kann als Grenzwert eines gleichmäßigen Pulses der Länge  $\alpha$  und der Amplitude  $1/\alpha$  für  $\alpha \rightarrow 0$  verstanden werden. Nach RICHARDS & JIA ist die beste formale Definition der Deltafunktion:

$$(2.25) \quad \delta(t) = 0, t \neq 0, \quad \int_{-\infty}^{\infty} \delta(t) dt = 1.$$

Entsprechend muss die Deltafunktion an der Stelle  $t=0$  unendlich sein.

Die Antwort des Systems auf ein beliebiges Signal wird mathematisch durch die Faltung (*convolution*) des Eingangssignals mit der Impulsantwort beschrieben (rechts Operatorschreibweise):

$$(2.26) \quad y(t) = \int_{-\infty}^{\infty} x(\tau)h(t - \tau)d\tau, \quad y(t) = x(t) * h(t).$$

Mathematische Voraussetzung für die Faltung ist, dass eine der beide Funktionen einen kompakten Träger hat, d.h. außerhalb eines kompakten Intervalls  $[-a,a]$  Null ist (KÖNIGSBERGER 1999: 310). Signaltheoretisch bedeutet dies, dass die Impulsantwort endlich sein muss. Ein solcher Filter wird auch FIR-Filter genannt (*finite impulse response*). Während das in der digitalen Signalverarbeitung kein Problem darstellt, ist die Implementierung analoger FIR-Filter nicht trivial. Durch die Laufzeitverzögerungen elektrischer Signale kann es bei analogen Filtern leicht zu Rückkopplungen und Eigenschwingungen kommen.

Die Faltung wirkt anschaulich, als würde die Impulsantwort entlang des ganzen Signals verschoben und das Signal dabei quasi mit der Impulsantwort gewichtet. Die Wichtungsfunktion wird auch als Faltungskern bezeichnet.

Da die Faltung bilinear ist (KÖNIGSBERGER 1999: 310) ist die Faltung mit einem bestimmten Kern linear und homogen, also ein LSI Operator (siehe 2.3.2). Damit ist sie insbesondere auch kommutativ, assoziativ und distributiv. Die Faltung mit der Diracschen  $\delta$ -Funktion ist die Identität, die  $\delta$ -Funktion ist also das neutrale Element bezüglich der Faltung.

Eine analoge Operation lässt sich auch für Bilder durchführen. Da das Signal hier zweidimensional und diskret ist, wird in der Bildverarbeitung vor allem die zweidimensionale diskrete Faltung verwendet. Diese ist in PINZ (1994: 59) definiert als:

$$(2.27) \quad A(x, y) = K(m, n) * E(x, y) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} K(m, n)E(x - m, y - n).$$

Dabei ist  $K(m,n)$  der Faltungskern, der in der BV oft auch Faltungsmaske oder Filtermaske genannt wird, und  $M$  und  $N$  sind die Anzahl der Spalten bzw. Zeilen des Bildes. Da der Kern, üblicherweise deutlich kleiner ist als das Bild, handelt es sich um eine Nachbarschaftsoperation.

Der Filterkern kann auch hier als elementare Antwort des Systems begriffen werden, die aber statt Impulsantwort meist Punktantwort (*point spread function*, PSF) genannt wird (JÄHNE 2005: 120). Die PSF beschreibt also, wie das System auf einen einzelnen Punkt reagiert und diesen ‚verteilt‘. Darin lässt sich auch die Linearität der Faltung veranschaulichen. Sei  ${}^m n \mathbf{P}$  zunächst ein Satz von Bildern, bei denen jeweils nur der Punkt  $(m,n)$  1 ist und alle anderen 0 sind. Diese bilden eine Basis für alle Bilder, d.h. jedes Bild lässt sich als Linearkombination von solchen Basisbildern

$$(2.28) \quad G = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \lambda_{mn} \mathbf{P}^{mn}$$

darstellen (JÄHNE 2005: 43), wobei die  $\lambda_{mn}$  der Grauwert an der Stelle  $(m,n)$  ist. Das Bild kann also vor der Faltung in seine einzelnen Bildpunkte zerlegt werden, deren Antwort auf das System ist jeweils die verschobene PSF, und anschließend werden die Einzelantworten wieder zusammengesetzt. Darum ist das Resultat einer Faltungsoperation auch völlig durch das Resultat eines Basisbildes bestimmt. Weiterhin lässt sich damit zeigen, dass jeder LSI-Operator einer Faltung entspricht (siehe etwa JÄHNE 2005:127).

Ein Problem liegt im endlichen Definitionsbereich von Bildern, da am Bildrand  $E(x-m,y-n)$  nicht für alle  $(m,n)$  definiert ist. Die Beschränkung auf einen endlichen Definitionsbereich entspricht der Multiplikation des vermeintlich unendlichen Bildsignals (die Welt hört ja am Bildrand nicht auf) mit einer Recheckfunktion, die 1 innerhalb des Bildes und 0 außerhalb ist. Oft werden die Werte außerhalb des Bildes daher als 0 angenommen, was aber das Ergebnis der Faltung am Bildrand verzerrt. Eine andere Möglichkeit ist, das Bild als periodisch wiederkehrende Struktur zu betrachten.

Nachdem nun allgemein auf die Faltung eingegangen wurde, sollen im folgenden Abschnitt wichtige lineare Filter in der Bildverarbeitung und ihre Auswirkungen vorgestellt werden.

### 2.3.5 Standardfilter und ihre Effekte

#### Glättungsfilter

Glättungsfilter werden verwendet um (etwa durch Rauschen verursachte) Inhomogenitäten herauszufiltern (*smoothing* oder *blurring*) und mitteln meistens über einen bestimmten Bereich in der Nachbarschaft jedes Pixels, der durch die Filtermaske bestimmt wird. Da die Bildinhalte dabei nicht verschoben werden sollen, müssen die Masken symmetrisch sein. Daher werden üblicherweise Filtermasken ungerader Größe verwendet, z.B.  $3 \times 3$ ,  $5 \times 5$  oder  $7 \times 7$ . Um den Mittelwert zu erhalten, muss die Summe der Filtermaske 1 sein.

Die einfachste Glättung ist die durch einen Rechteckfilter. Dieser addiert alle Punkte in der Filtermaske und teilt das Ergebnis durch die Anzahl:

$${}^3R = \frac{1}{9} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}, \quad {}^3R_x = \frac{1}{3} (1 \quad 1 \quad 1), \quad {}^3R_y = {}^3R'_x, \quad {}^3R = {}^3R_x * {}^3R_y.$$

Die Filtermaske kann selbst als Faltung der  $3 \times 1$  Maske  $R_x$  mit dem  $1 \times 3$  Filter  $R_y$  dargestellt werden. Da die Faltung assoziativ ist, kann die Faltung mit  ${}^3R$  durch die sequentielle Faltung mit  ${}^3R_y$  und  ${}^3R_x$  ersetzt werden, und da sie kommutativ ist, sogar in beliebiger Reihenfolge. Diese Eigenschaft wird Separierbarkeit des Filters genannt und kann erheblich Rechenzeit sparen.

Ein großer Nachteil von Rechteckfiltern ist, dass sie nicht isotrop sind, die Glättung hängt also von der Richtung ab. Außerdem gewichtet er alle Pixel in der Nachbarschaft unabhängig von ihrer Entfernung gleich stark.

Wesentlich bessere Eigenschaften in dieser Hinsicht haben Gaußkerne als Filter, die daher in der Bildverarbeitung weit verbreitet sind. Sie nähern eine 2-dimensionale Normalverteilung (Gaußsche Glockenkurve) an (vgl. (2.8) und PINZ 1994: 61):

$$(2.29) \quad N(x, y, \sigma) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right)$$

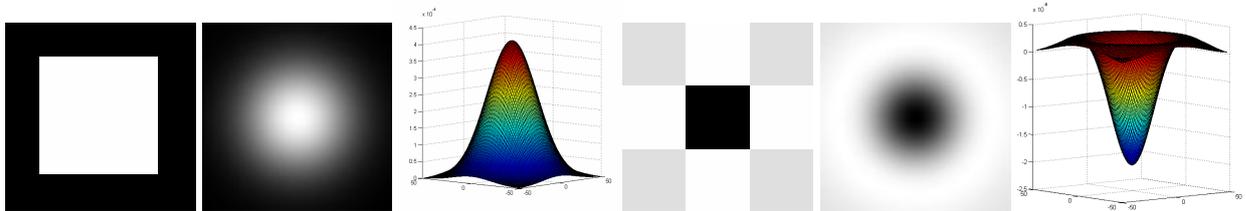


Abb. 2.14: Verschiedene Filtermasken als Bild und in 3D: Rechteck, Gauß, 3D-Gauß, Laplace, DoG, 3D-DoG

Die diskrete Näherung der Gaußschen Glockenkurve sind Binomialverteilungen. Der einfachste Binomialfilter ist:  $\mathbf{B} = \frac{1}{2} \begin{pmatrix} 1 & 1 \end{pmatrix}$ .

An dieser Stelle sei darauf hingewiesen, dass streng genommen stets zwischen dem Operator  $\mathcal{B}$  und der Filtermaske  $\mathbf{B}$  zu unterscheiden ist. Die Anwendung des Operators auf ein Bild entspricht dabei dessen Faltung mit der Filtermaske:  $\mathcal{B}I = \mathbf{B} * I$ . In der Literatur werden aber häufig die Filtermasken selbst unpräzise als Operatoren bezeichnet. Erschwerend kommt hinzu, dass sowohl die Masken als auch die Operationen als Filter bezeichnet werden. In dieser Arbeit wurde zwar eine klare Terminologie angestrebt, manchmal ist es jedoch umständlich zwischen beiden zu unterscheiden. Wird eine Filtermaske also vereinzelt als Operation bezeichnet, so ist darunter stets die Faltung mit der Filtermaske zu verstehen.

Alle weiteren Binomialfilter sind mehrfache sequentielle Anwendungen von  $\mathbf{B}$ . Die n-te Filtermaske  $\mathbf{B}^n$  ist damit die (n-1) fache Faltung des ersten Filters mit sich selbst. Der 2D- Binomialfilter ergibt sich wiederum durch Faltung eines horizontalen 1D-Binomialfilters mit einem vertikalen (JÄHNE 2005: 321):

$$\mathbf{B}^2 = \frac{1}{4} \begin{pmatrix} 1 & 2 & 1 \end{pmatrix} * \frac{1}{4} \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} = \frac{1}{16} \begin{pmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{pmatrix}.$$

Binomialfilter sind gut geeignet um statistisches, also normalverteiltes, Rauschen zu unterdrücken.

Weniger effektiv sind sie bei binärem Rauschen, bei dem zufällig verteilte Pixel falsche Grauwerte haben (z.B. durch Übertragungsfehler). Für dieses eignet sich zur Glättung der Medianfilter, der allerdings weder linear ist noch eine Faltungsoperation. Der Medianfilter ist ein Rangordnungsfiler, der alle Grauwerte aus einer durch eine Maske definierten Umgebung der Größe nach sortiert. Dann wird dem Pixel der Median der Grauwerte zugewiesen. Im Vergleich zum statistischen Mittelwert beim Rechteckfilter ist der Median so weniger empfindlich gegenüber Ausreißerwerten. Im Falle von

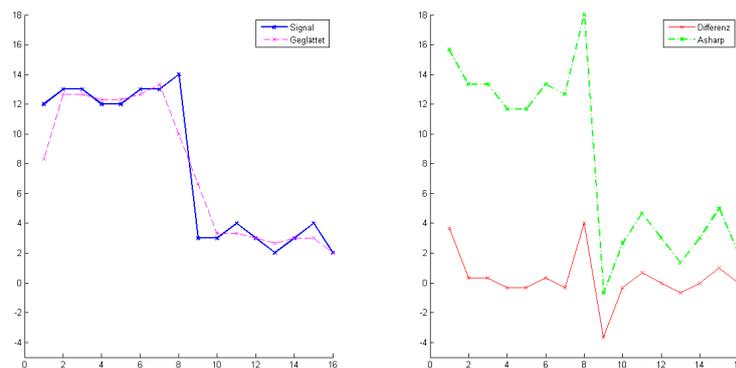
binärem Rauschen verändern die stark abweichenden Fehler-Werte das Ergebnis also nicht maßgeblich.

### Scharfzeichner

Eine dem Glätten entgegengesetzte Aktion ist das Scharfzeichnen. Während die Glättungsfilter zur Entfernung von Rauschen dienen, reduzieren sich gleichzeitig auch die geometrische Detailliertheit an Stellen, wo dies nicht erwünscht ist. So werden starke Grauwertübergänge im Bild (Kanten, *edges*), die verschiedene Objekte voneinander abgrenzen, verwischt.

Eine einfache Technik zur Schärfung ist, ein geglättetes Bild vom Original abzuziehen. Wird dieses Ergebnis zum Originalbild addiert, werden die Kanten hervorgehoben.

Abb. 2.15 zeigt die Wirkungsweise von Schärfungsfiltern: Links ist in Blau das Signal einer Kante mit Rauschen und in Magenta das gleiche Signal mit einem  $3 \times 1$  Rechteckfilter gedopt (das Rauschen wird unterdrückt, die Kante ist aber auch deutlich schwächer ausgeprägt). Außerdem entstehen Probleme am Bildrand) zu sehen. Rechts ist in Rot die Differenz zwischen Original und geglättetem Bild zu sehen und in Grün das geschärfte Bild (die Kante wird hervorgehoben, aber das Rauschen wird ebenfalls stärker).

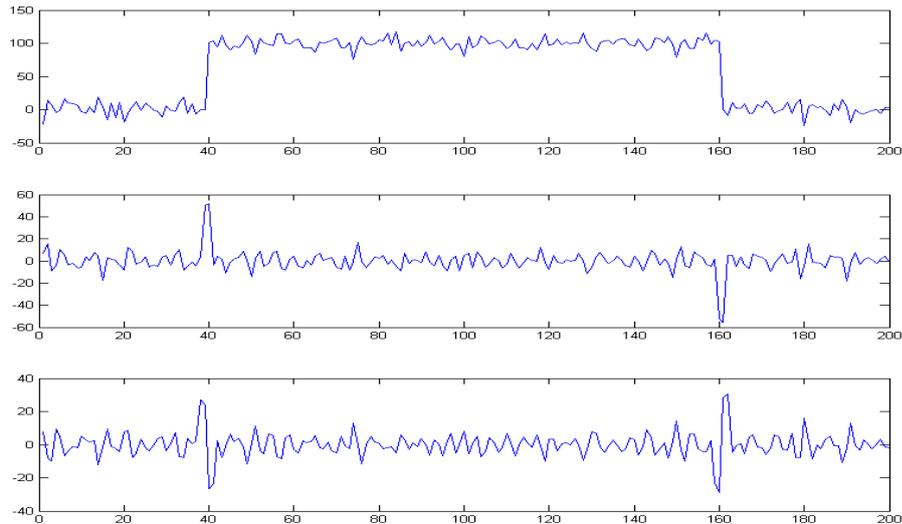


**Abb. 2.15: Wirkungsweise von Schärfungs- und Glättungsfiltern**

### Kantenfinder

Eine weitere wichtige Gruppe von Filtern sind die Kantenfinder. Eine ideale Kante ist eine Diskontinuität in der Grauwertfunktion (JÄHNE 2005: 345). Kantenfinder sollen Veränderungen entdecken bzw. verstärken und Bereiche konstanter Werte unterdrücken.

Auch sie sollen verschiebungsfrei sein, was durch einen antisymmetrischen Faltungskern erreicht wird. Da sie auf konstante Bereiche nicht reagieren dürfen, müssen sie eine Summe von Null haben (JÄHNE 2005: 350). Idealerweise sollte der Filter außerdem isotrop sein, was jedoch bei den diskreten Filtern immer höchstens annähernd erfüllt wird. Geeignete Kantenfinder sind Ableitungen, da die erste Ableitung an Kanten Extremwerte und die zweite Nulldurchgänge hat (siehe Abb. 2.16).



**Abb. 2.16: Diskretes Signal mit zwei Kanten und leichtem Rauschen, erste und zweite Ableitung.**

Um im diskreten Fall Ableitungen bilden zu können, muss aber zunächst eine Approximation für den Ableitungsoperator gefunden werden. Hierfür werden die partiellen Ableitungen durch Differenzen erster Ordnung angenähert (JÄHNE 2005: 355). Der einfachste Differenzenoperator ist  $(1, -1)$  bzw.  $(-1, 1)$ , erfüllt aber die Symmetriebedingung nicht. Der erste symmetrische Differenzenoperator ist:

$\mathbf{D}_x = \frac{1}{2} \begin{pmatrix} 1 & 0 & -1 \end{pmatrix}$  und kann als Faltung von  $(1,-1)$  mit dem ersten Binomialfilter aufgefasst werden:

$\mathbf{D}_x = (1 \ -1) * B$ . Um die Verschiebungsfreiheit zu erhalten, wird die Ableitung also leicht geglättet. Das führt an idealen Kanten aber dazu, dass die maximale Steigung halbiert wird (siehe Abb. 2.16).

Im mehrdimensionalen Fall kommt die Schwierigkeit hinzu, die totale Ableitung zu finden. In der Bildverarbeitung wird daher häufig der Gradient verwendet (GONZALEZ, WOODS 1993: 198), der Vektor aller Richtungsableitungen entlang der Koordinatenachsen:

$$(2.30) \quad \text{grad}(f) = \overset{\rho}{\nabla} f(x, y) = \begin{pmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{pmatrix}, \quad \overset{\rho}{\nabla} = \begin{pmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \end{pmatrix}.$$

Der Gradient kann in Operatorschreibweise mit dem Nabla-Operator  $\overset{\rho}{\nabla}$  ausgedrückt werden. Die erste Ableitung in eine bestimmte Richtung ergibt sich aus dem Skalarprodukt des Gradienten mit dem Einheitsvektor in die entsprechende Richtung  $(\cos \theta, \sin \theta)^t$ . Der Betrag des Gradienten ist

$$(2.31) \quad |\overset{\rho}{\nabla} f| = \|\overset{\rho}{\nabla} f\|_2 = \sqrt{\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2} \approx \left|\frac{\partial f}{\partial x}\right| + \left|\frac{\partial f}{\partial y}\right|$$

Er ist unabhängig von der Richtung des Koordinatensystems und bestimmt die Stärke der Steigung. Somit ist er gut als Kantenfiter geeignet. Sein Nachteil ist, dass er relativ aufwendig zu berechnen ist. Die meisten klassischen Kantenfiter können daher als Versuch betrachtet werden, die partiellen Ableitungen und den Betrag des Gradienten anzunähern.

Der Prewitt-Operator entspricht beispielsweise der Faltung mit den Masken  $\mathbf{P}$ :

$$\mathbf{P}_x = \begin{pmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{pmatrix}, \quad \mathbf{P}_y = \mathbf{P}_x^t, \quad \mathbf{S}_x = \begin{pmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{pmatrix}, \quad \mathbf{S}_y = \mathbf{S}_x^t, \quad \mathbf{R}_x = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \mathbf{R}_y = \mathbf{R}_x^t.$$

Analog ist der Sobel-Operator die Faltung mit  $\mathbf{S}$ . Der Roberts-Operator  $\mathbf{R}$  bildet Differenzen entlang der Diagonalen ab, ist aber nicht verschiebungsfrei.

Alternativ zu den gradientenbasierten Verfahren kann auch nach den Nulldurchgängen der zweiten Ableitung gesucht werden. Der Differenzenoperator für die zweite Ableitung ergibt sich durch zweifache Anwendung des einfachen Operators erster Ordnung (1,-1).

$D_x^2 = (1 \ -1) * (1 \ -1) = (1 \ -2 \ 1)$ . Die Summe der zweiten Ableitungen ist damit:

$$(2.32) \quad \Delta = \sum_n \frac{\partial^2}{\partial x_n^2}, \quad \Delta \approx D_x^2 + D_y^2 = (1 \ -2 \ 1) + \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{pmatrix},$$

und wird in Operatorschreibweise durch den Laplaceoperator  $\Delta$  ausgedrückt. In (2.32) rechts ist eine diskrete Näherung für  $\Delta$  angegeben. Da die Nulldurchgänge in der zweiten Ableitung auch durch Rauschen verursacht sein können, müssen die Signalspitzen vor bzw. nach dem Nulldurchgang deutlich über dem Rauschpegel liegen. Daher ist der Laplaceoperator als Kantenfiter rauschempfindlicher als gradientenbasierte Verfahren (JÄHNE, 2005: 360).

Um dem zu begegnen, kann das Bild vorher durch einen Gaußfilter geglättet werden:

$\Delta(B^2 * f) = (\Delta B^2) * f$ . Dieser kombinierten Filter wird *Laplacian of Gaussian* (LOG) genannt. Mit dem Parameter  $s$  des Gauskerns lässt sich der Maßstab für die gesuchten Nulldurchgänge steuern (PINZ 1994: 149). Ähnliche Ergebnisse liefert der *difference of Gaussian* Filter (DOG), siehe dazu beispielsweise JÄHNE (2005: 367).

### 2.3.6 Fouriertransformation

#### Vom Ortsraum zum Frequenzraum

Alle bisherigen Transformationen wurden direkt im Bild vorgenommen. Das Bild wurde dabei als räumliche Anordnung von Grauwerten verstanden, und wird daher auch als Ortsraum (*spatial domain*) bezeichnet. In 2.3.4 wurde eine orthonormale Basis für diesen Raum eingeführt, mit der sich jedes Bild aus Basisbildern mit nur je einem Punkt zusammensetzen lässt. Das Bild kann also auch als Vektor eines  $M \times N$ -Vektorraums aufgefasst werden. Die Elemente der Bildmatrix sind dann die Koeffizienten der Linearkombination der Basisvektoren  ${}^{mn}\mathbf{P}$ .

Da diese Basis aber beliebig gewählt wurde, gibt es daneben viele weitere Möglichkeiten, das Bild zu repräsentieren. Wird eine Koordinatentransformation des  $M \times N$ -Raum durchgeführt, bleibt das Bild dasselbe. Alle Bilddarstellungen sind einander also äquivalent und können durch Transformationen ineinander überführt werden (JÄHNE 2005: 43). Eine andere sehr gebräuchliche Repräsentation ist die Überlagerung von periodischen Funktionen unterschiedlicher Frequenz, insbesondere komplexer

Exponentialfunktionen. Daher wird der entsprechende Raum Frequenzraum (*frequency domain*) genannt. Die wichtigste Transformation, die ein Bild in seine periodischen Anteile zerlegt, ist die Fouriertransformation (FT).

Zur Vereinfachung soll zunächst wieder der eindimensionale Fall betrachtet werden. Zur Steigerung der Übersichtlichkeit wird im Folgenden soweit möglich versucht, Funktionen im Ortsraum mit kleinen und Funktionen im Frequenzraum mit großen Buchstaben zu bezeichnen. Da dies aber im Falle von Vektoren und Matrizen mit der üblichen Notation der linearen Algebra konkurriert, werden Elemente im Frequenzraum auch mit Dach gekennzeichnet, z.B.  $\hat{g}$ .

Die Fouriertransformation kann als Skalarprodukt einer Funktion mit der komplexen Exponentialfunktion aufgefasst werden. Die komplexe Exponentialfunktion ist gemäß der Eulerformel definiert als:

$$(2.33) \quad f(a) = a e^{i\phi} = a \cos \phi + ia \sin \phi.$$

Der Realteil der komplexen Exponentialfunktion ist also ein Kosinus und der Imaginärteil ein Sinus. Das Skalarprodukt für quadratintegrierbare Funktionen sei:

$$(2.34) \quad \langle g(x) | h(x) \rangle = \int_{-\infty}^{\infty} g^*(x) h(x) dx,$$

wobei \* für die komplexe Konjugation steht.

Die eindimensionale Fouriertransformation ist dann definiert als

$$(2.35) \quad F(u) = \int_{-\infty}^{\infty} f(x) e^{-2\pi i u x} dx = \langle e^{2\pi i u x} | f(x) \rangle = \langle w^{u x} | f(x) \rangle.$$

Die Konstante  $w = e^{2\pi i}$  wird als Kern der Fouriertransformation bezeichnet. Jeder Punkt der Fouriertransformierten  $F(u)$  entspricht also dem Skalarprodukt der Funktion  $f(x)$  mit einer komplexen Exponentialfunktion einer bestimmten Frequenz. Die inverse Fouriertransformation ist entsprechend:

$$(2.36) \quad f(x) = \int_{-\infty}^{\infty} F(u) e^{2\pi i u x} du.$$

Die Fouriertransformierte einer reellen Funktion ist grundsätzlich komplex. Zweckmäßiger als die Darstellung in Real- und Imaginärteil ist jedoch oft die Darstellung durch Betrag (auch Fourierspektrum genannt) und Phase  $F(u) = |F(u)| e^{i\phi(u)}$ , mit  $|F(u)| = (\Re(F(u))^2 + \Im(F(u))^2)^{1/2}$ , und  $\phi(u) = \arctan(\Im(F(u)) / \Re(F(u)))$ .

### Diskrete Fouriertransformation (DFT)

Wird von der kontinuierlichen Funktion  $f(x)$  zur diskreten  $f(n) = f(n\Delta x)$  übergegangen, dann vereinfacht sich das Integral in (2.35) zur Summe:

$$(2.37) \quad F(u) = \int_{-\infty}^{\infty} f(x) e^{-2\pi i u x} dx = \Delta x \sum_{n=0}^N f(n) e^{-2\pi i u n \Delta x}$$

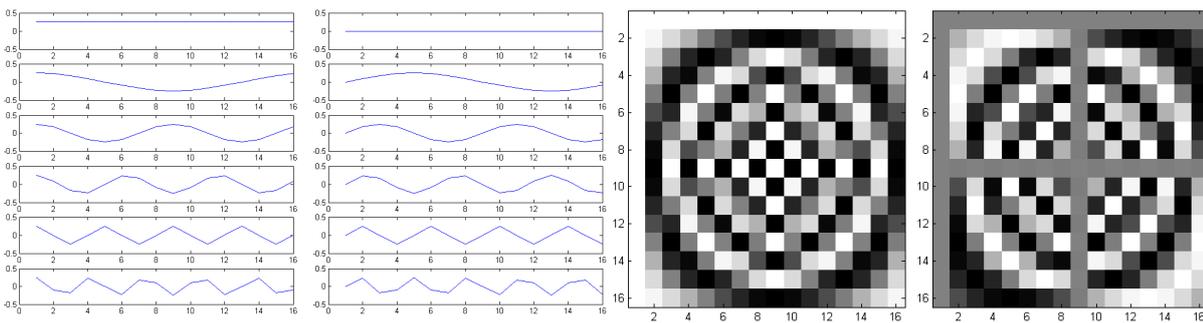
Wird weiterhin zu diskreten Frequenzen  $u = k\Delta u$  übergegangen, ist die diskrete Fouriertransformation (DFT) gegeben durch:

$$(2.38) \quad F(k) = \frac{1}{N} \sum_{n=0}^{N-1} f(n) \exp\left(\frac{-i2\pi kn}{N}\right) = \frac{1}{N} \sum_{n=0}^{N-1} f(n) w_N^{-kn}, \quad k = 0 \text{ K } N-1,$$

mit  $x = n\Delta x, u = k\Delta u, \Delta x\Delta u = 1/N$ . Die gesampelte Funktion  $f(n)$  kann dabei als auch als Spaltenvektor der Länge  $N$ :  $\mathbf{f} = (f_0 \ f_1 \ \text{K} \ f_{N-1})^t$  betrachtet werden. Der Kern der diskreten Fouriertransformation wurde (analog zu (2.35)) als  $w_N = w^{1/N} = \exp\left(\frac{2\pi i}{N}\right)$  abgekürzt. Es lässt sich zeigen, dass die Vektoren

$$(2.39) \quad \mathbf{b}_k = \frac{1}{\sqrt{N}} \left( w_N^0 \quad w_N^k \quad \text{K} \quad w_N^{(N-1)k} \right)^t, \quad k = 0 \text{ K } N-1$$

eine Orthonormalbasis des  $\mathbb{C}^N$  (JÄHNE 2005: 47) bilden. Die einzelnen Basisvektoren entsprechen dabei der gesampelten komplexen Exponentialfunktion einer bestimmten Frequenz, die jeweils so gewählt ist, dass  $k$  ganze Wellen ins Intervall  $[0, N]$  passen. Daher wird  $k$  auch Wellenzahl genannt. Abb. 2.17 zeigt die Basis für  $N = 16$ . Links sind dabei die ersten sechs Basisvektoren als gesampelte Exponentialfunktion dargestellt (Real und Imaginärteil), rechts die komplette Basis als Bild ( $k$  von oben nach unten,  $n$  von links nach rechts). Die Zeilen sind also die Basisvektoren, da  $w_N^{nk} = w_N^{kn}$  gilt, sind sie aber identisch mit den Spalten. Es ist klar zu erkennen, dass immer der Realteil (Kosinus) symmetrisch und der Imaginärteil (Sinus) antisymmetrisch zur Mitte ist.



**Abb. 2.17:** Links: die ersten sechs Basisvektoren der DFT für  $N=16$ , Realteil und Imaginärteil. Rechts: die komplette Basis als Bild, Real- und Imaginärteil

Die DFT kann damit auch als Skalarprodukt verstanden werden und zwar zwischen dem Vektors  $\mathbf{f}$  und den einzelnen Basisvektoren.

$$(2.40) \quad F(k) = \frac{1}{N} \sum_{n=0}^{N-1} \exp\left(\frac{-i2\pi nk}{N}\right) f(n) = \frac{1}{N} \sum_{n=0}^{N-1} w_N^{-nk} f(n) = \frac{1}{\sqrt{N}} \langle \mathbf{b}_k | \mathbf{f} \rangle$$

Das Skalarprodukt entspricht in der linearen Algebra der Projektion des Vektors auf die neuen Basisvektoren, die  $F(k)$  sind also die Komponenten bezüglich der neuen Basis. Die DFT ist also eine normale Koordinatentransformation wie beispielsweise eine Rotation, nur dass die neue Basis komplex ist (JÄHNE 2005: 48).

Die inverse diskrete Fouriertransformation (IDFT) ist gegeben durch

$$(2.41) \quad f(n) = \sum_{k=0}^{N-1} F(k) \exp\left(\frac{i2\pi kn}{N}\right).$$

**Zweidimensionale DFT**

Die DFT kann wie folgt auf 2D-Signale (Bilder) verallgemeinert werden. Wie die Vektoren in (2.39) bilden auch die Matrizen

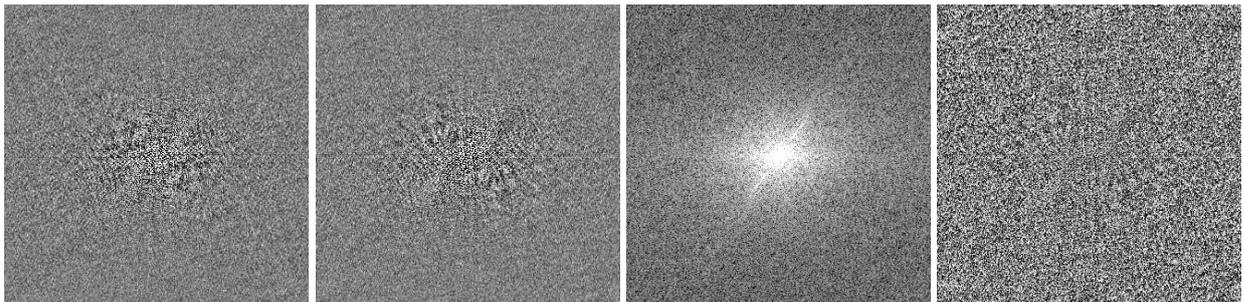
$$(2.42) \quad \mathbf{B}_{uv} = \frac{1}{\sqrt{MN}} \begin{pmatrix} w_M^0 \\ w_M^u \\ \mathbf{M} \\ w_M^{(M-1)u} \end{pmatrix} \begin{pmatrix} w_N^0 & w_N^k & \mathbf{K} & w_N^{(N-1)k} \end{pmatrix}$$

eine Basis für den Raum der komplexen  $M \times N$  Matrizen. Die Matrizen sind hierbei als äußeres Produkt von zwei Basisvektoren der 1D-DFT dargestellt, woran die Separierbarkeit der DFT gut zu erkennen ist (JÄHNE 2005: 50). Mit dem Skalarprodukt für Matrizen:  $\langle \mathbf{G} | \mathbf{H} \rangle = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} g_{mn}^* h_{mn}$ , ist die 2D-DFT analog zu (2.40):

$$(2.43) \quad g_{mn} = \frac{1}{\sqrt{MN}} \langle \mathbf{B}_{uv} | \mathbf{G} \rangle = \frac{1}{MN} \sum_{m=0}^{M-1} \left( \sum_{n=0}^{N-1} g_{mn} w_N^{-nv} \right) w_M^{-mu} = \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} g_{mn} \exp\left(-\frac{i2\pi mu}{M}\right) \exp\left(-\frac{i2\pi nv}{N}\right)$$

**Eigenschaften der Fouriertransformation**

Abb. 2.18 zeigt die Fouriertransformierte des Bildes in Abb. 2.13. Alle Bilder wurden so verschoben, dass die Null-Frequenz in der Mitte ist. Nachdem die mathematischen Grundlagen der FT ausführlich dargestellt wurden, soll noch kurz auf ihre Eigenschaften eingegangen werden.



**Abb. 2.18:** Fouriertransformierte des Bildes in Abb. 2.13. Real- und Imaginärteil (Histogramm beschnitten), Spektrum (gamma-korrigiert), Phase

Die FT ist eine globale Operation und damit sehr rechenaufwendig. Die Rechenzeit kann jedoch mit dem *Fast Fourier Transform* Algorithmus erheblich gesenkt werden (siehe dazu RICHARDS & JIA 2006: 179 ff).

Da die FT hier mit als Skalarprodukt hergeleitet wurde, folgt unmittelbar, dass sie linear ist. Weiterhin ist sie periodisch, d.h.  $g_{m+kM, n+lN} = g_{mn}$  und  $\hat{g}_{u+kM, v+lN} = \hat{g}_{uv}$ ,  $\forall k, l \in \mathbb{Z}$ . In (2.42) wurde bereits gezeigt, dass die 2D-DFT separierbar ist, es können also zuerst die Zeilen und dann die Spalten transformiert werden.

Daneben hat die FT eine ganze Reihe von Symmetrieeigenschaften (siehe JÄHNE 2005: 52), von denen hier nur einige wichtige erwähnt werden sollen. Ist die Funktion im Ortsraum, wie bei Bildern im Allgemeinen der Fall, reell, dann ist die Fouriertransformierte hermitesch, d.h. es gilt:  $\hat{f}(-k) = \hat{f}^*(k)$ . Wird die Periodizität  $\hat{g}_{M-m, N-n} = \hat{g}_{-m, -n} = \hat{g}_{mn}^*$  hinzu genommen, dann folgt, dass die DFT auch ‚hermitesch bezüglich der Bildmitte‘ ist. Dies ist gut an den Basisvektoren in Abb. 2.17 zu erkennen. Weiterhin ist die Transformierte einer geraden Funktion gerade und die einer ungeraden ungerade.

Wird eine Funktion im Ortsraum gedreht, so ist die Fouriertransformierte ebenso gedreht. Wird sie mit einem Skalar gestaucht, so wird die Transformierte entsprechend gedehnt und umgekehrt. Wird die Ortsfunktion verschoben, so ändert sich das Fourierspektrum nicht sondern lediglich die Phase (JÄHNE 2005: 54f). Umgekehrt entspricht eine Phasenverschiebung im Ortsraum um  $e^{i\pi(x+y)} = (-1)^{x+y}$  im Frequenzraum gerade einer Verschiebung um  $N/2$ , was zur besseren Darstellung des Fourierspektrums verwendet werden kann.

Für eine Funktion im Ortsraum und ihre Fouriertransformierte gilt eine Unschärferelation: Ist eine Funktion im Ortsraum genau lokalisiert, so ist sie im Frequenzraum stark ausgedehnt und umgekehrt (JÄHNE 2005: 58). Insbesondere gibt es eine Beschränkung der Glattheit einer Funktion und der Kompaktheit ihres Spektrums, d.h. das Spektrum einer stetigen Funktion fällt mindestens mit  $|k|$  ab, das einer stetig differenzierbaren mit  $|k|^2$  usw. (JÄHNE 2005: 58).

### Faltungstheorem

Seien  $f(x)$  und  $g(x)$  Funktionen im Ortsraum und  $F(k)$  und  $G(k)$  ihre Fouriertransformierten, dann gilt:

$$(2.44) \quad h = f * g \leftrightarrow H = FG \quad \text{und} \quad h = fg \leftrightarrow H = F * G.$$

Für Matrizen gilt ähnlich:  $\mathbf{H} = \mathbf{F} * \mathbf{G} \leftrightarrow \hat{\mathbf{H}} = \mathbf{M}\hat{\mathbf{N}}\hat{\mathbf{F}}\hat{\mathbf{G}}$ .

Die Faltung einer Funktion mit einer anderen entspricht also der Multiplikation ihrer Transformierten im Frequenzraum und umgekehrt. Dies kann darauf zurückgeführt werden, dass die Basisfunktionen des Fourierraums, also die komplexen Exponentialfunktionen in (2.42), Eigenfunktionen aller Faltungsoperationen sind (JÄHNE 2005: 56).

#### 2.3.7 Filter im Frequenzraum

Das Faltungstheorem ist für die BV von zentraler Bedeutung. Daran lässt sich erkennen, dass jede Operation im Ortsraum ein Äquivalent im Frequenzraum hat. So haben die oben vorgestellten Glättungsfilter die Eigenschaft, schnelle Änderungen im Bild zu unterdrücken, während homogene

Bereiche nicht beeinflusst werden. Dies entspricht einer Elimination von hohen Frequenzen im Frequenzraum. Da die niedrigen Frequenzen den Filter „passieren“ können, werden Glättungsfilter auch als Tiefpass-Filter (*low-pass*) bezeichnet. Demgegenüber eliminieren Kantenerfilter homogene Bereiche, also langsame Veränderungen im Bild und werden daher auch Hochpass-Filter (*high-pass*) genannt. Eine Kombination aus Hoch- und Tiefpass-Filtern sind Bandpassfilter, die einen nach oben und unten begrenzten Frequenzbereich durchlassen.

Oft ist einer der beiden Räume besser geeignet, eine bestimmte Operation durchzuführen (z.B. ist eine Multiplikation wesentlich schneller zu berechnen, als eine Faltung). Daher lohnt es sich, die Wirkung von räumlichen Filtern im Frequenzraum näher zu betrachten.

Die Fouriertransformierte der PSF oder Faltungsmaske eines LSI-Filters wird als Transferfunktion (TF) bezeichnet. Anschaulich enthält sie für jede Frequenz einen Faktor, mit dem diese multipliziert wird (JÄHNE 2005: 121). Da dieser komplex ist, wird nicht nur der Betrag verändert, sondern auch die Phase verschoben.

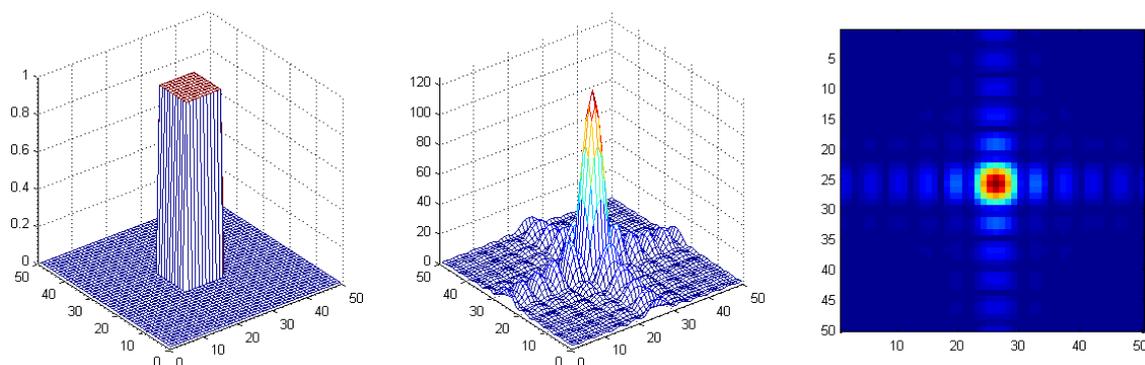
Eine überraschend einfache Transferfunktion ergibt sich beispielsweise für den Ableitungsoperator. Wird in (2.35) die Integration mit der Ableitung  $d/dx$  vertauscht, so ergibt sich die Transferfunktion zu  $2\pi i u$ . Im mehrdimensionalen Fall ist die Transferfunktion des Gradienten  $2\pi i \mathbf{k}$ . Da der Laplaceoperator das Quadrat des Gradienten ist, ist seine Transferfunktion entsprechend  $-4\pi \mathbf{k}^2$ . Diese Eigenschaft kann zur Lösung von Differentialgleichungen eingesetzt werden, da diese unter der FT zu algebraischen Gleichungen vereinfacht werden.

Da die  $\delta$ -Funktion (der Impuls) das neutrale Element bezüglich der Faltung ist, ergibt sich aus dem Faltungstheorem, dass ihre Transferfunktion das neutrale Element bezüglich der Multiplikation, also 1, sein muss. Analog ist die Transferfunktion einer Konstanten bis auf einen Faktor die Deltafunktion (RICHARDS & JIA 2006: 170). Anschaulich wird also das komplette Spektrum benötigt, um einen Impuls zu erzeugen während eine Konstante nur die Nullfrequenz enthält. Im Sinne der Unschärferelation ist das Signal im Ortsraum genau lokalisiert und im Frequenzraum gar nicht oder umgekehrt.

Offensichtlich ebenfalls nur aus einer Frequenz besteht der Kosinus. Seine Fouriertransformierte enthält aber noch die negative Frequenz und setzt sich daher aus zwei Delta-Funktionen zusammen.

Die Fouriertransformierte einer Rechteckfunktion ist die Sinus cardinalis Funktion  $\text{sinc}(k) = \frac{\sin(k)}{k}$ .

Das Paar ist in Abb. 2.19 zu sehen.



**Abb. 2.19: Rechteckfunktion, Spektrum der sinc-Transferfunktion als Plot und Bild**

Die Beschneidung eines Bildes durch Multiplikation mit einer Rechteckfunktion bewirkt also eine Veränderung des Spektrums, die im Mittel nur mit  $1/k$  abfällt. Peaks im Spektrum werden dadurch zu sternähnlichen Mustern (JÄHNE 2005: 287). Der Grund hierfür ist die Unstetigkeit der Rechteckfunktion. Daher kann es sinnvoll sein, ein Bild vor der Fouriertransformation mit einer anderen Fensterfunktion zu multiplizieren, deren Transferfunktion bessere Eigenschaften hat. Nach JÄHNE (2005: 288) ist dafür beispielsweise ein Kosinusfenster geeignet:  $W_{mn} = \sin\left(\pi \frac{m}{M}\right) \sin\left(\pi \frac{n}{N}\right)$ .

Eine besondere Stellung nimmt die Gaußfunktion ein, da sie invariant gegenüber der FT ist, d.h. die Transformierte einer Normalverteilung ist wieder eine Normalverteilung, die Varianzen unterliegen dabei der Unschärferelation. Folglich ist die Gaußfunktion auch kein idealer *low-pass* Filter, sondern vielmehr ein Kompromiss zwischen Lokalisation im Ortsraum und Frequenzraum. Ein idealer *low-pass* Filter entspräche der Multiplikation mit einer Kreismaske im Frequenzraum. Dafür wäre die entsprechende Faltungsmaske im Ortsraum nur sehr ungenau lokalisiert, um große Verluste zu vermeiden, müsste also eine sehr große Maske gewählt werden. Das Design von Filtern ist folglich ein komplexes Optimierungsproblem (JÄHNE 2005: 319). Kleine Filtermasken haben schlechte Transferfunktionen, ideale Filter benötigen zu große Masken.

### 2.3.8 Mathematische Morphologie

Die mathematische Morphologie kommt zwar ursprünglich aus der Mengenlehre, der Begriff deutet aber schon an, dass sie die Analyse von Formen und damit räumlichen Strukturen zum Ziel hat (SOILLE 1998: 1). Mit Hilfe der Morphologie können die Gestalt und Form von Objekten analysiert und verändert oder relevante Bildstrukturen extrahiert werden. Morphologische Operationen sind Nachbarschaftsoperationen und im Allgemeinen nicht linear.

In der Morphologie werden Bilder als Mengen von Koordinaten betrachtet. Daher wurde sie zunächst für binäre Bilder entwickelt, mittlerweile aber auf Grauwertbilder erweitert (ANDERSEN *et al.* 2001: 15). Im binären Fall sind die Mengen der schwarzen und weißen (auch Hintergrund und Vordergrund) Pixel in einem Bild duale Mengen, d.h. jedes Element gehört zu genau einer der beiden. Grauwertbilder werden als dreidimensionale Höhenmodelle betrachtet, was die begrifflichen Ähnlichkeiten mit der Geomorphologie erklärt. Die Menge aller Punkte der Oberfläche des Höhenmodells wird dabei als Graph des Bildes bezeichnet, der Graph und alle Punkte unter ihm als Untergraph (SOILLE 1998: 19). ANDERSEN (2001: 15) bezeichnet die Mengen als *top surface* und *umbra*.

Um relevante Strukturen extrahieren zu können, wird das Bild mit bestimmten Formelementen, die aus *a priori* Wissen über die relevanten Strukturen gewonnen werden, geprobt. Diese Formelemente werden meistens als Strukturelemente (*structuring elements*) bezeichnet, manchmal auch als Masken. Die Strukturelemente (SE) ähneln dabei in ihrer Form oft den Objekten von Interesse und haben jeweils einen Bezugspunkt. Die SE können maximal so viele Dimensionen haben, wie das Bild (für Grauwertbilder also drei). Allerdings wird die Verwendung von nicht-flachen, also nicht-binären, Strukturelementen meist nicht empfohlen (SOILLE 1998: 53). Das SE wird beim Proben quasi über das ganze Bild verschoben und damit verglichen.

### Erosion, Dilatation, Öffnung und Schließung

Die Basisoperationen der Morphologie sind die Erosion und die Dilatation. Das erodierte Bild ist anschaulich die Menge aller Punkte, mit denen als Bezugspunkt das SE vollständig in die Ausgangsmenge passt. Mathematisch ausgedrückt ist die Erosion einer Menge  $X$  mit einem Strukturelement  $B$  dessen Bezugspunkt sich am Punkt  $x$  befindet:

$$(2.45) \quad \varepsilon_B(X) = \{x \mid B_x \subseteq X\}.$$

Für ein Grauwertbild gilt also: Der Wert eines erodierten Pixels ist der kleinste Wert im Fenster des Strukturelements, wenn dessen Bezugspunkt an dem entsprechenden Pixel ist (SOILLE 1998: 54):

$$[\varepsilon_B(f)](x) = \min_{b \in B} f(x+b).$$

Die der Erosion entgegen gesetzte Operation ist die Dilatation. Die dilatierte Menge ist die Menge aller Punkte, an denen das SE die Ausgangsmenge berührt, also:

$$(2.46) \quad \delta_B(X) = \{x \mid B_x \cap X \neq \emptyset\}$$

Der dilatierte Wert an einem Punkt ist also der größte Wert in dem Fenster der SE.

$$[\delta_B(f)](x) = \max_{b \in B} f(x+b).$$

Beide Gleichungen können einfach für nichtflache Strukturelemente erweitert werden. In Operator-schreibweise werden Erosion und Dilatation häufig als  $X \ominus B$  bzw.  $X \oplus B$  geschrieben. Sie sind duale Transformationen in Hinblick auf die Komplementierung (Negativtransformation), d.h. die Komplementierung der Dilatation der Komplementierung eines Bildes entspricht der Erosion und umgekehrt.

Bei binären Bildern entfernt die Erosion mit den Nachbarschaftsmasken

$$(2.47) \quad N_4 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{pmatrix} \text{ bzw. } N_8 = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

für 4N bzw. 8N den Rand des Bildes, die Dilatation entfernt den Rand des Hintergrundes. Da die Dilatation eines Bildes jeweils den größten und die Erosion den kleinsten Wert in der Maske annimmt, ist die Differenz von beiden ein Maß für die Veränderung im Bild. Diese wird daher auch als elementarer morphologischer Gradient bezeichnet (SOILLE 1998: 72):  $\rho(f) = \delta_B(f) - \varepsilon_B(f)$ .

Durch die Erosion eines Bildes verschwinden alle Strukturen, die kleiner sind als das SE. Gleichzeitig werden aber auch alle größeren Strukturen geschrumpft. (SOILLE 1998: 96) Dies kann jedoch einfach rückgängig gemacht werden, indem anschließend mit demselben SE wieder dilatiert wird.

Diese kombinierte Operation wird morphologische Öffnung (*opening*) genannt:

$$(2.48) \quad \gamma_B(f) = \delta_B[\varepsilon_B(f)]$$

Manchmal wird sie auch über eine Dilatation mit dem gespiegelten Strukturelement definiert, da in der Praxis aber meist symmetrische SE zum Einsatz kommen, ist dieser Unterschied von begrenzter Bedeutung. Die Öffnung eliminiert Strukturen, in die das SE nicht hineinpasst, erhält jedoch Form und Größe aller anderen Objekte.

Die der Öffnung entgegengesetzte Operation ist die morphologische Schließung (*closing*). Diese besteht aus einer Dilatation und einer anschließenden Erosion mit demselben (oder gespiegelten) SE

$$(2.49) \quad \phi_B(f) = \varepsilon_B[\delta_B(f)].$$

Die Schließung schließt Löcher, die vollständig durch das Strukturelement abgedeckt werden und erhält andere Strukturen, vergrößert sie also insbesondere nicht wie die Dilatation.

Öffnung und Schließung sind idempotente (siehe 2.3.2) Operatoren. Sie sind ebenfalls dual, d.h. eine Öffnung bewirkt das gleiche wie eine Schließung auf dem Komplement und umgekehrt. In Operator-schreibweise werden sie als  $X \circ B = (X \ominus B) \oplus B$  bzw.  $X \bullet B = (X \oplus B) \ominus B$  geschrieben.

Für die durch morphologische Operationen erzeugten Bilder gilt folgende Anordnung der Größe nach:  $\varepsilon_B \leq \gamma_B \leq f \leq \phi_B \leq \delta_B$  bzw.  $\varepsilon_B \subseteq \gamma_B \subseteq X \subseteq \phi_B \subseteq \delta_B$  für die Mengen.

Operationen, die alle Eigenschaften von Öffnungen bzw. Schließungen erfüllen, jedoch nicht durch Erosion und Dilatation mit einem SE erzeugt werden können, sind algebraische Öffnungen oder Schließungen. Es kann gezeigt werden, dass sie immer als Vereinigung von morphologischen Öffnungen definiert werden können (SOILLE, 1998: 103). Ein Beispiel für eine algebraische Schließung ist die konvexe Hülle.

### Weitere morphologische Operationen

Um ein passendes SE für morphologische Filter zu wählen, wird *a priori* Wissen über die Form der irrelevanten Bildstrukturen benötigt, die herausgefiltert werden sollen. Oftmals dürfte aber eher Wissen über die relevanten Bildstrukturen vorliegen. Daher werden durch die Zylinderhut-Transformationen relevante Bereiche ausgefiltert und anschließend durch Subtraktion vom Originalbild wieder hergestellt.

Die Differenz zwischen einem Bild und seiner Öffnung wird weißer Zylinderhut (*top-hat*) genannt, die zwischen seiner Schließung und dem Bild schwarzen Zylinderhut (*bottom-hat*). Wie der Name vermuten lässt, findet erstere helle Strukturen letztere dunkle. Durch die Kombination von beiden lässt sich beispielsweise der Kontrast verstärken, indem die hellen Strukturen addiert und die dunklen abgezogen werden.

Die Hit-Miss (oder auch Alles oder Nichts-) Transformation erlaubt auf Binärbildern die Extraktion von Objekten einer bestimmten Form bzw. Vordergrund-Hintergrund- Konfiguration. Dafür werden zwei Strukturelemente  $B_1$  und  $B_2$  mit gemeinsamem Bezugspunkt benötigt und es wird für jedes Pixel geprüft, ob das eine SE in den Vordergrund und das andere in den Hintergrund (also das Komplement  $X^c$ ) passt (daher der Name *hit and miss*):  $ANT_B(X) = \varepsilon_{B_1}(X) \cap \varepsilon_{B_2}(X^c) = (X \ominus B_1) \cap (X^c \ominus B_2)$ .

Eine andere wichtige Klasse von morphologischen Operationen sind Skelette, die Objekte zu idealisierten Linien verdünnen, dabei aber gewisse topologische Eigenschaften erhalten (siehe dazu SOILLE 1998: 142 ff).

### Distanztransformation

Distanztransformationen weisen jedem Pixel des Vordergrundes eine Distanz zum Hintergrund zu. Bei Grauwertbildern geben sie die Distanz zum nächsten nullwertigen Pixel an

$$(2.50) \quad [D(f)](x) = \min\{d(x, y) \mid f(y) = 0\}.$$

Das Ergebnis der Distanztransformation hängt natürlich entscheidend von der verwendeten Metrik ab. Gebräuchliche Abstandsmaße sind in der Bildverarbeitung nach SOILLE (1998: 37) die Cityblock-Metrik:

$$(2.51) \quad d_4[(x_1, y_1), (x_2, y_2)] = |x_2 - x_1| + |y_2 - y_1|,$$

die Schachbrett-Metrik:

$$(2.52) \quad d_8[(x_1, y_1), (x_2, y_2)] = \max\{|x_2 - x_1|, |y_2 - y_1|\},$$

und die euklidische Distanz:

$$(2.53) \quad d_e[(x_1, y_1), (x_2, y_2)] = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}.$$

Die ersten beiden entsprechen dabei den Konzepten von 4- und 8-Nachbarschaft und können als Summe von sukzessiven Erosionen mit den Nachbarschaftsmasken  $N_4$  und  $N_8$  berechnet werden. Ihr großer Nachteil gegenüber der euklidischen Distanz ist aber, dass sie nicht isotrop sind.

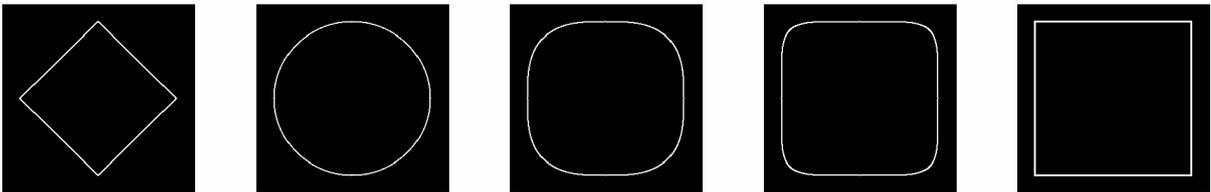


Abb. 2.20: Abstandsmaße: Einheitskreis in verschiedenen Normen: L1 (cityblock), L2 (euklidisch), L3, L6 und L∞ (chessboard)

### 2.3.9 Multiskalenrepräsentationen

#### Maßstabsräume

Ein wesentliches Problem in der Bildverarbeitung ist der Maßstab. Die meisten Konzepte sind nur in einer bestimmten Größenordnung sinnvoll. So wird ein Baum im Bereich von einigen Metern bis zu einigen zehn Metern lokalisiert sein, im Zentimeter- und Kilometerbereich werden dagegen eher Blätter oder ein Wald zu finden sein. Aber selbst das Konzept Baum variiert in seiner Größe in einem gewissen Bereich.

Um Signale auf verschiedenen Skalen handhabbar zu machen, schlägt LINDBERG (1994: 10) einen linearen Maßstabsraum (LMR, auch Skalenraum, *scale space*) als Multiskalenrepräsentation vor. Dieser ist eine von einem Parameter  $t \in \mathbb{R}^+$  abhängige Familie von abgeleiteten Signalen, bei der die feineren Details sukzessive unterdrückt werden. Per Definition soll der Maßstabsraum für  $t = 0$

identisch mit dem ursprünglichen Signal sein. Mit dem Skalenparameter  $t$  als zusätzlicher Achse hat der Maßstabsraum eine Dimension mehr als das ursprüngliche Signal.

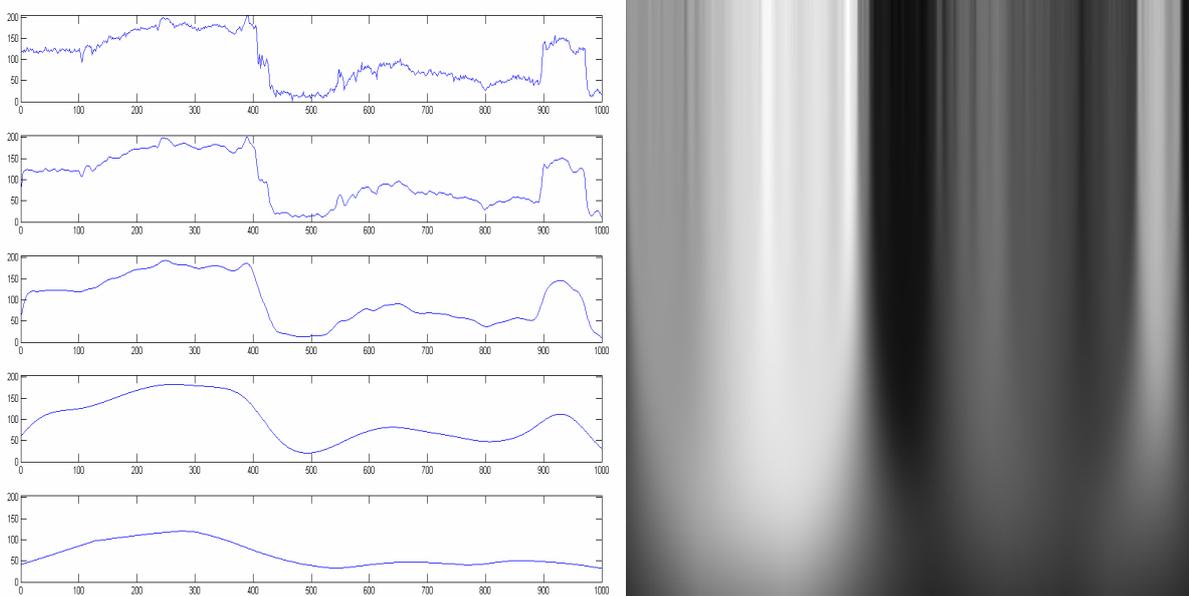
Ein solcher Raum kann durch Faltung mit Gaußkernen steigender Breite (siehe 2.3.5) erzeugt werden.

$$(2.54) \quad L(\mathbf{x}, t) = g(\mathbf{x}, t) * f(\mathbf{x}) ,$$

$\mathbf{x}$  ist dabei ein beliebiger Vektor im  $\mathbb{R}^N$ . Der Gaußkern ist im  $n$ -dimensionalen Fall

$$(2.55) \quad g(\mathbf{x}, t) = \frac{1}{(2\pi t)^{N/2}} \exp\left(\frac{-\mathbf{x}^t \mathbf{x}}{2t}\right).$$

Der Parameter  $t$  entspricht also der Varianz  $\sigma^2$  der Normalverteilung.



**Abb. 2.21:** Links: 1D-Signal (Bildzeile) im Maßstabsraum für  $t = 0, 2, 8, 32$  und  $128$ . Rechts: dasselbe Signal im ‚kontinuierlichen‘ Maßstabsraum (logarithmisch skaliert)

Es kann gezeigt werden, dass diese Funktionen  $L$  (die einzigen) Lösungen der Diffusions-Differentialgleichung sind (LINDBERG 1994: 11, JÄHNE 2005: 153).

$$(2.56) \quad \frac{\partial}{\partial t} L(\mathbf{x}, t) = \frac{1}{2} \nabla^2 L(\mathbf{x}, t)$$

Diese Gleichung wird auch Fouriersche Differentialgleichung der räumlichen Wärmeleitung genannt und bietet damit eine physikalische Veranschaulichung für den Maßstabsraum. Er entspricht der Entwicklung einer ursprünglichen Wärmeverteilung mit der Zeit  $t$ : die Wärme verteilt sich gleichmäßig über den Raum, die Entropie nimmt zu und der Informationsgehalt ab.

Darin zeigt sich auch eine wichtige Eigenschaft des Maßstabsraums, die Kausalität: Alle Merkmale im groben Maßstab müssen eine Ursache im feinen besitzen (STRAUB 2003A: 13). Da der Gaußkern homogen und isotrop ist, hat der Skalenraum ebenfalls diese Eigenschaften.

## Bildpyramiden

Der vorgestellte Skalenraum hat die Vorteile, dass der Skalenparameter frei wählbar ist und alle Bilder dieselbe räumliche Auflösung haben. Entsprechend hat er aber auch den Nachteil, dass der Rechenaufwand, um Operationen in den größeren Maßstäben durchzuführen, ebenfalls gleich bleibt. Daher ist es manchmal wünschenswert, dass die Bilder mit geringerem Informationsgehalt auch entsprechend kleiner sind.

Ein älteres Konzept, das dem Rechnung trägt, ist das der Bildpyramiden. Diese bestehen aus einer Menge von Bildern mit unterschiedlicher räumlicher Auflösung und gleichem Inhalt, die entsprechend der Auflösung in Ebenen angeordnet werden (PINZ 1994: 74). Dabei ist der Reduktionsfaktor zwischen zwei Ebenen jeweils gleich und das kleinste Bild besteht nur aus einem Bildpunkt. Die gesamte Pyramide benötigt beim üblichen Reduktionsfaktor von vier nur ein Drittel zusätzlichen Speicherplatz und eignet sich bestens zur parallelen Verarbeitung.

Da bei der Reduktion der Auflösung hohe Frequenzen Aliasing-Effekte erzeugen würden (siehe Abb. 2.12), müssen beim *resampling* gleichzeitig hohe Frequenzen ausgefiltert werden. Gemäß dem „Klassiker“ unter den *low pass*-Filtern wird die entsprechende Pyramide Gaußpyramide genannt. (JÄHNE 2005: 147, PINZ 1994: 77) und entspricht einer Reihe von tiefpass-gefilterten Bildern, bei denen die Grenzwellenlänge (bei einem Reduktionsfaktor von 4 um die Hälfte, also eine Oktave) abnimmt (JÄHNE 2005: 148). Die Differenz zwischen je zwei Ebenen (wozu natürlich das kleinere Bild wieder expandiert werden muss), ergibt eine Reihe bandpassgefilterter Bilder, die Laplacepyramide.

Bei dieser Bandpasszerlegung enthält jedes Bild nur noch einen bestimmten Frequenzbereich (wieder etwa eine Oktave) und hat eine Auflösung, die für diese Frequenzen ausreichend ist. Daher stellen Laplacepyramiden eine effiziente Datenstruktur dar und eignen sich zur Bildkompression (PINZ 1994: 78). Der Preis dafür ist ein höherer Rechenaufwand, da zur Wiederherstellung des Originalbildes rekursiv alle Ebenen expandiert und aufsummiert werden müssen.

## 2.4 Segmentierung

Die Segmentierung (*segmentation*) ist die Einteilung des Bildes in homogene Regionen (LEHMANN *et al.* 1997: 360). Sie ist bereits ein Teil der Bildanalyse, da sie eine erste Interpretation bietet, und wird auch auf Grund ihrer herausragenden Stellung für diese Arbeit als eigener Abschnitt geführt. Auch wenn die eigentliche Bildinterpretation als nächster Schritt folgen muss, geht bei der Segmentierung oft schon *a priori* Wissen über die Größe und Beschaffenheit der gesuchten Objekte ein. Im Sinne der Mustererkennung (siehe 2.2) ist sie die Suche nach den charakteristischen Stellen im Signal, also der Übergang zu einfachen Mustern (PAULUS & HORNEGGER 2003: 39).

Nach BÖHNER *et al.* (2006: 29) sind anspruchsvolle Segmentierungsverfahren notwendig, um semantische Informationen aus Bildern zu erhalten. Die grundlegende Idee ist dabei, das Bild vor der weiteren Klassifikation oder Verfeinerung zunächst in homogene Regionen einzuteilen, die Objekte oder Segmente genannt werden (BÖHNER *et al.* 2006: 30). Es geht also darum, das Bild in sinnvolle Bestandteile zu zerlegen. Was dabei sinnvoll ist, hängt in hohem Maße von der Aufgabenstellung ab. Nach BÖHM (2005: 38) dient die Segmentierung im Kontext der Objekterkennung hauptsächlich der Merkmalsextraktion. Die erzeugten Regionen sollten also möglichst genau mit den Objekten im Bild übereinstimmen.

Im Allgemeinen werden nach der Herangehensweise drei große Klassen von Segmentierungsverfahren unterschieden: pixelbasierte, kantenbasierte und regionenbasierte. Da letztere in der Fernerkundung dominieren (BÖHNER *et al.* 2006: 30), sollen sie hier den Schwerpunkt einnehmen.

#### 2.4.1 Pixel- und kantenbasierte Verfahren

Pixelbasierte Verfahren sind Punktoperationen im Sinne von 2.3.2, betrachten also nur den Grauwert eines einzelnen Pixels und nicht seine Nachbarschaft. Das einfachste Verfahren dieser Art ist der globale Schwellenwert (*threshold*). Dabei werden alle Pixel mit Grauwerten über dem Schwellenwert  $t$  dem Vordergrund und alle anderen dem Hintergrund zugeordnet.

$$(2.57) \quad f(g) = \begin{cases} 0, & \text{für } g \leq t \\ 1, & \text{für } g > t \end{cases}$$

Im Idealfall kann der optimale Schwellenwert aus dem Histogramm ermittelt werden. Besteht das Histogramm aus zwei Peaks mit einem ausgeprägten Minimum dazwischen, wird von einer bimodalen Verteilung gesprochen und davon ausgegangen, dass Vordergrund und Hintergrund jeweils einem Peak entsprechen. Der optimale Schwellenwert ist dann das Minimum der Grauwertverteilung. Dieses Verfahren versagt allerdings, wenn der Hintergrund inhomogen ist, beispielsweise bei ungleichmäßiger Beleuchtung (JÄHNE 2005: 474). Eine Möglichkeit dies auszugleichen sind dynamische oder lokale Schwellenwertverfahren (HABERÄCKER 1989: 238, LEHMANN *et al.* 1997: 365).

Werden die Grauwerte der Pixel nicht nur als Ortsfunktion sondern als Vektor in einem beliebig-dimensionalen Merkmalsraum betrachtet, so wird deutlich, dass die in Abschnitt 2.2.5 vorgestellte Klassifikation nichts als eine Verallgemeinerung der pixelbasierten Segmentierung ist.

Kantenbasierte Verfahren (*edge detection*) beruhen auf der Annahme, dass Grauwertkanten im Bild Grenzen zwischen Objekten sind. Gegenüber der Segmentierung mittels Schwellenwert haben sie den Vorteil, dass die Größe der Objekte nicht von ihrer Helligkeit verfälscht wird. Zur Kantenfindung können die in Operatoren in 2.3.5 verwendet werden. Die kantenbasierten Verfahren beruhen im Allgemeinen darauf, gewonnene „Kantenstückchen“ durch Konturverfolgung zu ganzen Objekten zu vervollständigen. Darauf soll hier aber nicht näher eingegangen werden.

#### 2.4.2 Regionenbasierte Segmentierung

Regionenbasierte Verfahren erweitern pixelorientierte um den entscheidenden Aspekt, dass Objekte im Bild zusammenhängend sind. Neben ihrem Grauwert oder allgemeiner der Position im Merkmalsraum ist also auch die relative räumliche Anordnung der Pixel von Bedeutung. Daher wird versucht, homogene Regionen im Bild zu identifizieren.

Bei der gängigen formalen Definition für eine vollständige Segmentierung (vgl. GONZALES & WOODS 1993: 458, BÖHM 2005: 38, HOOVER *et al.* 1996) bilden alle Punkte des Bildes die Region  $R$ , die in Teilregionen  $R_i$  zerlegt werden soll.  $H$  sei ein Homogenitätskriterium über eine Menge von Pixeln.

$$\begin{aligned}
 (2.58) \quad & \bigcup_i R_i = R \\
 & R_i \text{ ist zusammenhängend } \forall i \\
 & R_i \cap R_j = \emptyset \text{ für } i \neq j \\
 & H(R_i) = \text{WAHR } \forall i \\
 & H(R_i \cup R_j) = \text{FALSCH falls } R_i, R_j \text{ benachbart}
 \end{aligned}$$

Die erste Bedingung fordert gerade die Vollständigkeit der Segmentierung, die dritte die Überdeckungsfreiheit – jedes Pixel muss also genau einer Region angehören. Die Regionen müssen weiterhin zusammenhängend sein und das Homogenitätskriterium erfüllen, die Vereinigung von zwei angrenzenden Regionen darf das Homogenitätskriterium aber nicht erfüllen. Da nicht notwendigerweise jedes Pixel einem Objekt angehört, kann es sinnvoll sein, die Vollständigkeit dadurch zu erreichen, dass solche Pixel einem Hintergrund oder einer objektlosen Menge zugewiesen werde. Diese ist dann aber nicht notwendigerweise zusammenhängend und homogen im Sinne der obigen Definition.

Wichtig ist weiterhin festzustellen, dass das Ergebnis einer Segmentierung nicht eindeutig ist. Bei einem gegebenen Bild und Homogenitätskriterium kann es (und wird es im Allgemeinen auch) verschiedene Segmentierungen geben, die alle Bedingungen erfüllen. Es besteht also die zweifache Herausforderung einerseits ein für die Aufgabenstellung geeignetes Homogenitätskriterium zu finden und andererseits einen Algorithmus, der möglichst die beste Lösung zum gegebenen Kriterium findet.

Gerade bei den Algorithmen besteht aber eine sehr große Vielfalt. HOOVER *et al.* haben bereits 1996 siebzehn verschiedene verglichen, seitdem sind natürlich viele weitere hinzu gekommen. Da eine vollständige Übersicht demnach nahezu unmöglich wäre, sollen hier nur die Grundzüge von relevanten Gruppen erläutert werden. Konkrete Algorithmen, die zur Baumextraktion verwendet werden, finden sich in Kapitel 3.

### 2.4.3 Split-and-merge, regionsplitting und pyramid-linking

Bei *split-and-merge* Algorithmen wird zunächst eine Übersegmentierung erzeugt. Diese muss bereits die ersten vier Bedingungen in (2.58) erfüllen, nicht aber die Maximalitätsbedingung. Im zweiten Schritt werden dann so lange benachbarte Regionen verschmolzen, die gemeinsam das Homogenitätskriterium erfüllen, bis keine solchen Paare mehr vorhanden sind. Damit ist auch die fünfte Bedingung erfüllt.

Eine Übersegmentierung entsteht beispielsweise, indem jedes Pixel als eigene Region definiert wird. Alternativ kann vom Gesamtbild ausgehend so lange weiter unterteilt werden, wie das interne Homogenitätskriterium nicht erfüllt wird. Diese Vorgehensweise wird *region splitting* genannt. Werden die inhomogenen Bildregionen jeweils in Quadranten geteilt, wird von *quadtrees* gesprochen (GONZALEZ & WOODS 1993: 461).

Das *pyramid-linking*-Verfahren eignet sich zum Segmentieren verrauschter Bilder. Zugrunde liegt eine Gauß-Pyramide. Werden verrauschte Bilder vor dem Segmentieren geglättet, werden damit natürlich auch die Grenzen zwischen den Objekten unschärfer. Es wäre also wünschenswert, beim Glätten nur diejenigen Pixel mit einzubeziehen, die demselben Objekt angehören, was wiederum erst nach der Segmentierung bekannt ist. Die Lösung ist ein rekursiver Algorithmus, bei dem das Ergebnis einer vorläufigen Segmentierung an den Glättungsoperator zurückgegeben wird. Mit diesem Zwischener-

gebnis werden im zweiten Rekursionsschritt eine bessere Glättung und eine bessere Segmentierung erzeugt. So wird weiter verfahren, bis ein stabiler Zustand erreicht ist, und der Algorithmus abbricht.

#### 2.4.4 Region Growing

Region-Growing-Verfahren (RG) ähneln der zweiten Phase des *split-and-merge*-Verfahrens. Der Ausgangspunkt ist aber kein übersegmentiertes Bild sondern eine Liste von Punkten, die Saatpunkte (*seeds*) genannt werden. Den Saatpunkten werden sukzessive Nachbarpunkte hinzugefügt, solange das interne Homogenitätskriterium erfüllt bleibt. Dabei steuert die Kontrollstruktur, in welcher Reihenfolge Saatpunkte und angrenzende Regionen bearbeitet werden. Die einfachste Möglichkeit ist, Saatpunkte und Nachbarpixel jeweils der Reihe nach abzuarbeiten, was aber kein optimales Ergebnis produziert. Da das Homogenitätskriterium üblicherweise eine Metrik beinhaltet, ist es sinnvoller, immer das Paar aus Region und Nachbarpixel zu vereinigen, das den geringsten Abstand bezüglich dieser Metrik hat.

Neben der Metrik werden RG-Verfahren anhand ihrer *linkage*-Strategie charakterisiert, also den Entscheidungskriterien wann zwei Regionen verbunden werden. Beim *single-linkage* werden beispielsweise nur einzelne Pixelpaare betrachtet, während beim *centroid-linkage* der Abstand zum Mittelwert aller Pixel der Regionen gemessen wird. Eine vollständige Aufstellung von *linkage*-Strategien findet sich in LEHMANN (1997: 375).

Ein Nachteil von RG ist, dass die erzeugte Segmentierung nicht notwendigerweise vollständig ist, also die erste Bedingung in (2.58) erfüllt. Ist dies der Fall, werden weitere Saatpunkte benötigt. Da RG-Verfahren abhängig von der Wahl der Saatpunkte gegen verschiedene Segmentierungen konvergieren, hängt die Qualität der Segmentierung entscheidend von der Wahl der Saatpunkte ab.

#### 2.4.5 Morphologische Segmentierung: Wasserscheidentransformation

Bei der Wasserscheidentransformation (WST) wird das Bild wieder als Grauwertgebirge betrachtet, weshalb sie hier auch als morphologische Segmentierung bezeichnet wird. Das Verfahren ist sehr anschaulich zu erklären: Fällt ein Wassertropfen auf das Grauwertgebirge, so fließt er nach unten, bis er ein lokales Minimum erreicht hat. Das Einzugsgebiet eines Minimums sei sein Auffangbecken und die Grenzen zwischen benachbarten Becken Wasserscheiden (SOILLE 1998: 239). Oft wird die WST auch mit Hilfe sukzessiver Flutung des Bildes definiert, da der entsprechende Algorithmus so aufgebaut ist.

Die Wasserscheidentransformation gruppiert Pixel um die regionalen Minima. Es empfiehlt sich somit, das Bild derart zu transformieren, dass die Minima relevante Objekte markieren. Das so transformierte Bild wird Segmentierungsfunktion genannt (SOILLE 1998: 248). Wird die WST beispielsweise auf ein Gradientenbild angewendet, so markieren die Minima homogene Bereiche und die Wasserscheiden Kanten im Bild.

Eine weitere Schwierigkeit ist, dass die WST meist zu einer Übersegmentierung führt, die entweder durch nachträgliches Verschmelzen von Regionen (LEHMANN *et al.* 1997: 374) oder durch Vorfiltern der Segmentierungsfunktion (SOILLE 1998: 250) behoben werden muss. Ein Vorteil der WST gegenüber kantenbasierten Verfahren ist, dass sie geschlossene Linienzüge produziert (LEHMANN *et al.* 1997).

#### 2.4.6 Clustering und Segmentierung

Zum Schluss soll noch darauf auf den Zusammenhang zwischen den in 2.2.5 dargestellten Verfahren zur unüberwachten Klassifikation und Verfahren zur Segmentierung hingewiesen werden. Beide gruppieren die Pixel eines Bildes, aber auf unterschiedliche Weise: Beim *clustering* wird im Merkmalsraum gruppiert, bei der Segmentierung im Merkmals- und im Bildraum. Dabei wird eine überlappungsfreie und zusammenhängende Einteilung im Bildraum erzeugt, die jedoch nicht unbedingt einer überlappungsfreien Einteilung im Merkmalsraum entspricht. Umgekehrt wird beim *clustering* der Merkmalsraum überlappungsfrei eingeteilt, was jedoch meist nicht zu zusammenhängenden Regionen im Bildraum führt (BÖHM 2005: 42).

Ein solcher Zusammenhang kann aber durch nachverarbeitende Schritte nachträglich hergestellt werden, womit *clustering* Verfahren ebenfalls zur Segmentierung einsetzbar sind. Gleichzeitig können auch die Ergebnisse des *clusterings* verbessert werden, da benachbarte Pixel mit hoher Wahrscheinlichkeit derselben Klasse angehören. RICHARDS & JIA (2006: 209ff) führen mehrere solche Verfahren wie beispielsweise den ECHO-Klassifikator (*Extraktion and Classification of Homogenous Objects*) unter dem Stichwort Kontext-Klassifikation auf.

Es bleibt festzuhalten, dass sowohl in den Merkmalen (Grauwerten oder radiometrischen Daten) einerseits als auch in der räumlichen Anordnung andererseits, Informationen stecken. Ein geschickter Algorithmus nutzt also beide Domänen.

### 3 Objektextraktion von Bäumen aus Fernerkundungsdaten

In diesem Kapitel wird eine Übersicht über die verfügbaren Verfahren zur automatischen Baumextraktion und Bestimmung von Kronenbegrenzung und weiteren Parametern aus Fernerkundungsdaten gegeben. Dabei soll aber nicht nur der Stand der Forschung wiedergegeben werden, sondern insbesondere eine Bewertung der Verfahren durchgeführt und ihre Verwendbarkeit und Verallgemeinerbarkeit im Hinblick auf ein Expertensystem untersucht werden. Für ein solches System, das zu jeder gegebenen Situation und Datengrundlage automatisch das beste der jeweils verfügbaren Verfahren auswählt und kalibrieren soll, ist eine vorangehende Bestandsaufnahme und systematisch vergleichende Evaluation von Verfahren als vorbereitender Schritt unabdingbar. Dies soll im folgenden Kapitel geleistet werden.

Das Kapitel gliedert sich wie folgt: In 3.1 werden zunächst einige Grundlagen der Objektextraktion eingeführt, die über das allgemeine Lehrbuchwissen in Kapitel 2 hinausgehen, wobei für Bäume relevante Aspekte besonders berücksichtigt werden. Da ein großer Teil der bisherigen Arbeiten das Objekt Straße zum Gegenstand hat, nehmen Beispiele aus diesem Bereich auch einen gewissen Raum ein.

In den darauf folgenden Abschnitten werden dann einzelne Verfahren zur Extraktion von Bäumen und Parametern von Bäumen näher vorgestellt und einer kritischen Bewertung unterzogen. Da eine einheitliche Kategorisierung in der Literatur fehlt, wurde dabei auf eine übergeordnete Strukturierung verzichtet. Zur Erhöhung der Übersichtlichkeit und um eine gewisse Einteilung vorzunehmen, werden aber zu Beginn jedes Verfahrens die wichtigsten Eigenschaften tabellarisch aufgeführt.

Eine wichtige Unterscheidung ist die zwischen Einzelbaum-basierten Verfahren (*individual tree crown*, ITC) und Bestands-basierten (*stand-level*), die jeweils unterschiedliche Extraktionsparameter mit sich bringen. Hier werden vor allem ITC-Verfahren behandelt. Um ein künftiges Expertensystem möglichst universell zu halten, werden jedoch sowohl Verfahren, die auf Höhenmodellen arbeiten, berücksichtigt als auch solche, denen radiometrische Bilder zu Grunde liegen.

Die einzelnen Verfahren werden wie folgt untergliedert: Sofern diese explizit aufgeführt sind, werden die eingehenden Modellannahmen und Modelle vorgestellt. Anschließend werden das Verfahren selbst, also einzelne Algorithmen bzw. Methoden, sowie die Ergebnisse näher erläutert. Dabei wird entsprechend dem Fokus der Arbeit das größte Augenmerk auf die Segmentierungs-Algorithmen gelegt. Abschließend wird das Verfahren bewertet.

Die einzelnen Verfahren werden dabei unterschiedlich ausführlich behandelt. Einerseits sollen bei Überschneidungen überflüssige Wiederholungen vermieden werden, andererseits liegt dem aber teilweise auch schon eine Vorabbewertung hinsichtlich ihrer Verwendbarkeit zugrunde, so dass den vermeintlich nützlichsten Verfahren der größte Raum eingeräumt wird. Teilweise wurde auf die Darstellung von relevanten Verfahren verzichtet, wenn diese Sonderwege gehen oder als wenig verallgemeinerbar angesehen wurden. So wird beispielsweise der Ansatz von BRANDBERG & WALTER (1998) nicht diskutiert, da er als einziges Verfahren eine kantenbasierte Segmentierung vornimmt.

Vorab soll eine grobe Übersicht über bisherige Forschungsansätze und beteiligte Gruppen gegeben werden: Die ersten Arbeiten zur automatischen Extraktion von Bäumen entstanden bereits vor über 20 Jahren. Während in den frühen Arbeiten ausschließlich auf Farb- und Falschfarb-Luftbildern (etwa PINZ 1989, BRANDBERG & WALTER 1998, LARSEN & RUDEMO 1997) gearbeitet wurde, die in den

Neunziger Jahren durch Digitale optische Sensoren (*multispectral mapper*) ergänzt wurden (GOUGEON 1995), kamen seit Ende der Neunziger Jahre und im neuen Jahrtausend vorwiegend Laserscanner zum Einsatz. Voraussetzungen hierfür waren die Verfügbarkeit von genauem Differential-GPS und Entwicklung und Integration von INS (CAMPBELL 2002: 237, NÆSSET *et al.* 2004: 484), von denen die Photogrammetrie insgesamt erheblich profitiert hat.

Viele wichtigste Impulse bei der Objektextraktion von Bäumen stammen aus den skandinavischen Ländern. Dabei sind unterschiedliche Strategien zu identifizieren: Während sich die Finnen und Schweden bereits früh auf Einzelbaumverfahren konzentrierten (etwa HYYPPÄ & INKINEN 1999, PERSSON *et al.* 2002), die auch in dieser Arbeit im Vordergrund stehen, entwickelten die Norweger vorwiegend anwendungsorientierte Verfahren zur Bestimmung von Parametern ganzer Bestände aus Laserscanner-Daten (etwa NÆSSET 1997), die in Norwegen seit 2002 kommerziell angeboten werden. Eine sehr gute Übersicht über alle skandinavischen Arbeiten gibt NÆSSET *et al.* (2004).

Auch im deutschsprachigen Raum haben sich in den vergangenen zehn Jahren verschiedene Forschungseinrichtungen mit dem Problem beschäftigt. Hervorzuheben ist dabei vor allem die Universität Freiburg (DIEDERSHAGEN *et al.* 2003A, 2003B & 2004, HERRERA *et al.* 2004), die im Rahmen des europäischen HIGHSCAN Projektes auch vergleichende Studien mit den Finnen durchgeführt haben (HYYPPÄ *et al.* 2001). Ebenfalls im HIGHSCAN-Projekt involviert war das Joanneum Research in Graz (SCHARDT *et al.* 2002). Weiterhin zu nennen sind die FH Hildesheim/Holzwinden (KÄTSCH 2002, KRUMMHEUER 1998), die TU München (HEURICH *et al.* 2004, HEURICH & KENNEL 2007), die Universität Zürich (MORS DORF *et al.* 2004) und die Universität Trier (ATZBERGER & SCHLERF 2002). Daneben gibt es einige interessante Arbeiten von der Universität Hannover (STRAUB 2002, 2003A & 2004, STRAUB & HEIPKE 2001 & 2004), die jedoch mehr auf 3D-Stadtmodelle denn auf forstwirtschaftliche Anwendungen abzielen.

## 3.1 Grundlagen der Objektextraktion

### 3.1.1 Steuerung

Generell werden im Bildverstehen zwei Ansätze unterschieden, die als *Top-Down* und *Bottom-Up* bezeichnet werden. Diese Begriffe beziehen sich auf verschiedene Repräsentationsebenen eines Bildes, die auf DAVID MARR (1982) zurückgehen. Ganz unten steht dabei das Bild selbst, also die radiometrischen Daten. Daraus erstellt MARR zunächst die erste Skizze (*primal sketch*), die unterschiedliche Informationen enthalten kann. Noch eine Ebene höher ist die  $2\frac{1}{2}$ D-Repräsentation, mit der Konturen, Tiefeninformationen sowie Orientierung von Flächen und Kanten dargestellt werden. Dafür werden die sichtbaren Oberflächen mit Normalenvektoren versehen. Das 3D-Objektmodell ist dann eine volle 3D Repräsentation auch der verdeckten Teile. Es soll Zerlegung in Komponenten, beispielsweise Zylinder, enthalten (MARR 1982: 37, PINZ 1994: 113 ff).

Bei *Bottom-Up* Strategien wird von den niedrigen Repräsentationsebenen, also dem Bild, ausgegangen und versucht in verschiedenen Verarbeitungsschritten zu einer höheren Repräsentationsebene zu gelangen. Diese kann dann beispielsweise mit in einer Datenbank gespeichertem Wissen verglichen werden. *Top-Down* Strategien beginnen dagegen auf einer hohen Repräsentationsebene und versuchen ausgehend von einem bestimmten Modell, beispielsweise einem 3D Baummodell, einzelne Objekte im Bild zu finden. Daneben gibt es bidirektionale Ansätze, die beide Strategien zu verbinden versuchen. *Bottom-Up* und *Top-Down* Ansätze werden auch der *recognition school* bzw. *reconstruction school*

zugeordnet, die versuchen Bildobjekte zu beschreiben bzw. deren physikalischen Eigenschaften zu modellieren (PINZ 1994: 117). STRAUB (2003A, 28f) unterteilt die Verfahren zur Baumextraktion in modell-getriebene und daten-getriebene Ansätze, die mit *Top-Down* und *Bottom-Up* gleichgesetzt werden können.



**Abb. 3.1: Modellgetriebener Ansatz zur Baumextraktion nach STRAUB (2003A, S. 28)**

Die modell-getriebenen Ansätze erzeugen dabei ausgehend von einem 3D-Baummodell ein oder mehrere 2D Muster, die im Bild gesucht werden. Das Objektmodell wird also ins Bild projiziert und daraus eine Objekthypothese aufgestellt und bewertet (HINZ 2003: 12). Bei detaillierter Modellbildung ist die Bildverarbeitung vergleichsweise einfach (z.B. durch Autokorrelation).



**Abb. 3.2: Datengetriebener Ansatz zur Baumextraktion nach STRAUB (2003A, S. 28)**

Bei datengetriebenen Ansätzen werden zunächst Merkmale wie Kanten aus dem Bild extrahiert und erst in weiteren Verarbeitungsschritten zu höherwertigen Strukturen zusammengefasst, bis eine Zuordnung zum Modell möglich ist (HINZ 2003: 12). Teile des Modells verbergen sich hierbei in den komplexen Verarbeitungsregeln, die schwer zu steuern sein können. Dafür findet der Vergleich zwischen Modell und Bild auf einer hohen Repräsentationsebene statt und ist daher abstrakter.

Jedem Extraktionsverfahren liegen eine Strategie und ein Modell (bei datengetriebenen Verfahren teilweise implizit) zu Grunde: „Die Extraktion von Objektbeschreibungen aus Bildern basiert in der Regel auf einem mehr oder weniger detaillierten Modell des Objektes und einer Strategie zur Extraktion, Gruppierung und Bewertung von Merkmalen“ (STRAUB 2003A: 7). Das Modell bezieht sich dabei auf die Eigenschaften des Objektes im Bild, hebt also aussagekräftige Merkmale hervor und vereinfacht zugleich. Die Strategie ist der Plan zur Erkennung von Objekten in Bildern (MEYER 1998 nach STRAUB 2003A: 7), also die Abfolge von Arbeitsschritten, und wird im Folgenden als Verfahren bezeichnet.

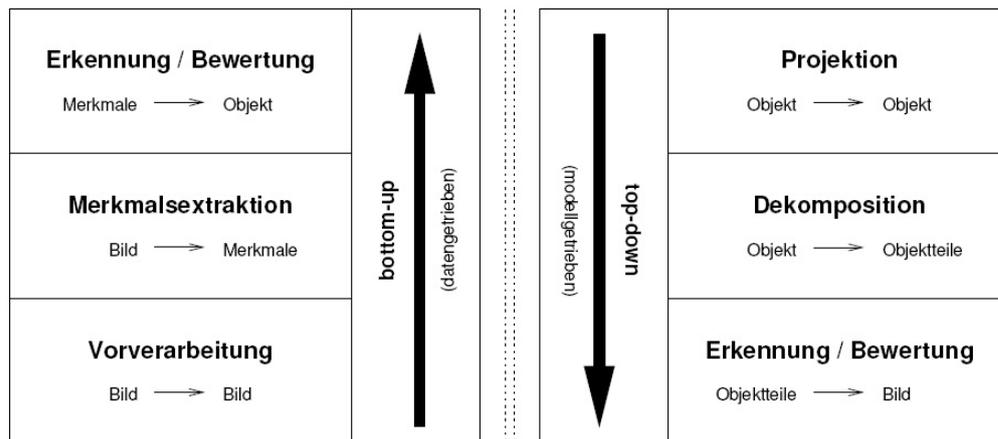


Abb. 3.3: Vorgehensweisen bei der Bildanalyse nach HINZ (2003: 12)

### 3.1.2 Modell

Die Modelle können aus unterschiedlichen Komponenten bestehen, wobei die geometrischen (also die Form), sich als die Wichtigsten herausgestellt haben. Daneben auch bedeutsam sind radiometrische Komponenten, die sich aus dem Material ergeben.

Die geometrischen Merkmale unterteilt STRAUB (2003A: 42) in Form, Grob- und Feinstruktur. Die Grobstruktur entspricht dabei der Struktur der Primäräste und die Feinstruktur Details innerhalb der Grobstruktur. In Luftbildern äußert sich die Struktur v.a. durch typische Texturen. Die Form ist die Kontur der Baumkrone, also quasi die Hülle des Baumes.

Generell besteht die Schwierigkeit, dass Grob- und Feinstruktur von Bäumen nur schwer zu beschreiben sind. Außerdem sind sie natürlich vom phänologischen Zustand (der Jahreszeit) abhängig. In der Forstwirtschaft ist daher die Verwendung von Skizzen und Beispielbildern zur Baumbestimmung üblich. Werden die feineren Strukturen aber vernachlässigt, so ist die Form vergleichsweise einfach anzunähern. Sehr häufig wird dafür das Pollock-Modell verwendet, das in 3.2 vorgestellt wird.

Weitere mögliche Modellkomponenten sind Funktion und Kontext, die vorwiegend dann relevant sind, wenn weitere Objekte zur Extraktion von Bäumen herangezogen werden oder umgekehrt. So wird häufig auf die Bedeutung von Baumgruppen bei der Extraktion von Straßen hingewiesen (BAUMGARTNER *et al.* 1997: 300, HINZ 2003: 16f, BUTENUTH 2003: 15), die etwa durch Schattenwurf oder Verdeckung das erwartete Erscheinungsbild der Straße verändern können. Dies wird als lokaler Kontext bezeichnet und bezieht sich auf räumlich beschränkte Relationen zwischen benachbarten Objekten (BAUMGARTNER *et al.* 1997: 300, HINZ 2003: 11) einer Größenordnung oder „*Abstraktionsebene*“ (STRAUB 2003A: 51).

Davon wird der globale Kontext abgegrenzt. Dieser modelliert spezielle Attribute und Aspekte, die von Objekten in größeren Regionen abhängen, also räumlich topologischen Relationen auf einer höheren Abstraktionsebene (BAUMGARTNER *et al.* 1997: 300, BUTENUTH 2003: 13, STRAUB 2003A: 51). So unterscheidet sich etwa die erwartete Erscheinungsform von Straßen in Wald, Stadt oder offener Landschaft, die durch globale Kontextmodelle beschrieben werden können (BAUMGARTNER *et al.* 1997: 302).

BUTENUTH *et al.* (2003: 169f) benutzen die Erkennung von Baumreihen in offener Landschaft, um Lücken im extrahierten Straßennetz zu schließen. Zur Erkennung benutzen sie Texturparameter, die

sie aus Bäumen im globalen Kontext Wald gewinnen. Umgekehrt könnte natürlich auch das Wissen um die Existenz einer Allee, etwa aus einem GIS, zur Baumerkennung genutzt werden.

Oftmals wird in den Ansätzen kein explizites Modell entworfen, sondern nur von relativ unspezifischen Annahmen ausgegangen, etwa der Art: jeder Baum bildet ein Helligkeitsmaximum oder zwischen zwei Bäumen ist ein (Helligkeits-)Tal. Problematisch ist dies insbesondere dadurch, dass meistens nicht gekennzeichnet wird, welche Annahmen in welche Verarbeitungsschritte eingehen. Daher werden im Folgenden eingehende Annahmen gekennzeichnet und in Kapitel 4 werden die implementierten Algorithmen genau auf implizite Voraussetzungen hin untersucht.

### 3.1.3 Maßstab

Eine weitere Schwierigkeit, die alle Verfahren in der einen oder anderen Art adressieren müssen, ist die des Maßstabs. Das grundlegende Problem besteht darin, dass die Größe der gesuchten Objekte erst nach der Extraktion bekannt ist, für die Extraktion selbst aber eigentlich als Parameter benötigt wird. STRAUB spricht darum von einem „*chicken-and-egg problem*“ (2004: 42), den optimalen Maßstab zu finden. Dieser zeichnet sich dadurch aus, dass die Grob- und Feinstrukturen gerade unterdrückt werden, die Form aber erkennbar bleibt (STRAUB 2003A, 49).

In der Literatur werden unterschiedliche Lösungen für das Problem vorgeschlagen: DIEDERSHAGEN *et al.* (2003A) führen beispielsweise vorab eine grobe Segmentierung in junge und alte Waldbestände durch, für die anschließend unterschiedliche Glättungsfilter und damit Skalenparameter zur Anwendung kommen. PERSSON *et al.* (2002) segmentieren dagegen in drei unterschiedlichen Maßstabsebenen und entscheiden anschließend durch beste Anpassung an Paraboloiden, welche Maßstabsebene lokal am besten geeignet ist. In STRAUB (2003a, 2003b, 2004) wird sogar auf zahlreichen Ebenen segmentiert und anhand des Modells bewertet, welche Baumhypothesen im Maßstabsraum am besten geeignet sind, wobei eine gewisse Überlappung toleriert wird. ANDERSEN *et al.* (2001) testen in einer Art morphologischem Maßstabsraum verschiedene Größen von Strukturelementen. Dabei schlagen sie zwar vor, die Ergebnisse verschiedener Maßstäbe zu addieren, geben aber keine Antwort auf die Frage, wie anschließend die besten Hypothesen ausgewählt werden sollen. Oftmals wird dem Problem auch komplett aus dem Weg gegangen, indem im Voraus manuell ein bestimmter Glättungsfilter ausgewählt wird, so etwa bei vielen *Template-matching* Verfahren.

### 3.1.4 Bewertung der Ergebnisse

Leider hat sich bis heute noch kein einheitlicher Standard zur Bewertung der Ergebnisse durchgesetzt. Daher sind die Ergebnisse unterschiedlicher Autoren und Verfahren auch nur sehr bedingt vergleichbar, hinzu kommt die sehr unterschiedliche Komplexität der Waldsituationen.

Ein allgemeines Problem ist, die Identität von zwei Objekten aus zwei verschiedenen Datensätzen festzulegen. In STRAUB & HEIPKE (2004) wird ein allgemeines Evaluationsverfahren vorgeschlagen, das auf den möglichen topologischen Relationen zwischen Regionen (gilt für beliebige Mengen, bezieht sich hier aber auf eine Region aus der Segmentierung und eine Region aus dem Referenzdatensatz zur Bewertung) basiert. Wird der Rand (also Berührungen von Regionen) vernachlässigt, kann eine Region A mit einer anderen Region B fünf topologische Relationen haben: A und B können disjunkt oder identisch sein, A kann B enthalten, in B enthalten sein oder mit B überlappen. STRAUB & HEIPKE unterscheiden dabei starke und schwache Überlappung. Zur Diskriminierung dient ein *overlap factor*, der wie folgt definiert ist:

$$(3.1) \quad OF_{AB} = \frac{A \cap B}{\min(|A|, |B|)}$$

Ist dieser Faktor kleiner als 0.5 werden die Mengen als ‚eher disjunkt‘ betrachtet und damit als verschieden, ist er größer als 0.5, werden sie ‚eher gleich‘ betrachtet und damit als Übereinstimmung.

Weiterhin werden zur Quantifizierung der Extraktionsergebnisse verschiedene Qualitätsmaße benutzt, deren unterschiedliche Bezeichnungen leider oft zur Verwirrung beitragen. Dafür werden folgende Fälle unterschieden: *true positives* seien die richtig identifizierten Objekte, also diejenigen, die sowohl im Extraktionsergebnis als auch im Referenzdatensatz sind. *false positives* seien diejenigen, die fälschlicherweise extrahiert wurden, also im Extraktionsergebnis sind, aber nicht in der Referenz. *false negatives* seien schließlich diejenigen Fälle, die nicht gefunden wurden, also nicht im Extraktionsergebnis sind, wohl aber in der Referenz.

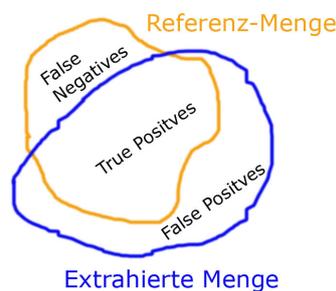


Abb. 3.4: Visualisierung der Begriffe *false negatives*, *true positives* und *false positives*

Davon ausgehend werden üblicherweise folgende Qualitäts- und Fehlermaße unterschieden (siehe etwa RICHARDS & JIA 2006: 304): Die Vollständigkeit (auch Herstellergenauigkeit, *producer accuracy*) ist der Anteil der richtig erkannten Objekte an der Referenz, also *true positives* / (*true positives* + *false negatives*). Das Gegenstück zur Vollständigkeit ist der *error of omission*, also der Anteil der nicht gefundenen Objekte *false negatives* / (*true positives* + *false negatives*).

Als Korrektheit (auch Anwendergenauigkeit, *user accuracy*) wird der Anteil der richtig erkannten Objekte an allen erkannten Objekten *true positives* / (*true positives* + *false positives*) bezeichnet. Ihr Gegenstück ist der *error of commission*, also der Anteil der falsch extrahierten Objekte *false positives* / (*true positives* + *false positives*)

Alle richtig und falsch zugeordneten Objekte können in einer *confusion matrix* (LILLESAND & KIEFER 1979: 475) dargestellt werden, was insbesondere bei mehr als zwei Klassen (bisher wurde ja nur zwischen *positives* und *negatives* unterschieden) die Übersichtlichkeit erhöht.

## 3.2 Template-matching nach POLLOCK, LARSEN & OLOFSSON

### Literatur

POLLOCK 1996, LARSEN & RUDEMO 1997, LARSEN 1999, ERIKSON & OLOFSSON 2005, STITELER & HOPKINS 2000.

**Weitere:** LARSEN 1998, OLOFSSON 2002, OLOFSSON *et al.* 2006.

### Zum Verfahren

Steuerung: Top-Down

Modell:	Pollock-Modell
Segmentierung:	keine
Maßstabsrepräsentation:	Template-Größe
Aggregation:	Einzelbaum
Extrahierte Parameter:	Position der Baumspitze
Datentyp:	Farb(-infrarot)-Luftbilder ( <i>colour infrared</i> , CIR)
----- <b>Zum Datensatz</b> -----	
Baumarten:	Fichte ( <i>Picea abies</i> L.)

### Modell

Insbesondere der geometrische Teil des in POLLOCK (1996) eingeführten Baum-Modells wird bis heute vielfältig verwendet und kopiert und hat sich als quasi-Standard etabliert. Dabei wird die Kronenform durch einen Ellipsoid mit drei Formparametern angenähert:

$$(3.2) \quad \frac{z^n}{a^n} + \frac{(x^2 + y^2)^{n/2}}{b^n} = 1.$$

Der Parameter  $a$  entspricht der Baumhöhe,  $b$  dem Kronendurchmesser und  $n$  ist ein Parameter für die Form. Üblicherweise wird nur die obere Hälfte des Ellipsoids als Modell verwendet (ERIKSON & OLOFSSON 2005: 260). Typische Werte sind nach GONG *et al.* (2002 nach STRAUB 2003a: 12)  $a = 25$ ,  $b = 5$  und  $n = 1.2$  für einen Nadelbaum bzw.  $a = 7$ ,  $b = 7$  und  $n = 2.0$  für einen Laubbaum.

### Verfahren

In LARSEN & RUDEMO (1997) werden aus dem geometrischen Modell mit Hilfe eines Raytracers kreisförmige oder elliptische *templates* (also quasi radiometrische Muster) für die Baumkronen mit der Spitze in der Mitte erstellt. In das Modell gehen etliche Parameter ein, neben der Ellipsoidform auch die relative Position von Sonne, Kamera und Objekt, der Anteil der einfallenden Strahlung von Sonne und Streulicht sowie die Reflektanz der einzelnen Materialien. Diese *templates* werden dann mit einem *matching*-Algorithmus im Bild gesucht. Dafür wird das *template* an jeder Stelle mit dem Bild korreliert (entspricht der Faltung mit dem gespiegelten *template*). Die lokalen Maxima des Korrelationsbildes sind Kandidaten für Baumspitzen. Da es üblicherweise zu viele Kandidaten gibt, stellt sich die Frage, wie die richtigen ausgewählt werden. Eine Methode ist die Anzahl der Bäume zu schätzen oder als *a priori* bekannt voraus zu setzen und dann unter Ausschluss gewisser Nachbarschaften die besten Kandidaten auszuwählen. In ERIKSON & OLOFSSON (2005: 260) wird der Schwellenwert manuell festgelegt.

In LARSEN 1999 wird alternativ vorgeschlagen, das Verfahren mit einer Reihe von leicht veränderten Parametern zu wiederholen (unterschiedliche Größe und Form des *templates* und Position der Spitze darin). Das Verfahren wird *template voting* genannt, da die lokalen Maxima, die über einer bestimmten Korrelation liegen, addiert werden, also die einzelnen Durchgänge quasi für ihre Kandidaten abstimmen. Die Annahme besteht darin, dass richtige Baumspitzen weniger fehleranfällig gegenüber Variation der Parameter sind. Das hat zum einen den Vorteil, dass diese nicht genau kalibriert werden müssen, zum anderen lässt sich der Schwellenwert für die Bestimmung der richtigen Kandidaten aus der Verteilung der Abstimm-Ergebnisse ermitteln (eigentlich ein moderner Ansatz im *automated*

*decision making*, der auch als *committee of classifiers* oder *ensemble based system* bezeichnet wird, siehe POLIKAR 2006). Damit wird die Anzahl jedoch tendenziell unterschätzt.

Weiterhin sei darauf hingewiesen, dass die Größe des *templates* auch als Skalenparameter betrachtet werden kann, da von ihr abhängig über eine größere oder kleinere Fläche mit dem relativ glatten *template* korreliert wird.

### Bewertung

Das Verfahren ist insgesamt vergleichsweise unhandlich, da viele Parameter sorgfältig gewählt werden müssen. Weiterhin werden nur die Positionen der Baumspitzen extrahiert. Die Anzahl der Bäume muss entweder vorher bekannt sein (womit der Nutzen für die Forstwirtschaft gegen Null geht), oder wird beim *template voting* unterschätzt. Für die genaue Baumposition bei schlechter Vollständigkeit sind zwar photogrammetrische Anwendungen denkbar, etwa um stereoskopische Bilder zuzuordnen, für forstwirtschaftliche Anwendungen ist die genaue Anzahl der Bäume aber weit wichtiger als die Genauigkeit der Position.

Nach wie vor sehr aktuell ist aber das Baumkronen-Modell von POLLOCK, auf dessen geometrischen Teil beinahe alle jüngeren DOM-basierte Verfahren beruhen.

## 3.3 Blob-detection: PINZ, KÄTSCH, DRALLE & RUDEMO

----- Literatur -----	
	PINZ 1989, BISCHOFF & PINZ 1989, DRALLE & RUDEMO 1997
Projekt:	<i>Vision Expert System</i>
----- Zum Verfahren -----	
Steuerung:	Top-Down
Modell:	lokales Maximum, kreisförmig Grauwertverteilung (helle, kreisförmige Bildobjekte)
Segmentierung:	teilweise anhand radialer Helligkeitsverteilung
Maßstabsrepräsentation:	manuell durch verschiedene Glättungsfilter
Datentyp:	CIR-Fotos
Auflösung:	10cm

### Verfahren

In PINZ (1989: 105ff) wird das Farbinfrarotbild auf Graustufen reduziert, indem der Mittelwert aus den drei Kanälen gebildet wird. Danach wird mit einem Rechteckfilter geglättet, dessen Größe aus der erwarteten Baumgröße und der Auflösung berechnet wird. Anschließend werden lokale Maxima ermittelt (also helle Flecken im Bild, auch *blobs* genannt), im Falle von Plateaus wird der Schwerpunkt genommen. Von diesen ausgehend, wird die radiale Helligkeitsverteilung berechnet, indem die Mittelwerte entlang von Kreisen in verschiedenen Abständen gebildet werden. Der Radius der Krone wird dort gesetzt, wo die Helligkeit 50% unterschreitet, anschließend wird geprüft, ob der gefundene Radius im zulässigen Bereich liegt.

Zum *tuning* wird das Verfahren mit verschiedenen zu erwartenden Baumgrößen (und somit verschiedenen Glättungsfiltern) wiederholt. Dabei werden zusätzliche Maxima gefunden, die radiale Helligkeitsverteilung wird aber für alle auf dem Graustufenbild berechnet. Sind zwei Baumhypothesen zu

nahe beieinander, wird die größere entfernt. Außerdem werden widersprechende Objekte wie Straßen gesucht.

Dafür wird eine sehr umfangreiche Wissensrepräsentation vorgeschlagen, bei der zwischen Bildobjekten und natürlichen Objekten unterschieden wird. Widerspricht eine Hypothese über ein natürliches Objekt dem Vorwissen über ein anderes (etwa eine Straße oder ein anderer Baum), so wird das natürliche Objekt entfernt, das Bildobjekt bleibt aber erhalten. Weiterhin wird zwischen generischen (prototypischen) und individuellen Objekten unterschieden.

In KÄTSCH (2002) werden ebenfalls regionale Maxima nach einer Glättung gesucht. Zur besseren Separierbarkeit von geschlossenen Kronendecken werden daneben morphologische Filter verwendet. Parallel zur radiometrischen Auswertung wird durch automatische Bildregistrierung ein Kronenoberflächenmodell erzeugt. Dies stellte sich jedoch in den meisten Fällen als unbrauchbar heraus, da der *matching*-Algorithmus bei komplizierten Texturen wie Waldoberflächen versagt.

In BISCHOF & PINZ (1989) wird gezeigt, dass sich neuronale Netze eignen, um die Baumart zu klassifizieren. Bei der Unterscheidung zwischen Fichte und Kiefer wurden damit sehr gute Ergebnisse erzielt und selbst die Klassifikation von fünf Arten funktionierte noch erstaunlich gut.

### Bewertung

Alle *blob*-basierten Verfahren sind relativ einfach. So wird im Wesentlichen nur die Baumspitze extrahiert. Der geschätzte Kronenradius bei einem gewissen Grauwertabfall (etwa 50%) ist dagegen heuristisch bestimmt und ist sicher sehr abhängig vom Bildmaterial. Ebenso werden die verwendeten Glättungsfilter manuell bestimmt und somit das Problem des optimalen Maßstabs umgangen. Schlecht ist außerdem, dass oft rechteckige Filtermasken verwendet werden, da diese anisotrop sind. Nicht zu verstehen ist auch, warum bei PINZ (1989) die Kanäle gemittelt werden, da der NIR-Kanal die meiste Information beinhaltet.

Durchaus brauchbar scheinen die Wissensrepräsentation BISCHOF & PINZ (1989) sowie die Unterscheidung zwischen Bildobjekten und natürlichen Objekten. Außerdem kann überprüft werden, ob sich das Verfahren eignet, Saatpunkte für komplexere Segmentierungsalgorithmen zu bestimmen.

## 3.4 Valley-following approach nach GOUGEON

-----	<b>Literatur</b>	-----
		GOUGEON 1995, GOUGEON & MOORE 1988
-----	<b>Zum Verfahren</b>	-----
Steuerung:		Bottom-Up
Modell:		Lokale Maxima, Täler zwischen Bäumen
Segmentierung:		zweistufig: Valley-Following und regelbasierte Nachverarbeitung
Maßstabsrepräsentation:		keine
Aggregation:		Einzelbaum
Extrahierte Parameter:		Anzahl, Kronenumriss
Datentyp:		Multispektral, insbesondere NIR
-----	<b>Zum Datensatz</b>	-----
Scanner:		MEIS II
Auflösung:		31cm

Untersuchungsgebiet:	Hudson Plantation (46°00'N, 77°23'W)
Baumarten:	Rotkiefer ( <i>Pinus resinosa</i> ), Amerikanische Rot-Fichte ( <i>Picea rubens</i> Sarg.), Weiß-Fichte ( <i>Picea glauca</i> ) und Gemeine Fichte ( <i>Picea abies</i> L.)

### Verfahren

Die Segmentierung erfolgt in zwei Durchgängen. Der erste Durchgang nennt sich *Valley-Following* und arbeitet auf dem leicht geglätteten NIR-Kanal. Für eine erste grobe Segmentierung wird ein Schwellenwert gebildet, um die Kronen vom Hintergrund zu trennen. Dann werden zunächst alle lokalen Minima markiert und bei vier aufeinander folgenden Durchgängen entlang der Diagonalen zu Tälern erweitert. Als Tal-Pixel gelten dabei solche, deren Nachbarn links und rechts in Bearbeitungsrichtung höhere Werte haben. Diese Prozedur wird wiederholt, bis keine Täler mehr gefunden werden.

Das Zwischenergebnis wird an den zweiten Schritt, die regelbasierte Ermittlung der Kronenbegrenzung, übergeben. Deren Ziel ist es, den Rändern der einzelnen Kronen zu folgen, um sie zu schließen. Dafür wird gemäß fester Regeln an jeder Krone ‚möglichst‘ im Uhrzeigersinn entlanggegangen. Je stärker vom Uhrzeigersinn abgewichen werden muss, desto mehr wird überprüft, ob durch Überspringen einiger Pixel die Bewegungsrichtung beibehalten werden könnte. Bei einer Drehung um 135° gegen den Uhrzeigersinn können beispielsweise bis zu 3 Pixel in Bewegungsrichtung übersprungen werden. Auf diese Weise sollen fälschlicherweise verbundene Kronen getrennt werden. Der zweite Schritt wird ebenfalls mehrfach wiederholt.

### Bewertung

Der frühe Ansatz von GOUGEON wurde viel beachtet und war somit erfolgreich. Ein Nachteil ist aber, dass das Verfahren schwer an andere Waldsituationen und Bildauflösungen anzupassen sein dürfte. So sind nicht nur die Parameter für eine Situation optimiert, sondern es stecken auch viele Vorannahmen in den Regeln der Kronenverfolgung (bei höherer Auflösung müssten vielleicht fünf statt nur drei Pixel übersprungen werden). Das macht das Verfahren sehr schwer zu steuern. Insbesondere ist das Ergebnis stark von der Topologie der Bäume und damit vom Sonnenstand abhängig.

## 3.5 Fouriertechniken nach KRUMMHEUER

----- <b>Literatur</b> -----	
	KRUMMHEUER 1998
Software:	ERDAS Imagine, MATHEMATICA
----- <b>Zum Verfahren</b> -----	
Steuerung:	Bottom-Up
Modell:	räumlich wiederkehrende (periodische) Strukturen durch farbiges Rauschspektrum und Autokovarianzfunktion; Annahme einer dominanten Skalenlänge.
Segmentierung:	keine
Aggregation:	Bestand
Extrahierte Parameter:	mittlere räumliche Skalenlängen
Datentyp:	Roter Kanal aus CIR-Fotos
Auflösung:	0.5m
----- <b>Zum Datensatz</b> -----	
Untersuchungsgebiet:	bei Göttingen

Landschaftstyp:	Buchenwald (50 bzw. 100 Jahre); „quasi-homogene Bestände“
Baumarten:	Buche ( <i>Fagus</i> )

### Verfahren

Dieses Verfahren dient nicht der Baumextraktion, sondern der Erfassung von räumlich periodisch wiederkehrenden Mustern. Deren Wellenlängen werden durch ‚Fouriertechniken‘ erfasst. Da sich bei älteren Buchenbeständen die ursprüngliche Struktur durch Differentiationsprozesse und waldbauliche Eingriffe aufgelöst hat, gibt es keine räumliche Vorzugsrichtung (Isotropie). Daher lohnt eine 2D-Betrachtung nicht und die Analyse erfolgt nur entlang ausgewählter Profile, wofür Bildzeilen und Spalten verwendet werden.

Die Zeilen- bzw. Spaltenvektoren werden zentriert und auf die Standardabweichung eins normiert. Anschließend werden sie fouriertransformiert. Betrachtet werden im Folgenden die quadratischen Spektren, die Aussagen über die Energie in den verschiedenen Frequenzbereichen zulassen, die Phaseninformationen werden dagegen vernachlässigt.

Zunächst werden verschiedene idealtypische Fälle betrachtet. Für ein rein periodisches Signal hat das Spektrum nur einen Wert ungleich Null, da die gesamte spektrale Energie in einer Wellenlänge liegt. Beim „weißen Rauschen“ sind die Energien im Mittel gleichmäßig auf alle Wellenlängen verteilt (Die Begriffe weißes und farbiges Rauschen stammen aus der Optik. Es sei darauf hingewiesen, dass diese Analogie nur im übertragenen Sinne zutreffend ist, da es sich hier bei den Wellenlängen nicht um Farben von Licht handelt, sondern um räumlich wiederkehrende Muster). In der Realität wird sich oft ein „farbiges Rauschen“ vorfinden, bei dem ein bestimmter Wellenlängenbereich überwiegt, dieser jedoch mehr als eine einzelne Wellenlänge umfasst.

Diese farbigen Rauschspektren werden durch Funktionen der Art:

$$(3.3) \quad \text{spec}(s) = \frac{a}{a^2 + \left(b - \frac{2\pi s}{N}\right)^2}$$

angenähert. Dabei sind  $a$  und  $b$  anzupassende Parameter größer Null und  $N$  ist die Länge des Vektors. Da die Anpassung der Parameter im Wellenzahlbereich vergleichsweise aufwendig ist, wird stattdessen die Autokovarianzfunktion im Ortsbereich betrachtet.

$$(3.4) \quad \text{akv}(\tau) = \frac{1}{N - \tau - 1} \sum_{r=0}^{N-\tau-1} \gamma_r \gamma_{r+\tau}$$

Diese kann durch Kosinustransformation aus dem quadratischen Spektrum berechnet werden. Zu den farbigen Rauschspektren gehören entsprechend Autokovarianzfunktionen des Typs  $\text{akv}(\tau) = \exp(-a\tau) \cos(b\tau)$ . Zudem hängt die erste Nullstelle  $\tau_0$  dieser Funktion direkt linear mit der maximalen Wellenlänge des Spektrums zusammen:  $\lambda_{\max} = 4\tau_0$ . Da die Nullstelle wesentlich leichter zu bestimmen ist als das Maximum des quadratischen Spektrums, werden die Parameter  $a$  und  $b$  anhand der Autokovarianzfunktion derart angepasst, dass neben der Nullstelle auch für  $\tau = 1$  der richtige Wert angenommen wird.

Auf diese Weise werden für alle bis auf wenige aussortierte Bildzeilen und Spalten die jeweiligen Nullstellen von  $akv(\tau)$  bestimmt. Die Verteilung der Nullstellen wird durch verschiedene parametrische Wahrscheinlichkeitsverteilungen angenähert, von denen sich die Log-Normalverteilung am besten eignet.

### Bewertung

Auch wenn keine einzelnen Bäume extrahiert werden, ist dieses Verfahren sehr interessant in Hinblick auf ein Expertensystem. Da die dominante Skalenlänge im Bild unmittelbar mit der zu erwartenden Baumgröße zusammenhängt, könnte hiermit ein wertvoller Parameter zur Steuerung anderer Verfahren gewonnen werden. Natürlich wäre vorher zu überprüfen, wie sich das Verfahren in abweichenden Fällen verhält, insbesondere bei inhomogenen Baumgrößen und anisotropen Situationen.

Dennoch wirkt das Verfahren noch nicht vollkommen ausgereift. Die für die Anpassung gewählten Funktionstypen sind ebenso heuristisch bestimmt wie die Fixierung der Autokovarianzfunktion an der Nullstelle und der Stelle  $\tau = 1$ . Die Bestimmung der Funktion scheint ohnehin überflüssig, da anschließend nur noch die Nullstellen weiter verwendet werden. Ebenfalls fragwürdig ist die Annahme, dass die Nullstellen der einzelnen Bildzeilen und Spalten log-normalverteilt sind. Trotz dieser Erklärungsdefizite ist das Verfahren aber ein gelungener Beweis, dass sich Skalenlängen automatisch bestimmen lassen, und damit ein äußerst wertvoller Beitrag.

## 3.6 Morphologischer Ansatz nach ANDERSEN *et al.*

### ----- Literatur -----

ANDERSEN *et al.* 2001

### ----- Zum Verfahren -----

Steuerung:	Bottom-Up
Modell:	konische oder ellipsoide Baumspitzen
Segmentierung:	morphologische <i>top-hat</i> Transformation
Maßstabsrepräsentation:	verschieden große Strukturelemente
Aggregation:	Einzelbaum
Extrahierte Parameter:	Baumspitzen, Höhe
Datentyp:	LIDAR, DGM

### ----- Zum Datensatz -----

Scanner:	TopEye
Auflösung:	3.5 Pulse/m <sup>2</sup>
Untersuchungsgebiet:	Capitol State Forest, WA, USA
Landschaftstyp:	Wald, hügelig
Baumarten:	Douglasie ( <i>Pseudotsuga menziesii</i> )

### Verfahren

Zur Erstellung des Digitalen Kronenmodells (DKM) wird durch einen Maximum-Filter die größte Pulsantwort in einem bestimmten Bereich gesucht. Dabei ist die Größe des Extraktionsfilters entscheidend: Ist sie zu klein, so werden Bodenpunkte als Oberfläche extrahiert, ist sie zu groß, wird die Genauigkeit reduziert. Aus den gefilterten Punktdaten wird mit Hilfe von Delaunay-Triangulation

ein Raster-Höhenmodell erzeugt. Dabei wurden Filtergrößen von 91cm bis 152cm bei einer Rasterweite von 30cm getestet.

Die Baumspitzen werden durch eine morphologische *top-hat* Transformation (siehe 2.3.8) mit einem kreisförmigen Strukturelement extrahiert. Dabei wird die Größe der Strukturelemente ebenfalls variiert, um Bäume unterschiedlicher Größen zu finden. Dies kann als ‚morphologischer Maßstabsraum‘ angesehen werden und lässt sich veranschaulichen, als würde eine Abfolge von Sieben mit gröberen und feineren Löchern verwendet. Daher bezeichnet (SOILLE 1998: 117ff) auf diese Art erzeugte Menge auch in Anlehnung an mineralogische Korngrößen als Granulometrie.

Von den so transformierten Bildern wird jeweils ein Schwellenwert gebildet, um binäre Bilder zu erzeugen. Die Binärbilder werden anschließend morphologisch geöffnet, um Rauschen zu entfernen. Dafür wird ein etwas kleineres, ebenfalls kreisförmiges Strukturelement verwendet. Zur Schätzung der Höhe der Bäume wird aus den LIDAR-Daten außerdem ein DGM generiert.

Die Referenzdaten wurden manuell-photogrammetrisch aus stereoskopischen Luftbildern gewonnen. Die Ergebnisse schwanken beträchtlich je nach Wahl der Parameter Filtergröße für das Kronenmodell und Größe der Strukturelemente. Je größer die Strukturelemente, desto geringer die Vollständigkeit, aber desto höher die Korrektheit, da weniger Objekte erkannt werden. Der beste erreichte Wert sind 83% Vollständigkeit und 89% Korrektheit. Die Höhe wurde etwa auf einen Meter genau bestimmt.

### **Bewertung**

Positiv an diesem Ansatz ist vor allem die grundlegende Verarbeitung mit Hilfe morphologischer Funktionen hervorzuheben, die bei Höhenmodellen ja eigentlich nahe liegend ist. Daher wurde dieser Ansatz in die Auswahl von Verfahren aufgenommen, auch wenn er nicht sonderlich ausgereift ist und in der jüngeren Literatur nicht weiter verfolgt wurde.

Eine Verbesserung könnte die Verwendung nicht-flacher Strukturelemente sein, die es ermöglichen, bestimmte Formen zu extrahieren. Eine deutliche Schwäche des Verfahrens ist, dass die Frage des optimalen Maßstabs nicht befriedigend gelöst ist. Es wird zwar für einen einzelnen Datensatz bestimmt, welche Größe von Strukturelementen die besten Ergebnisse bringt, da dies aber nicht verallgemeinerbar ist, muss es für jedes Bild neu kalibriert werden. Außerdem wird die Frage offen gelassen, wie bei Segmentierung in verschiedenen morphologischen Maßstabsebenen die besten Hypothesen ausgewählt werden sollen.

## **3.7 WST im Maßstabsraum nach STRAUB**

### **Literatur**

am ausführlichsten in STRAUB 2003A; Daneben auch STRAUB 2002, STRAUB 2003B, STRAUB 2004, STRAUB & HEIPKE 2004 und WOLF & HEIPKE 2007

### **Zum Verfahren**

Steuerung:	Top-Down
Modell:	Größe, Kreisförmigkeit, Konvexität und Vitalität
Segmentierung:	WST in verschiedenen Maßstabsebenen
Maßstabsrepräsentation:	linearer Maßstabsraum
Aggregation:	Einzelbaum
Extrahierte Parameter:	Kronenbegrenzung und -fläche

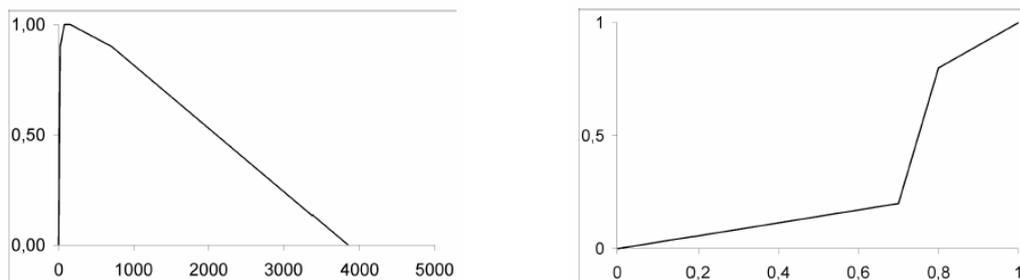
Datentyp:	LIDAR & radiometrisch
----- Zum Datensatz -----	
Scanner:	verschiedene (siehe STRAUB 2003A: 69)
Auflösung:	Oberflächenmodelle: 0.2-1.0m, optisch: 0.1-0.5m
Landschaftstyp:	vorwiegend urban, aber auch Wald (Nadelbäume verschiedenen Alters)

### Modell

Das Modell von STRAUB besteht aus drei geometrischen und einer radiometrischen Eigenschaft. Für diese werden Zugehörigkeitsfunktionen aufgestellt, die mit Fuzzy-Logik-Operatoren verknüpft werden.

STRAUB verwendet zwar nicht das Pollock-Modell selbst, leitet daraus aber wesentliche Annahmen über die Geometrie von Bäumen ab. Diese drücken sich in seinem Modell in den Kenngrößen Kreisförmigkeit und Konvexität aus.

Zur Erstellung der Zugehörigkeitsfunktionen werden *a priori* Grenzen für wahrscheinliche und weniger wahrscheinliche Ausprägungen der einzelnen Merkmale festgesetzt. So bekommt die ‚typische‘ Größe (Fläche des jeweiligen Segments) von 80-150m<sup>2</sup> den Zugehörigkeitswert 1, die ‚maximale‘ Größe von 700m<sup>2</sup> den Wert 0.75 und die Obergrenze von 3850m<sup>2</sup> (entspricht einem Kreis mit 35 m Radius) den Wert 0, dazwischen wird interpoliert. Die resultierende Zugehörigkeitsfunktion für das erste Merkmal Größe  $a_1$  ist in Abb. 3.5 (links) zu sehen.

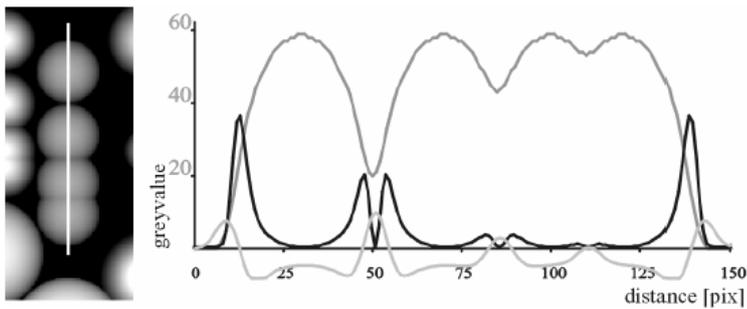


**Abb. 3.5: Zugehörigkeitsfunktionen für Größe und Kreisförmigkeit einer Baumkrone nach STRAUB (2003A: 43f)**

Das zweite Merkmal  $a_2$  ist die Kreisförmigkeit eines Segmentes. Diese wird definiert als Verhältnis der Fläche des Segments zum kleinsten Kreis um seinen Schwerpunkt, in den das gesamte Segment hinein passt:  $a_2 = \frac{F}{\max(r)^2 \pi}$ , wobei  $\max(r)$  den größten Abstand zwischen Schwerpunkt und Rand

des Segments bezeichnet. Für einen Kreis ergibt dies den Wert 1, für ein Quadrat den Wert  $2/\pi \approx 0.6$ . Die Zugehörigkeitsfunktion für die Kreisförmigkeit wurde empirisch ermittelt.

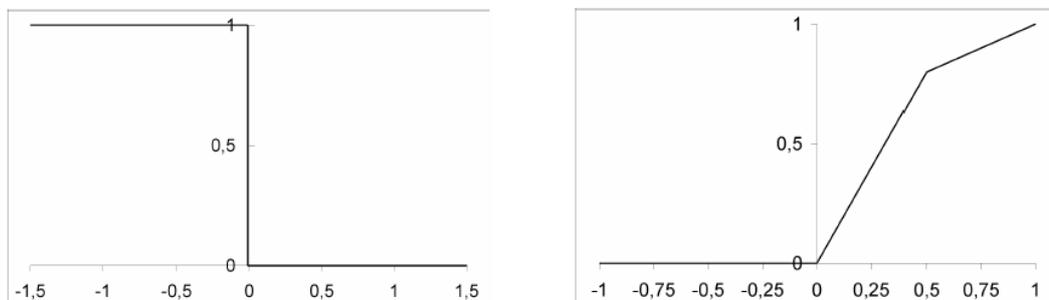
Das dritte Merkmal von Straub beinhaltet differentialgeometrische Eigenschaften der Baumkrone. Eine Eigenschaft des Pollockmodells ist, dass die Baumkrone durchgehend konvex ist, d.h. die partiellen zweiten Ableitungen der Höhe  $H_{xx}$  und  $H_{yy}$  negativ sind. Daraus folgt, dass auch ihre Summe (das Ergebnis des Laplace-Operators) negativ ist:  $\Delta H(\mathbf{x}) < 0$ . Zwischen zwei Bäumen liegt ein Sattel, so dass an dieser Stelle die Ableitung quer zur Ausrichtung der Achse zwischen den Bäumen zwar negativ, die längs aber größer als Null ist.



**Abb. 3.6: Schnitt durch idealisierte Pollockbäume und Ableitungen. Dunkelgrau: Höhe; schwarz: Betrag des Gradienten; hellgrau:  $\Delta$ -Operator nach STRAUB (2003A: 45)**

Die Konvexität  $a_3$  wird als normalisierter Mittelwert der Summe der zweiten Ableitungen in einem Segment definiert. Die Zugehörigkeitsfunktion fordert, dass sie für Bäume negativ ist. Diese differentialgeometrische Eigenschaft ist von der Höhe des Baumes und damit auch von der darunter liegenden Geländehöhe unabhängig.

Als vierte und einzige radiometrische Eigenschaft wird der NDVI (siehe 2.1.1) verwendet. Das entsprechende Merkmal Vitalität  $a_4$  für ein Segment ist der Mittelwert des NDVI aller Pixel im Segment. Dafür werden neben den LIDAR-Daten auch Bilder in Rot und nahem Infrarot benötigt. Diese werden wie das jeweilige Segment im linearen Maßstabsraum skaliert. Die Vitalität dient vor allem der Unterscheidung von Bäumen und anderen Objekten mit baumähnlichen Geometrien, etwa großen Autos. STRAUB (2003A: 48) weist darauf hin, dass auf das Merkmal in Waldgebieten gegebenenfalls verzichtet werden kann. Alternativ könne auch auf Texturmerkmale zurück gegriffen werden.



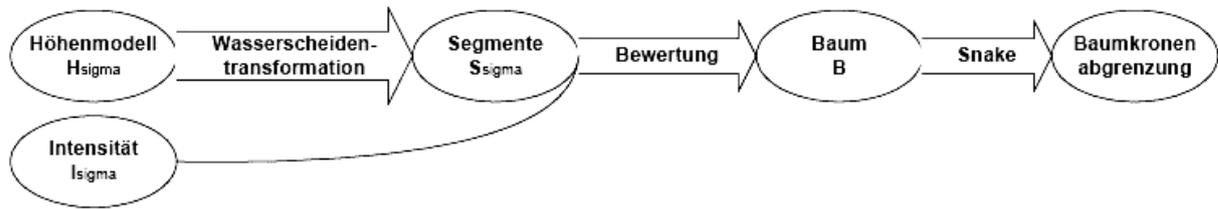
**Abb. 3.7: Zugehörigkeitsfunktionen für die Merkmale Konvexität und Vitalität nach STRAUB (2003A: 44, 46)**

Auf Kontextmerkmale wird bei der Extraktion verzichtet. Dies wird dadurch begründet, dass „Aussehen und Vorkommen von einzelnen Bäume meist unabhängig von anderen Objekten sind“ (STRAUB 2003A: 54). Damit können Bäume zuerst extrahiert und anschließend zur Extraktion anderer, kontextabhängiger Objekte verwendet werden.

## Verfahren

Die Strategie besteht darin, in verschiedenen Maßstabsebenen zu segmentieren und dann jeweils diejenigen Segmente auszuwählen, die am ehesten einem Baum entsprechen. Dafür werden alle Segmente mit den Zugehörigkeitsfunktionen gemäß den vier vorgestellten Merkmalen Größe, Konvexität, Kreisförmigkeit und Vitalität bewertet. Da es konkurrierende Hypothesen in verschiede-

nen Maßstäben geben kann, werden zunächst die topologischen Relationen zwischen den Segmenten untersucht und anschließend jeweils die besten ausgewählt und als Baumkrone klassifiziert. Dann wird mit Hilfe einer aktiven Kontur (*Snake*) die Kronenbegrenzung durch einen geschlossenen Polygonzug noch genauer bestimmt, als dies durch die Segmentierung erfolgt ist.



**Abb. 3.8: Verfahren nach STRAUB (2003A: 55)**

### *Segmentierung*

Zur Bestimmung der optimalen Maßstabsebene wird das Ausgangsbild in den linearen Maßstabsraum (siehe 2.3.9) transformiert. Der relevante Bereich des Maßstabsraums wird dabei aus den minimalen und maximalen Kronendurchmessern geschätzt. Da zwei Impulse im Maßstabsraum erst dann verschmelzen, wenn  $\sigma$  größer ist als der halbe Abstand zwischen den Impulsen, liegt der interessante Bereich zwischen dem kleinsten und dem größten Radius einer Baumkrone.

Die Segmentierung wird in allen Maßstabsebenen durchgeführt. Als Segmentierungsverfahren wird die Wasserscheidentransformation verwendet (siehe 2.4.5), jedes Segment enthält also ein lokales Minimum. Da die Baumkronen im Oberflächenmodell  $H(\mathbf{x}, \sigma)$  aber lokalen Maxima entsprechen, werden drei mögliche Segmentierungsfunktionen vorgeschlagen: Das invertierte Oberflächenmodell  $I^0 H(\mathbf{x}, \sigma)$ , der Betrag des Gradienten  $|\nabla H(\mathbf{x}, \sigma)|$  und das invertierte Quadrat der Summe der zweiten Ableitungen  $I^0((\Delta H(\mathbf{x}, \sigma))^2)$ .

Erstere hat den Nachteil, dass die Segmente neben den Bäumen noch Bereiche der Umgebung enthalten, wenn zwischen den Bäumen Lücken sind. Daher sind die Segmentierungsfunktionen mit Maxima direkt an den Baumbegrenzungen besser geeignet. Der Betrag des Gradienten hat sein Maximum an der Stelle der stärksten Steigung. Die zweite Ableitung zeigt dagegen die Krümmungseigenschaften an. Durch das Quadrieren wird erreicht, dass sowohl konkave als auch konvexe Gebiete positiv sind, der Übergang zwischen beiden bleibt aber Null (also ein lokales Minimum, das durch Invertierung zur Wasserscheide wird). Diese Funktion hat sich als am besten geeignet herausgestellt.

### *Bewertung und Hypothesenbildung*

Zur Bewertung werden die Zugehörigkeitsfunktionen gemäß der vier Merkmale des Modells für die einzelnen Segmente berechnet. Die Zugehörigkeitswerte werden mit einem UND verknüpft. Diejenigen Segmente, deren Zugehörigkeitsfunktionen einen Mindestwert, etwa 0.75, haben, werden als gültige Baumhypothesen ausgewählt.

Da alle Objekte über mehrere Maßstabsebenen bestehen bleiben, werden auf den verschiedenen Ebenen konkurrierende gültige Hypothesen aufgestellt, von denen anschließend die beste ausgewählt werden muss. Dafür werden die topologischen Relationen der Segmente untersucht. Sind diese nicht disjunkt (oder überlappen nur wenig – siehe 3.1.4), so wird davon ausgegangen, dass sie zum selben

Baum gehören, und es wird jeweils das Segment mit dem höchsten Zugehörigkeitswert als beste Hypothese im Maßstabsraum ausgewählt.

### *Bestimmung der Kronenbegrenzung*

Trotz Wahl einer geeigneten Segmentierungsfunktion werden die Kronenbegrenzungen für die besten Hypothesen noch genauer bestimmt. Dafür werden aktive Konturen (*snakes*) eingesetzt. Diese gehen auf KASS *et al.* (1988) zurück und dienen der Bestimmung von Konturen, der Objektverfolgung oder der Gruppierung von Merkmalen bekannter Form. Dabei wird die Kontur durch eine parametrische Kurve, etwa ein Polygonzug, beschrieben, die von inneren und äußeren Energien gesteuert wird.

Die innere Energie wird hier durch zwei Parameter gesteuert, die die Spannung der Kurve und ihre Krümmung modellieren. Je nach Wahl der Parameter verhält sich die Kurve eher wie ein Gummizug oder wie ein Stahlseil und lässt die Ausbildung von Ecken zu oder nicht. Die äußere Energie ergibt sich durch die Grauwerte des Bildes, als ob ein Kraftfeld die *Snake* zu niedrigen Werten bewegen wollte. Bei Höhenmodellen lässt sich dies leicht als Schwerkraft veranschaulichen.

Der Algorithmus minimiert nun die Gesamtenergie des Systems. Dafür wird die *Snake* mit einem Startwert initialisiert und über einen gewissen Zeitraum dem Kraftfeld überlassen. Um die Reichweite nach außen zu vergrößern, wird weiterhin eine nachlassende, nach außen gerichtete Kraft eingeführt, die die *Snake* aufbläst wie einen Ballon. Die Druckkraft wird durch einen weiteren Parameter gesteuert. Um die Geschwindigkeit der Entwicklung zu steuern, wird das Ganze schließlich in ein viskoses Medium eingebettet, das ebenfalls durch einen Parameter gesteuert wird. Als externes Energiefeld wird die Summe der Gradientenbeträge in allen Maßstabsebenen verwendet.

Da sich durch die Konturbestimmung die topologische Relation zu anderen Bäumen verändern kann, werden diese anschließend neu bestimmt. Abschließend werden alle Baumkronen noch einmal überprüft und diejenigen mit den schlechtesten Zugehörigkeitswerten verworfen.

### **Ergebnisse**

Das Verfahren wurde an vier Szenen unterschiedlichen Inhalts und unterschiedlicher Auflösung getestet, die nach den Aufnahmeorten Grangemouth, Hohentauern, Paris und Ravensburg benannt sind. Drei davon beinhalteten urbane Gebiete und Hohentauern ein Nadelwaldgebiet. Die Zugehörigkeitsfunktionen für Größe, Kreisförmigkeit und Konvexität wurden für alle Datensätze verwendet, nur die Vitalität wurde an das vorliegende Bildmaterial angepasst. Die Oberflächenmodelle wurden in zwei Fällen aus LIDAR-Daten und in zwei Fällen photogrammetrisch aus stereoskopischen Bildern gewonnen.

Die Referenzdaten wurden von einem Operateur manuell aus den Oberflächenmodellen unter punktueller Einbeziehung der radiometrischen Bilder ermittelt. Die Vollständigkeit der besten Hypothesen im Maßstabsraum lag zwischen 50% und 93%, die Korrektheit bei 54-86%. Nach Anwendung der aktiven Konturen verbesserte sich die Korrektheit auf 66-100%, die Vollständigkeit sank dagegen auf 41-81%. Der Datensatz mit der höchsten Auflösung (Grangemouth: 0.1m optisch und 0.2m Oberflächenmodell) lieferte dabei deutlich bessere Werte als die übrigen.

Die Genauigkeit der bestimmten Position lag im Mittel 1m vom Referenzwert entfernt. Der mittlere bestimmte Radius stimmte bei einer Standardabweichung von etwa 1m mit der Referenz überein. Erstaunlicherweise wurde die Genauigkeit durch die Verwendung von aktiven Konturen nicht erhöht.

## Bewertung

Das Verfahren von Straub ist sehr vielversprechend und hat deutliche Stärken. Dazu ist in jedem Fall das objektspezifische, parameterfreie und kontextunabhängige Modell zu zählen, das bis auf das Merkmal Vitalität für die verschiedenen Datensätze nicht angepasst werden musste. Der Verzicht auf das Merkmal Höhe macht ein genaues Geländemodell überflüssig, dafür wird dieser forstwirtschaftlich bedeutende Parameter allerdings auch nicht extrahiert, obwohl er für die Bestimmung des Holzvolumens wichtiger ist als etwa der Kronendurchmesser (HYYPÄ *et al.* 2005: 144). Aber auch unabhängig davon ist der Verzicht nicht ganz nachvollziehbar, da auch das Klassifikationsergebnis (hier Hypothesenauswahl) durch das Merkmal deutlich verbessert werden könnte. Im urbanen Umfeld wäre er beispielsweise geeignet Bäume von Autos zu unterscheiden, was alleine auf Grund der Differentialgeometrie schwierig sein dürfte.

Eine weitere Stärke ist die Segmentierung im linearen Maßstabsraum, durch die das Problem der optimalen Maßstabebene adressiert wird. Der anschließende Vergleich von allen Hypothesen ist zwar zusätzlicher Aufwand, lohnt sich aber auf jeden Fall im Hinblick auf die universelle Einsetzbarkeit. Auch die Verwendung der WST macht auf Oberflächenmodellen Sinn und bringt gute Ergebnisse. Die Wahl der Segmentierungsfunktion ist ebenfalls überzeugend.

Weniger überzeugend ist dagegen die Verwendung von Fuzzy-Logik bei den Zugehörigkeitsfunktionen. Da der geforderte Zugehörigkeitswert für einen Baum nicht aus dem Modell hervorgeht, sondern heuristisch ermittelt wird, könnten als Zugehörigkeitsfunktion ebenso gut einfache Schwellenwerte verwendet werden. Wird beispielsweise eine Zugehörigkeit von 0.75 gefordert, so wird trotz Fuzzy-Logik ein Segment der Größe 700m<sup>2</sup> in diesem Merkmal immer als Baum gewertet, eines der Größe 701m<sup>2</sup> dagegen nie. Dem könnte durch eine andere Verknüpfung der Merkmale begegnet werden. Weiterhin ließen sich sicher die Zugehörigkeitsfunktionen selbst optimieren.

Nur geringe Verbesserungen bringt der Einsatz der aktiven Konturen. Zum einen sind diese durch die Vielzahl von Parametern schwer zu handhaben. Zum anderen wurde der Gewinn an Korrektheit durch einen Verlust an Vollständigkeit erkauft. Die Genauigkeit wurde durch sie nicht verbessert. Dies liegt möglicherweise zum Teil daran, dass die Referenzbäume immer kreisförmig sind, und damit eine genauere Bestimmung der Kontur im Vergleich keinen Vorteil bringt. Hierin liegt aber ein generelles Problem von Referenzdaten.

Weiterhin ist festzuhalten, dass sehr gute Ergebnisse nur mit sehr hoch auflösenden Daten (sowohl optisch als auch LIDAR) erzielt wurden, auch die übrigen Ergebnisse bewegen sich aber im Rahmen der Vergleichswerte aus der Literatur.

## 3.8 NATSCAN: DIEDERSHAGEN *et al.*

----- Literatur -----	
	DIEDERSHAGEN <i>et al.</i> 2003A, DIEDERSHAGEN <i>et al.</i> 2003B, DIEDERSHAGEN <i>et al.</i> 2004
Projekt:	NATSCAN
Software:	ERDAS Imagine, HALCON Image Processing
----- Zum Verfahren -----	
Steuerung:	Bottom-Up
Modell:	lokale Maxima, Größe und Formparameter (Anisometrie, Kompaktheit), Abstand

Segmentierung:	<i>pouring</i> -Algorithmus, anschließend <i>ray</i> -Algorithmus
Maßstabsrepräsentation:	nur Unterscheidung von jungen und alten Beständen
Aggregation:	Bestand und Einzelbaum
Extrahierte Parameter:	Bestand: Mittlere Höhe, Kronendichte; Einzelbaum: Höhe, Kronenfläche
Klassifikation:	Laubbaum, Nadelbaum
Datentyp:	LIDAR (first pulse, last pulse), RBG-NIR, FOGIS ( <i>Forestal Geographical Information System</i> Baden- Württemberg)

---

**Zum Datensatz**

---

Scanner:	Toposys FALCON
Auflösung:	1m ( <i>footprint</i> )
Untersuchungsgebiet:	nahe Freiburg (Rheintal, Schwarzwald)
Landschaftstyp:	Mischwald, unterschiedliche Höhen

## Verfahren

### *Bestandsweise Segmentierung*

Zunächst wird aus Oberflächenmodell und DGM ein normalisiertes Oberflächenmodell bzw. Kronenmodell erzeugt. Es wird darauf hingewiesen, dass das verwendete Verfahren sich nicht für sehr steiles Gelände eignet.

Die Bestände werden durch sukzessives Filtern extrahiert. Dabei werden einmal erkannte Bestände aus dem Bild gelöscht. Als erster Schritt wird eine Segmentierung mit einem Schwellenwert von 15m durchgeführt, um junge Bestände mit einer homogenen, geschlossenen Kronendecke zu finden. Als nächstes wird mit einem  $15 \times 15$  Mittelwert-Filter geglättet, um Bestände zu finden, die von Laubbäumen über 25m Höhe dominiert werden. Ein weiterer Filter berechnet die lokale Maßabweichung, um die homogene Textur von jungen Nadelbäumen als Extraktionskriterium zu verwenden. In DIEDERSHAGEN *et al.* (2004: 209) wird alternativ vorgeschlagen, den Wertebereich erst durch eine unüberwachte ISODATA-Klassifizierung auf 5 Grauwerte zu verkleinern und anschließend eine überwachte Klassifikation und kontextabhängige Nachverarbeitung durchzuführen.

Für die Bestände werden die Parameter Durchschnittshöhe und Kronendichte gemessen. Hierzu wird aber auf die Ergebnisse der Einzelbaumextraktion zurück gegriffen.

### *Baumweise Segmentierung*

Zur Vorverarbeitung wird zunächst mit einem Gaußfilter geglättet. Dabei wird zwischen jungen und alten Beständen unterschieden, die durch einen Schwellenwert von 20m (DIEDERSHAGEN *et al.* 2003A) bzw. 22m (DIEDERSHAGEN *et al.* 2003B) unterschieden werden. Die alten Bestände werden stark gefiltert, die jungen gar nicht oder nur wenig.

Die Segmentierung wird mit einem *pouring*-Algorithmus durchgeführt, der Ähnlichkeiten mit der WST hat. Zunächst werden alle regionalen Maxima (bezüglich 4-Nachbarschaft) markiert. Von diesem aus wird dann in alle Richtungen ‚bergab‘ expandiert, bis ein Talboden erreicht ist. Dies lässt sich durch Regentropfen, die einen Berg hinab fließen, veranschaulichen. Da die Täler zu mehreren Maxima gehören können, werden die unsicheren Punkte abschließend durch eine gleichmäßige Expansion aller betroffenen Regionen zugeordnet.

Da es einerseits falsche Segmente gibt und die Segmente andererseits auch Teile des Bodens enthalten (es wurde ja bis zum Talboden segmentiert), werden verschiedene Nachverarbeitungsschritte angewandt, deren Reihenfolge sich jedoch zwischen den einzelnen Artikel unterscheidet.

Bei den Formparametern und wechselseitigen Beziehungen handelt es sich um eine Reihe von Operationen, um die Segmentierung zu verbessern und falsche Segmente (solche die keine Bäume darstellen) zu entfernen. Dies geschieht anhand bestimmter Merkmale damit geht in die Nachverarbeitung Modellwissen ein. So werden alle Regionen unter einer bestimmten Größe mit der Nachbarregion vereinigt, mit der sie die größte gemeinsame Grenze haben. Das Verhältnis zwischen den beiden orthogonalen Radien in den Hauptachsen der Region wird als Anisometrie definiert. Weicht dieser Wert zu sehr vom Wert eins für einen Kreis ab, so wird das Segment einem anderen hinzugefügt oder verworfen. Ebenso wird die Kompaktheit als Verhältnis zwischen Quadrat des Umfangs und  $4\pi$  mal der Fläche definiert, die für einen Kreis ebenfalls den Wert eins annimmt. Schließlich werden Maxima, die zu nah beieinander liegen, vereinigt.

Der *ray*-Algorithmus dient dazu, die Form der Segmente zu verbessern und die Bodenteile zu entfernen. Dazu werden ausgehend von der Spitze Strahlen in verschiedene Richtung berechnet. Unterschreitet die Steigung entlang dieser Strahlen den Grenzwert von  $-2m/0.5m$ , so wird die Stelle als neue Begrenzung markiert, sonst bleibt der Endpunkt erhalten. Der neue Umriss ergibt sich aus dem Polygonzug, der alle so markierten Punkte verbindet.

Anschließend wird versucht, zwischen Laub- und Nadelbäumen zu klassifizieren. Als Merkmal wird das Verhältnis von *light crown height* zu *light crown area* benutzt. Dieses ist für Nadelbäume größer als für Laubbäume. Es wird einschränkend darauf hingewiesen, dass dies bei jungen Laubbäumen nicht zutrifft.

### **Bewertung**

Bemerkenswert an diesem Verfahren ist, dass es bei vergleichsweise schwierigen Waldverhältnissen noch brauchbare Ergebnisse liefert. Ebenfalls positiv hervorzuheben ist die Klassifikation in Nadel- und Laubwald.

Der Algorithmus selbst ist aber mit einigen Nachteilen behaftet. Zum einen ist das Maßstabsproblem durch die Unterscheidung von jungen und alten Waldbeständen nur sehr unzureichend gelöst, die Wahl des Glättungsfilters erfolgt weiterhin von Hand. Die Segmentierung durch die Kombination von *pouring* und *ray*-Algorithmus ist umständlich und bringt eigentlich keine Vorteile gegenüber der WST auf einer geeigneten Segmentierungsfunktion. Die größte Schwäche sind aber sicher die zahlreichen Nachverarbeitungsschritte, die vielfältig implizites Modellwissen enthalten und den Anschein erwecken, als sei das Verfahren für den vorliegenden Datensatz ‚getunt‘ worden. Die Übertragbarkeit auf andere Datensätze wird dadurch aber erheblich erschwert. Daher wundert es nicht, dass der Algorithmus auf anderen Datensätzen in der vergleichenden Studie HYYPPÄ *et al.* (2001) eher schlecht abgeschnitten hat.

### 3.9 HIGHSCAN: HYYPPÄ *et al.*

----- <b>Literatur</b> -----	
	HYYPPÄ <i>et al.</i> 2001, HYYPPÄ & HYYPPÄ 1999, SCHARDT <i>et al.</i> 2002, ZIEGLER <i>et al.</i> 2001
Projekt:	HIGHSCAN (Österreich)
----- <b>Zum Verfahren</b> -----	
Steuerung:	Bottom-Up
Modell:	lokale Intensitätsmaxima (blobs); Mindesthöhe der Kronenpunkte: 6m
Segmentierung:	WST auf invertiertem Oberflächenmodell
Maßstabsrepräsentation:	<i>auto-scale selection</i>
Aggregation:	Bestand, Einzelbaum
Extrahierte Parameter:	Direkt: Höhe, Kronenbegrenzung und -fläche, Waldbodenmodell; Indirekt: Stammdurchmesser, Holz- volumen
Datentyp:	LIDAR, first & last pulse
----- <b>Zum Datensatz</b> -----	
Scanner:	TOPOSYS I
Auflösung:	83KHz, 1-20 Pulse/m <sup>2</sup> , Raster: 1m
Landschaftstyp:	alpiner Nadelwald
Baumarten:	hauptsächlich Fichte

Das von der EU finanzierte HIGHSCAN-Projekt lief von 1998-2001. In seinem Rahmen wurden Verfahren zur Forstinventur mittels Laserscannerdaten und hoch auflösenden Satellitenbildern entwickelt und erstmals europaweit verglichen. Dafür wurden insgesamt acht Testgelände in Finnland, Deutschland, der Schweiz und Österreich im Feld erfasst und ausgewertet.

#### Modell

Die Autoren stellen kein explizites Modell auf, ihr Verfahren beruht aber auf der Annahme, dass jeder Baum ein einzelnes Intensitätsmaximum im Oberflächenmodell hat. Da diese Annahme in der Realität nicht erfüllt wird, weil Baumkronen aus zwei oder mehreren Maxima bestehen können, wird vorher mit einem Gaußfilter geglättet (HYYPPÄ *et al.* 2001, SCHARDT *et al.* 2002). Alles was höher als 6m ist, wird als Baum betrachtet.

#### Verfahren

In HYYPPÄ *et al.* (2001: 61f) werden zwar die Ergebnisse von 3 Segmentierungsverfahren verglichen, die Verfahren selbst allerdings nicht näher vorgestellt. Das im Folgenden vorgestellte Verfahren stammt aus SCHARDT *et al.* (2002).

Die Extraktion der Parameter wird auf zwei Aggregationsebenen durchgeführt: Für Bestände und für Einzelbäume. Als Vorbereitung für beide Stufen wird zunächst ein „Waldbodenmodell“ (DGM) erstellt. Dafür kommt ein *multi-resolution*-Verfahren zum Einsatz, bei dem das Modell zunächst in einer groben Auflösung von 10m angelegt und dann mit feineren Auflösungen bis zu 1m verglichen wird. Es beruht auf der Annahme, dass selbst in dichtem Wald in der groben Auflösung wenigstens ein Puls vom Boden kommt, stehen jedoch mehrere zur Verfügung, wird die Auflösung erhöht. Zur

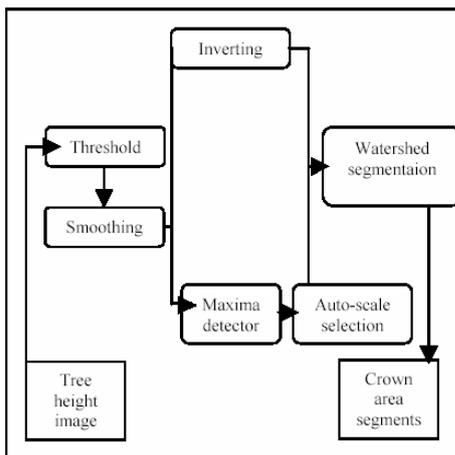
weiteren Verarbeitung wird dann aus der Differenz von DOM (*first pulse*) und Waldbodenmodell ein Kronenmodell erzeugt.

### *Vorverarbeitung*

Zur baumweisen Extraktion wird zunächst ein Schwellenwert von 6m gebildet, da die kleineren Bäume nur in geringerem Maße zum Holzvolumen beitragen (eine sehr starke Modell-Annahme!). Anschließend wird mit einem Gaußfilter geglättet, um nach Möglichkeit doppelte Maxima in Kronen aufzulösen. Da dies aber nicht vollständig möglich ist, werden alle Maxima gesucht – „*maxima-detection*“ (SCHARDT *et al.* 2002: 305) – und auf Signifikanz und Durchmesser überprüft. Der Algorithmus sucht in einem vorgegebenen Bereich für Durchmesser von 1-10m nach dem lokal am besten geeigneten Maßstab – „*auto-scale selection*“ (SCHARDT *et al.* 2002: 305).

### *Segmentierung*

Die Segmentierung wird auf dem invertierten Kronenmodell durchgeführt. Da die WST Probleme mit nicht-minimalen Plateaus hat, wird zunächst ein „*lower completion algorithm*“ (SCHARDT *et al.* 2002: 306) angewendet, der die Plateaus zu Hügeln macht, indem er sie gleichmäßig von der Grenze zum Zentrum hin erhöht. Anschließend wird mit WST segmentiert.



**Abb. 3.9: Verfahren nach SCHARDT *et al.* (2002: 305)**

### *Ableitung weiterer Parameter*

Aus den Segmenten werden direkt die Parameter Kronenfläche (Anzahl der Pixel) und Baumhöhe (Maximale Höhe in einem Segment) ermittelt. Weiterhin ergibt sich die Anzahl der Bäume direkt aus der Segmentierung. Für den Stammdurchmesser auf Brusthöhe wird empirisch anhand von Feldmessungen ein lineares Regressionsmodell mit den beiden bekannten Größen entwickelt. Das Holzvolumen wird schließlich mit einem weiteren Regressionsmodell mit Stammdurchmesser und Höhe abgeleitet.

### **Ergebnisse**

Das Verfahren wurde an der Testfläche Hohentauern evaluiert. Als Referenzdatensatz wurde die Position von 197 Bäumen mit Differential-GPS in einer Genauigkeit von 0.1m aufgenommen. Damit wurde eine Vollständigkeit von 50% erzielt. Die Hauptfehlerquelle war mit 45% der Fälle, dass mehrere Kronen zu einer zusammengefügt wurden. Dass eine Krone als mehrere erkannt wurde, kam

dagegen nur in knapp 5% der Fälle vor. Die Anzahl der Bäume wurde bei einer Analyse von 23 Beständen systematisch unterschätzt, so dass im Mittel nur 54% gefunden wurden. Die Ergebnisse für die abgeleitete Größe Holzvolumen sind aber deutlich besser, da sich die Fehler teilweise kompensieren.

### Bewertung

Ausgereift erscheint das *multi-resolution*-Verfahren zur Erstellung des DGM. Die Segmentierung mit WST auf dem invertierten Höhenmodell als Segmentierungsfunktion ist vergleichsweise einfach und führt zu Fehlern, wenn zwischen den einzelnen Bäumen Lücken sind. Das Verfahren zur Auswahl des optimalen Maßstabs scheint brauchbar, ist in SCHARDT *et al.* (2002) aber leider nur sehr knapp dokumentiert.

Die vorverarbeitende Segmentierung in Baum und Nicht-Baum durch Bildung eines Schwellenwertes von 6m mag zwar für die Schätzung des Holzvolumens vertretbar sein (da kleinere Bäume hierzu nicht maßgeblich beitragen), als Modell ist es aber unzureichend.

Die Systematik der Fehler (viele zusammengefügte Kronen, selten doppelte Extraktion einer Krone) könnte zum einen darauf hindeuten, dass die Auflösung der Laser-Daten nicht ausreichend war, um alle Maxima einzeln zu erkennen. Zum anderen könnte es aber auch für eine zu starke Glättung sprechen. Bei der Volumenschätzung gleichen die größeren Kronen die fehlenden Bäume größtenteils aus. Da Kronenbedeckung und Höhe ins Regressionsmodell zur Vorhersage des Holzvolumens eingehen, könnte dieses aber auch entsprechend falsch geeicht sein. Aus SCHARDT *et al.* (2002: 304) geht nicht hervor, wie die entsprechenden Formeln gewonnen wurden.

Die Ergebnisse sind im Vergleich mit anderen Verfahren eher unterdurchschnittlich, was jedoch auch an der Auflösung der DKM liegen kann. HYYPPÄ ET AL (2001: 64f) legen allerdings nahe, dass die Pulsdichte keine all zu großen Auswirkungen auf die Genauigkeit der Ergebnisse hat (wohl aber auf die Vollständigkeit).

## 3.10 Hill-Climbing nach PERSSON

### ----- Literatur -----

PERSSON 2001, PERSSON *et al.* 2002, HEURICH *et al.* 2004

### ----- Zum Verfahren -----

Steuerung:	Bottom-Up
Modell:	lokale Maxima, Paraboloid (ähnlich Pollock), Höhe des Kronenrandes bei 2m
Segmentierung:	<i>Hill-Climbing</i>
Maßstabsrepräsentation:	3 Maßstabsebenen. Beste Ebene wird durch lokale Annäherung mit einer parabolischen Fläche bestimmt.
Aggregation:	Einzelbaum
Extrahierte Parameter:	Direkt: Position, Höhe, Kronendurchmesser, abgeleitet: Stammdurchmesser, Holzvorrat
Datentyp:	LIDAR (First Pulse, Last Pulse)

### ----- Zum Datensatz -----

Scanner:	TopoEye, 7000 Hz, 1064nm
Auflösung:	26-368cm ( <i>footprint</i> ); Raster: 1/3m

---

Untersuchungsgebiet:	Südschweden (58°30'N, 13°40'O, 120-145m ü. NN)
Landschaftstyp:	borealer Wald
Baumarten:	Gemeine Fichte ( <i>Picea abies</i> L.), Waldkiefer ( <i>Pinus sylvestris</i> L.) und Birke ( <i>Betula</i> )

## Verfahren

### *Vorverarbeitung*

Zunächst werden aus den höchsten und niedrigsten Punkten in jeder Rasterzelle die Oberflächenmodelle  $DSM_{\min}$  und  $DSM_{\max}$  gebildet. Aus  $DSM_{\min}$  wird mit Hilfe von aktiven Konturen (siehe auch 3.7) ein DGM gebildet. Derselbe Algorithmus wird verwendet, um die Oberfläche genauer zu bestimmen. Damit sollen Pulsantworten aus dem Inneren des Baumes entfernt werden. Aus der Differenz der beiden wird ein Kronenmodell erstellt. Von diesem werden mit Hilfe eines Schwellenwertes alle Pixel unter 2m verworfen. Da bei der Erstellung des DOM bei baumnahen Pixeln zwischen Boden und Krone interpoliert wurden, wird mit einem 5x5 Median-Filter nachbearbeitet.

Anschließend wird die Oberfläche mit drei verschiedenen Gaußfiltern ( $\sigma = 4/\pi$ ,  $6/\pi$  und  $8/\pi$ ) geglättet. Die Maßstäbe sind so gewählt, dass im groben Maßstab kein Baum mehr als ein Maximum hat und im feinen Maßstab möglichst jeder Baum wenigstens eines.

### *Segmentierung*

Zur Segmentierung werden lokale Maxima gesucht. Dafür werden in allen Pixeln über 2m Saatpunkte gesetzt. Von dort steigen sie entlang der stärksten Steigung auf, bis sie ein lokales Maximum erreichen. Alle Punkte, die zu einem Maximum führen, bilden ein Segment. Die Segmentierung wird in allen drei Maßstabsebenen durchgeführt.

Um die optimale Maßstabsebene in verschiedenen Bereichen des Bildes auszuwählen, werden die 30% höchsten Punkte eines Segmentes durch einen Paraboloid angenähert.

Hat in einem Segment der gröberen Ebene die jeweils feinere Maßstabsebene auch nur ein Maximum, so wird der Baum als korrekt betrachtet. Hat die feinere Ebene mehrere Maxima, bestehen die Möglichkeiten, dass mehrere Maxima in einem Baum detektiert wurden oder dass in der größeren Ebene mehrere Bäume verschmolzen sind. Um zu entscheiden, ob zusätzliche Maxima auf der feineren Ebene als einzelne Bäume gewertet werden, werden für beide Fälle Paraboloid angepasst. Liegt die Summe der Residuen auf der feineren Ebene um mindestens 8% unter der gröberen, wird ein zusätzlicher Baum gewertet. Auf diese Weise werden zunächst die grobe mit der mittleren und anschließend die mittlere mit der feinen Ebene verglichen.

### *Schätzung der Parameter*

Als Baumhöhe wird die maximale Höhe in einem Segment angesetzt. Der Kronendurchmesser wird aus der Fläche ermittelt, als sei diese kreisförmig. Daraus werden die weiteren Parameter Stammdurchmesser und Holzvorrat mit Hilfe gängiger Modelle abgeleitet.

## Ergebnisse

Zum Vergleich wurden als Referenzdaten 12 Plots von 50 mal 20m im Feld erhoben, in denen die relative Position der Einzelbäume zu zwei genau bekannten Referenzpunkten gemessen wurde. Für eine zufällige Auswahl von etwa 15 Bäumen je Plot wurden außerdem Höhe und Kronendurchmesser

erfasst. Lag ein Referenzbaum in einem Segment, wurde er diesem zugeordnet, waren es mehrere wurde der nächstliegende zugeordnet. Für nicht zugeordnete Referenzbäume wurde in einem Radius von zwei Pixeln nach freien Segmenten gesucht. Zugeordnete Bäume, deren Höhen zu stark abweichend, wurden als falsch zugeordnet aussortiert.

Insgesamt wurden 71% der Bäume erkannt, wobei vorwiegend kleinere Bäume in Nachbarschaft größerer unerkannt blieben. Die schlechteste Rate ergab sich im jungen, dichten Kiefernwald, die zweitschlechteste im Kiefernwald mit zwei Stockwerken. Damit wurden 91% des Stammvolumens gefunden. Die durchschnittliche Abweichung der Position betrug 51cm. Etwas schlechter wurde der Kronendurchmesser extrahiert, bei der Höhe kam es wie üblich zu einer systematischen Unterschätzung von 1.13m. Beim abgeleiteten Stammdurchmesser lag der Fehler etwas höher aber immer noch unter 10%.

Außerdem wurde der Einfluss der Strahlweite auf die Ergebnisse untersucht. Während mit dem kleinsten Strahl die Erkennungsrate etwas höher war, wurde die Genauigkeit von Höhe und Kronendurchmesser kaum beeinträchtigt.

In HEURICH *et al.* (2004) wurde das Verfahren in Kooperation mit der TU München und der Nationalparkverwaltung Bayrischer Wald auf andere Waldarten übertragen.

---

#### Literatur

---

HEURICH *et al.* 2004

---

#### Zum Datensatz

---

Scanner:	TopoSys II, 1560nm
Auflösung	0.8m ( <i>footprint</i> )
Untersuchungsgebiet:	Bayrischer Wald
Landschaftstyp:	subalpiner Fichtenwald, über 1100m ü. NN [1], Mischwald, 600-1100 m ü. NN [2] und Fichtenwald [3]
Baumarten:	Gemeine Fichte ( <i>Picea abies</i> L.: 1,2,3) und Eberesche ( <i>Sorbus aucuparia</i> L.: 1,3), Weißtanne ( <i>Abies alba</i> L.: 2), Rotbuche ( <i>Fagus sylvatica</i> L.: 2) und Bergahorn ( <i>Acer Pseudplatanus</i> : 2); und Birken ( <i>Betula pendula</i> Roth, <i>Betula pubescens</i> Ehrh.: 3)

In diesen Gebieten wurden 44% aller Bäume richtig erkannt, wobei Nadelbäume mit 50.6% gegenüber 38.4% bei Laubbäumen deutlich besser abschnitten. Noch deutlicher ist die Differenzierung nach Schichten, wo in der oberen 67.9% gegenüber 5.9% und 2.5% in der mittleren bzw. unteren Schicht entdeckt wurden. Das geschätzte Holzvolumen lag mit 80.5% für Laubbäume und 88.1% für Nadelbäume deutlich besser. Die Höhe wurde im Schnitt nur um 55cm unterschätzt.

### Bewertung

Das Verfahren ist insgesamt überzeugend. Besonders bemerkenswert ist die Bestimmung der Oberflächenmodelle durch aktive Konturen. Die Segmentierung ist dem *pouring*-Algorithmus ähnlich aber einfacher zu berechnen. Eine Abweichung ist, dass die Bäume in einer Höhe von 2m abgeschnitten werden. Das mag für viele Waldsituationen funktionieren, ist aber wie schon bei HIGHSCAN eine starke Modellannahme, die nicht explizit angegeben wird. Immerhin bedeutet dies, dass alles über 2m als Baumkrone angesehen wird und alles darunter als Nicht-Baumkrone. Insbesondere haben

Kronenränder immer hie Höhe 2m. Da die Segmente nicht weiter klassifiziert werden, ist diese Segmentierung außerhalb von dichten Wäldern sicherlich unbrauchbar.

Das Problem des optimalen Maßstabs ist mit der lokalen Anpassung sehr überzeugend gelöst, die Beschränkung auf drei Maßstabsebenen scheint ausreichend. Allerdings müsste überprüft werden, ob durch die Wahl der Parameter der aktiven Konturen bei der Erstellung des Oberflächenmodell (beispielsweise Elastizität) nicht schon bestimmte Maßstäbe bevorzugt werden.

Die Vollständigkeit der Extraktionsergebnisse ist im mittleren Bereich und für die schwierigeren Waldsituationen in HEURICH *et al.* (2004) zumindest noch vielversprechend. Die Zuordnung der Bäume zu den Referenzbäumen ist dabei allerdings wesentlich großzügiger, als in 3.1.4 vorgeschlagen. So wurden auch Referenzbäume die bis zu 2 Pixeln außerhalb des extrahierten Segments liegen, als Übereinstimmung gewertet. Aus forstwirtschaftlicher Sicht ist dies nachvollziehbar, da die genaue Position der Bäume sicher nicht der wichtigste Parameter ist. Es erschwert aber eine vergleichbare Evaluation. Die Genauigkeit der extrahierten Positionen ist ausgesprochen hoch. Die durchschnittliche Abweichung betrug nur 51cm, was insbesondere in Anbetracht der Tatsache, dass hier Stammpositionen mit Positionen von Baumspitzen verglichen wurden, als ausgezeichnet bewertet werden kann.

Sehr gelungen ist außerdem die Auswertung der Charakteristik nicht erkannter Bäume (PERSSON *et al.* 2002: 929).

### 3.11 Bandpassfilter nach KÖRBER, MÖLLER & KÄTSCH

----- <b>Literatur</b> -----	
	KÖRBER <i>et al.</i> 2005, MÖLLER <i>et al.</i> 2005
Software:	MATLAB
----- <b>Zum Verfahren</b> -----	
Steuerung:	Bottom-Up
Modell:	homogen Textureigenschaften innerhalb von Baumkronen
Segmentierung:	im Merkmalsraum
Maßstabsrepräsentation:	keine
Aggregation:	Einzelbaum
Extrahierte Parameter:	Kronenbegrenzung
Klassifikation:	k-Means Clustering
Datentyp:	Schwarz-Weiß-Foto als Graustufen-Bild
----- <b>Zum Datensatz</b> -----	
Untersuchungsgebiet:	Süd-West-Niedersachsen
Landschaftstyp:	Mischwald, 40-120 Jahre
Baumarten:	v.a. Eiche, Buche und Fichte

#### Verfahren

Der Ansatz basiert auf einem zweistufigen Filteralgorithmus. Im ersten Durchgang, der *windowed multi-band signature analysis*, wird zu jedem Pixel eine Signatur erzeugt, die die Frequenzverteilung in einem Fenster um das Pixel repräsentiert. Im zweiten Schritt wird mit einem *k-means clustering*

klassifiziert, so dass eine Überpartitionierung entsteht. Diese wird anschließend in einem manuellen Schritt zu Klassen zusammengefasst.

### *Texturanalyse und Signaturerzeugung*

Die Textureigenschaften werden im Frequenzraum (siehe 2.3.7) untersucht. Hierzu wird zunächst um jedes Pixel ein Fenster mit 7 bis 25 Pixeln ausgeschnitten, das in die Analyse für diesen Bildpunkt eingeht. Das ausgeschnittene Fenster wird mit einem Filter multipliziert, der am Rande auf Null abfällt und bessere Eigenschaften bezüglich der Fouriertransformation hat, als ein Rechteckfenster (die Multiplikation mit einem Rechteck im Ortsraum entspricht der Faltung mit einer sinc-Funktion im Frequenzraum, siehe 2.3.7). Zusätzlich wird ein *zero-padding* durchgeführt, um numerische Fehler zu vermeiden.

Das so gefilterte Fenster wird 2D diskret fouriertransformiert. Aus dem räumlichen Spektrum werden anschließend mit einer Reihe von idealen Bandpass-Filtern verschiedene Frequenzbereiche ausgeschnitten. Der Radius jedes Filters ist  $\sqrt{2}$  mal der seines Vorgängers, was ein weiterer Grund ist, warum das Spektrum durch vorheriges *zero-padding* vergrößert wurde. Aus den gefilterten Bändern wird die Signatur des Pixels erzeugt, indem die Mittelwerte der einzelnen Bänder zu einem Vektor zusammengefügt werden.

### *Klassifikation (Segmentierung)*

Im zweiten Schritt wird mit Hilfe der gewonnenen Texturmerkmale eine Klassifikation durchgeführt. Dazu wird zunächst mit einem unüberwachten *k-Means-Clustering* Algorithmus (siehe 2.2.5) eine bewusste Überpartitionierung des Merkmalsraums erzeugt. Es werden also quasi natürliche Klassen im Merkmalsraum gesucht, die aber feiner unterteilt sind, als die gesuchten Klassen. Anschließend werden die gefundenen Klassen manuell zu größeren verschmolzen, die dann reale Klassen repräsentieren sollen (also verschiedene Baumarten sowie eine Nicht-Baum Klasse). Es wird vorgeschlagen, den halb-automatischen Schritt perspektivisch durch ein neuronales Netz zu ersetzen. Es ist jedoch fraglich, ob damit vergleichbare Ergebnisse erzielt werden könnten, und der Aufwand für den manuellen Schritt ist unter Voraussetzung einer einfach zu bedienenden Benutzerschnittstelle nicht sehr groß.

Die Klassifikation im Merkmalsraum erzeugt automatisch auch eine Segmentierung im Bild, bei der die räumlichen Eigenschaften der Pixel jedoch keine Berücksichtigung finden (siehe 2.4.6). Dies wird durch menschliche Intelligenz bei der Verschmelzung der Klassen ausgeglichen, indem im Bild räumlich nahe beieinander liegende Klassen, die offensichtlich zu einem Objekt gehören, zusammengefügt werden.

### **Bewertung**

Bemerkenswert an dem Verfahren ist die Texturanalyse im Frequenzraum, die wie die Ergebnisse nahe legen, herkömmlichen Texturparametern (Semivarianz, Texturenergien, etc.) überlegen ist. Dies müsste weiter überprüft werden. Damit werden mächtige Merkmale gewonnen, die auch in anderen Segmentierungs- oder Klassifikations-Algorithmen Verwendung finden könnten.

Zur Segmentierung selbst ist das Verfahren allerdings nicht optimal, da ausschließlich im Merkmalsraum gruppiert wird und die räumlichen Eigenschaften erst mit Hilfe eines manuellen Verarbeitungs-

schritts integriert werden. Die Texturanalyse könnte beispielsweise in ein auch im Ortsraum arbeitendes Segmentierungsverfahren wie RG integriert werden. Da es hier aber auch andere aussichtsreiche Verfahren gibt, scheint es sinnvoller, diese Merkmale für den nachfolgenden Schritt der Klassifizierung nach Baumarten aufzuheben. Hierfür scheint die Texturanalyse im Frequenzraum ein Erfolg versprechendes Diskriminierungspotential zu bieten.

### 3.12 Region-Growing nach ERIKSON

----- <b>Literatur</b> -----	
	ERIKSON 2003, ERIKSON 2004A, ERIKSON 2004B, ERIKSON & OLOFSSON 2005, ERIKSON <i>Et Al.</i> 2006
----- <b>Zum Verfahren</b> -----	
Steuerung:	Bottom-Up
Modell:	Farbvariation innerhalb von Kronen ist geringer als zwischen verschiedenen Kronen; dunkle Täler zwischen Kronen; <i>six tree conditions</i> (ERIKSON 2003: 1559)
Segmentierung:	verschiedene <i>Region-Growing</i> -Verfahren
Maßstabsrepräsentation:	keine; erwarteter Baumradius geht teilweise in Segmentierung ein; zur Bestimmung der Saatpunkte wird ein Glättungsfilter verwendet.
Aggregation:	Einzelbaum
Extrahierte Parameter:	genaue Kronenbegrenzung, -fläche
Klassifikation:	vier Baumarten
Datentyp:	CIR
----- <b>Zum Datensatz</b> -----	
Auflösung:	sehr hoch (3-15cm)
Untersuchungsgebiet:	Zentralschweden [1], Südwestschweden [2], Ecuador [3] (siehe ERIKSON 2004a: 16)
Landschaftstypen:	borealer Nadelwald [1,2], tropischer Wald [3]
Baumarten:	Gemeine Fichte ( <i>Picea abies</i> L.), Waldkiefer ( <i>Pinus sylvestris</i> ) [1,2]; Moor-Birke ( <i>Betula pubescens</i> Ehrh.) und Espe ( <i>Populus tremula</i> ) [1]

#### Verfahren

ERIKSON vergleicht insgesamt drei verschiedene *Region-Growing*-Verfahren miteinander. Relativ ähnlich ist dabei jeweils die Bestimmung der Saatpunkte. Dazu wird jeweils ein Schwellenwert des Infrarot-Bandes gebildet (0.3 bei ERIKSON 2004A ERIKSON 2003, 0.2 bei ERIKSON *et al.* 2006 und 0.15 bei ERIKSON & OLOFSSON 2005). Vom Schwellenwertbild wird eine Distanz-Transformation nach BORGEFORS (1986) gebildet, und die Maxima im Distanzbild sind die Saatpunkte. Diese werden in Reihenfolge ihrer Werte im Distanzbild sortiert, das quasi den Abstand zum nächsten Schatten darstellt.

In ERIKSON 2003 und ERIKSON 2004A werden die Saatpunkte anschließend mit Hilfe eines *Fuzzy-Threshold*-Verfahren nach CHI *et al.* (1996: 57ff) neu bestimmt. Damit wird ein minimal-entropischer Schwellenwert ermittelt, von dem ausgehend dann lange konvexe Kantenstücke extrahiert werden. Aus deren Krümmungseigenschaften ergeben sich wiederum die neuen Saatpunkte. In ERIKSON (2004B: 471) wird dagegen alternativ einfach das Distanzbild mit einem starken Gaußfilter (entspre-

chend der erwarteten mittleren Kronengröße, die damit als *a priori* Modellwissen ins Verfahren eingeht) geglättet, um die Anzahl der Saatpunkte zu reduzieren.

#### Segmentierung mit Fuzzy-Regeln

Bei der Segmentierung mit Fuzzy-Regeln wird ein Ähnlichkeitsmaß  $\mu$  fürs RG ermittelt.

$$(3.5) \quad \mu = \exp\left(-\frac{1}{2}\left(\frac{i_d^2 + r_d^2 + g_d^2}{\sigma_1^2} + \frac{x_d^2 + y_d^2}{\sigma_2^2}\right)\right)$$

Dabei sind  $i_d$ ,  $r_d$  und  $g_d$  die Abstände zwischen dem aktuellem Pixel und dem Saatpunkt im Farbraum und  $x_d$  und  $y_d$  die Abstände im Ortsraum. Der Parameter  $\sigma_1$  bestimmt die zulässige Farb-Varianz eines Segments. Der Parameter  $\sigma_2$  bestimmt seine maximale Größe und lässt sich aus der Krümmung der Krümmung der Kantenpixel beim Fuzzy-Schwellenwert ableiten (ERIKSON 2003: 1559).

Ist das Ähnlichkeitsmaß kleiner als ein vorgegebener Wert  $\alpha$ , wird das Pixel der Region zugeordnet. Regionen ab einer bestimmten Größe gelten als Baumkandidaten. Zur besseren Abgrenzung wird das RG mit jedem Pixeln der Region als Startpunkt wiederholt und somit weitere Baumhypothesen. Als beste Hypothese wird diejenige bestimmt, bei der der Mittelwert der Intensität Randpixel am niedrigsten ist (hier geht die Annahme von Tälern zwischen den Bäumen ein). In ERIKSON 2003 (1559f) werden zudem verschiedene Bedingungen an die Größenverhältnisse und Überdeckungsgrade zwischen dem Segment, seiner konvexen Hülle und dem ursprünglichen Schwellenwertbild gestellt („*six tree conditions*“: ERIKSON 2003: 1559). Anschließend werden alle Saatpunkte innerhalb der Region aus der Liste gelöscht.

#### Segmentierung mit Brownscher Bewegung

Bei der diesem Segmentierungsverfahren (ERIKSON 2004A: 26ff, ERIKSON 2004B: 471f) wird die Brownsche Bewegung eines Moleküls in einer Flüssigkeit simuliert. Dafür werden ausgehend vom Saatpunkt normalverteilte Vektoren addiert und gezählt, wie oft das virtuelle Teilchen die einzelnen Pixel besucht. Um zu verhindern, dass es sich in den dunklen Hintergrund bewegt, wird der Vektor mit einer Funktion des Grauwertes des eigentlichen Zielpunktes gewichtet. Aus dem bis zu mehreren Millionen Schritten langen Pfad wird ein *numPass*-Bild erstellt. Von diesem wird ein Schwellenwert bei eins gebildet.

Darauf aufbauend wird mit einem Expansions-Algorithmus segmentiert. Dafür werden die Regionen mit den Saatpunkten initiiert, dann wird jeder Region ein Kreis von Punkten zugeordnet, der gerade ins Schwellenwertbild passt. Der entsprechende Radius wird wieder durch eine Distanz-Transformation nach BORGEFORS (1986) bestimmt. Die weiteren Pixel werden in Reihenfolge ihrer *numPass*-Werte den Segmenten zugeordnet. Dafür wird jeweils eine Liste aller Pixel geführt, die momentan genau eine Nachbarregion haben.

#### Segmentierung mit Random Walks

Eine leichte Abwandlung der Brownschen Bewegung stellt die Methode der Random Walks (ERIKSON 2004A: 28f, ERIKSON *et al.* 2006) dar. Das Ziel ist dabei, den Kontrast zwischen Tälern und Baumkronen zu erhöhen. Dafür werden statt den normalverteilten nun gleichverteilte Zufallsvektoren verwendet und nur ganzzahlige Schritte zugelassen. Weiterhin wird vor jedem Schritt eine Nachbarschaft von

potentiellen Zielen gefunden, aus der eine intendierte Position zufällig ausgewählt wird. Der Schritt wird aber nur ausgeführt (und der *numPass*-Wert des Zielpixels inkrementiert), wenn der Grauwert des Zielpixels unter einer weiteren gleichverteilten Zufallszahl zwischen Null und dem höchsten Grauwert der Nachbarschaft liegt. Die Nachbarschaft wird so bestimmt, dass die Summe ihrer Grauwerte gerade unter einer vorgegebenen Grenze  $T$  liegt:

$$(3.6) \quad \min_v \sum_{q \in N_v} f(q) > T.$$

Je größer diese Grenze  $T$  ist, desto dunkler wird das *numPass*-Bild am Rand der Segmente. Bei ERIKSON *et al.* 2006 wird weiter spezifiziert, dass die Nachbarschaft symmetrisch sein soll.

Die Segmentierung durch Expansion erfolgt wie bei der Brownschen Bewegung.

### Klassifikation

Im Gegensatz zu den meisten anderen Verfahren nimmt ERIKSON nach der Segmentierung auch eine Klassifikation der vier wichtigsten Baumarten Gemeine Fichte (*Picea abies L.*), Waldkiefer (*Pinus sylvestris L.*), Moor-Birke (*Betula pubescens Ehrh.*) und Espe (*Populus tremula L.*) vor. Dabei wird jedes Segment klassifiziert, es wird also davon ausgegangen, dass alle gefundenen Segmente Bäume sind.

Zu diesem Zweck werden Merkmale verwendet, die auf bestimmten Eigenschaften der Erscheinung der jeweiligen Baumart beruhen. Diese werden in einer festgelegten Reihenfolge abgearbeitet. Erfüllt ein Segment also die Kriterien für Birke, wird es als Birke klassifiziert. Andernfalls wird als nächstes überprüft, ob es die Kriterien für Espe erfüllt, dann Fichte und zuletzt Kiefer. Nicht klassifizierte Bäume werden zunächst nach Größe und dann mit einem MED-Klassifikator zugeordnet.

Für die genauen Kriterien sei auf ERIKSON (2004B: 472ff) verwiesen, an dieser Stelle sollen nur einige Grundzüge wiedergegeben werden. Für Birke wird ein gewisser Anteil weißer und (infra-)roter Pixel vorausgesetzt. Für Espe werden weiße *blobs* bei verschiedenen Schwellenwerten gesucht. Bei Kiefern wird nach dem Schatten gesucht, den der Baum durch seine konische Form auf sich selbst wirft. Dafür wird das Segment zu seiner konvexen Hülle erweitert und die größte Ergänzung gesucht. Liegt diese gegenüber der Sonneneinfallrichtung, wird sie als Schatten gewertet. Geht ein solcher Schatten bis nah an die Baumspitze, wird der Baum als Fichte klassifiziert, anderenfalls als Kiefer.

### Ergebnisse

Die Verfahren wurden an verschiedenen Szenen mit sehr hoher Auflösung getestet, unter anderem Hulten in Zentral-Schweden, Remningstorp in Süd-West-Schweden, und Tiputini Biodiversity Station in Ecuador. Die Ergebnisse sind sehr gut: Während Fuzzy-Segmentation beim Remningstorp-Bild schon fast 80% der Bäume korrekt erkannte, schafften Brownian-Motion und Random-Walks jeweils 95%. Random Walks lieferte die beste Abgrenzung zwischen Baumkronen.

Ebenfalls als gut zu bewerten sind die Klassifizierungsergebnisse, die im Schnitt bei 76% Korrektheit lagen aber nach Baumarten differierten. Da sich falsch klassifizierte Bäume dabei noch teilweise gegenseitig aufheben, lag die geschätzte Artenzusammensetzung sogar sehr nah an der Realität.

## Bewertung

Die Segmentierungsverfahren gehören sicherlich zu den Besten, die für optische Bilder verfügbar sind. Ein großer Vorteil ist, dass die Segmentierung direkt im Bild und nicht auf einer Maßstabsebene durchgeführt wird. Dadurch lassen sich Merkmale für die anschließende Klassifikation gewinnen, die sonst verloren gingen.

Das Problem des optimalen Maßstabs wird aber auch hier nicht gelöst. Auch wenn es für das RG selbst keine Rolle spielt, wird es auf die vorangehende Bestimmung der Saatpunkte verschoben. Da aus jedem Saatpunkt ein Baum wird, ist deren zuverlässige Bestimmung elementar. Bei den einfachen borealen Waldsituationen, genügt es, manuell einen geeigneten Glättungsfilter zu wählen. Für die komplizierte Situation in Ecuador werden jedoch die kompletten Saatpunkte manuell ausgewählt, worin ein Großteil der benötigten Intelligenz liegt.

Weiterhin könnten Details verbessert werden. Unverständlich ist beispielsweise, warum die Distanztransformationen in einer 3-4-Metrik statt der üblichen euklidischen durchgeführt werden. Das Kanten-basierte *fuzzy threshold* Verfahren zur Bestimmung der Saatpunkte in ERIKSON 2004A(23f) ist umständlich und unverständlich, wurde aber in den folgenden Veröffentlichungen bereits abgeändert. Außerdem könnte überlegt werden, inwiefern ein abgewandeltes RG-Verfahren, etwa mit einer *centroid linkage* Strategie zu Verbesserungen führen könnte.

Das Objekt-Modell für die Bäume ist nicht sonderlich ausgereift. Allerdings gehen Annahmen über die Baumgröße implizit in die Parameterwahl mit ein. Weiterhin werden die Annahmen getroffen, dass die Variation der Farbe innerhalb der Baumkrone geringer als zwischen Baumkronen und das dunkle Täler zwischen den Kronen existieren. Die Modelle der einzelnen Baumarten für die Klassifikation sind dafür aber sehr ausgereift, wenn allerdings auch von äußeren Parametern – z.B. Sonnenstand – abhängig. Trotzdem ist es als absolute Besonderheit hervorzuheben, dass neben der Extraktion auch ein vielversprechender Algorithmus zur Klassifikation von vier Baumarten geliefert wird.

### 3.13 Vergleichende Ansätze

Alle vorgestellten Verfahren arbeiten auf spezifischem Bildmaterial mit spezifischen Waldsituationen, daher sind ihre Ergebnisse nur in den wenigsten Fällen übertragbar. In jüngster Zeit setzt sich aber die Erkenntnis durch, dass einzelne Verfahren nicht ausreichen und daher vergleichende Studien notwendig sind: *“It is most likely that none of today’s methods are [sic] alone able to handle all types of forests. Comparative studies of different segmentation methods with different types of forests are therefore of importance in order to clarify how much a method is reliable with respect to a certain type of forest. [...] Comparative studies among the existing methods are rare so far”*. (ERIKSON *et al.* 2006: 1)

ERIKSON kann als Wegbereiter auf diesem Gebiet bezeichnet werden. In ERIKSON & OLOFSSON (2005) werden zwei seiner RG Verfahren mit einem *Template-matching* Verfahren nach OLOFSSON verglichen und in ERIKSON *et al.* (2006) werden drei verschiedene Extraktionsverfahren an unterschiedlichen Datensätzen getestet.

Im LIDAR-Bereich ist hier vor allem das HIGHSCAN-Projekt zu nennen, bei dem verschiedene Segmentierungs-Algorithmen an insgesamt acht Testgeländen in vier Ländern getestet wurden. In

HEURICH *et al.* (2004) wird der von PERSSON für skandinavische Nadelwälder entwickelte Segmentierungsalgorithmus auf verschiedene Waldtypen in Deutschland übertragen.

Dennoch besteht erheblicher Bedarf an weiteren vergleichenden Studien und Forschungsarbeiten auf diesem Gebiet. Dabei geht es um zweierlei: Erstens muss die Leistungsfähigkeit von bestehenden Verfahren an anderen Waldsituationen und Bildmaterial getestet und verglichen werden. Wäre zu jeder gegebenen Situation der beste Algorithmus bekannt, ließen diese sich in einem Expertensystem zusammenfassen, „*which is supposed to be able to extract individual tree crowns in any type of forests*“ (ERIKSON *et al.* 2006: 1).

Zweitens ist neben vergleichenden Studien mit dem Ziel der der bloßen *cross-validation* von Verfahren auch eine konzeptionelle Analyse der Übertragbarkeit von Verfahren nötig. Dazu soll diese Arbeit einen Beitrag leisten. Es ist die Überzeugung des Autors, dass die oftmalige Konzeption von Verfahren an konkreten Datensätzen eine Hypothek für deren Übertragbarkeit darstellen wird. Daher ist es dringend nötig zu explizieren, welche Modellannahmen an welcher Stelle in ein Verfahren eingehen. In diesem Kapitel wurde bereits darauf hingewiesen, dass dies von vielen Autoren vernachlässigt wird und an einzelnen Stellen wurde es bereits nachgeholt. Denn nur so ist zu überprüfen, ob die Annahmen für einen neuen Datensatz ebenfalls erfüllt sind und ob das Verfahren somit prinzipiell für den Datensatz verwendbar ist.

Daher soll im folgenden Kapitel den Fragen nachgegangen werden, wie ein zukünftiges Expertensystem aufgebaut und gesteuert werden sollte, welche Annahmen eingehen, welche Parameter zur Steuerung und welche Verfahren zur Integration und Weiterentwicklung geeignet sind. Anschließend wird exemplarisch gezeigt, wie Verfahren an neue Situationen angepasst werden können, und verschieden Segmentierungsverfahren werden an neuen Datensätzen verglichen.

## 4 Ein Expertensystem zur Baumextraktion

Ein Expertensystem ist „*ein Computersystem, das in einem spezialisierten Bereich über ausreichendes Wissen verfügt und fähig ist, dieses Wissen zur Lösung von Aufgabenstellungen eines Benutzers auf einem einem [sic] Experten entsprechenden Niveau anzuwenden*“ (HORN 1984 nach PINZ 1994).

In diesem Kapitel sollen grundlegende Vorschläge für den Aufbau und die Architektur eines Expertensystems zur automatischen Extraktion von Bäumen gemacht werden. Allgemein wird unter einem Expertensystem ein Software-System verstanden, das mit Hilfe von Expertenwissen bestimmte Problemstellungen löst. Im konkreten Fall soll im Folgenden darunter ein System zur automatischen Extraktion von Bäumen sowie relevanter Parameter aus Fernerkundungsdaten verstanden werden.

Damit ein solches Expertensystem gegenüber den vorhandenen Verfahren einen Mehrwert bietet, ist die Hauptanforderung, dass es möglichst universell sein soll, oder anders ausgedrückt, eine größere Menge von Aufgabenstellungen (im Folgenden auch Problemstellung) bearbeiten kann. Unter einer Problemstellung werden im Folgenden ein neuer Fernerkundungsdatensatz mit Bäumen plus die Spezifikation, also die Anforderungen an das gewünschte Ergebnis und die zu extrahierenden Größen, verstanden. Die zu extrahierenden Größen werden im Folgenden auch Extraktions-Parameter genannt, sind aber nicht mit den Steuerungsparametern der Verfahren zu verwechseln.

Die Anforderungen an das Ergebnis hängen wesentlich von der Anwendung des Systems ab. Für die Forstwirtschaft sind die wichtigsten Parameter nach ERIKSON (2004: 8) Stammvolumen, Alter, Artenzusammensetzung und ökologische Werte wie Schlüsselbiotope und Lebensräume. Daneben sind aber ganz unterschiedliche Anwendungen von Umwelt-Monitoring über Feuerbekämpfung bis hin zu 3D-Stadtmodellen denkbar, die jeweils eigene Parameter erfordern. Je nachdem, welcher Parameter extrahiert werden soll, stellen sich unterschiedliche Anforderungen an das Expertensystem. Bei der Konzeption sollte vorab bedacht werden, welche Bereiche abgedeckt werden können und sollen.

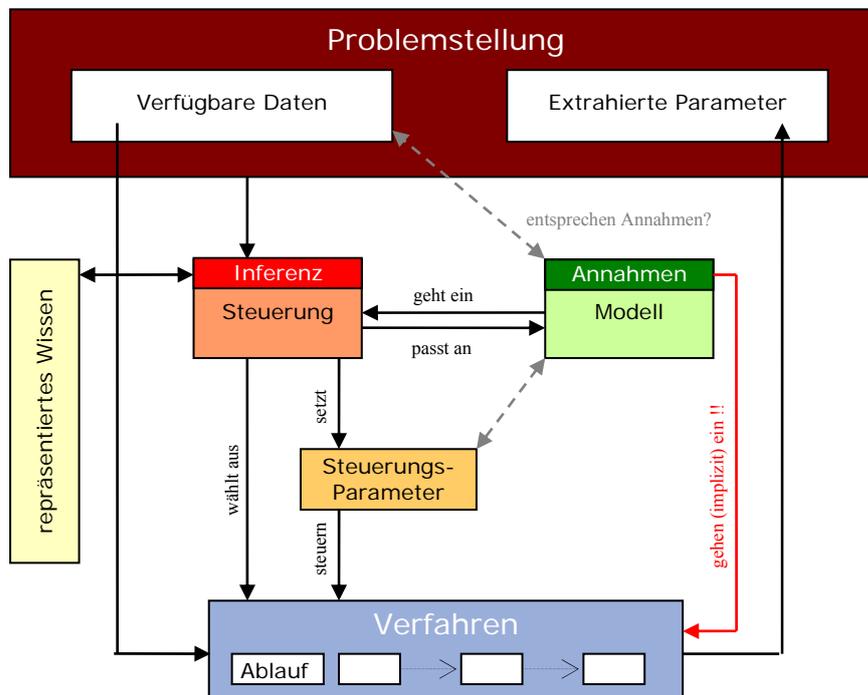
In der künstlichen Intelligenz werden anhand Ihrer Strategie zur Problemlösung verschiedene Arten von Expertensystemen unterschieden. Fallbasierte Systeme versuchen neue Probleme auf bekannte, schon gelöste Probleme zurückzuführen und analog zu lösen. In Bezug auf die Baumextraktion hieße dies, aus einem Katalog von bekannten Verfahren (siehe Kapitel 3) das für ein neues Problem am besten geeignete auszuwählen und anzuwenden. Regelbasierte Systeme (*rule-based reasoning*) bestehen dagegen aus Regeln und einem Schlussfolgerungsalgorithmus. Die Regeln werden mit Elementarlogik abgearbeitet und führen zu einer Entscheidung. Ein Beispiel für ein solches System ist das VES in PINZ (1989).

### 4.1 Aufbau des Expertensystems

Da das Expertensystem aus verschiedenen Komponenten besteht, die voneinander abhängig sind und aufeinander zugreifen, wird in diesem Abschnitt zunächst der schematische Aufbau skizziert, bevor im Weiteren auf die einzelnen Komponenten eingegangen wird. Dennoch lassen sich auf Grund der Komplexität des Systems Querverweise, Vorwegnahmen und Wiederholungen nicht ganz vermeiden. Abb. 4.1 zeigt den schematischen Aufbau des Systems und die Relationen der einzelnen Komponenten.

Ganz oben steht die Problemstellung. Die erste Aufgabe des Expertensystems ist es, eine vergleichbare Problemstellung als solche zu erkennen und das bekannte Lösungsverfahren anzuwenden. Davon

ausgehend stellt sich zunächst die Frage, wie eine bekannte oder „vergleichbare“ Problemstellung identifiziert werden kann. In Abschnitt 4.1.1 wird daher versucht, Variablen zur Abgrenzung aufzuzeigen. Im weiteren Verlauf der Untersuchung wird sich jedoch zeigen, dass die Domäne auch anhand der erfüllten Modell-Annahmen eingeteilt werden kann.



**Abb. 4.1: Schematischer Aufbau des Expertensystems**

Um die Vielseitigkeit zu erhöhen, sollte das Expertensystem weiterhin für Fälle, in denen keine vergleichbare Problemstellung samt Lösung bekannt ist, in der Lage sein, eine bestmögliche Lösung anzubieten. Dadurch rückt neben der Vergleichbarkeit die Übertragbarkeit in den Blickpunkt. Die Frage lautet also: Welche nicht-gleichen Problemstellungen können dennoch mit einem Verfahren gelöst werden, oder was ist das relativ am besten geeignete Verfahren für eine gegebene Problemstellung?

Die unsichere Domäne des Systems ist damit die Gesamtheit der möglichen Problemstellungen von der auf die Gesamtheit der enthaltenen Verfahren mit den jeweils möglichen Parametern geschlossen wird. Die Komponente, die dies leistet, wird als Inferenz-Komponente bezeichnet, und benötigt eine Wissensbasis, die repräsentiertes Wissen enthält. Hier übernimmt das System die Rolle eines Experten, der anhand seiner Erfahrung auswählt, was ein geeignetes Lösungsverfahren für das vorliegende Problem ist. Da das System seine Wissensbasis selbstständig vergrößern (also lernen) soll, ist die Relation von Wissensbasis und Inferenz bidirektional.

Die Inferenz-Komponente wird hier als Teil der Steuerungskomponente des Systems angesehen, da das System ja nicht nur ein Verfahren auswählen, sondern dieses auch durchführen soll. Dafür muss einerseits die Abfolge von einzelnen Verfahrensschritten (im Folgenden auch Module genannt) gesteuert werden. Zum anderen müssen aber auch die für die Module erforderlichen Steuerungsparameter gesetzt werden. Ideal wäre, wenn auch Steuerungsparameter automatisch gewonnen werden könnten. Da dies aber sicher nicht für alle Parameter möglich sein dürfte, müssen die Steuerungspara-

meter ebenfalls Teil der Wissensbasis sein und in die Inferenz eingeschlossen werden, womit sich die Dimensionalität der Domäne erhöht.

Weiterhin stellt sich die Frage, wie die einzelnen Verfahren optimal parametrisiert werden können, um einen möglichst großen Funktionalitätsumfang zu erreichen. Dieser Aspekt war in den bisherigen Arbeiten leider viel zu wenig im Blickpunkt. Im Gegenteil erwecken viele Verfahren den Eindruck, dass sie optimiert wurden, um bei einer einzelnen Problemstellung ein möglichst gutes Ergebnis zu erzielen. Damit nimmt die Performance bei veränderten Problemstellungen rapide ab und Informationen zur Übertragbarkeit fehlen fast gänzlich. Eine wichtige Aufgabe ist also, die Übertragbarkeit empirisch zu testen, wozu bereits erste Studien vorliegen (ERIKSON *et al.* 2006, ERIKSON & OLOFSSON 2005). In dieser Arbeit soll daneben exemplarisch versucht werden, die Übertragbarkeit von einzelnen Verfahren analytisch zu überprüfen. Dafür wird das Modell in zwei Teile geteilt. Im ersten Teil werden häufig gemachte Annahmen akkumuliert, die oft implizit in Verfahren eingehen, im zweiten Teil werden diese zu quantifizierbaren Merkmalen erweitert. In Kapitel 5 wird dann für eine Auswahl von Segmentierungsverfahren beispielhaft gezeigt, wo die Annahmen in Verfahren eingehen und die Übertragbarkeit beeinträchtigen.

#### 4.1.1 Kategorisierung der Problemstellungen

Es wurde bereits die Frage aufgeworfen, inwiefern Problemstellungen vergleichbar sind. Dabei wurde bewusst neben den eingehenden Daten auch die Anforderungen an das Ergebnis einbezogen, da es beispielsweise überflüssig ist, die genaue Position eines Baumes zu bestimmen, wenn eigentlich nur der Kronendurchmesser von Belang ist. Andererseits dürften die Möglichkeiten einzelne Parameter zu extrahieren, stark vom Informationsgehalt der Eingangsdaten bestimmt sein.

Von den Daten ausgehend wäre es nahe liegend, die Datensätze nach Art der Daten und Szeneninhalte zu charakterisieren. In Tab. 1 sind Variablen zusammengefasst, die dafür herangezogen werden könnten.

**Tab. 1: Variablen zur Kategorisierung von Problemstellungen**

<b>Variable</b>	<b>Ausprägungen</b>
<b>----- Zum Datensatz -----</b>	
Datentyp	Bild, Höhenmodell, Punktdaten ( <i>first pulse</i> , <i>last pulse</i> , komplett)
Geometrische Auflösung	Maßstab, Bodenpixelgröße
Kanäle	R,G,B, NIR, andere
Zusätzliche Daten	DGM (Genauigkeit), GIS
<b>----- Zur Aufnahme -----</b>	
Sonnenstand	Azimut & Höhe
Abbildungs-Geometrie	Flughöhe, Brennweite Beobachtungswinkel, etc.
Sensor	Foto, Multimapper, LIDAR
<b>----- Zum Szeneninhalte -----</b>	
Globaler Kontext (Landschaftstyp):	Wald (Typ), urban, offene Landschaft
Lokaler Kontext:	alleinstehende Bäume, Gruppen, Reihen, Kronendecke (offen/geschlossen)
<b>----- Zu den Bäumen -----</b>	
homogen	ja/nein

	Laubbäume/Nadelbäume
Zusammensetzung	Alter, Spezies, Größe ...
----- <b>Zu anderen Objekten</b> -----	
Material	Spektrum, Textur
Form	Geometrie

Es soll nicht näher auf die einzelnen Variablen eingegangen werden, da das Grundproblem bereits deutlich wird: Jeder dieser Punkte ließe sich beliebig weiter differenzieren und würde im Endeffekt zu einer umfangreichen Kategorisierung aller möglichen Szeneninhalte und Datensätze führen. Damit würde die betrachtete Domäne unendlich komplex. Gleichzeitig sind viele dieser Informationen für die Übertragbarkeit gar nicht notwendig. So ist möglicherweise für ein Verfahren gar nicht der Sonnenstand relevant, sondern nur, ob es ausgeprägte Schatten zwischen den Kronen gibt.

Daher soll an dieser Stelle nicht versucht werden, die Variablen weiter zu differenzieren, um Problemstellungen abzugrenzen. Stattdessen wird im Folgenden betrachtet, welche grundlegenden Annahmen in die Verfahren eingehen, und ob sich die Problemstellungen anhand dieser kategorisieren lassen.

#### 4.1.2 Wissensrepräsentation und Inferenzmechanismus

Die Wissensrepräsentation (also die Frage, in welcher Form das Expertenwissen am besten abgelegt wird) und der Inferenzmechanismus (also die Frage, wie das System in der gegebenen Domäne das jeweils beste Verfahren auswählt) sollen hier nicht vertieft werden. Es sei aber kurz darauf hingewiesen, dass Bayes'sche Netze sich als Inferenzmechanismus eignen würden. Diese stellen probabilistische Zusammenhänge mit gerichteten azyklischen Graphen dar, bei dem die Knoten Zufallsvariablen sind und die Kanten die Kausalitäten (bzw. bedingte Abhängigkeiten) zwischen den Variablen kodieren (WITTIG 2002: 15ff). Somit werden alle Variablen, die nicht in einer bestimmten Art verbunden sind, als unabhängig betrachtet: „*Bayesian networks have built-in independence assumptions*“ (CHARNIAK 1991: 53). Dadurch wird die Anzahl der Zustände, die die Zufallsvariablen einnehmen können, erheblich reduziert. Gleichzeitig bleibt die Modellierung durch den Graphen anschaulich und nachvollziehbar.

## 4.2 Steuerung

Um das Verfahren möglichst allgemein zu halten und die Übertragbarkeit auf andere Problemstellungen zu erhöhen, sollte die Steuerung vorzugsweise Top-Down operieren, also vom abstrakten Modell zum Bild hin. Bei Bottom-Up-Strategien steckt dagegen oft schon sehr viel Modellwissen in den Verfahren, so dass diese im Nachhinein nur sehr schwer anzupassen sind. Viele Parameter machen das Verfahren zudem unhandlich.

Es soll aber nochmals ausdrücklich darauf hingewiesen werden, dass immer Annahmen und Modellwissen implizit ins Verfahren eingehen. Beispielsweise macht eine einfache Glättung mit einem Gauskern einen Baum (im Sinne eines kreisförmigen, homogenen Objekts einer bestimmten Größe) noch ‚baumförmiger‘ im selben Sinne. Ein linienartiges Objekt, etwa eine Straße, würde durch das gleiche Verfahren aber geschwächt oder gar verschwinden, also ‚unstraßenförmiger‘ werden. Ebenso macht ein rechteckiger Glättungsfilter, wie er vielfach eingesetzt wird (z.B. PINZ 1989, DIEDERSHAGEN *et al.* 2003A), ein Objekt in gewisser Hinsicht ‚unbaumförmiger‘. So wird es zwar homogener, durch die Anisotropie (also Vorzugsrichtung des Filters) aber unter Umständen weniger kreisförmig.

Dieses einfache Beispiel zeigt, dass solch komplexe Verfahren praktisch nie eine Steuerungsrichtung in Reinform haben, da schon während der Vorverarbeitung Annahmen über die gesuchten Objekte eingehen. Auch wird es ohne einen Blick ins Bild nie gelingen, Modelle zu entwickeln, zu verfeinern und zu diversifizieren. Ebenso lässt es sich kaum vermeiden, dass die Modelle jeweils an das vorliegende Bildmaterial angepasst werden müssen. In gewissem Umfang muss das Modell sogar an das jeweilige Verfahren angepasst werden, wie in Abschnitt 4.4.5 gezeigt wird.

### 4.3 Steuerungsparameter

An dieser Stelle werden einige allgemeine Überlegungen zu Anzahl und Art der Steuerungsparameter aufgestellt. Die konkreten Parameter für die einzelnen Verfahren werden in Kapitel 5 vorgestellt.

Bezüglich der Anzahl der Steuerungsparameter gibt es einen grundsätzlichen Zielkonflikt: Einerseits sind parameterfreie Verfahren wünschenswert, da diese am einfachsten zu bedienen bzw. zu steuern sind. Auf der anderen Seite sind sie aber nicht besonders flexibel. Im Sinne der vielseitigen Verwendbarkeit sollten dagegen möglichst viele Steuerungsparameter in die Verfahren integriert werden, um sie entsprechend einer gegebenen Problemstellung justieren und abstimmen zu können. Damit wachsen aber wieder die Komplexität und die benötigte Expertise (oder das repräsentierte Wissen in einem Expertensystem), um das Verfahren zu bedienen. Grundsätzlich sollte auch das Expertensystem noch die Möglichkeit vorsehen, dass ein Operateur einzelne Parameter selbst bestimmt. Damit dies möglichst leicht durchführbar ist, sind Parameter mit einer realen Bedeutung abstrakten vorzuziehen. So kann sich ein Operateur unter dem Abstand zwischen zwei Kronen sicherlich leichter etwas vorstellen als unter dem  $\sigma$  eines Gaußfilters.

Das Ziel der optimalen Parametrisierung eines Verfahrens besteht also darin, einen möglichst großen Funktionsumfang mit einer möglichst überschaubaren Anzahl von Parametern zu realisieren. Später könnte dann untersucht werden, inwiefern sich einzelne Steuerungsparameter ihrerseits automatisch gewinnen ließen, um sie somit wieder zu einem Teil des automatisierten Verfahrens zu machen.

Ein Teil der Steuerungsparameter ergibt sich direkt aus dem Modell. So kann beispielsweise Vorwissen über die Größe der erwarteten Bäume ins Segmentierungsverfahren eingehen. Es sollte aber gut abgewogen werden, welche Modell-Parameter in welchem Verfahrensschritt eingesetzt werden. Wird gemäß dem obigen Beispiel die Größe zur Segmentierung genutzt (und damit nur Segmente in der gesuchten Größe erzeugt), steht sie anschließend als Merkmal zur Hypothesenauswahl nicht mehr zur Verfügung.

Weiterhin können zwei Arten von Parametern unterschieden werden: Diskrete Parameter, die nur eine bestimmte Anzahl von Zuständen einnehmen können, beschreiben Verzweigungen des Programmflusses (also quasi algorithmische Varianten). Kontinuierliche Parameter können dagegen beliebige Werte (üblicherweise aus einem begrenzten, abgeschlossenen Wertebereich) annehmen. Anforderungen an solche Parameter sind, dass die Verfahren möglichst monoton auf sie reagieren, was insbesondere bedeutet, dass die Algorithmen gegenüber kleinen Änderungen der Parameter möglichst unempfindlich sein sollten. Weiterhin ist es erforderlich, dass Veränderungen des Parameters in eine Richtung das Ergebnis ebenfalls immer in eine Richtung verändern. Anders ausgedrückt sollte das Ergebnis als Funktion des Parameters stetig und monoton sein.

## 4.4 Modell

„Ziel der Modellbildung ist es, eine strukturierte, möglichst widerspruchsfreie und im Computer formal repräsentierbare Beschreibung eines Objektes zu liefern“ (HINZ 2003: 8).

Für Menschen ist es eine relativ leichte Aufgabe, einen Baum auf einem Bild zu erkennen, da das menschliche Gehirn hochgradig auf Mustererkennungsaufgaben spezialisiert ist. Schon etwas schwerer ist es, zu sagen, woran ein Baum eigentlich zu erkennen ist. ‚Ein Baum ist grün‘, könnte eine spontane Antwort sein. Aber was ist mit einem Laubbaum im Herbst? Der mag vielleicht gerade rot oder gelb sein, aber der Betrachter verfügt über das konzeptionelle Wissen, dass dieses Muster dennoch zur Klasse Baum gehört. Oder ein grünes Auto – das würde kaum jemand fälschlicherweise für einen Baum halten. Das Auto hat eine andere Form wäre ein denkbare Argument. Aber würde das Auto von oben nicht ähnlich aussehen? Ja, vielleicht, doch ein Auto hat eine Windschutzscheibe und ein Baum nicht. Der Betrachter muss also auch wissen, was eine Windschutzscheibe ist und in welcher Relation sie zum Auto steht. Dieses Beispiel ließe sich beliebig fortsetzen, aber es ist schon zweierlei klar geworden: Zwei wichtige Merkmale zur Erkennung sind Farbe (Radiometrie) und Form (Geometrie). Diese Merkmale alleine reichen jedoch nicht aus, da ein wichtiger Teil der menschlichen Vorstellung von einem Objekt der Kontext ist: Während ein Baum beim sonntäglichen Waldspazierpunkt sofort richtig erkannt würde, wäre die Wahrnehmung des gleichen sensorischen Musters in der U-Bahn sehr irritierend – schließlich gehört der Baum „einfach nicht dort hin“.

Da ein Expertensystem solch komplexes Kontextwissen nicht erlangen kann, muss die Domäne stark eingeschränkt werden, beispielsweise insofern, dass nur Aufnahmen von Bäumen von oben verarbeitet werden. Das Expertensystem muss also beispielsweise mit einer Röntgenaufnahme nichts anfangen können und daher auch nichts über Knochen wissen. Diese Beschränkung wird auch *closed world assumption* genannt (ähnlich dem Problemkreis  $\Omega$ , siehe 2.2.1). Je geschlossener die Welt ist, desto einfacher ist es, ein Modell zu entwickeln.

In diesem Abschnitt soll gezeigt werden, wie ein geeignetes Modell entwickelt werden kann und welche Merkmale dafür in Frage kommen. Das Modell kann mehrere Repräsentationsebenen überspannen. Insbesondere sind 3D und 2D-Modell zu unterscheiden. Eine weitgehend identische Einteilung ist die in Objektmodell als Repräsentation des Objektes in der realen Welt und Bildmodell als Repräsentation des Objektes im Bild. Dabei ist es üblich, mit Hilfe eines Sensormodells vom Objektmodell zum Bildmodell überzugehen, das die physikalischen Prozesse und geometrischen Transformationen bei der Aufnahme beschreibt (HINZ 2003: 9). Da ein eigenes Sensormodell die Modellbildung erheblich erschwert, soll hier darauf verzichtet und direkt ein Bildmodell entwickelt werden, im Hinblick auf die unterschiedlichen Aufnahmeverfahren wäre ein solches aber perspektivisch wünschenswert.

Die Modellbildung erfolgt hier in zwei Stufen: Erstens werden häufig gemachte qualitative Annahmen zusammengetragen, die auf verschiedenen Repräsentationsebenen angesiedelt sind. Der Übergang zwischen Objekt und Bild-Modell sollte aber stets nachvollziehbar sein. Im zweiten Schritt werden die Annahmen zu quantifizierbaren Merkmalen erweitert. Weiterhin werden einige Überlegungen zum Kontext sowie zum Verhalten des Modells im Maßstabsraum angestellt.

#### 4.4.1 Annahmen

Die folgende Zusammenstellung von Modell-Annahmen und impliziten Annahmen ist überwiegend den in Kapitel 3 vorgestellten Verfahren entnommen und aus ihnen abgeleitet. Auch wenn es sich dabei überwiegend um einfache Annahmen handelt, wird sich zeigen, dass sie nur unter bestimmten Bedingungen erfüllt sind. Neben der Modellbildung werden die Annahmen später auch zur analytischen Betrachtung der Übertragbarkeit von Verfahren verwendet. Das Modell wird in einen geometrischen und einen radiometrischen Teil untergliedert.

##### Geometrische Annahmen

Ausgehend vom geometrischen (3D, Objekt-) Pollock-Modell einer Baumkrone (im Folgenden Krone genannt) lässt sich feststellen:

*Annahme 1: Die Grobstruktur (3D-Modell) einer Krone ist ein Ellipsoid.*

Daraus lassen sich verschiedene weitere Eigenschaften ableiten. Zunächst gilt für das 2D-Modell von oben betrachtet:

*Annahme 2: Eine Krone ist von oben (Bildmodell) ebenfalls elliptisch.*

Eigentlich nur ein Spezialfall davon, der aber sehr häufig eingeht, ist:

*Annahme 3: Die Krone ist im Bildmodell kreisförmig.*

Weiterhin folgen:

*Annahme 4: Die 3D-Hülle der Krone ist konvex,*

*Annahme 5: der 2D-Rand einer Krone ist konvex.*

Diese Annahme ist richtig für das Bildmodell eines DOM, bei optischen Bildern ist sie allerdings vom Sonnenstand abhängig. Insbesondere stark konische Bäume wie Fichten können sich selbst beschatten und es entstehen (durch die geometrische Transformation bei der Abbildung) konkave Einbuchtungen in der Kronenfläche (ERIKSON 2004B: 472).

Weitere Annahmen, die sich aus der Form ergeben, sind:

*Annahme 6: Es gibt eine Baumspitze und die ist der höchste Punkt (lokales Maximum).*

*Annahme 7: Die Baumhöhe ist symmetrisch bezüglich des höchsten Punktes. Im Kreisfalle fällt sie in alle Richtungen gleichmäßig ab.*

*Annahme 8: Die Krone beginnt erst in einer gewissen Höhe über dem Boden.*

Dies ist relativ einsichtig, problematischer ist dagegen der Umkehrschluss, der implizit beispielsweise von SCHARDT *et al.* (2002) oder PERSSON *et al.* (2002) gemacht wird:

*Annahme 9: Alles über einer gewissen Höhe ist Teil einer Krone.*

Daraus folgt insbesondere:

*Annahme 10: Der Rand einer Krone hat immer die gleiche Höhe.*

Daneben haben die Kronen topologische Eigenschaften:

*Annahme 11:* Eine (2D & 3D-)Krone ist kompakt (d.h. beschränkt und abgeschlossen).

*Annahme 12:* Eine (2D & 3D-)Krone ist einfach zusammenhängend (d.h. frei von Löchern).

Scheinbar trivial, aber nicht zu vergessen ist die Größe der Krone:

*Annahme 13:* Der Durchmesser einer Krone ist bestimmt. Insbesondere gibt es eine kleinste und eine größte Krone im Bild.

*Annahme 14:* Die Kronen haben einen bestimmten Abstand voneinander. Der Abstand ist größer oder gleich ihrer Größe.

### **Radiometrische Annahmen**

*Annahme 15:* Eine Baumkrone ist hell. Insbesondere ist sie heller als die Umgebung (Schatten).

*Annahme 16:* Eine Baumkrone hat das Spektrum grüner Vegetation.

Dies ist offensichtlich nur in der Vegetationsperiode zutreffend. Es folgt insbesondere:

*Annahme 17:* Eine Krone ist im sichtbaren Spektrum grün:  
 $DoG > 0, G > R, G > B$ .

*Annahme 18:* Eine Krone hat hohe Reflektanz im nahen Infrarot.

*Annahme 19:* Das Spektrum hat ein Red-Edge, d.h.  $NDVI > 0, NIR \gg R$ .

Sehr wichtig sind die Annahmen bezüglich der Homogenität der Kronen. Eine relativ schwache Form ist:

*Annahme 20:* Die Farbvariation innerhalb einer (2D-)Krone ist geringer als zwischen verschiedenen Kronen.

Insbesondere und expliziert:

*Annahme 21:* Es gibt dunkle Täler (Schatten) zwischen den Kronen.

Formalisiert bedeutet dies:

*Annahme 22:* Die Krone ist radiometrisch homogen bezüglich eines bestimmten Homogenitätsmaßes. Eine Krone plus Umfeld sind nicht homogen.

Trifft diese Annahme zu, dann müssten sich RG-Verfahren prinzipiell eignen, um Kronen zu segmentieren (siehe 2.4.4). Unter bestimmten Erwartungen an das Homogenitätskriterium kann außerdem angenommen werden:

*Annahme 23:* Innerhalb der Krone gibt es keine dunklen Pixel.

Ergänzend zur den einfachen radiometrischen Merkmalen kann die Textur betrachtet werden:

*Annahme 24: Eine Krone hat eine Textur, die bezüglich eines bestimmten Maßes homogen ist.*

Aus dem 3D-Modell unter der Annahme einer einzigen parallelen Lichtquelle (also der Sonne) lässt sich weiterhin ableiten:

*Annahme 25: Es gibt eine Baumspitze und die ist der hellste Punkt (lokales Maximum) in einem geeigneten Maßstab,*

wobei diese Annahme offensichtlich nicht für jeden Sonnenstand zutrifft. Daraus folgen:

*Annahme 26: Die Intensität der Randpixel ist niedriger als die Intensität der Krone. Insbesondere ist der Rand dunkler als die Baumspitze.*

*Annahme 27: Die Grauwert-Verteilung ist konvex.*

Dies gilt wieder nur in einem bestimmten Maßstab, da lokal in der Feinstruktur weitere Helligkeitsmaxima auftreten können. Sehr stark ist die Annahme einer symmetrischen Grauwertverteilung:

*Annahme 28: Die Helligkeit ist symmetrisch bezüglich des hellsten Punktes. Im Kreisfalle fällt sie in alle Richtungen gleichmäßig ab, die Helligkeitsverteilung ist radial.*

Obwohl dies offensichtlich nur im Falle eines hohen Sonnenstands und senkrechten Blickwinkels zutrifft, wird es häufig vorausgesetzt. Die radiometrische Symmetrie ist damit schwieriger zu garantieren als die geometrische.

#### 4.4.2 Merkmale

Um aus den obigen Annahmen ein 2D-Bildmodell für die gesuchten Kronen im Bild zu erzeugen, müssen Merkmale gefunden werden, die die Annahmen formalisieren und quantifizieren. Da in Kapitel 5 Segmentierungsverfahren für optische Bilder verglichen werden, beziehen sich die ausgewählten Merkmale ebenfalls auf diese Art von Daten. Auf ähnliche Weise ließen sich aber auch Merkmale für Höhenmodelle konstruieren. Die Merkmale werden für alle Segmente erhoben.

##### Geometrische Merkmale

AREA	Fläche eines Segments in Pixeln oder m <sup>2</sup> .
SOLIDITY	Ein Maß für die Konvexität des Randes. Anteil der Fläche des Segments an seiner konvexen Hülle.
RADIUS	Größter Abstand vom Schwerpunkt eines Segments zum Rand.
CIRCLENESS	Kreisförmigkeit des Segments: $\frac{AREA}{\pi \cdot r^2}$ .

##### Radiometrische Merkmale

B, G, R, NIR	mittlere Intensität im jeweiligen Kanal.
BRIGHT	mittlere Intensität (aller Kanäle oder eines gewählten Kanals).
DOA	mittlerer <i>Degree of Artificiality</i> (siehe 2.1.1)

NDVI	mittlerer <i>normalized difference vegetation index</i> .
CONVEX	mittlere (radiometrische) Konvexität, berechnet mit dem $\Delta$ -Operator.
SPOT	Intensität des hellsten Punktes im Segment.
BOUND_BRIGHT	Median der Intensität der Randpixel.
SPOT2BOUND	Verhältnis: $\frac{SPOT}{BOUND\_BRIGHT}$ .

#### 4.4.3 Klassen und Wahrscheinlichkeitsdichten

Zu jedem Merkmal werden nun noch Informationen zu den klassenspezifischen Ausprägungen benötigt. Dabei lassen sich leicht qualitative Aussagen machen. So muss AREA für die Klasse Baum in einem bestimmten Bereich liegen und SOLIDITY, CIRCLENESS und SPOT2BOUND sollten „groß“ sein. Bei manchen Merkmalen bieten sich auch offensichtliche Grenzen an. So macht es Sinn, zu fordern, dass DOA und NDVI > 0 und CONVEX < 0 sein sollten.

Was aber zur Klassifikation eigentlich benötigt würde, wären die genauen Wahrscheinlichkeitsdichten (PDF) jeder Klasse für jedes Merkmal. Dabei stellen sich zwei Fragen: Erstens: Was ist eine geeignete Klasseneinteilung und zweitens: Woher kommen die PDF?

Bei der Klasseneinteilung sind zwei Wege denkbar. Entweder ein Modell für jede Baumart oder ein einzelnes Baum-Modell für alle Arten. Ersteres würde es erlauben, direkt Bäume einer bestimmten Art zu extrahieren. Dafür würde das Modell aber sehr komplex und das Diskriminierungspotential der vorgeschlagenen Merkmale wäre wahrscheinlich nicht ausreichend. Außerdem ist eine Klassifizierung nach Baumarten für viele Anwendungen gar nicht erforderlich.

Daher scheint es angemessen, sich zunächst auf ein Baum-Modell zu beschränken, womit es zunächst nur die Klassen Baum und Nicht-Baum gibt. Daneben kann es aber abhängig vom Kontext sinnvoll sein, weitere Klassen einzuführen (siehe 4.4.4).

Zur Gewinnung der PDF sind ebenfalls verschiedene Wege denkbar. Steht ein genügend großer Referenz-Datensatz zum Training zur Verfügung, können sie im Prinzip mit jedem überwachten Klassifikator (siehe 2.2.3) erzeugt werden (Vorwissen über die Artenzusammensetzung könnte dann als *a priori* Wahrscheinlichkeiten in die Klassifikation eingehen). Der Vorteil dieser Methode ist, dass sie an das jeweilige Bildmaterial angepasst sind. Andererseits bringt dies den Verlust der gewünschten Top-Down-Steuerungsrichtung, da das Modell aus dem Datensatz erzeugt wird.

Besser wäre es also, wenn sich das Modell aus sich heraus erklären würde. Einen solchen Versuch macht STRAUB (2003A: 41ff), wobei die von ihm verwendeten Fuzzy-Zugehörigkeitsfunktionen bis auf einen konstanten Faktor als *a priori* festgesetzte Wahrscheinlichkeitsdichten und der gewählte Schwellenwert als Entscheidungsgrenze im Merkmalsraum (also quasi als Box-Klassifikator) interpretiert werden können. Dieser Ansatz hat wiederum den Nachteil, dass er unmöglich allen Situationen und Datensätzen gerecht werden kann. Demnach müssten die Modelle diversifiziert und nach Situationen differenziert werden. Solange keine sehr guten Sensormodelle vorliegen, wird eine Anpassung insbesondere der radiometrischen Merkmale an einzelne Datensätze unvermeidlich sein (eine solche wird übrigens auch von STRAUB vorgenommen).

Dies kann im Rahmen dieser Arbeit nicht geleistet werden. In Abschnitt 6.6 wird aber gezeigt, dass die vorgeschlagenen Merkmale prinzipiell zur Hypothesenauswahl geeignet sind.

#### 4.4.4 Kontext und Komplexität

*“The automatic extraction of objects from digital imagery is a very complex task. It is widely accepted that the complexity can be reduced by integrating context information into the extraction process [...] context means that there exists knowledge not only about the object of interest but also about other relevant facts and their relations to the object of interest”* (BAUMGARTNER *et al* 1997: 1).

So schwer es sein mag, ein selbsterklärendes Modell für alle Situationen und Baumarten zu schaffen, so einfach kann es für einen Einzelfall sein. Von SCHARDT *et al.* (2002) und PERSSON *et al.* (2002) unterscheiden Bäume und Nicht-Bäume etwa allein anhand von Modell-Annahme 9: Alles über einer gewissen Höhe ist Baum. Aus Mustererkennungssicht verwenden sie also einen eindimensionalen Merkmalsraum mit einer festgelegten Entscheidungsgrenze. Obwohl dieses Modell unglaublich simpel ist, funktioniert es erstaunlich gut. Dies beruht auf einer einfachen Begebenheit: Der Szeneninhalt ist in den vorliegenden Fällen ein geschlossener Wald und der enthält (außer vielleicht Hochsitzen) keine weiteren Objekte, die so hoch sind wie Bäume.

Daran wird deutlich, dass die für das Modell benötigte Komplexität entscheidend vom Kontext abhängt. Je komplexer der Kontext ist, desto komplexer muss auch das Modell sein. So wird es beispielsweise in einer urbanen Umgebung mit vielen verschiedenen Objekten sehr viel mehr Objekte geben, die einem Baum ähneln – selbst in einer Welt mit zehn Merkmalen. Und auch in einer offenen Landschaft könnten andere Objekte, beispielsweise Büsche, leicht für Bäume gehalten werden.

Daher kann es sinnvoll sein, das Modell um weitere Klassen zu erweitern, welche als Widerspruch zur Baumhypothese fungieren. Gäbe es beispielsweise bei einer offenen Kronendecke eine eigene Klasse für Lichtflecken auf dem Boden, würden diese seltener fälschlicherweise als Bäume interpretiert.

Es ist also angebracht, das Modell nach Kontexten zu differenzieren. Dafür ist es aber erforderlich, dass bei einer Hypothesenauswahl auch nur ein Kontext-Typ auftritt, was eine Vor-Segmentierung in homogene Kontextbereiche erforderlich macht.

#### 4.4.5 Modell und Maßstab

Eine weitere Schwierigkeit bei der Modellbildung liegt in der Maßstabsproblematik. So beschreiben die meisten der aufgestellten Modell-Annahmen und abgeleiteten Merkmale die Grobstruktur des Baumes in einer Größenordnung von einigen Metern. Wird der Baum aber in einem feineren Maßstab betrachtet, so sind die einzelnen Äste, Blätter und Früchte mit Sicherheit nicht symmetrisch, ellipsoid und radiometrisch homogen. Das Modell ist also immer nur für einen bestimmten Maßstab geeignet (siehe Abb. 6.5) oder muss den Maßstab selbst als Parameter enthalten. Daraus folgt zum einen, dass die Eignung eines Modells wesentlich von der Auflösung des verwendeten Bildmaterials abhängt. Zum anderen hängt sie aber auch vom verwendeten Segmentierungsverfahren ab, da die Segmentierung in unterschiedlichen Maßstabsebenen durchgeführt wird. Es müsste also auch untersucht werden, wie das Modell an Maßstab und Verfahren angepasst werden kann.

### 4.5 Verfahren

In diesem Abschnitt sollen der Ablauf und die Einzelschritte der in Abb. 4.2 als Verfahren bezeichneten Komponente beschrieben werden. Dabei handelt es sich um eine Standardabfolge, von der im Einzelfall abgewichen werden kann, die aber für die überwiegende Zahl von Problemstellungen zweckmäßig erscheint. Für jedes Modul können unterschiedliche Algorithmen zum Einsatz kommen,

die in geeigneter Weise kombiniert werden können. Daher umfasst dieser Abschnitt nur eine grobe Erläuterung der Funktionalität, nicht jedoch der genauen Algorithmen. Für einzelne Segmentierungsverfahren wird dies exemplarisch in Kapitel 5 geleistet.

Abb. zeigt zunächst eine Übersicht über die einzelnen Verfahrensschritte, die im Folgenden auch Module genannt werden.

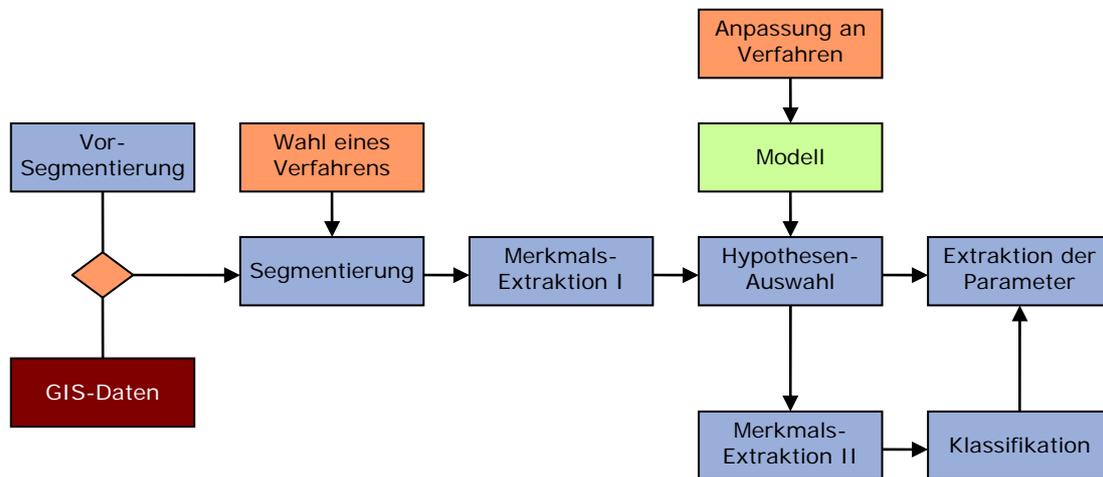


Abb. 4.2: Ablauf des Verfahrens bei der Baumextraktion, Einzelschritte (Module)

### Vor-Segmentierung

In Abschnitt 4.4.4 wurde bereits darauf hingewiesen, dass es sinnvoll ist, das Modell nach globalen Kontexten zu differenzieren. Dies setzt voraus, dass vorab eine Segmentierung in einheitliche Kontextgebiete erfolgt, etwa zunächst in die Landschaftstypen urban, offene Landschaft und Wald. Bei Wald wäre es weiterhin sinnvoll, diesen vorab in möglichst homogene Bestände zu segmentieren. In STRAUB (2002) wird mit gutem Erfolg die Eignung von Gaborfiltern aus der MPEG-7 Textur-Beschreibung zur Unterscheidung von Wald und urbanen Bildern getestet. Demnach scheinen Textur-Merkmale prinzipiell geeignet, eine Segmentierung nach Landschaftstypen vorzunehmen.

Alternativ kann auch auf vorhandene Daten der Raumordnung, etwa in einem GIS zurückgegriffen werden. Umfasst der vorliegende Datensatz ohnehin nur einen Kontext-Typ, so entfällt dieser Schritt.

### Segmentierung

Bei der Segmentierung wird das Bild in homogene Regionen unterteilt. Dabei sollten die Algorithmen derart beschaffen sein, dass einzelne Kronen nach Möglichkeit ein Segment bilden. Unter Segmentierung soll hier etwas allgemeiner ein Verfahren zur Abgrenzung von Kronen (*delineation*) verstanden werden. In den meisten Fällen handelt es sich dabei um ein Segmentierungsverfahren in Sinne der Definition in (2.58). Dadurch ist insbesondere festgelegt, dass die Kronen disjunkt sind, also nicht überlappen. Auch wenn Überlappungen in der Natur vorkommen, wird diese Einschränkung für eine große Anzahl von Fällen akzeptabel sein, da diese ohnehin nur sehr schwer zu erkennen sein dürften. In einzelnen Fällen können aber auch andere Verfahren zur Kronenabgrenzung den Segmentierungsschritt ersetzen.

Eine weitere Bedingung vollständiger Segmentierung nach (2.58), die von einigen Algorithmen verletzt wird, ist die Vollständigkeit. Diese kann aber im Allgemeinen durch die Einführung eines Hintergrundes wieder hergestellt werden, der dann aber nicht notwendiger Weise homogen und zusammenhängend ist.

Die Segmentierung kann in einer oder in mehreren Maßstabsebenen erfolgen. Damit kann das Maßstabsproblem im Segmentierungsschritt umgangen werden, allerdings auf Kosten eines erhöhten Aufwandes bei der Hypothesenauswahl. Die Steuerung muss in diesem Schritt ein geeignetes Verfahren auswählen und die Steuerungsparameter bestimmen.

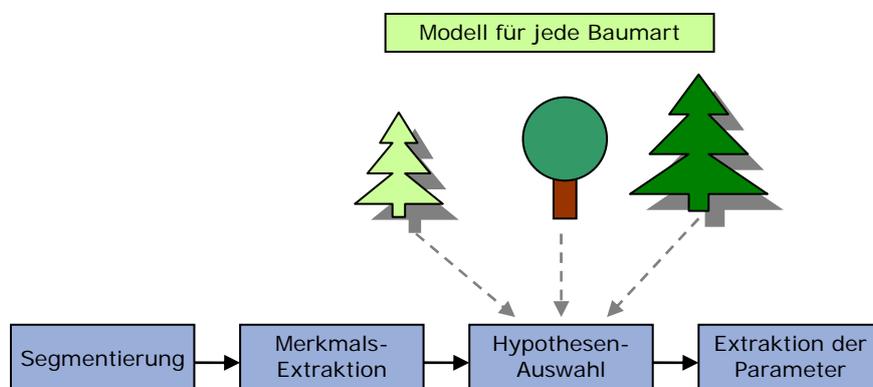
### **Merkmalsextraktion I und Hypothesenauswahl**

Gemäß dem Vorschlag in 4.4.3 soll in diesem ersten Klassifikationsschritt zunächst nur in die Klassen Baum und Nicht-Baum (oder gegebenenfalls verschiedene Nicht-Baum-Klassen wie Waldboden, Strauch, etc.) klassifiziert werden. Dafür werden zunächst von allen Segmenten Merkmale extrahiert, die sich zur Unterscheidung der Klassen eignen, und anschließend klassifiziert. Da die als Baum klassifizierten Segmente auch als Baum-Hypothesen bezeichnet werden, wird der Schritt in Abgrenzung zur folgenden Klassifikation nach Baumarten Hypothesenauswahl genannt. Damit wird auch die erwünschte Top-Down-Steuerung betont.

Die Steuerung muss in diesem Schritt das für den Kontext geeignete Modell auswählen, dieses an das benutzte Segmentierungsverfahren und den entsprechenden Maßstab anpassen sowie dafür sorgen, dass die richtigen Merkmale für das spezifische Modell extrahiert werden. Erfolgte die Segmentierung im Maßstabsraum, muss die Hypothesenauswahl auf mehreren Ebenen erfolgen.

### **Merkmalsextraktion II und Klassifikation**

Falls die Baumart als Parameter benötigt wird, werden in diesem zweiten Klassifikationsschritt weitere Merkmale extrahiert, die sich zur Unterscheidung von Baumarten eignen, und anschließend in Baumarten klassifiziert. Natürlich wird auch hierfür Modellwissen oder eine Stichprobe benötigt. Die Trennung der Merkmalsextraktion in zwei Teile spart Rechenzeit, da die weiteren Merkmale nur noch für die Baumhypothesen erhoben werden. Alternativ wäre es aber möglich, wie in Abschnitt 4.4.3 vorgeschlagen, gleich ein eigenes Modell für jede Baumart vorzusehen und so die beiden Klassifikationsschritte zusammenzufassen.



**Abb. 4.3: Alternatives Verfahren bei mehreren Baumarten**

---

Abschließend werden die Extraktions-Parameter erhoben (soweit sie nicht schon als Merkmale vorliegen) und als Ergebnis zurückgegeben.

Eine wichtige Frage bei der Steuerung der Abfolge von Modulen ist, welche Merkmale in welche Schritte eingehen. Wird ein Merkmal (etwa die Größe oder die Konvexität) zur Steuerung der Segmentierung verwendet, steht es anschließend bei der Hypothesenauswahl nicht mehr oder nur noch eingeschränkt zur Verfügung. Ebenso kann die Maßstabsproblematik in verschiedenen Schritten adressiert werden, insbesondere bei der Vor-Segmentierung, der Segmentierung oder der Hypothesenauswahl. Darum wird bei der folgenden Vorstellung von Einzelverfahren jeweils darauf eingegangen, welches Wissen in welchen Schritt eingeht und wo das Maßstabsproblem adressiert wird.

## 5 Segmentierungsverfahren

In diesem Kapitel werden die implementierten Verfahren beschrieben, die alle Teil des Moduls Segmentierung sind und auf optischen Bildern arbeiten. Die Konzentration auf die Segmentierung wurde als sinnvoll erachtet, da es sich dabei um den grundlegenden ersten Schritt handelt, der Voraussetzung für alle weiteren ist. So wird eine genaue Segmentierung beispielsweise als Voraussetzung für die Klassifikation gesehen (ERIKSON 2004: 8). Optischen Bildern wurde der Vorzug gegeben, da ihr Anwendungspotenzial aufgrund der größeren Verfügbarkeit und der gegenüber der Laserbefliegung immer noch deutlich geringeren Kosten als sehr groß eingeschätzt wird. Dennoch spielen sie in jüngeren Veröffentlichungen eher eine zweitrangige Rolle.

```

PREF
|
|SERIES_NAME:      goett{3}
|SERIES_PATH:
|SEG_TYPE:        BROWN_MO
|PRESET:  1
|
|-----
|MODELL
|
|ANISO:  2
|
|-----
|SIZE
|
|RAD:      40
|RAD_MIN:  25
|RAD_MAX:  70
|
|-----
|SEED
|
|SAATGUT:  DISTRANS
|CHANNEL:  2
|SIGMA:    16
|FSIZE:    81
|THRESH:   0.35
|
|-----
|BM
|
|NRUNS:      10
|NSTEPS:    100000
|WEIGHT:    RGB_mean
|WEIGHT_POT:  1.75
|MIN_PASS:   5
|INIT_CIRCLE: 1

```

Abb. 5.1: Alle Steuerungsparameter werden als **PREF** gespeichert

Im Gegensatz zu vielen Arbeiten geht es hier ausdrücklich nicht darum, für eine bestimmte Problemstellung ein optimiertes Verfahren zu finden. Vielmehr stehen zwei Aspekte im Vordergrund. Zum einen sollen die Verfahren auf eingehende Annahmen und implizite Voraussetzungen überprüft werden, um Anhaltspunkte für ihre Übertragbarkeit und Einsetzbarkeit bei neuen Problemstellungen zu gewinnen. Zum anderen wird analysiert, an welchen Stellen geeignete Steuerungsparameter eingeführt werden können, um die Vielseitigkeit zu erhöhen.

Alle Verfahren wurden für diese Arbeit in MATLAB unter Verwendung der Image Processing Toolbox (IPT) implementiert. Bezeichner aus dem MATLAB-Programmcode (inklusive der Steuerungsparameter) werden im Folgenden in einem Schriftsatz mit fester Laufweite gesetzt. Die Steuerungsparameter werden in einem struct Datenfeld namens `PREF` abgelegt, wie es in Abb. 5.1 zu sehen ist. Vor der Verarbeitung werden alle Bilder zum Typ `double` konvertiert und auf den Wertebereich `[0...1]` normiert.

### 5.1 Wasserscheidentransformation im Maßstabsraum (WST)

Das Verfahren Wasserscheidentransformation im Maßstabsraum (WST) ist dem Segmentierungsverfahren in STRAUB (2003A und andere) nachempfunden und wurde aus zwei Gründen ausgewählt. Erstens bietet es einen interessanten Ansatz für das Maßstabsproblem, da es gleich auf verschiedenen

Maßstabsebenen segmentiert. Zum anderen scheint es von den diversen Verfahren mit WST oder ähnlichem (z.B. *pouring* nach DIEDERSHAGEN *et al.* 2003A oder *hill-climbing* nach PERSSON *et al.* 2002) besonders geeignet, insbesondere da verschiedene Segmentierungsfunktionen getestet werden. Außerdem soll demonstriert werden, dass ein Verfahren, das eigentlich für Höhenmodelle entwickelt wurde, unter gewissen Umständen auch auf optische Bilder angewendet werden kann.

Das Verfahren selbst ist vergleichsweise einfach. Da die WST nur auf einem Kanal arbeitet, kann zunächst mit dem Parameter `PREF.WST.COLOUR_CHANNEL` ein entsprechendes Band ausgewählt werden. Möglich sind ein Graustufenbild 'GREY', ein einzelnes Farbband wie 'G' oder eine Linearkombination wie '2G-R'. Daneben kann auch auf Indizes segmentiert werden ('NDVI' oder 'DOA'), soweit die entsprechenden Bänder vorhanden sind.

Anschließend wird eine Maßstabsebene im Maßstabsraum berechnet, indem mit einem Gaußfilter mit den Parametern `PREF.WST.SIGMA` und `PREF.WST.FSIZE` gefaltet wird. `FSIZE` bestimmt im Wesentlichen die Anisotropie des Filters, die aber hier durchaus von Bedeutung ist. Es ist zu beachten, dass das Bild bei der Faltung kleiner wird, da es nicht periodisch ist und der Filter somit am Rand nicht fehlerfrei arbeiten kann.

Durch die Faltung verändern die Objekte im Bild ihre Form und werden dabei dem Faltungskern ähnlicher, da jedes Pixel ja gerade auf die PSF, also den Faltungskern, abgebildet wird (siehe 2.3.4). Damit geht insbesondere Annahme 3 (Kreisförmigkeit) ein. Da diese Annahme für Höhenmodelle und optische Bilder gleichermaßen zutrifft (bzw. nicht zutrifft), ist das Verfahren prinzipiell auf radiometrische Daten übertragbar. Weiterhin wirkt die Glättung wie ein Sieb für bestimmte Größen, womit auch Annahme 13 (Größe) eingeht. Dies gilt aber nur für eine einzelne Maßstabsebene. Durch Segmentierung auf verschiedenen Ebenen können Objekte beliebiger Größe extrahiert werden.

Da die WST entlang von Wasserscheiden segmentiert, muss das Bild durch eine geeignete Transformation in eine Segmentierungsfunktion (SOILLE 1998, siehe 2.4.5) überführt werden, die Maxima zwischen den gesuchten Objekten hat. STRAUB (2003A: 57) kommt zu dem Schluss, dass das invertierte Quadrat der Summe der zweiten Ableitungen  $\Delta I$ , hier `'-laplace^2'`, sich am besten als Segmentierungsfunktion eignet. Dennoch wurden auch das invertierte Bild `'inverted'` (etwa bei SCHARDT *et al.* 2002 oder MEI & DURRIEU 2004) und der Betrag des Gradienten `'mog'` (siehe GONZALEZ *et al.* 2004: 420) als Segmentierungsfunktionen implementiert. Die Segmentierungsfunktion wird dann mit Wasserscheidentransformation segmentiert.

Die Segmentierungsfunktion  $N\left((\Delta I)^2\right)$  hat Maxima an Nullstellen der zweiten Ableitung, es wird also quasi entlang der „Wendepunkte“ segmentiert.  $N$  ist dabei die Negativtransformation, die bei STRAUB(2003A: 66) als  $I^0$  bezeichnet wird. Da hier aber auch das (Intensitäts-)Bild als  $I$  bezeichnet wird, könnte dies zu Unübersichtlichkeit führen. Das Homogenitätsmaß der Segmentierung ist damit eine einheitliche Krümmung in einer bestimmten Maßstabsebene, jedes Segment ist also konvex oder konkav. Hierbei geht Annahme 4 ein, die im Fall von optischen Bildern durch Annahme 27 ersetzt wird. Weiterhin ist die Segmentierungsfunktion so konstruiert, dass sie an Stellen starker Krümmung besonders klein ist. Als Maß für die Krümmung kann sie also auch in die Merkmalsextraktion einbezogen werden.

Durch die Segmentierung im Maßstabsraum wird das Problem des optimalen Maßstabs auf einen späteren Schritt, in diesem Fall die Hypothesenauswahl, verschoben. Ein Nachteil des Verfahrens ist,

dass nicht nur starke Annahmen bezüglich der Form in die Segmentierung eingehen, sondern bei der Glättung die Feinstruktur aufgelöst und damit die Form selbst verändert wird. Daran muss einerseits das Modell angepasst werden, da beispielsweise die Merkmale SOLIDITY und CIRCLENESS durch die Eliminierung hoher Frequenzen größer werden. Weiterhin gehen wichtige Merkmale verloren, die später für die Klassifikation nach Baumarten nicht mehr zur Verfügung stehen.

## 5.2 Saatgut

Die Ergebnisse der RG-Verfahren sind wesentlich von der Qualität der Saatpunkte abhängig. Daher sollen verschiedene Verfahren zur Gewinnung von Saatpunkten verglichen werden. Das verwendete Verfahren wird als Steuerungsparameter in `PREF.SEED.SAATGUT` gespeichert.

### 5.2.1 Saatpunkte durch regionale Maxima: (REGMAX)

Dieses Verfahren basiert von der Grundidee auf den *blob-detection* Verfahren nach PINZ, KÄTSCH und anderen (siehe 3.3), nur dass hier nicht direkt die Position von Baumspitzen extrahiert wird, sondern die ermittelten Maxima als Saatpunkte in einem RG-Verfahren weiter verarbeitet werden.

Die Maxima werden auf einem einzelnen Bildkanal ermittelt, der durch den Steuerungsparameter `PREF.SEED.CHANNEL` bestimmt wird. Im Allgemeinen eignet sich dafür der grüne oder der NIR-Kanal, die für Vegetation besonders kontrastreich sind. Die Verwendung von Farbinformationen erscheint bei der Bestimmung der regionalen Maxima dagegen überflüssig. Der gewählte Kanal wird mit einem Gaußfilter geglättet. Die Größe des Filters wird durch die Parameter `PREF.SEED.FSIZE` und `PREF.SEED.SIGMA` bestimmt. Die Filtergröße bestimmt wieder seine Anisotropie – je größer der Filter desto geringer. Da das Bild im Allgemeinen begrenzt und nicht periodisch ist, führen große Filter aber auch zu größeren Fehlern am Bildrand. Da ohnehin nur die Maxima des geglätteten Bildes betrachtet werden, ist eine gewisse Anisotropie hier eher in Kauf zu nehmen.

Die regionalen Maxima bezüglich 8-Nachbarschaft im geglätteten Bild sind die gesuchten Saatpunkte. Das Verfahren beruht also auf Annahme 15 (*blobs*) und Annahme 25 (Baumspitze), weiterhin geht Annahme 14 (Abstand zwischen Kronen ist bestimmt) ein. Die Position der Saatpunkte wird gemeinsam mit den Intensitätswerten im gewählten Kanal als `seed.list` gespeichert. Es besteht die Möglichkeit, zur Erhöhung der Anwendergenauigkeit bereits in diesem Schritt ein bestimmtes Perzentil der dunkelsten Saatpunkte auszufiltern. Dadurch wird zwar Rechenzeit gespart, da für schlechte Saatpunkte keine Regionen erzeugt werden müssen, andererseits geht hier Modellwissen ein und somit ist die Selektion besser im Modul Hypothesenauswahl angesiedelt (insbesondere haben Segmente mehr Merkmale als Punkte).

Das Verfahren ist vergleichsweise robust, da die eingehenden Annahmen sehr grundlegend und damit fast immer erfüllt sind. Das Maßstabsproblem wird nicht adressiert: REGMAX erfordert, dass der Abstand zwischen den Bäumen relativ einheitlich und vorab bekannt ist.

### 5.2.2 Saatpunkte durch Distanz-Transformation (DISTRANS)

Dieses Verfahren ist eine Weiterentwicklung des Verfahrens nach ERIKSON (2004B: 471, vgl. 3.12). Wie bei REGMAX wird der Kanal `PREF.SEED.CHANNEL` ausgewählt. Davon wird ein Schwellenwert in der Annahme gebildet, dass von jedem Baum Teile darüber liegen, dunkle Schatten zwischen Bäumen aber darunter. Hier geht also ebenfalls Annahme 15 (Bäume als helle *blobs*) ein. Die Höhe

des Schwellenwerts wird durch den Parameter `PREF.SEED.DISTRANS_THRESH` bestimmt. Der gültige Wertebereich ist  $]0..1[$ , sinnvolle Werte liegen unter 0.5. ERIKSON selbst verwendet Werte zwischen 0.15 (ERIKSON & OLOFSSON 2005) und 0.3 (ERIKSON 2004A). Dieser Parameter ermöglicht eine gewisse Anpassung an die radiometrische Qualität und Auflösung, da beispielsweise bei viel Streulicht oder geringer geometrischer Auflösung die Schatten ‚zulaufen‘, also heller werden. Das Schwellenwert-Bild wird dann einer Distanztransformation unterzogen (im Gegensatz zur 3-4-Distanztransformation im Original wird hier das euklidische Abstandsmaß verwendet), deren Ergebnis ein Distanz-Bild ist, in der jedes Pixel seinen kleinsten Abstand zu einem dunklen Bereich enthält, in der Annahme, dass es sich bei dem dunklen Bereich um Schatten (Annahme 21) zwischen den Bäumen handelt (Achtung: Hier geht außerdem Annahme 23: Homogenität der Krone, keine dunklen Flecken innerhalb der Krone ein). Diese ist jedoch nicht immer erfüllt, siehe Abb. 6.5) sind die lokalen Maxima, also die Punkte mit dem größten Abstand zum Schatten die Baumspitzen. Dabei geht die geometrische Symmetrie (Annahme 7) stärker ein als die radiometrische. Zur Bestimmung der Saatpunkte wird das Distanzbild mit einem Gaußfilter mit den Parametern `PREF.SEED.FSIZE` und `PREF.SEED.SIGMA` gefiltert und regionale Maxima gesucht. Diese werden gemeinsam mit ihren Helligkeiten und Distanzwerten in `seed.list` gespeichert.

Der große Vorteil von DISTRANS ist, dass für jeden Saatpunkt eine Schätzung der Größe des korrespondierenden Objektes geliefert wird. Das Maßstabsproblem wird somit zumindest teilweise in diesem Schritt adressiert, da anschließend Bäume unterschiedlicher Größe gesucht werden können. Durch die Verwendung eines Tiefpass-Filters werden aber wiederum relativ einheitliche Abstände zwischen den Objekten vorausgesetzt. Der Nachteil gegenüber REGMAX ist, dass es deutlich fehleranfälliger ist, da DISTRANS klar ausgeprägte Schatten und homogene Kronen voraussetzt.

### 5.3 Basic Region Growing (BRG)

Das Basic Region Growing (BRG) ist eine Weiterentwicklung der Segmentierung mit Fuzzy-Regeln nach ERIKSON (2003, 2004A, ERIKSON & OLOFSSON 2005, siehe 3.12) und wird hier so bezeichnet, da es sich um ein klassisches RG-Verfahren handelt. Ausgehend von einem Saatpunkt werden so lange Nachbarpunkte hinzugefügt, wie diese das Homogenitätskriterium erfüllen. Homogenität ist gegeben, wenn das Ähnlichkeitsmaß in (3.5) einen Mindestwert hat. Das Homogenitätskriterium kann also durch folgende Ungleichung ausgedrückt werden:

$$(5.1) \quad \exp\left(-\frac{1}{2}\left(\frac{(\mathbf{f}_p - \mathbf{f}_s)(\mathbf{f}_p - \mathbf{f}_s)}{\sigma_1} + \frac{(\mathbf{r}_p - \mathbf{r}_s)(\mathbf{r}_p - \mathbf{r}_s)}{\sigma_2}\right)\right) > \alpha.$$

Dabei ist  $\mathbf{r}_p$  die Position des Kandidaten-Pixels im Ortsraum,  $\mathbf{f}_p$  die Position im Farbraum,  $\mathbf{r}_s$  und  $\mathbf{f}_s$  sind die jeweiligen Positionen des Saatpixels,  $\sigma_1$  und  $\sigma_2$  sind Varianzen. Die Ähnlichkeit wird in der Implementation als `similarity` bezeichnet. Der Grenzwert  $\alpha$ , ab dem ein Pixel als ähnlich angesehen wird, wird im Parameter `PREF.RG.ALPHA` abgelegt. Die Ähnlichkeit kann als Produkt von zwei Funktionen verstanden werden (siehe ERIKSON 2003: 33), von denen eine nur vom räumlichen Abstand und die andere vom Farbabstand abhängig sind. Beide sind Normalverteilungen mit Mittelwert beim entsprechenden Wert des Saatpunktes. Entscheidend für die Bestimmung der Ähnlichkeit sind also die Varianzen, die bei gegebenem  $\alpha$  steuern, welche Abstände in Ortsraum und

Farbraum toleriert werden. Die Varianzen werden in den Parametern `PREF.RG.SIGMA1` und `PREF.RG.SIGMA2` abgelegt. Die Werte für  $\alpha = 0.15$  und  $\sigma_1 = 0.36$  wurden aus ERIKSON & OLOFSSON (2005: 262) übernommen, wobei ersterer frei gewählt ist, und letzterer derart angepasst, dass gemäß bei verschwindendem räumlichen Abstand um 40% der maximalen Intensität abweichende Farbwerte zugelassen werden:

$$\sigma_1 = \sqrt{\frac{3 * 0.4^2}{-2 \ln(\alpha)}} = 0.3557, \text{ vgl. (5.1).}$$

Ebenso wird der Parameter  $\sigma_2$  in Abhängigkeit vom größten zulässigen räumlichen Abstand  $\rho$  bestimmt. Dieser entspricht im Kreisfall dem Baumradius (nicht dem Durchmesser, wie in ERIKSON (2003: 1559) fälschlicher Weise angegeben). Es gilt also:

$$\sigma_2 = \sqrt{\frac{\delta^2}{-2 \ln(\alpha)}}.$$

Sind keine näheren Informationen über die Größen der einzelnen Objekte bekannt, wird der maximale Baumradius `PREF.MOD.SIZE.RAD_MAX` aus dem Modell entnommen, womit Modellwissen in den Segmentierungsschritt eingeht.

Zusätzlich wurde die Möglichkeit eingeführt,  $\rho$  aus den vom DISTRANS-Verfahren gelieferten Abständen schätzen zu lassen. Dafür muss der Parameter `PREF.RG.AUTO_SIZE` gesetzt werden. Der maximale Baumradius wird dann wie folgt aus dem DISTRANS-Abstand berechnet: Da dieser dem kleinsten Abstand vom Saatpunkt zu einem Schatten entspricht und somit nur als Schätzwert für den kleinsten Kronenradius gesehen werden kann, wird er mit dem Verhältnis zwischen größtem und kleinstem Kronendurchmesser (einer Krone) multipliziert. Dieser wird aus dem Modell entnommen, wo er als `PREF.MOD.ANISO` abgelegt ist. Weiterhin geht ein Korrekturfaktor in Abhängigkeit von `PREF.SEED.DISTRANS_THRESH` ein, da bei der Schwellenwertbildung ja ein Teil der Krone abgeschnitten und dem Schatten zugeordnet wurde. Zur Berechnung des Korrekturfaktors wird davon ausgegangen, dass die Intensität vom Saatpunkt zum Schatten linear auf Null abfällt. Das ist gegenüber dem Modell zwar eine Vereinfachung (die Höhe fällt beim Ellipsoid ja harmonisch ab, die Helligkeit richtungsabhängig unterschiedlich). Diese Über-Korrektur ist aber zu einem gewissen Grad erwünscht, da das Homogenitätskriterium in (5.1) ja so definiert wurde, dass ein Pixel im (minimal kleineren) Abstand  $\mathbf{r}$  nur bei völliger Farbgleichheit als ähnlich gilt. Abschließend wird noch kontrolliert, ob der berechnete Radius im zulässigen Bereich (Annahme 13) liegt.

Im Original wird nach der ersten Segmentierung das RG mit jedem Punkt der extrahierten Region als Saatpunkt wiederholt und anschließend nach gewissen Merkmalen die beste Region ausgewählt. Darauf wurde hier auf Grund exorbitant höherer Rechenzeiten verzichtet (bei Regionen mit tausenden Pixeln würde das Verfahren in der Größenordnung von  $10^3$ -mal so lange dauern, die Zeit für den Vergleich nicht eingeschlossen). Dadurch gewinnt aber die Qualität der Saatpunkte etwas an Bedeutung. Die Saatpunkte werden in Reihenfolge steigender Helligkeit abgearbeitet wobei bereits segmentierte Regionen überschrieben werden können. Dies könnte umgangen werden, indem zunächst für alle Pixel und Saatpunkte die Ähnlichkeiten berechnet würden und anschließend, ähnlich wie bei der Expansion in 5.4, in einem eigenen Schritt segmentiert würde. Auch darauf wurde auf Grund des höheren Rechenbedarfs und des vergleichsweise geringen Mehrwerts zunächst verzichtet (laufen zwei

Regionen stark ineinander, so sind meistens die Saatpunkte schlecht gewählt und beide Hypothesen fehlerhaft). Insbesondere bei einheitlichem  $\sigma_2$  für alle Saatpunkte (also ohne Option `AUTO_TREE_SIZE`) kann ein zu groß gewählter Radius zu zahlreichen Überschreibungen führen, was als Nachteil des Verfahrens zu werten ist. Ein zu kleiner Radius könnte dagegen eventuell durch eine Betrachtung der Verteilung der Farbwerte im Segment aufgespürt werden. Liegen diese wesentlich näher beisammen, als es bei der Berechnung von  $\sigma_1$  zu Grunde gelegt wurde, so war offensichtlich der räumliche Abstand die limitierende Größe und das Segment könnte zu klein geraten sein.

Auch in dieses Verfahren gehen implizit verschiedene Annahmen ein. Zunächst finden RG-Verfahren grundsätzlich kompakte Regionen, was Annahme 11 voraussetzt. In das Homogenitätskriterium geht Annahme 22 (radiometrische Homogenität) ein. Ausdrücklich nicht geht Annahme 15 (helle Baukronen) ein, da der Farbabstand von der Richtung im Farbraum unabhängig ist. Daher eignet sich das Verfahren prinzipiell auch, dunkle kompakte Regionen zu finden. Die hier im Speziellen gewählte Varianz  $\sigma_1$  macht allerdings nur Sinn, wenn die Intensität des Saatpunktes am Rande des Wertebereiches liegt (also hell oder dunkel), schließlich lässt es schon jetzt in allen Kanälen einen Abstand von 40% der Intensität oder von maximal  $\sqrt{3} * 0.4^2 = 69.3\%$  in einem einzelnen Kanal zu.

Da der Abstand im Ortsraum ebenfalls nicht von der Richtung abhängt, wächst die Krone isotrop. Damit geht auch hier die nur selten erfüllte Annahme 3 (Kreisförmigkeit) ein.

## 5.4 Brownsche Bewegung (BROWNMOTION)

Die Segmentierung mit Hilfe Brownscher Bewegung ist eine Weiterentwicklung der ‚*Segmentation via Brownian Motion*‘ nach ERIKSON (2004A: 26ff). Die Grundidee ist die Simulation der Bewegung eines virtuellen Teilchens im Bild. Dafür werden ausgehend von jedem Saatpunkt normalverteilte Zufallsvektoren zur Position des virtuellen Teilchens addiert. Dabei wird ausgehend von den einzelnen Saatpunkten gezählt, wie oft jedes Pixel vom virtuellen Teilchen besucht wird. Dies wird in einer 3D Matrix Namens `numPass` gespeichert (zwei Dimensionen für die Bildachsen, eine für die Saatpunkte). Da dieses Array relativ speicherintensiv ist, können die `numPass`-Bilder mit der Funktionen `bbDeflateNumPass` derart komprimiert werden, dass von jeder Ebene nur die Zeilen und Spalten, die Werte größer Null enthalten gespeichert werden.

Da das Teilchen nur innerhalb einer Krone umherlaufen soll, wird die Schrittweite wie folgt modifiziert: Je dunkler das Pixel, das eigentlich erreicht würde, desto kleiner der Schritt, das Teilchen wird also ausgebremst, wenn es ins Dunkle läuft. An dieser Stelle gehen somit Annahme 15 (*blobs*) und vor allem Annahme 21 (klare und dunkle Täler zwischen den Kronen) ein. Da diese Voraussetzung aber häufig nur unzureichend erfüllt ist, stellt sich die Frage, wie das Verfahren gegebenenfalls angepasst werden kann. Zu diesem Zweck wurde eine Reihe neuer Steuerungsparameter eingeführt.

Zunächst kann die Gewichtsfunktion, die die Schrittweite bestimmt, variiert werden. Die Position des Teilchens nach  $i$  Schritten wird bestimmt durch:

$$(5.2) \quad \mathbf{r}_i = \mathbf{r}_{i-1} + w(\mathbf{I}(\mathbf{r}_{i-1} + \mathbf{z}))\mathbf{z}.$$

Dabei ist  $\mathbf{r}_i$  die Position des Teilchens nach  $i$  Schritten,  $\mathbf{I}$  ist das Bild,  $\mathbf{z}$  ein Zufallsvektor und  $w$  die Gewichtsfunktion.

Da das Bild auf den Wertebereich [0..1] normiert ist, verwendet ERIKSON hierfür einfach einen zum Grauwert (in einem Kanal) des Zielpixels proportionalen Wert, also eine Gewichtsfunktion des Typs:

$$(5.3) \quad w(\mathbf{I}(\mathbf{r})) = c I_\lambda(\mathbf{r}).$$

Dabei ist  $c$  die Konstante und  $I_\lambda$  der gewählte Farbkanal.

Im Prinzip kann aber jede Funktion des Typs

$$(5.4) \quad w : \mathbf{I}(\mathbf{r}) \alpha [0 : 1],$$

die also den Farbraum eines Pixels auf das Intervall [0..1] abbildet, verwendet werden. Daher werden hier zu Testzwecken weitere Gewichtsfunktionen eingeführt und im Steuerungsparameter `PREF.BM.WEIGHT` abgelegt. Diese einfache Gewichtsfunktion wird nach dem benutzten Kanal bezeichnet, etwa 'G', die neuen Gewichtsfunktionen verwenden dagegen mehrere Kanäle. So wurde beispielsweise die Gewichtsfunktion 'RGB\_mean' implementiert, die das Mittel aus den 3 Farbkanälen R, G und B verwendet:

$$(5.5) \quad w_{RGB\_mean}(\mathbf{I}(\mathbf{r})) = \frac{I_R(\mathbf{r}) + I_G(\mathbf{r}) + I_B(\mathbf{r})}{3} = \frac{1}{n} \sum_{\lambda=1}^n I_\lambda(\mathbf{r}).$$

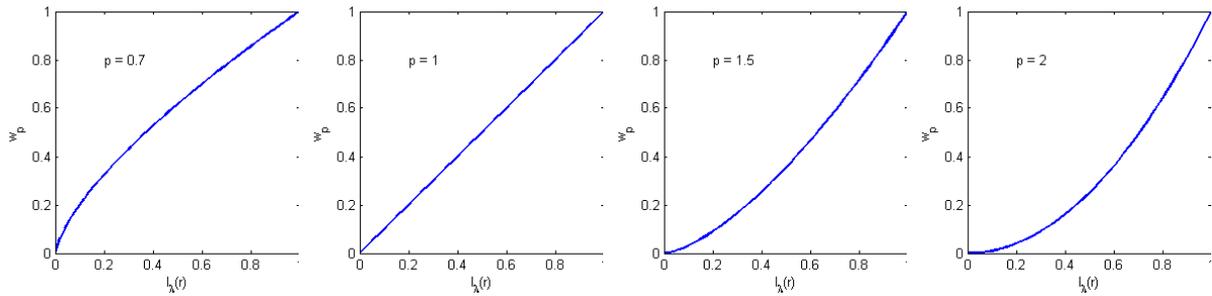
Dadurch wird ein Intensitätsabfall in anderen Farbkanälen berücksichtigt. Das macht beispielsweise Sinn, wenn neben den Kronen weitere Objekte zu sehen sind, die ebenfalls eine hohe Intensität im grünen Spektrum haben, aber in anderen Farben dunkler sind. Neben dem Mittel wäre natürlich auch eine Linearkombination von Kanälen der Art

$$w(\mathbf{I}(\mathbf{r})) = \frac{1}{n} \sum_{\lambda=1}^n a_\lambda I_\lambda(\mathbf{r}), \text{ mit } \sum_{\lambda} a_\lambda = 1 \text{ denkbar.}$$

Um die Schrittweite selbst zu modifizieren, können die bisherigen Gewichtsfunktionen potenziert werden. Damit werden sie zu:

$$w_p(\mathbf{I}(\mathbf{r})) = (I_\lambda(\mathbf{r}))^p \text{ für einen, bzw. } w_p(\mathbf{I}(\mathbf{r})) = \sum_{\lambda=1}^n a_\lambda (I_\lambda(\mathbf{r}))^p \text{ für mehrere Kanäle.}$$

Die Potenz  $p$  wird in einem neuen Steuerungsparameter namens `PREF.BM.WEIGHT_POT` gespeichert. Je höher die Potenz, desto langsamer läuft das Teilchen in dunklere Bereiche. So kann beispielsweise auf schlechte radiometrische Qualität oder Streulicht reagiert werden. Brauchbare Werte für `PREF.BM.WEIGHT_POT` liegen in etwa zwischen 0.7 und 2.5. Abb. 5.2 zeigt die Kennlinien für verschiedene Potenzen.



**Abb. 5.2:** Gewichtsfunktionen  $w_p$  für  $p = 0.7, 1, 1.5$  und  $2$

An der Darstellung wird auch die Verwandtschaft mit den homogenen Punkttransformationen in 2.3.3 deutlich. Der Algorithmus reagiert auf Grund der vielfachen Wiederholung sehr sensibel auf kleine Veränderungen von `PREF.BM.WEIGHT_POT`. Insbesondere soll darauf hingewiesen werden, dass die durchschnittliche Schrittweite durch die linke Gewichtsfunktion ( $p=0.7$ ) erhöht und durch die rechte ( $p=2.5$ ) verringert wird, was Auswirkungen auf die Größe der extrahierten Regionen hat (siehe unten).

ERIKSON variiert die Anzahl der Gesamtschritte von einigen Tausend bis hin zu Millionen (ERIKSON & OLOFSSON 2005: 262). Diese Anzahl wird nun im Steuerungsparameter `PREF.BM.NUM_STEPS` abgelegt. Dabei geht implizit Annahme 13 (Kronen haben eine bestimmte Größe) ein. Werden beispielsweise Kronen mit 3000 Pixeln erwartet, so ist es sehr unwahrscheinlich, dass bei 5000 Schritten jedes Pixel erreicht wird. Es muss also gelten: `PREF.BM.NUM_STEPS`  $\gg$  `AREA`.

Da es trotz angepasster Gewichtsfunktion je nach Qualität der Daten mehr oder weniger häufig geschieht, dass das Teilchen aus der jeweiligen Krone herausläuft, wurde das Verfahren in mehrere Einzelläufe unterteilt, d.h. das virtuelle Teilchen wird nach einer gewissen Anzahl von Teilschritten auf den Saatpunkt zurückgesetzt und startet erneut von dort. Die Anzahl der Teilläufe wird im zusätzlichen Parameter `PREF.BM.NUM_RUNS` abgelegt. Die Anzahl der Gesamtschritte beläuft sich damit auf `NUM_RUNS * NUM_STEPS`, aufgeteilt auf `NUM_RUNS` Einzelläufe. Jeder Einzellauf sollte dabei noch lang genug sein, um die ganze Region auszufüllen. Bricht das Teilchen jedoch aus der Region aus, bleibt der „Schaden“ dank Wiederholung kontrollierbar.

Zur Feinsteuerung wurde weiterhin der Parameter `PREF.BM.MIN_PASS` eingeführt, der festlegt, wie oft ein Pixel besucht werden muss, um im weiteren Verfahren berücksichtigt zu werden. Wird `MIN_PASS > 1` gewählt, so muss natürlich auch `NUM_STEPS` entsprechend angepasst werden. Es ist aber fraglich, ob durch diesen Parameter eine wirkliche Steuerung erzielt werden kann. Vielmehr wird durch seine Festlegung das Signifikanzniveau der Messung bestimmt.

Nach der Erzeugung des `numPass`-Bildes beginnt das eigentliche Region Growing, das bei ERIKSON Expansion genannt wird. Bei ERIKSON werden die Regionen mit einem Kreis initiiert, dessen Mittelpunkt der Saatpunkt ist, und der gerade ins `numPass`-Bild passt (wobei Annahme 3: Kreisförmigkeit eingeht). Dies spart Rechenzeit, setzt allerdings voraus, dass die Kreise sich nicht überlappen. Daher kann die Initiierung mit Kreisen durch den Parameter `PREF.BM.INIT_CIRCLE` ausgeschaltet werden. In diesem Fall werden die Regionen nur mit ihren Saatpunkten initiiert.

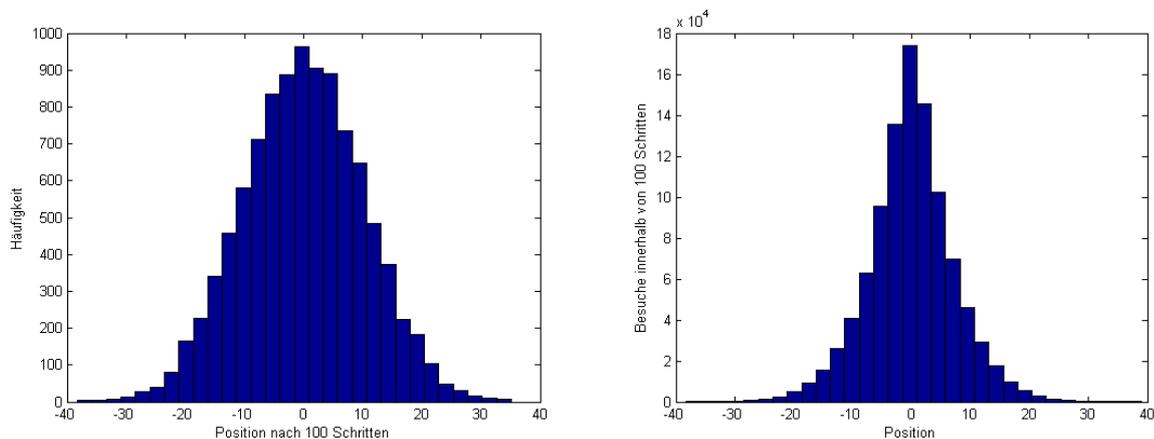
Zum RG werden dann zunächst alle Nachbar-Pixel von Regionen zusammen mit der entsprechenden Region und dem `numPass`-Wert des Pixels in der Region als Kandidaten in eine Liste geschrieben.

Als Nachbarn werden nur die vier unmittelbar angrenzenden Pixel gesehen, damit die erzeugten Regionen in 4-Nachbarschaft zusammenhängend sind. Die Regionen werden im Array `regions` gespeichert.

Die Liste wird dann in Reihenfolge der `numPass`-Werte abgearbeitet. Gehört ein Kandidatenpixel bislang keiner Region an und erfüllt die `MIN_PASS` Bedingung, so wird es der entsprechenden Region zugeordnet. Anschließend wird die Kandidatenliste mit allen Nachbarn des Pixels aktualisiert, die im Bild liegen, keiner Region angehören und einen ausreichenden `numPass`-Wert in der entsprechenden Ebene haben. Kandidaten, die eine der Bedingungen nicht erfüllen, werden verworfen, der Algorithmus terminiert, wenn die Kandidatenliste leer ist.

Der große Vorteil dieses Verfahrens ist, dass es „relativ“ unterschiedliche Formen zulässt (die geometrischen Annahmen also „nicht so stark“ eingehen), und in einem großen Maßstab die Kronenbegrenzung genau segmentieren kann. Im Folgenden soll aber gezeigt werden, dass es trotzdem nicht völlig frei von Annahmen über die Form ist. Zunächst finden Region Growing Verfahren immer kompakte Regionen, womit Annahme 11 eingeht. Um der Segmentierungsbedingung der Vollständigkeit zu genügen, muss übrigens der Hintergrund als Null-Region definiert werden.

Daneben wird die Form aber auch durch das Zufallsexperiment der Teilchenbewegung bestimmt. Es wurde bereits darauf hingewiesen, dass die Anzahl der Schritte die Größe der Region mitbestimmt. Dies liegt daran, dass die Aufenthaltshäufigkeit des Teilchens nach außen hin abnimmt. Zur Verdeutlichung soll hier ein vereinfachtes Gedanken-Experiment dienen. Dazu bewege sich das Teilchen nur auf einer Achse und in einem völlig homogenen weißen Bild, die Gewichtsfunktion seien also konstant 1. Abb. 5.3 zeigt links die Verteilung der Teilchenpositionen nach jeweils 100 Schritten (`NUM_STEPS`) bei 10000 Durchgängen (`NUM_RUNS`).

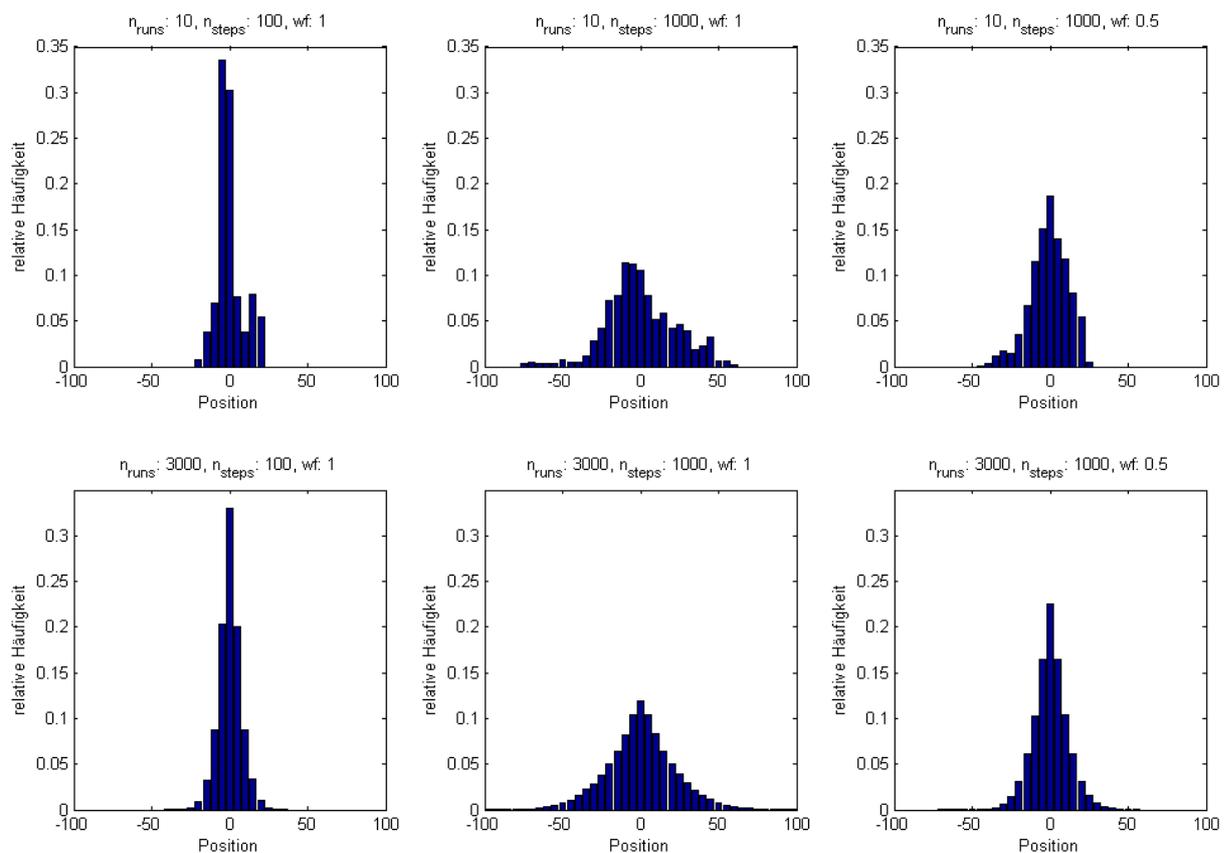


**Abb. 5.3: Idealisertes Experiment. Links: Position des Teilchens nach 1000 Schritten, Verteilung bei 10000 Versuchen. Rechts: Besuchshäufigkeit**

Das Ergebnis hat die Form einer Binomialverteilung, die im Grenzfall von unendlich vielen Durchgängen zur Normalverteilung wird. Da dies natürlich nicht nur für den Schritt  $i = 100$ , sondern auch für jeden anderen Zeitpunkt gilt, ist die Besuchshäufigkeit (Abb. 5.3, rechts: Summe der Häufigkeiten an den Positionen über alle Zeitschritte) eine Summe von Normalverteilungen mit gleichem Mittelwert.

Daraus lässt sich bereits zweierlei ablesen: Erstens fällt die Besuchshäufigkeit nach außen hin stark ab, und zweitens fällt sie in alle Richtungen gleichmäßig ab. Damit geht die Distanz vom Saatpunkt nachweislich ins Verfahren ein, da im Expansionsschritt ja hinterher die Besuchshäufigkeiten von verschiedenen Saatpunkten miteinander verglichen werden. Insbesondere werden kreisförmige Regionen erzeugt, wenn keine konkurrierenden Saatpunkte vorhanden sind und einfach der MIN\_PASS Schwellenwert gebildet wird. Insofern geht auch hier immer (also unabhängig von der Initiierung des Expansionsschritts) die meistens nicht zutreffende Annahme 3 (Kreisförmigkeit) ein.

Abb. 5.4 zeigt darüber hinaus, wie die Wahl der Steuerungsparameter das Ergebnis beeinflusst. Dafür wurde das obige Experiment mit verschiedenen Parameter-Werten wiederholt. Um die verschiedenen Einstellungen vergleichen zu können, wurde von absoluten Besuchszahlen zu relativen Aufenthaltshäufigkeiten übergegangen.



**Abb. 5.4: Relative Aufenthaltshäufigkeiten für verschiedene Parameter. Oben: 10 Durchläufe; Unten: 3000. Links: 100 Schritte; Mitte: 1000 Schritte; Rechts: 1000 Schritte bei konstanter Gewichtsfunktion von 0.5**

Die obere Zeile zeigt die Ergebnisse von jeweils 10 Durchläufen (NUM\_RUNS), die untere das von 3000. Es ist klar zu erkennen, dass dadurch kein grundlegend anderes Ergebnis produziert wird, sondern nur die Bestimmtheit (also die Messgenauigkeit) erhöht. Von links nach rechts werden der Parameter NUM\_STEPS sowie die Gewichtsfunktion variiert. Links werden pro Durchlauf 100 Schritte gemacht, in der Mitte 1000. Es ist klar erkennbar, dass die Varianz der Verteilung (also die Größe der erzeugten Region) durch eine größere Anzahl von Schritten zunimmt. In der rechten Spalte ist NUM\_STEPS ebenfalls 1000, dafür wurde aber die Gewichtsfunktion variiert. Hierfür wurde einfach ein konstanter Faktor gemäß (5.3) von  $wf = c = 0.5$  eingeführt. Dies entspricht vom Effekt her

einer höheren `WEIGHT_POT`, die ja die durchschnittliche Schrittlänge verkürzt (siehe oben). Erwartungsgemäß werden dadurch die Breite der Verteilung und damit die Größe der extrahierten Region reduziert.

Durch das Gedankenexperiment wurde also hergeleitet, welche impliziten Annahmen über die Form im Verfahren stecken und wie diese durch die Wahl der Parameter beeinflusst werden. Nichts desto trotz sind die Auswirkungen weitaus schwächer als bei den beiden anderen vorgestellten Verfahren und durch die starke Gewichtung der radiometrischen Charakteristik sind die Form und die Größe der Krone weiterhin relativ variabel. Das Maßstabsproblem wird also in gewissem Maße im Segmentierungsschritt adressiert, wenn auch weniger deutlich als beim BRG mit `DISTRANS` und `AUTO_SIZE`. Eine Möglichkeit könnte sein, mit Hilfe dieser Größeninformationen die jeweiligen Gewichtsfunktionen zu modifizieren.

Dafür ist aber die genaue Position der Saatpunkte weniger entscheidend als beim BRG. Sehr relevant ist dagegen die Anzahl der Saatpunkte, da aus jedem Saatpunkt eine eigene gleichberechtigte Region wird. So können falsche Saatpunkte beispielsweise das Wachstum richtiger Regionen verhindern.

Werden beide Aspekte zusammen genommen, lässt sich das Maßstabsproblem in zwei Teilprobleme trennen: Dem Problem der richtigen Baumgröße wird durch `BROWNMO` zumindest teilweise Rechnung getragen. Das Problem des richtigen Baumabstandes liegt dagegen bei der Erzeugung der Saatpunkte und wird hier nicht gelöst.

Ein Nachteil des Verfahrens ist, dass es nicht deterministisch ist, also nicht zwangsläufig bei jeder Instanzierung zum selben Ergebnis kommt. Der Ausgang kann zwar fixiert werden, indem der Zufallsgenerator mit einem bestimmten Wert initiiert wird, aber dann ist das Ergebnis von diesem Initialwert abhängig. Dieser Nachteil ist jedoch kontrollierbar, da das Verfahren bei der großen Anzahl von Schritten gemäß dem Gesetz der großen Zahlen zu einem Ergebnis konvergiert und dann stabil ist.

Ein weiterer Nachteil ist, dass der Algorithmus sehr rechenintensiv ist, da `NUM_RUNS * NUM_STEPS` mal der Farbwert des Zielpixel gefunden, die Gewichtsfunktion des Farbwerts ausgerechnet und der gewichtete Schritt gemacht werden muss. Auch das listenbasierte Region Growing ist eher langsam, da durch die Forderung, dass jeweils nur das Pixel mit dem höchsten `numPass`-Wert hinzugefügt wird, nach jeder Aktualisierung sortiert werden muss.

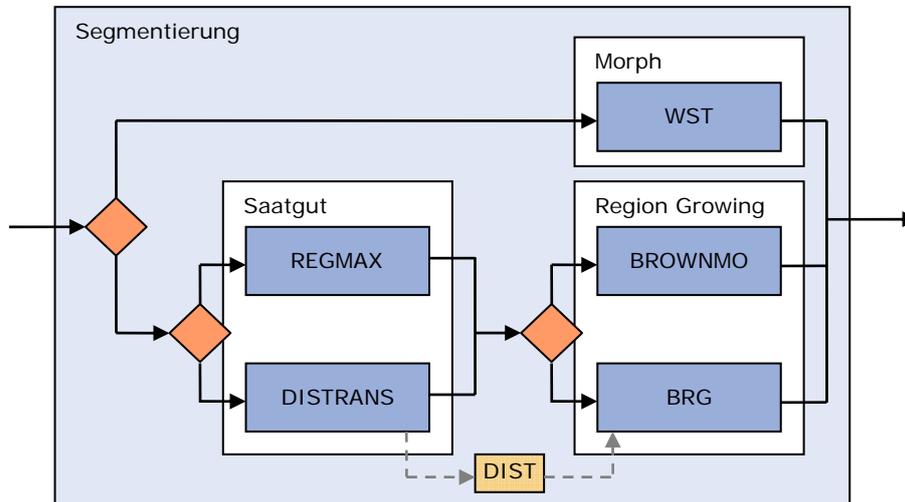
Daher stellt sich die Frage, ob das gleiche Ergebnis nicht auch einfacher zu erzielen wäre. Das obige Gedankenexperiment hat ja gezeigt, dass nach sehr vielen Durchgängen eine einfache parametrische Verteilung produziert wird. Zu klären wäre also nur, wie sich die Intensitäten der Pixel auf die Verteilung auswirken. Dabei bringt der Übergang zur relativen Aufenthaltshäufigkeit in Abb. 5.4 einen wertvollen Denkanstoß, da diese an die Aufenthaltswahrscheinlichkeit in der Quantenmechanik erinnert. Es wäre also zu überprüfen, ob sich die Intensitätsverteilung und die Gewichtsfunktionen nicht als Potential angeben ließen (in dem sich das virtuelle Teilchen bewegt), und wie dieses Potential beschaffen sein müsste. Wird das Bild morphologisch als Grauwertgebirge betrachtet, so beschleunigt das Teilchen ja bislang, wenn es bergauf läuft, und bremst bergab. Leider kann dieser Frage im Rahmen der vorliegenden Arbeit nicht weiter nachgegangen werden.

Ein weiterer Nachteil ist, dass die Parameter schwer zu handhaben sein könnten. So wurde zwar der Funktionsumfang des Algorithmus erweitert, insbesondere die Wahl einer geeigneten Gewichtsfunktion könnte aber Schwierigkeiten bereiten. Es wäre daher zu überprüfen, inwiefern die Bestimmung

dieses Parameters automatisiert werden könnte, etwa durch eine Analyse der Histogramme der verwendeten Bänder.

## 5.5 Übersicht

Abb. 5.5 gibt eine Übersicht über die implementierten Algorithmen und den internen Ablauf des Moduls Segmentierung aus Abb. 4.2.



**Abb. 5.5: Ablauf des Moduls Segmentierung und implementierte Algorithmen**

In Tab. 2 wird noch einmal zusammengefasst, welche Voraussetzungen und Modell-Parameter in die oben vorgestellten Verfahren eingehen. Da die formulierten Annahmen häufig mehrfach den gleichen Sachverhalt beschreiben, wurde sich dabei auf wesentliche Annahmen konzentriert. Um zum Ausdruck zu bringen, dass Voraussetzungen unterschiedlich stark in Verfahren eingehen, wurden neben „ja“ und „nein“ auch die Stufen „schwach“ und „ja, stark“ verwendet. „Ergebnis“ bedeutet, dass der entsprechende Parameter vom Verfahren selbst geliefert wird.

Manche Annahmen gehen auch indirekt ein, da sie von anderen abhängig sind. In einzelnen Fällen wurde dies in einer Fußnote festgehalten, insgesamt eignet sich diese Übersicht aber nicht, sämtliche Abhängigkeiten noch einmal darzustellen. Darum wurden auch jeweils nur die relevanten Felder gefüllt und es sei dafür noch einmal auf die Abschnitte 4.4.1 und 5.1 bis 5.4 verwiesen.

**Tab. 2: Voraussetzungen und benötigte Modell-Parameter für ausgewählte Segmentierungsverfahren**

Voraussetzung (Annahme)	Morphologie	Saatgut		Region Growing	
		WST	REGMAX	DISTRANS	BRG
<i>geometrisch</i>					
Kreisförmigkeit (3)	erzwungen	ja	nein	ja	schwach
3D-Hülle ist konvex (4)	ja (bei DOM)				
Symmetrie bzgl. Baumspitze (7)			ja		
Kompaktheit (11)				ja	ja
Größe ist bestimmt (13)	nein [2]	nein	Ergebnis	ja	schwach
Abstand ist bestimmt (14)		ja	ja	nein	indirekt [4]
<i>radiometrisch</i>					
Blobs (15)	ja	ja, stark	ja	nein	ja
dunkle Täler zwischen Kronen (21)			ja		ja, stark
radiometrische Homogenität (22)	nein			ja [1]	indirekt [6]
keine dunklen Pixel in der Krone (23)		nein	ja, stark	nein	schwach
Baumspitze ist Intensitätsmaximum (25)		ja			
radiometrische Konvexität (27)	ja (optisch)				
<b>Modell-Parameter</b>					
<i>Einzelkrone</i>					
Radius (MOD.SIZE.RAD)			Ergebnis	ja [5]	
Abstand	nein [2]	ja	ja		
Anisometrie (MOD.ANISO)				ja [3]	
<i>Bestand</i>					
minimaler & maximaler Radius				ja	

[1]: maximaler Abstand im Farbraum als Parameter

[2]: da verschiedene Maßstabsebenen

[3]: bei PREF.RG.AUTO\_SIZE

[4]: da Kronen nicht überlappen können

[5]: bei NOT\_PREF.RG.AUTO\_SIZE

[6]: helle Bereiche werden bevorzugt

## 6 Test der Segmentierungsverfahren und Diskussion

In diesem Kapitel werden die mit den implementierten Algorithmen erzielten Ergebnisse vorgestellt. Dabei geht es nicht darum, die Algorithmen für eine gegebene Situation zu optimieren, um möglichst gute Ergebnisse zu erzielen. Vielmehr soll exemplarisch gezeigt werden, wo die Grenzen der Übertragbarkeit von Verfahren liegen, und warum sie nicht bei allen Datensätzen funktionieren. Weiterhin soll untersucht werden, wie die oben eingeführten Steuerungsparameter das Extraktionsergebnis beeinflussen und ob damit der Funktionsumfang der Segmentierungsverfahren erweitert werden kann.

### 6.1 Datensätze und Untersuchungsgebiete

Zunächst werden die verwendeten Bilder vorgestellt. Dabei soll nicht nur auf die Aufnahmebedingungen und Untersuchungsgebiete eingegangen werden, sondern gleich überprüft werden, welche der in 4.4.1 aufgestellten Annahmen von den einzelnen Datensätzen erfüllt werden, um so einen analytischen Zugriff auf die Übertragbarkeit von Verfahren zu bekommen.

#### 6.1.1 La Motte Plantation, Western Cape Province, Südafrika

Der erste Datensatz ist ein Luftbild einer Kiefern-Plantage in der Nähe von Franschhoek in der Western-Cape-Province in Südafrika. Die Aufnahme zeigt einen Teil der La Motte Plantation der SAFCOL (South African Forestry Company Limited), die nördlich des Hottentotten-Naturschutzgebiets liegt. Die Aufnahme entstand im März 2006 und wurde mit einer digitalen Spiegelreflexkamera aus einem Ultraleichtflugzeug bei geringer Flughöhe (etwa 300m) gemacht. Das Bild wurden mit der *Leica Photogrammetry Suite* anhand mit Differential-GPS vermessener *ground control points* geometrisch korrigiert und hat eine Bodenauflösung von ca. 15 cm. Verarbeitet wurde ein mit hoher Qualität JPEG-komprimiertes Bild. Eine radiometrische Korrektur wurde nicht durchgeführt. Da das Bild relativ kontrastarm ist und über die spektrale Empfindlichkeit, den dynamischen Bereich und die Kalibrierung der Kamera keine weiteren Informationen vorliegen, ist die radiometrische Qualität des Bildes als eher niedrig zu bewerten. Als Beleg dafür kann die Verteilung der Farbwerte in den einzelnen Kanälen angeführt werden. So liegen im Ausschnitt *lamotte1* (siehe Abb. 6.2) nur 3% der Pixel unter 0.15 der maximalen Intensität, wobei deutliche Unterschiede zwischen den einzelnen Farbkanälen auszumachen sind: Während im roten Kanal 8.9% darunter liegen, sind es im grünen nur 0.36% und im blauen sogar nur rund 0.034%. Da davon ausgegangen werden kann, dass die Schatten im gesamten Spektrum schwarz sein müssten, und dass es auf Grund des eisenhaltigen Bodens im Zweifel im roten Kanal mehr Streulicht als im blauen geben müsste, scheint die Kamera im roten Spektrum unempfindlicher zu sein als im blauen.

Natürlich könnte der Kontrast mit einfachen Mitteln erhöht werden. Das Bild wurde aber gezielt gewählt, um zu untersuchen, ob auch mit Aufnahmen niedrigerer Qualität, die dafür günstiger zu akquirieren sind, gute Ergebnisse erzielt werden können. Es soll damit getestet werden, wie die Verfahren an die radiometrische Situation anzupassen sind.



**Abb. 6.1: Szene 1: Kiefern-Plantage in der Nähe von Franschhoek, Südafrika**

Zu sehen sind etwa 12-jährige *Pinus radiata*. Die Bestandssituation ist einfach, da alle Bäume gleich alt und in regelmäßigen Abständen angepflanzt sind. Damit sind Annahme 13 und Annahme 14 erfüllt. In der Bildmitte sind die Kronen relativ kreisförmig (Annahme 3), am Rande sind sie jedoch auf Grund der niedrigen Flughöhe schräg von der Seite zu sehen. Die Kronen sind kompakt (Annahme 11). Die Bäume sind klar als helle *blobs* zu erkennen und es gibt Schatten zwischen den Bäumen. Annahme 15 ist somit erfüllt, Annahme 21 jedoch nur bedingt, da neben den Schatten auch helle Bodenflecken an die Kronen angrenzen. Diese stellen eine Erschwernis bei der Extraktion dar, und das verwendete Modell muss in der Lage sein, helle Bodenflecken von Bäumen zu unterscheiden. Die Kronen sind aber radiometrisch relativ homogen (Annahme 22) und enthalten (abgesehen von Bäumen im Schatten) kaum dunkle Pixel. Annahme 23 kann also ebenfalls als erfüllt betrachtet werden.

Aus der Szene wurden zu Testzwecken zwei Teilbilder ausgeschnitten, die in Abb. 6.2 zu sehen sind.



**Abb. 6.2: Ausschnitte aus Szene 1: Bilder *lamotte1* (links) und *lamotte2* (rechts)**

„*lamotte1*“ ist aus der Szenenmitte und erfüllt Annahme 3, „*lamotte2*“ ist aus der rechten oberen Ecke, wo die Bäume schon geometrisch verzerrt zu sehen sind.

### 6.1.2 Göttinger Wald

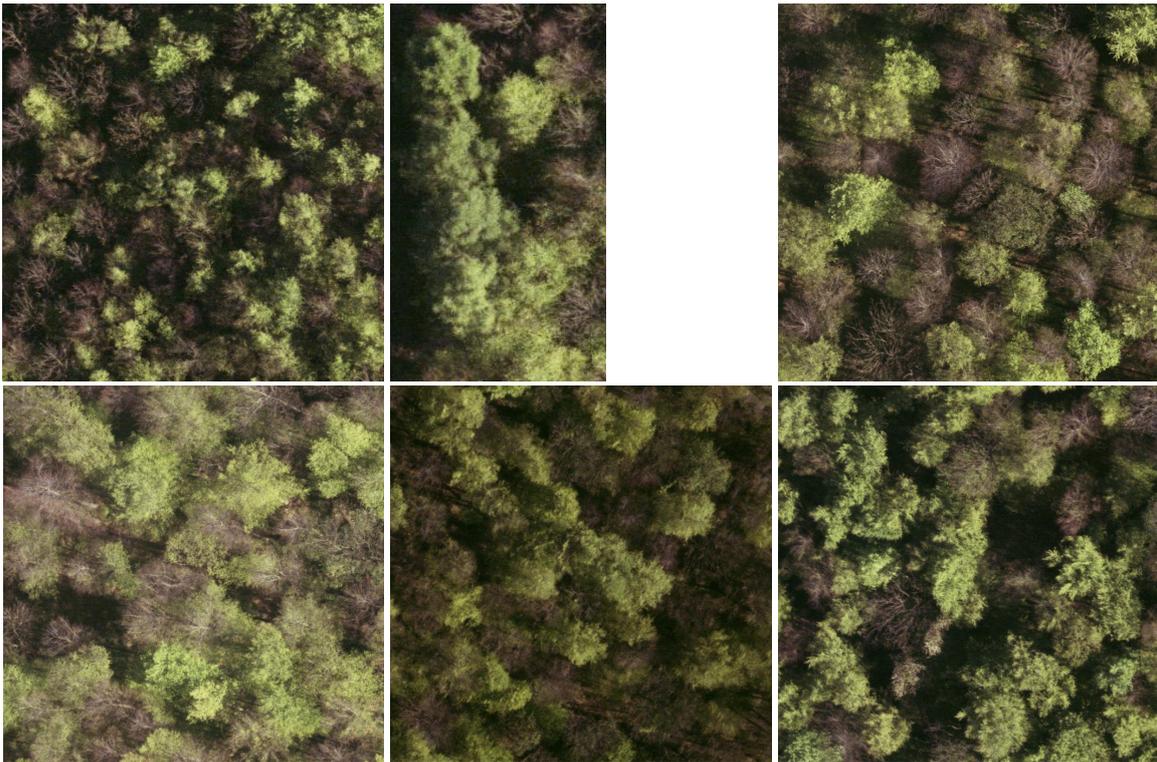
Bei der zweiten Szene handelt es sich ebenfalls um eine Luftaufnahme, allerdings von einer völlig anderen Wald-Bestandssituation. Die Aufnahme zeigt einen Buchen-Edellaubholz Mischwald im Alter von etwa 100 bis 140 Jahren aus dem Niedersächsischen Forstamt Reinhausen und wurde am frühen Morgen Ende April 1999 mit einer RMK Leica Typ Wild RC30 aus einer Cessna 404 gemacht. Der Maßstab des Fotos beträgt etwa 1:4900 und es wurde mit 25µm gescannt. Da es sich um eine Frühjahrsaufnahme handelt, sind Esche (*Fraxinus L.*) und Ahorn (*Acer L.*) noch nicht voll ausgetrieben, die Buche (*Fagus L.*) erscheint dagegen in saftigem Hellgrün.



**Abb. 6.3: Szene 2: Luftaufnahme aus dem Göttinger Wald, Niedersachsen**

Aus der Gesamtszene wurden 6 Ausschnitte mit unterschiedlichen Situationen ausgewählt, die in Abb. 6.4 zu sehen sind.

Im Vergleich zu Szene 1 handelt es sich um eine ausgesprochen komplexe Bestandsituation. Zunächst liefert ohnehin keines der bekannten Verfahren zur Baumextraktion bei Laubmischwäldern befriedigende Ergebnisse. Dazu kommt in diesem Fall noch erschwerend, dass zwischen den grünen Kronen winterlich kahle Kronen zu sehen sind, die praktisch keine der radiometrischen Annahmen erfüllen. Dies erschwert einerseits die Segmentierung, andererseits müsste das Modell in der Lage sein, grüne und nicht ausgetriebene Kronen zu unterscheiden. Das Bild wurde ausgewählt, um die Eignung der eingeführten Steuerungsparameter zu testen.



**Abb. 6.4:** Verwendete Ausschnitte aus Szene 2. Oben:  $goett\{1\}$ ,  $goett\{2\}$  und  $goett\{3\}$ ; Unten:  $goett\{4\}$ ,  $goett\{5\}$  und  $goett\{6\}$

Es ist klar zu erkennen, dass die elementare geometrische Voraussetzung der Kreisförmigkeit (Annahme 3) nur ansatzweise erfüllt ist, und selbst die Kompaktheit (Annahme 11) ist nicht bei allen Kronen gegeben, wie in Abb. 6.5 zu sehen ist. Die Größen der einzelnen Kronen sind zwar bestimmt, variieren aber über einen großen Bereich, ebenso der Abstand zwischen den Kronen.

Auch die radiometrischen Annahmen sind nur bedingt erfüllt. So sind die Kronen zwar hell und an vielen Stellen sind auch dunkle Täler dazwischen erkennbar, dafür sind sie aber keineswegs homogen und enthalten viele dunkle Pixel.

In Abb. 6.5 wird gezeigt, wie sich das Verhalten im Maßstabsraum verändert. Dazu wurde ein Ausschnitt aus ‚ $goett\{6\}$ ‘ bilinear auf die Hälfte und ein Viertel der Größe resampelt. Es ist deutlich zu erkennen, wie die Homogenität in der Krone zunimmt, während die dunklen Täler zwischen den Kronen verschwinden.

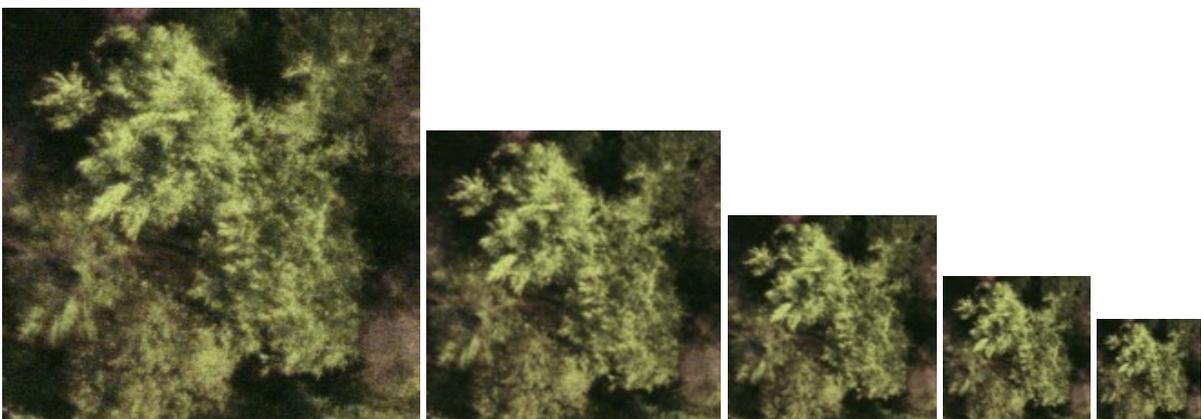




Abb. 6.5: Ausschnitt aus *goett{6}* (links); bilinear-resampelt auf halbe Größe (Mitte) und auf ein Viertel (rechts).

### 6.1.3 Hans Kayinga Forest, Kavango, Namibia

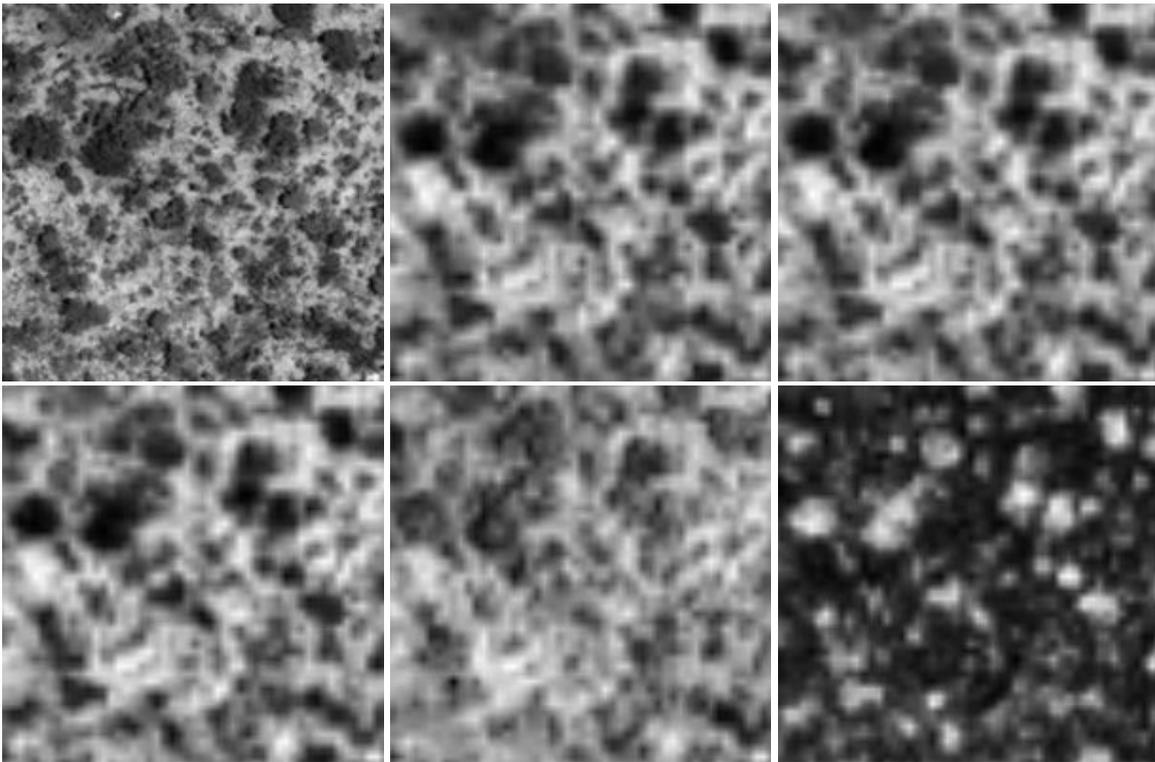
Die dritte Szene ist eine Quickbird-Aufnahme des Hans Kanyinga Community Forest in der Region Kavango im Norden Namibias.

Quickbird ist zurzeit unter den kommerziellen Satelliten derjenige mit der höchsten Auflösung (der Start des ebenfalls amerikanischen Satelliten GeoEye-1 wurde mehrfach verschoben und zuletzt für das Frühjahr 2008 angekündigt). Quickbird erreicht im panchromatischen Kanal eine Nadir-Auflösung von 60 cm. Daneben hat er vier multispektrale Kanäle in den Wellenlängenbereichen 450-520 nm (blau), 520-600 nm (grün), 630-690 nm (rot) sowie 760-900 nm (NIR). Die multispektralen Kanäle haben jeweils eine Nadir-Auflösung von 2.40 m. Ein Bildstreifen ist 16.5 km breit und die Quantisierung (also die radiometrische Auflösung) ist mit 11 Bit höher als bei den meisten Sensoren. Quickbird-Bilder werden von der Firma DigitalGlobe vertrieben und bilden beispielsweise an vielen Stellen den Hintergrund für den Planeten-Browser Google Earth.

Die Szene zeigt eine tropische Baumsavanne mit etwa 26 Baumarten, wobei *Baikiea plurijuga*, *Guibourtia coleosperma*, *Pterocarpus angolensis*, *Burkea africana* und *Dialium engleranum* die dominanten Baumarten sind (KAMWI 2007: 17, CURTIS & MANNHEIMER 2005). Das Klima ist für Namibia mit einem mittleren jährlichen Niederschlag von knapp 600mm (VERLINDEN & LAAMANEN 2006: 235) zwar vergleichsweise feucht, in den Monaten Mai bis September fällt aber kein Niederschlag (KAMWI 2007: 16).

Die Szene besteht aus zwei Teilbildern, die beide vom Satelliten Quickbird II stammen. Die Bilder wurden an unterschiedlichen Tagen zwischen Ende Oktober und Ende November 2004 jeweils früh morgens aufgenommen und sind radiometrisch korrigiert. Für vorangegangene wissenschaftliche Arbeiten wurden in dem Gebiet bereits Inventur-Stichproben an über 200 Stellen erhoben, die im Folgenden Plots genannt werden. Der Autor hatte selbst die Gelegenheit, im April 2007 an einer Forst-Inventur teilzunehmen, bei der einzelne Plots sehr genau erfasst wurden. Da in der Region aber kein Differential-GPS Signal zu Verfügung steht, ist die absolute Plot-Position nur in der üblichen GPS-Genauigkeit bekannt und somit für ein Einzelbaum-Matching ungenügend.

Aus der Szene wurden verschiedene Sample-Plots ausgewählt und ausgeschnitten, wobei die Spektralkanäle jeweils mit SAGA (siehe CONRAD 2006) bilinear auf die Auflösung des panchromatischen Kanals resampelt wurden. In Abb. 6.6 sind alle Kanäle der Quickbird-Szene und der NDVI für Plot 155 (*namb1*) zu sehen.



**Abb. 6.6:** Plot 155 aus der Namibia-Quickbird-Szene (*namb1*). Oben: panchromatisch, blau und grün; Unten: rot, NIR, NDVI

Das *resampling* hat in den Farbkanälen Artefakte geschaffen, die als helle horizontale und vertikale Linien erkennbar sind. Bei allen Kanälen wurde zur Darstellung der Grauwertbereich gespreizt.

Das Bild wurde vor allem ausgewählt, um zu demonstrieren, wie unterschiedlich Bäume im Bild aussehen können. Zu sehen sind teils einzelne Bäume, teils Gruppen von etwa 4 bis 12 Bäumen. Besonders auffällig ist, dass entgegen Annahme 15 der Hintergrund höhere Intensität hat als die Kronen. Dies liegt am hellen sandigen Boden und ist in den Farben rot, grün und blau besonders ausgeprägt, aber selbst im NIR noch deutlich. Im gespreizten NDVI ist die Vegetation zwar klar zu erkennen, aber auch die Werte für Nicht-Vegetation liegen sämtlich über Null. Damit sind beinahe sämtliche radiometrischen Annahmen nicht erfüllt und das entwickelte Baum-Modell ist vollkommen unbrauchbar. Auch das geometrische Bildmodell ist nicht zu verwenden, da (wie in Abb. 6.7 zu sehen ist) elementare Annahmen wie Kompaktheit oder Bestimmtheit des Abstandes nicht erfüllt sind.



**Abb. 6.7:** Geometrische Beschaffenheit verschiedener Bäume im Hans Kanyinga Forest (a-c); Team bei der Forstinventur im April 2007 (d)

Es ist zu erkennen, dass manche Bäume mehrere Kronen haben, die nicht zusammenhängend sind. Im zweiten Bild greifen mehrere Baumkronen ineinander, so dass sie überhaupt nicht zu trennen sind. Dadurch wird die Segmentierung einzelner Kronen praktisch unmöglich. Es soll dennoch überprüft werden, ob sich eines der vorgestellten Verfahren eignet, zumindest die Baumgruppen in ‚*namb1*‘ zu segmentieren.

#### 6.1.4 Übersicht

In Tab. 3 ist noch einmal zusammengefasst, welche der oben für die Algorithmen als relevant herausgearbeiteten radiometrischen und geometrischen Voraussetzungen von den einzelnen Datensätzen erfüllt werden. Um zu zeigen, dass dies in unterschiedlicher Abstufung der Fall sein kann, wurden neben den Kategorien ‚ja‘ und ‚nein‘ auch abgestufte Schulnoten vergeben, ‚teilweise‘ und ‚überwiegend‘ bedeuten, dass die entsprechende Annahme nicht überall im Bild zutreffend ist.

**Tab. 3: Erfüllte Voraussetzungen der Datensätze *lamotte1*, *lamotte2*, *goett* und *namb1***

Voraussetzung (Annahme)		<i>lamotte1</i>	<i>lamotte2</i>	<i>goett</i>	<i>namb1</i>
<i>geometrisch</i>					
Kreisförmigkeit	(3)	sehr gut	befriedigend	ausreichend	nein
3D-Hülle ist konvex	(4)				
Symmetrie bzgl. Baumspitze	(7)	ja	nein		nein
Kompaktheit	(11)	gut	gut	überwiegend	nein
bestimmte Größe	(13)	ja	ja	ausreichend	
bestimmter Abstand	(14)	ja	ja	ausreichend	nein
<i>radiometrisch</i>					
Blobs	(15)	ja	ja	ja	nein
dunkle Täler zwischen Kronen	(21)	überwiegend [1]	überwiegend [1]	teilweise	nein
radiometrische Homogenität	(22)	gut	gut	mangelhaft	ja
keine dunklen Pixel in der Krone	(23)	sehr gut	sehr gut	nein	nein
Baumspitze ist Intensitätsmaximum	(25)	gut	nein	teilweise	teilweise[2]
radiometrische Konvexität	(27)	sehr gut	gut		nein

[1]: helle Bodenflecken grenzen ebenfalls an Kronen an

[2]: sofern es eine Baumspitze gibt, ist die Intensität meist höher als die der Nachbapixel, nicht jedoch als die der Umgebung

Im Folgenden werden die Ergebnisse der einzelnen Verfahren vorgestellt. Dabei wird zunächst in der Reihenfolge der Verfahren aus Kapitel 5 vorgegangen und die `PARAMETER` beziehen sich auf die dort eingeführten Steuerungsparameter. Bei jedem einzelnen Verfahren wird zunächst eine tabellarische Übersicht über die wichtigsten Voraussetzungen gegeben. Anhand derer soll überprüft werden, ob das in 4.4.1 entwickelte Set von Annahmen zweckmäßig ist, die Eignung von Verfahren für einzelne Datensätze vorherzusagen. Sollte dies gelingen, wäre ein erster Schritt zu einem analytischen

Verständnis der Übertragbarkeit von Verfahren gemacht. Insbesondere wird darauf hingewiesen, wo Konflikte zwischen den Voraussetzungen der Datensätze und den eingehenden Annahmen der Verfahren bestehen. Im Anschluss werden dann die Segmentierungsverfahren aus Kapitel 5 an den ausgewählten Datensätzen getestet. Insbesondere soll dabei untersucht werden, wie sich die Steuerungsparameter auf die Ergebnisse auswirken, und ob durch sie eine Anpassung an den Datensatz möglich ist. Wäre dies der Fall, könnte durch die eingeführten Änderungen der Funktionsumfang der Verfahren erweitert werden.

Die Ergebnisse der einzelnen Tests werden dabei zunächst eher qualitativ (also visuell) bewertet, da es in der vorliegenden Arbeit nicht in erster Linie darum ging, für einzelne Datensätze besonders hohe Genauigkeiten zu erzielen. Um dennoch auch quantitative Ergebnisse zu liefern, werden die Ergebnisse der einzelnen Verfahren aber in 6.7 mit Referenzdatensätzen verglichen, und Kennziffern für Vollständigkeit, Korrektheit und Genauigkeit der extrahierten Bäume ermittelt. Da mit der Segmentierung aber zunächst nur der erste elementare Schritt untersucht wurde, sind hierfür weitere Module aus dem Ablauf in Abb. 4.2 notwendig. Darum wurde die Hypothesenauswahl anhand eines einfachen Modells implementiert, das in 6.6 vorgestellt wird.

## 6.2 Wasserscheidentransformation im Maßstabsraum (WST)

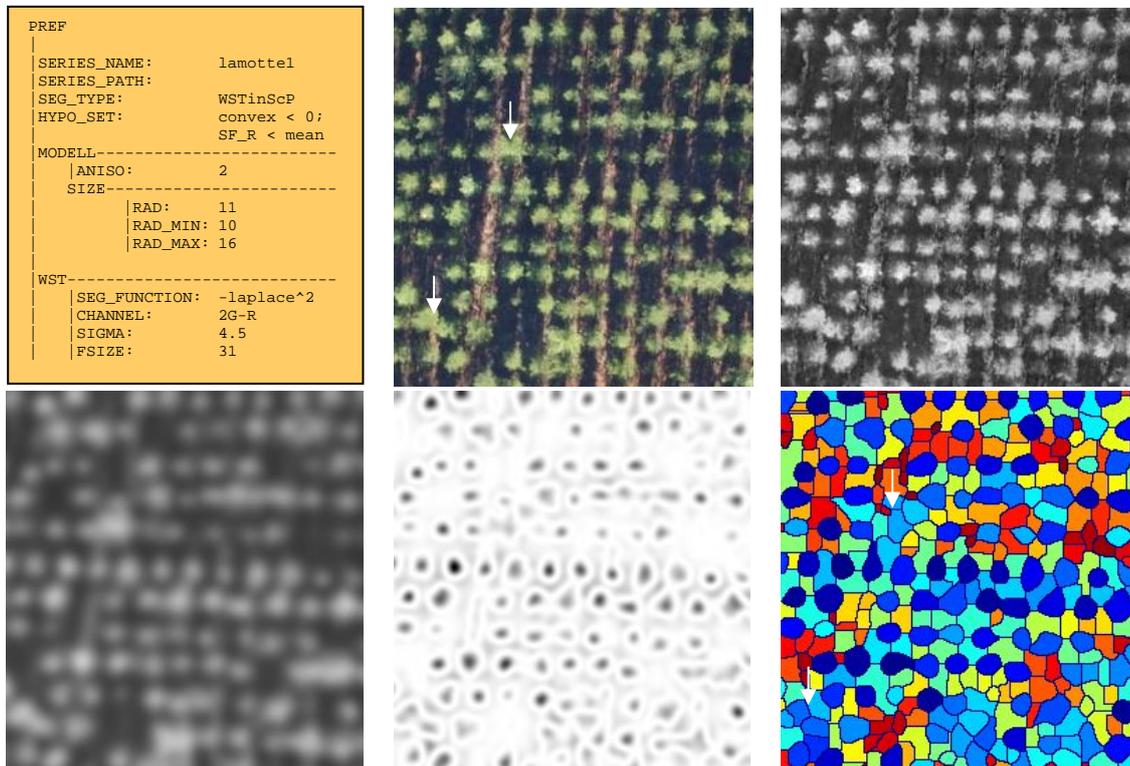
Hier werden die Ergebnisse der Wasserscheidentransformation im Maßstabsraum vorgestellt. Tab. 4 zeigt zunächst eine Übersicht der herausgearbeiteten Voraussetzungen des Verfahrens und der Eignung der ausgewählten Datensätze.

**Tab. 4: Eignung der verwendeten Datensätze für WST**

Voraussetzung	AN	WST	<i>lamotte1</i>	<i>lamotte2</i>	<i>goett</i>	<i>namb1</i>
		<i>nötig?</i>	<i>erfüllt?</i>			
Kreisförmigkeit	(3)	erzwungen	sehr gut	befriedigend	ausreichend	nein
3D-Hülle konvex	(4)	ja (bei DOM)				
Symmetrie	(7)		ja	nein		nein
Kompaktheit	(11)		gut	gut	überwiegend	nein
Größe	(13)	für 1 Ebene	ja	ja	ausreichend	
Abstand	(14)		ja	ja	ausreichend	nein
Blobs	(15)	Ja	ja	ja	ja	nein
dunkle Täler	(21)		überwiegend	überwiegend	teilweise	nein
Homogenität	(22)	nein	gut	gut	mangelhaft	ja
keine dunklen Pixel	(23)		sehr gut	sehr gut	nein	nein
Intensitätsmaximum	(25)		gut	nein	teilweise	teilweise
radiom. konvex	(27)	ja (optisch)	sehr gut	gut		nein

Danach ist zu vermuten, dass das Verfahren bei den *lamotte*-Bildern anwendbar sein müsste, da sie die zentrale Anforderung radiometrischer Konvexität gut erfüllen. Für die *goett*-Bilder ist diese Bedingung noch offen, Ausschlusskriterien für die Verwendbarkeit bestehen anhand der Übersicht aber nicht.

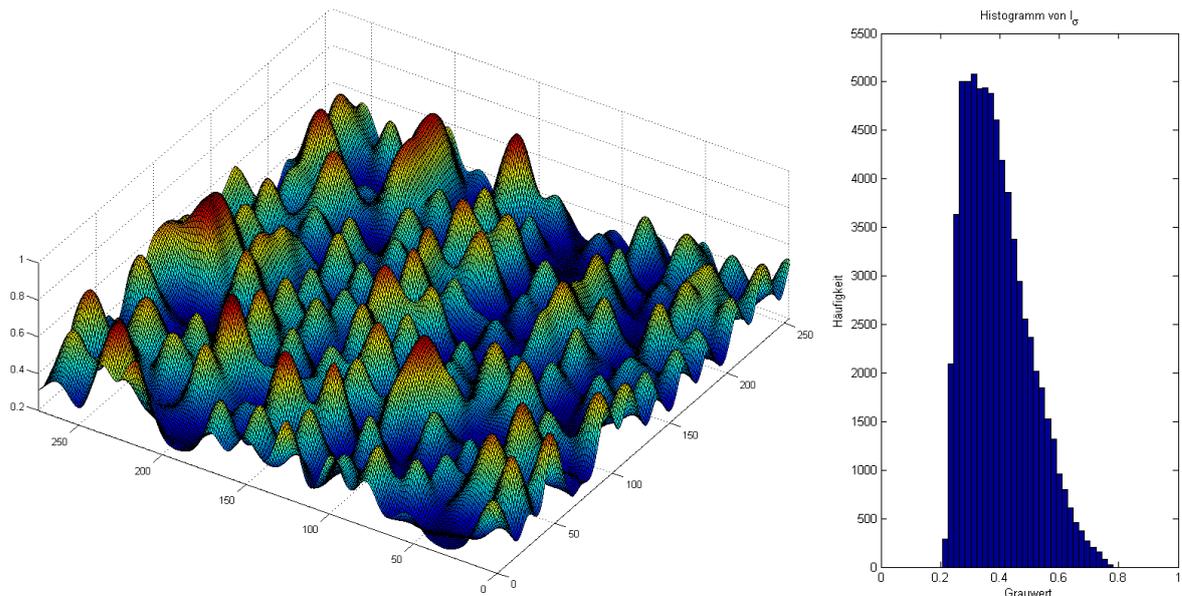
Abb. 6.8 zeigt den Ablauf des Verfahrens in einer Sequenz von Bildern. Oben sind die eingestellten Parameter dokumentiert. Daneben sind das Originalbild und der gewählte Farbkanal 2G-R (also eine Linearkombination aus den Kanälen Grün und Rot, siehe unten) zu sehen. Unten wird links das geglättete Bild (also die Maßstabsebene im LMR) gezeigt, daneben die Segmentierungsfunktion  $N((\Delta I)^2)$  und das Segmentierungsergebnis. Es ist zu beachten, dass der Bildausschnitt durch die Faltung etwas kleiner wird.



**Abb. 6.8:** Ergebnisse von WST auf *lamotte1*. Oben: Steuerungsparameter, Bild, gewählter Kanal 2G-R; Unten: Maßstabsebene, Segmentierungsfunktion, Segmentierungsergebnis

Die Bereiche mit den niedrigen Werten (dunkel, unten Mitte) auf der Segmentierungsfunktion werden zu den Segmenten mit den niedrigsten Nummern (dunkelblau, unten rechts). Von diesen sind die meisten eindeutig Bäumen im Bild zuzuordnen, der Maßstab scheint also gut gewählt. Besonders große Bäume (weiße Pfeile) werden aber nicht gut getroffen. Die Tatsache, dass die besten Hypothesen die niedrigsten Werte auf der Segmentierungsfunktion haben (also besonders gekrümmt sind), kann auch als Merkmal zur Hypothesenauswahl genutzt werden. Weiterhin fällt auf, dass die Baumsegmente runder sind als die Zwischensegmente auf dem Boden, was daran liegt, dass die Anisotropie des Filters bei den Maxima schwächer wirkt als bei den Minima.

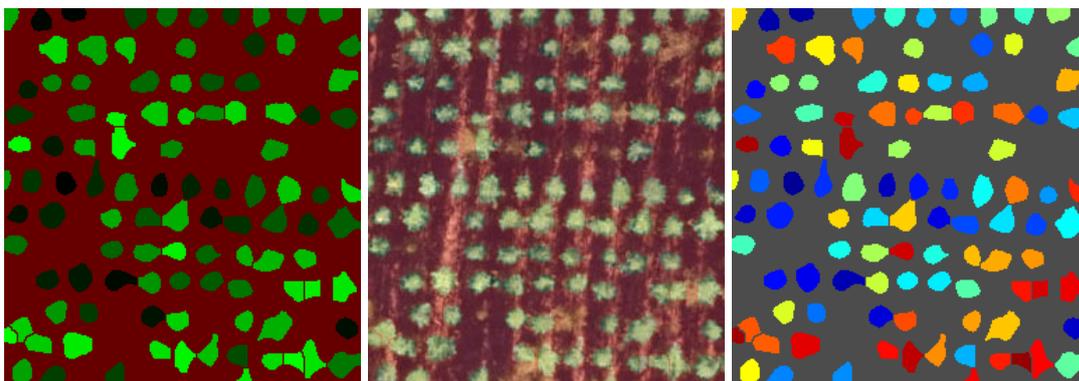
Die gelungene Segmentierung kann als Beleg dafür betrachtet werden, dass dieses Verfahren für Oberflächenmodelle auf radiometrische Bildern übertragbar ist. Dies soll mit Abb. 6.9 (links) unterstrichen werden, wo die obige Maßstabsebene als 3D-Funktion aufgetragen ist. Die Beschaffenheit der Intensitätsfunktion erinnert unmittelbar an ein Höhenmodell und hat insbesondere dieselben differentialgeometrischen Eigenschaften. Daher lässt sich das Bild „morphologisch“ segmentieren.



**Abb. 6.9:** Links: Maßstabsebene des Kanals in Abb. 6.8 (rechts unten) als 3D-Grauwertfunktion; Rechts: Histogramm des Kanals

Abb. 6.9 (rechts) zeigt das Histogramm der Grauwerte desselben Kanals. Es ist auffällig, dass weder ganz dunkle noch ganz helle Pixel vorhanden sind, der Kontrast ist also schwach. Auch wenn dies teilweise am Filter liegt, kann es als weiterer Beleg für die schlechte radiometrische Qualität gesehen werden (siehe 6.1.1).

Da das Segmentierungsergebnis relativ unübersichtlich ist, soll hier bereits ein einfaches Modell zur Auswahl von Baumhypothesen bei WST entwickelt werden. Dabei werden zwei Merkmale für die Klassifikation als Baum herangezogen: Erstens muss das Segment überdurchschnittlich gekrümmt ein (d.h. die Segmentierungsfunktion im Segment muss unter dem Mittelwert liegen) und zweitens muss es radiometrisch konvex sein (also  $CONVEX > 0$ ). Aus Sicht der ME entspricht das einem Box-Klassifikator, wobei eine Entscheidungsgrenze ( $CONVEX = 0$ ) direkt festgelegt und die andere aus dem Histogramm gewonnen wird. Das Ergebnis der Hypothesenauswahl ist in Abb. 6.10 zu sehen.

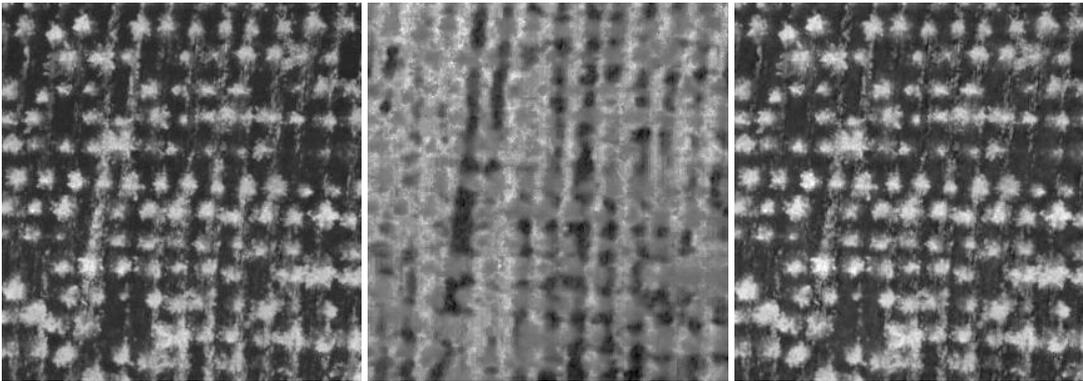


**Abb. 6.10:** Ergebnis der Hypothesenauswahl mit einfachem Modell, unterschiedliche Darstellungen

Zur Darstellung wurden verschiedene *colormaps* entwickelt. Die linke Darstellung mit grünen Segmenten vor rotbraunem Hintergrund wird im Folgenden standardmäßig benutzt. Die rechte Darstellung dient der besseren Unterscheidbarkeit zwischen Regionen, die mittlere zeigt das Extraktionsergebnis als Maske vor dem Originalbild: Extrahierte Regionen sind transparent, die übrigen

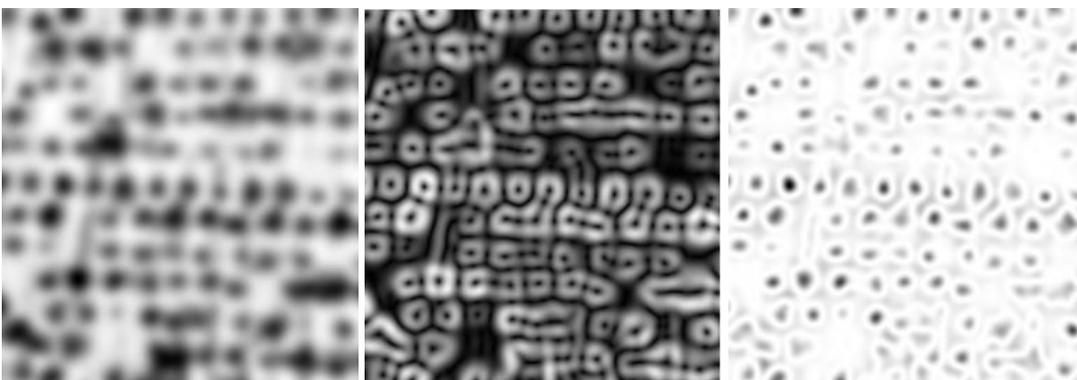
Bereiche mit einem halbtransparenten roten Layer überdeckt. Das einfache Modell für WST liefert bereits ein gutes Extraktionsergebnis, was die Überlegungen in 4.4.4 unterstreicht. Da die Hypothesenauswahl später ohnehin im LMR stattfinden soll, ist es überflüssig, an dieser Stelle ein noch besseres Modell für Bäume in einer Maßstabsebene zu entwickeln.

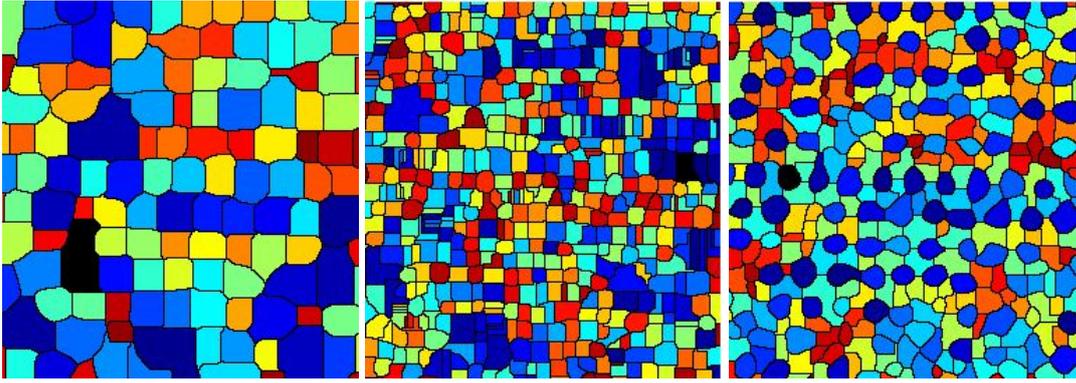
Im Folgenden wird untersucht, wie sich die Steuerungsparameter auf die Ergebnisse auswirken. Abb. 6.11 zeigt zunächst verschiedene Kanäle (Parameter `PREF.WST.COLOUR_CHANNEL`), die zur Segmentierung genutzt werden können.



**Abb. 6.11: Verschiedene Kanäle (`PREF.WST.COLOUR_CHANNEL`) von *lamotte1*: G, DOA, und 2G-R**

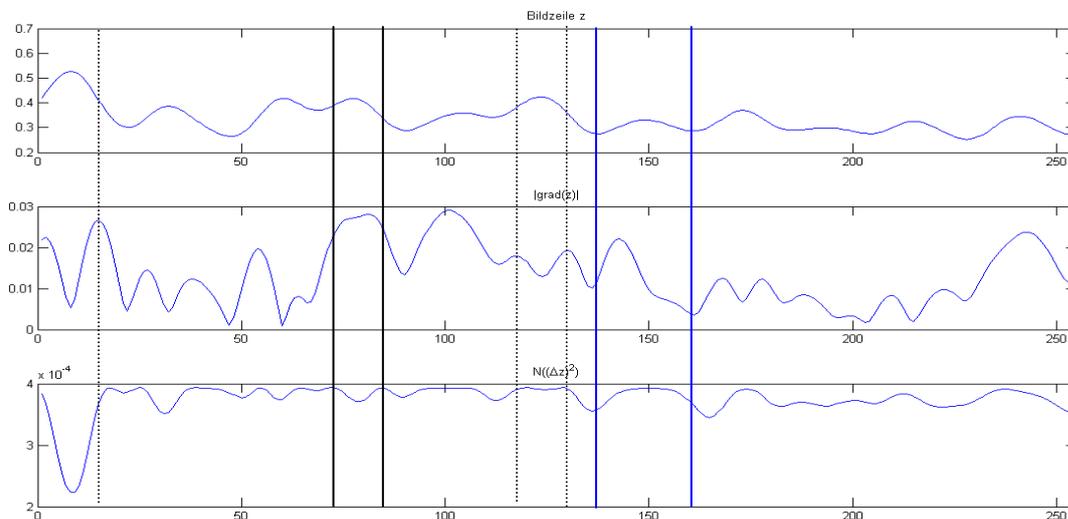
Im linken Bild wurde der grüne Kanal gewählt: Die Bäume haben klaren Kontrast gegenüber den schattigen Bereichen, der Boden ist aber ähnlich hell wie die Kronen. Demgegenüber werden beim DOA (Mitte) die Bodenbereiche zwar wirksam unterdrückt, dafür werden die Schatten aber heller als die Kronen, was der schlechten radiometrischen Qualität (unterschiedliche Empfindlichkeit in den Spektralbereichen, siehe 6.1.1) geschuldet sein mag. Grundsätzlich scheint das Merkmal zumindest geeignet, zwischen Vegetation und Boden zu unterscheiden. Rechts ist die Linearkombination 2G-R des grünen und roten Kanals dargestellt. Im Vergleich zum linken Bild ist der Kontrast gegenüber dem Schatten etwa gleich, der gegenüber dem hellen Boden aber größer. Da der Einfluss des Bodens durch die Subtraktion des roten Kanals abgeschwächt wird, wurde dieser Kanal für die folgenden Tests ausgewählt.





**Abb. 6.12: Test von Segmentierungsfunktionen (PREF.WST.SEG\_FUNCTION) auf *lamotte1***

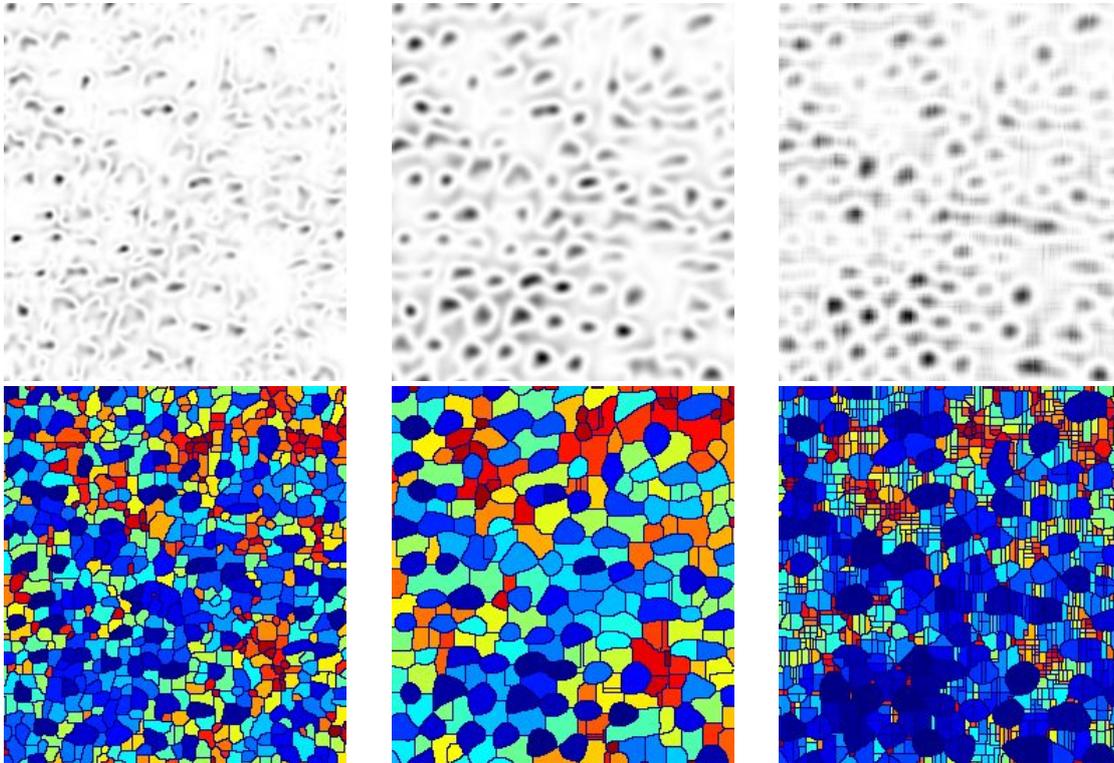
Abb. 6.12 zeigt verschiedene Segmentierungsfunktionen (Parameter WST.SEG\_FUNCTION, siehe 5.1) im Vergleich. Oben sind die Segmentierungsfunktionen  $-I$  (inverted),  $|\vec{\nabla}I|$  (mog) und  $N((\Delta I)^2)$  ( $-\text{laplace}^2$ ) zu sehen, darunter die entsprechenden Segmentierungsergebnisse. Beim invertierten Bild (links) wurden Teile der Bereiche zwischen den Kronen mit segmentiert. Die maximale Steigung (Betrag des Gradienten, Mitte) wird dagegen bereits zu nah am Zentrum der Krone erreicht, so dass die Kronen-Segmente zu klein werden. Beim invertierten Quadrat der Summe der zweiten Ableitungen (rechts) ist die Abgrenzung der Baumkronen am besten getroffen. Damit bestätigt sich das Ergebnis von STRAUB (2003A: 66), dass  $N((\Delta I)^2)$  gut als Segmentierungsfunktion geeignet ist.



**Abb. 6.13: Schnitt durch eine Bildzeile aus *lamotte1* (oben) und entsprechende Segmentierungsfunktionen Betrag des Gradienten (Mitte) und invertiertes Quadrat der Summe der zweiten Ableitungen (unten)**

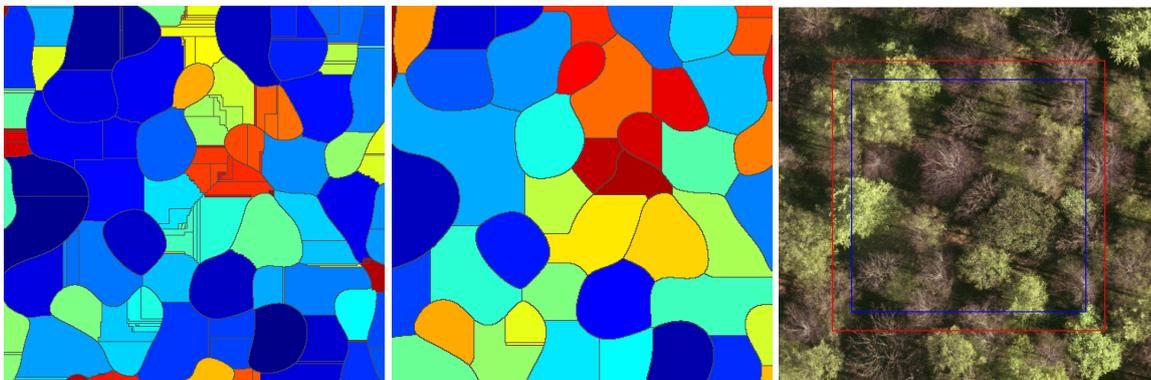
Zur Verdeutlichung zeigt Abb. 6.13 die Intensitätsfunktion einer einzelnen Bildzeile (oben) sowie die Segmentierungsfunktionen  $|\vec{\nabla}I|$  und  $N((\Delta I)^2)$  derselben Zeile darunter. Es sei darauf hingewiesen, dass es sich dabei nicht um die (Richtungs-)Ableitung der Bildzeile handelt. Die Linien verdeutlichen,

welche Bereiche einzelner Intensitätsmaxima bei den Segmentierungsfunktionen  $-I$  (blau),  $|\vec{\nabla}I|$  (schwarz) und  $N((\Delta I)^2)$  (gestrichelt) in dieser Zeile abgegrenzt würden.



**Abb. 6.14:** Auswirkung des Parameters `WST.SIGMA` bei `lamotte2`:  $\sigma = 3, 5, 7$ , Segmentierungsfunktionen (oben) und Segmentierungsergebnis (unten)

In Abb. 6.14 sind die Auswirkungen des Parameters  $\sigma$  dargestellt. An der unterschiedlichen Größe der resultierenden Segmente wird deutlich, dass in jeder Maßstabsebene Objekte einer bestimmten Größe optimal abgegrenzt werden. Diese Größe wächst mit dem Parameter  $\sigma$ . Zwischen dem mittleren und rechten Bild nimmt die Größe der Segmente jedoch sprunghaft wieder ab. Dies ist auf die Anisotropie des Filters zurückzuführen, die bei gleich bleibender Filtergröße mit  $\sigma$  wächst. Scheinbar gibt es also ein kritisches Maß, ab dem Artefakte in der Segmentierungsfunktion (rechts oben) zum Zerfall homogener Krümmungsbereiche führen. Die dabei zusätzlich entstehenden Kanten verlaufen parallel zur Begrenzung der Filtermaske (also auch parallel zu den Bildachsen).



**Abb. 6.15:** Verschiedene Filtergrößen bei `goett{3}`,  $\sigma = 26$ , `WST.FSIZE`: 201(links) und 271 (Mitte), rechts: Gültiger Bereich für `FSIZE`: 201(rot) und 271(blau)

Weiterhin ist wie schon in Abb. 6.8 zu erkennen, dass die Baum-Segmente von den Artefakten weniger stark betroffen sind als die Zwischen-Segmente. Bei Variation von  $\sigma$  bleiben die Maxima also länger stabil als die Minima, was möglicherweise als Merkmal nutzbar gemacht werden könnte. Die Einführung der Filtergröße  $FSIZE$  als Steuerungsparameter hat sich somit bewährt. Der Effekt ist noch einmal in Abb. 6.15 dargestellt, in der *goett{3}* mit Filtern gleicher Varianz aber unterschiedlicher Größe gefaltet wurde: Die hellen Regionen sind relativ ähnlich während die dunkleren beim kleineren Filter (links) zerfallen. Ein Nachteil größerer Filter ist, dass der ungültige Bereich am Bildrand wächst (rechts). Außerdem nimmt der Rechenaufwand quadratisch mit der Filtergröße zu, was aber durch eine Transformation in den Frequenzraum umgangen werden kann. In Abb. 6.16 sind die wichtigsten Auswirkungen der Steuerungsparameter zusammengefasst.

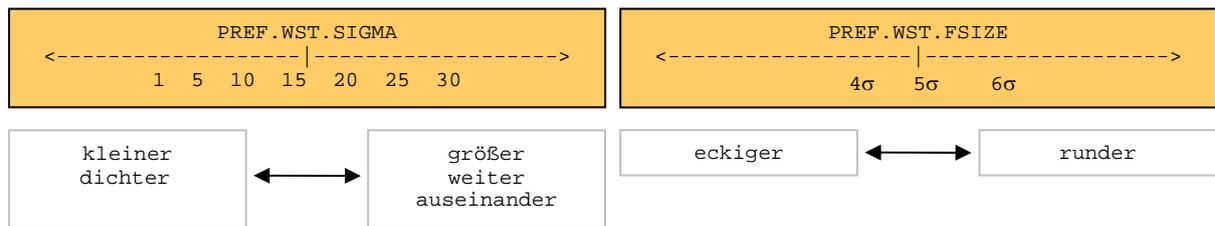


Abb. 6.16: Auswirkungen der Steuerungsparameter bei WST

### 6.3 Saatpunkte

Im Folgenden werden die Ergebnisse der Saatpunkt-Verfahren vorgestellt. Da diese nur eine Vorstufe des RG fungieren, werden sie etwas knapper behandelt. In Tab. 5 wird wieder eine Übersicht über die in Kapitel 5 herausgearbeiteten Voraussetzungen gegeben.

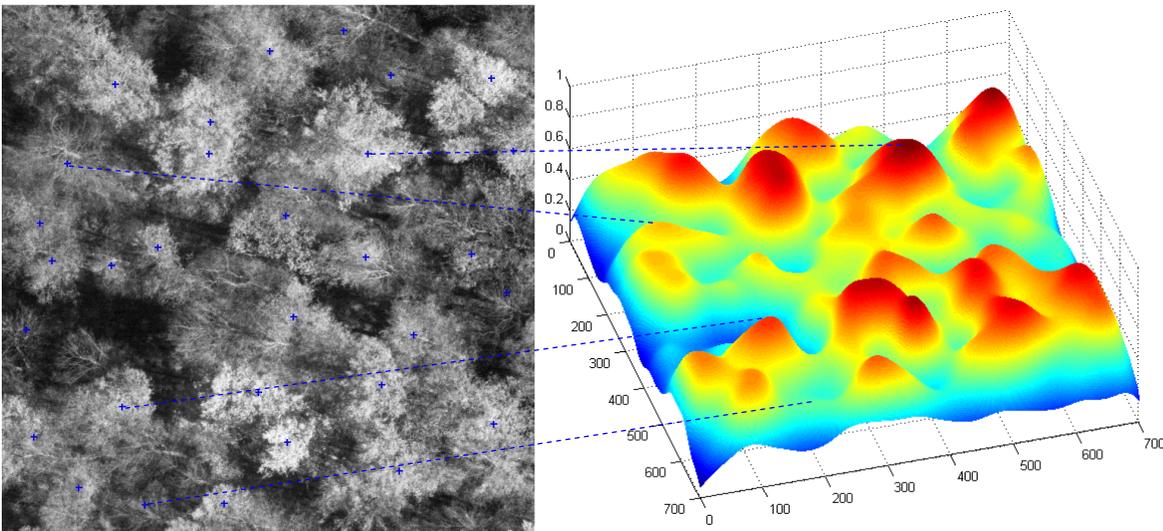
Tab. 5: Eignung der verwendeten Datensätze für die Saatgut-Verfahren REGMAX und DISTRANS

Voraussetzung	AN	REGMAX	DISTRANS	<i>lamotte1</i>	<i>lamotte2</i>	<i>goett</i>	<i>namb1</i>
		nötig?		erfüllt?			
Kreisförmigkeit	(3)	ja	nein	sehr gut	befriedigend	ausreichend	nein
3D-Hüllekonvex	(4)						
Symmetrie	(7)		ja	ja	nein		nein
Kompaktheit	(11)			gut	gut	überwiegend	nein
Größe	(13)	nein	Ergebnis	ja	ja	ausreichend	
Abstand	(14)	ja	ja	ja	ja	ausreichend	nein
Blobs	(15)	ja, stark	ja	ja	ja	ja	nein
dunkle Täler	(21)		ja	überwiegend	überwiegend	teilweise	nein
Homogenität	(22)			gut	gut	mangelhaft	ja
keine dunklen Pixel	(23)	nein	ja, stark	sehr gut	sehr gut	nein	nein
Intensitätsmaximum	(25)	ja		gut	nein	teilweise	teilweise
radiom. konvex	(27)			sehr gut	gut		nein

Die Voraussetzungen für REGMAX (insbesondere AN 3: Kreisförmigkeit, AN 14: Abstand und AN 15: *blobs*) werden vom Datensatz *lamotte1* gut erfüllt, und auch bei *lamotte2* müssten für die

einzelnen Bäume Saatpunkte gefunden werden. Da die Bäume hier auf Grund der geometrischen Verzerrung aber nicht symmetrisch sind, liegen diese möglicherweise nicht bei den Baumspitzen. Bei den *goett*-Datensätzen kann angenommen werden, dass auf Grund des ungleichmäßigen Abstands zwischen den Kronen Probleme auftreten. Dies gilt ebenso für die Distanztransformation mit DISTRANS, wobei auch die fehlende Homogenität (AN 23) Schwierigkeiten bereiten könnte.

Abb. 6.17 (links) zeigt ein Ergebnis des REGMAX-Verfahrens auf *goett{4}*. Für den unbelaubten Zustand mancher Bäume fällt dieses gut aus, es treten aber die vermuteten Schwierigkeiten durch den inhomogenen Abstand zwischen Kronen auf: Während einzelne große Kronen mehrere Saatpunkte enthalten, enthalten andere kleine gar keinen. Rechts ist der geglättete Kanal als Grauwertfunktion aufgetragen, die regionalen Maxima sind orange und rot. Die gestrichelten Linien dienen der Orientierung bei der Zuordnung.



**Abb. 6.17: Saatpunkte durch REGMAX auf *goett{4}*, SEED.SIGMA = 18, Maßstabsebene als Grauwertfunktion**

Abb. 6.18 zeigt den Ablauf des DISTRANS-Verfahrens (siehe 5.2.2) auf *goett{2}* in Einzelschritten. Links ist der gewählte Kanal (PREF.WST.COLOUR\_CHANNEL) zu sehen, daneben das Schwellenwert-Bild (SEED.DISTRANS\_THRESH = 0.35), dann das geglättete Distanzbild (SEED.SIGMA = 14) und schließlich die extrahierten Saatpunkte als Binärbild.



Abb. 6.18: Ablauf des DISTRANS-Verfahrens auf *goett{2}* in Einzelschritten

Am Schwellenwertbild (Mitte-links) ist zu erkennen, welche Probleme dunkle Pixel in den Kronen verursachen: Innerhalb der Kronen werden einzelne Pixel dem Hintergrund zugeordnet. Dadurch werden bei der Distanztransformation die Saatpunkte nicht optimal platziert und die extrahierten Größeninformationen werden verfälscht, da der Abstand zu diesen Pixeln statt zum Schatten berechnet wird. Allerdings ist der Schwellenwert (`SEED.DISTRANS_THRESH`) mit 0.35 auch relativ groß gewählt (siehe unten).

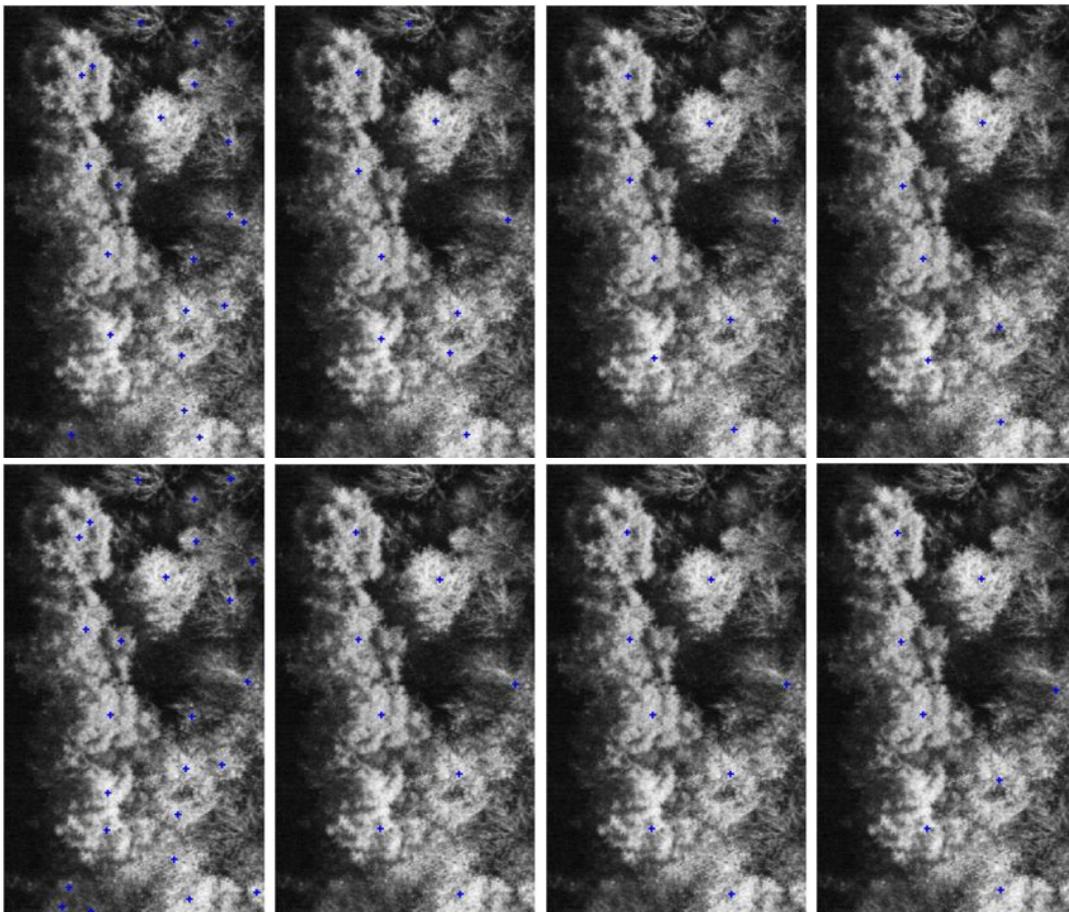


Abb. 6.19: Ergebnisse für  $\sigma$  7, 10, 13, 16 für REGMAX (oben), und DISTRANS (unten)

Im Folgenden werden die Auswirkungen der eingeführten Steuerungsparameter untersucht. In Abb. 6.19 sind die Ergebnisse von REGMAX (oben) und DISTRANS (unten) in Abhängigkeit von SEED.SIGMA dargestellt: Von links nach rechts beträgt der Parameter 7, 10, 13 und 16.

Erwartungsgemäß nimmt die Zahl der Saatpunkte (blaue Kreuze) mit der Varianz des Faltungskerns ab (je mehr Frequenzen eliminiert werden, desto weniger Maxima bleiben, siehe Abb. 2.21). Zwischen dem zweiten und dem vierten Bild ist die Anzahl der Saatpunkte aber relativ stabil (bei REGMAX nimmt sie von 10 zu 7 ab, bei DISTRANS liegt sie sogar konstant bei 8), was als Beleg für die Eignung des Parameters betrachtet werden kann (siehe 4.3). REGMAX und DISTRANS liefern relativ ähnliche Ergebnisse. DISTRANS ist hier also trotz der fehlenden radiometrischen Homogenität anwendbar.



**Abb. 6.20:** Auswirkungen des Parameters `DISTRANS_THRESH` auf `goett{2}`: Schwellenwert bei 0.15, 0.25, 0.35 und 0.45

Abb. 6.20 zeigt die Auswirkungen der Parameters `DISTRANS_TRESH`. Zu sehen ist das binäre Schwellenwert-Bild bei Schwellenwerten zwischen 0.15 und 0.45. Beim niedrigsten Schwellenwert werden kaum Kronen-Pixel dem Hintergrund zugeordnet, dafür jedoch viele Bereiche zwischen den Kronen dem Vordergrund. Im Gegensatz zur einfachen Segmentierung durch Schwellenwertbildung (vgl. 2.4.1) ist es hier aber auch nicht erforderlich, dass alle Pixel dem richtigen Bereich zugeordnet werden, sondern nur, dass jeweils ein Pixel am Kronenrand zum Hintergrund gehört. Je größer der Schwellenwert, desto besser sind die einzelnen Kronen getrennt, desto mehr Punkte innerhalb der Kronen werden aber auch dem Hintergrund zugeordnet. In 5.3 wurde zwar für die verlorenen Bereiche am Rand ein Korrekturfaktor eingeführt, die schwarzen Pixel innerhalb der Krone machen die gelieferten Abstandsinformationen jedoch unbrauchbar. Die errechneten Distanzen zum Schatten sind entsprechend instabil gegen Variation des Parameters und betragen (unkorrigiert) 20.2, 11.6, 5.6, 4.0 Pixel bei den Schwellenwerten 0.15, 0.25, 0.35 bzw. 0.45. Möglicherweise könnte dem begegnet werden, indem die Schwellenwertbildung bereits in einer größeren Maßstabsebene ausgeführt wird, was jedoch Gegenstand weiterer Untersuchungen werden müsste.

### 6.4 Basic Region Growing (BRG)

Im Folgenden werden exemplarisch repräsentative Ergebnisse des BRG-Algorithmus vorgestellt. In Tab. 6 findet sich wieder eine Übersicht der wichtigsten Voraussetzungen des Verfahrens und der Eignung der verwendeten Datensätze.

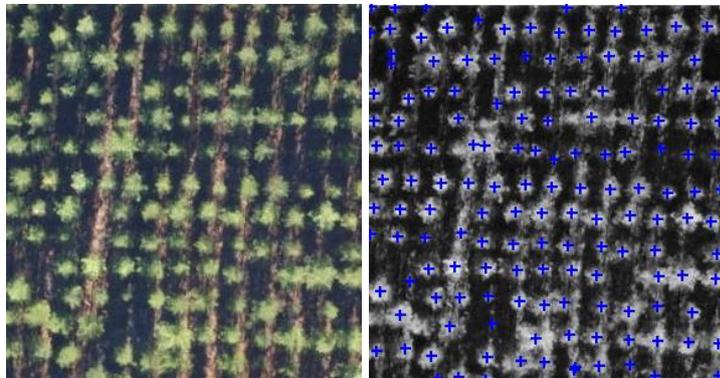
**Tab. 6: Eignung der verwendeten Datensätze für BRG**

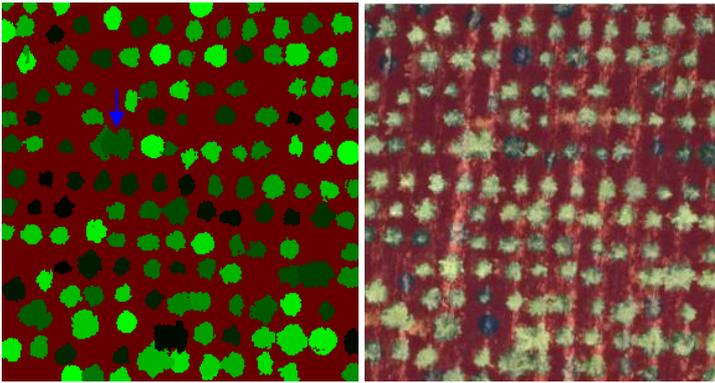
Voraussetzung	AN	BRG	<i>lamotte1</i>	<i>lamotte2</i>	<i>goett</i>	<i>nabl</i>
		nötig?	erfüllt?			
Kreisförmigkeit	(3)	ja	sehr gut	befriedigend	ausreichend	nein
3D-Hüllekonvex	(4)					
Symmetrie	(7)		ja	nein		nein
Kompaktheit	(11)	ja	gut	gut	überwiegend	nein
Größe	(13)	ja	ja	ja	ausreichend	
Abstand	(14)	nein	ja	ja	ausreichend	nein
Blobs	(15)	nein	ja	ja	ja	nein
dunkle Täler	(21)		überwiegend	überwiegend	teilweise	nein
Homogenität	(22)	ja	gut	gut	mangelhaft	ja
keine dunklen Pixel	(23)	nein	sehr gut	sehr gut	nein	nein
Intensitätsmaximum	(25)		gut	nein	teilweise	teilweise
radiom. Konvex	(27)		sehr gut	gut		nein

Auch die Voraussetzungen für BRG werden am besten vom *lamotte1*-Bild erfüllt. Auch *lamotte2* sollte prinzipiell geeignet sein, offen ist allerdings, wie die Abweichung von der Kreisform die Genauigkeit beeinträchtigt. Für die *goett*-Bilder scheint das Verfahren dagegen eher ungeeignet, da zentrale Voraussetzungen nicht erfüllt werden. So ist die absolute Voraussetzung Kompaktheit zwar zumindest beinahe überall gegeben, insbesondere die große radiometrische Varianz dürfte das Ergebnis aber negativ beeinflussen. Bei der Abgrenzung zwischen den Kronen spielt der Abstand (also die Kreisform) eine wichtige Rolle.

```

PREF
|SERIES_NAME:      lamotte1
|SERIES_PATH:      L:
|SEG_TYPE:         RG_BASIC
|HYPO_SET:         all
|PRESET:           1
-----
MODELL
|ANISO:            2
|SIZE
|RAD:              11
|RAD_MIN:          10
|RAD_MAX:          16
-----
SEED
|SAATGUT:          DISTRANS
|THRESH:           0.35
|CHANNEL:          2
|SIGMA:            4
|FSIZE:           31
-----
RG
|ALPHA:            0.15
|SIGMAL:           0.36
|AUTO_SIZE:        1
    
```





**Abb. 6.21: BRG auf *lamotte1*: Parameter, Bild, Saatpunkte, erzeugte Regionen (Segmente und Maske)**

Abb. 6.21 zeigt den Ablauf des BRG-Verfahrens auf *lamotte1*. Auf der letzten Seite unten sind links die verwendeten Steuerungsparameter zu sehen, in der Mitte das Originalbild und rechts das Ergebnis des Saatpunkt-Verfahrens DISTRANS. Auf dieser Seite ist das Ergebnis des BRG als Regionen und Maske dargestellt. Das Verfahren wurde mit Option `RG.AUTO_SIZE` durchgeführt, und die extrahierten Regionen haben klar erkennbar unterschiedliche Größen. Damit wurde ein Teil des Maßstabsproblems erfolgreich in diesem Schritt adressiert. Objekte, die nicht klar voneinander getrennt sind, können bei der Segmentierung ineinander wachsen (blauer Pfeil).

Die Darstellung als Maske verdeutlicht, dass die erzeugte Segmentierung bereits ohne Hypothesenauswahl eine hohe Güte hat, da fast alle Segmente Bäume sind. Es ist also offensichtlich bereits Wissen über die gesuchten Objekte in den Segmentierungsschritt eingegangen, was im Folgenden reflektiert werden soll: Neben den Bäumen wurden einige helle Bodenbereiche, vereinzelt aber auch schattige Bodenbereiche segmentiert. Letzteres verdeutlicht, dass in das BRG-Verfahren nur die radiometrische Homogenität, nicht aber die Helligkeit des Segments eingeht (daher wird auch AN 15, helle *blobs*, nicht vorausgesetzt). Die Helligkeit geht ausschließlich in den Saatpunkt-Schritt ein (maximale regionale Helligkeit bei REGMAX bzw. maximaler Abstand von einem dunklen Bereich bei DISTRANS), d.h. die Saatpunkte haben zwangsläufig eine hohe Intensität nicht aber die Segmente. Somit steht das radiometrische Merkmal BRIGHT für die Hypothesenauswahl noch zur Verfügung. Insbesondere würde es sich eignen, die schattigen Segmente zu eliminieren.

Umgekehrt dürften geometrische Merkmale nur sehr bedingt zur Hypothesenauswahl nutzbar sein, da sie bereits in hohem Maße ins Verfahren eingehen. So werden als Voraussetzungen Kreisförmigkeit (AN 3), Kompaktheit (AN 11) und bestimmte Größe (AN 13) gefordert. Entsprechend weist das Segmentierungsergebnis im Wesentlichen runde, kompakte Regionen einer bestimmten Größe auf, die nur deshalb variiert, weil durch die zusätzlichen Informationen aus dem DISTRANS-Verfahren eine bestimmte Größe für jedes einzelne Objekt statt für das gesamte Bild festgelegt wurde.

Es bleibt also festzuhalten, dass erstens (implizit) Wissen ins Segmentierungs-Verfahren eingeht und dieses somit gemäß dem Fokus dieser Arbeit nur für Datensätze mit bestimmten Voraussetzungen benutzt werden kann. Zweitens stehen Merkmale, die bereits in die Segmentierung eingehen bei der Hypothesenauswahl nur noch eingeschränkt zur Verfügung. Im Fall von BRG bleiben daher vor allem radiometrische Merkmale, ein einfaches Modell hierfür wird in Abschnitt 6.6 entwickelt.

Im Folgenden wird untersucht, wie das Verfahren bei schlechter erfüllten Voraussetzungen reagiert. Abb. 6.22 zeigt das Ergebnis desselben Verfahrens auf *lamotte2*, bei dem Kreisförmigkeit nur in eingeschränktem Maße gegeben ist.

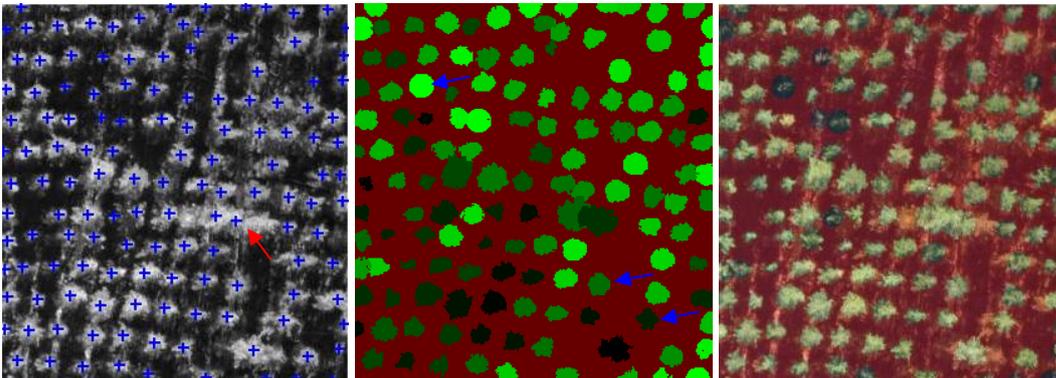


Abb. 6.22: BRG auf *lamotte2*: Saatpunkte und erzeugte Regionen (2 mal)

Zum Teil bestätigen sich die bisherigen Ergebnisse, da auch hier Regionen unterschiedlicher Größe erzeugt wurden. Es zeigt sich aber auch, dass das Verfahren nur eine gewisse Abweichung von der Kreisform toleriert und die extrahierten Segmente teilweise runder sind als die Objekte (blaue Pfeile). Außerdem liefert DISTRANS bei uneinheitlichen Abständen und verwachsenen Kronen keine belastbaren Ergebnisse (roter Pfeil).

Diese Grenzen treten noch deutlicher zu Tage, wenn mehrere grundlegende Voraussetzungen nicht mehr erfüllt sind. Abb. 6.23 zeigt den Ablauf von BRG auf *goett{6}*. Da das DISTRANS-Verfahren an den dunklen Pixeln in der Krone scheitert (siehe 6.3), wurden die Saatpunkte mit REGMAX bestimmt. Ohne Größeninformationen zu den einzelnen Saatpunkten ist das BRG-Verfahren offensichtlich nicht in der Lage, dicht stehende Bäume voneinander abzugrenzen (roter Pfeil). Dagegen werden Grenzen zwischen Kronen und radiometrisch unterschiedlichen Objekten wie Schatten aber sehr gut getroffen (blaue Pfeile).

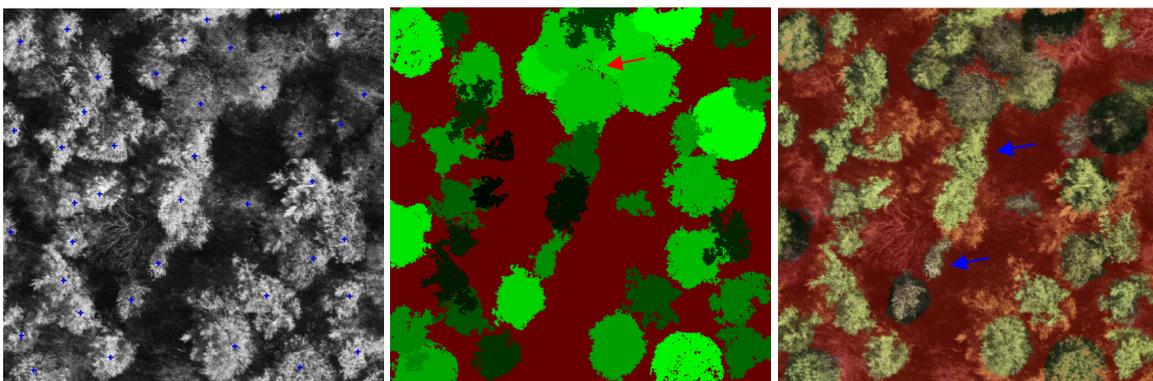


Abb. 6.23: BRG auf *goett{6}*: Saatpunkte mit REGMAX, und erzeugte Regionen (2 mal)

Das Verfahren findet weiterhin relativ kreisförmige, kompakte Objekte einer bestimmten Größe (AN 13), nur dass diese Größe nunmehr ohne `auto_size` Option für das gesamte Bild einheitlich ist und durch den Parameter `MOD.SIZE.RAD_MAX` bestimmt wird. Seine Wirkungsweise ist in Abb. 6.24 dargestellt, wo das obige Verfahren mit einem maximalen Radius von 40 (links), 60 (Mitte) bzw. 80 (rechts) Pixeln wiederholt wurde.

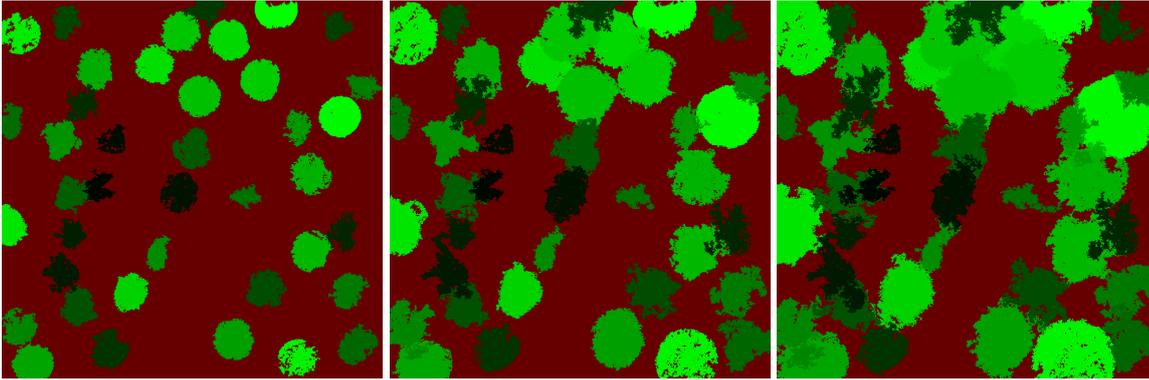


Abb. 6.24: BRG auf *goett{6}* für verschiedene Baumgrößen. `MOD.SIZE.RAD_MAX = 40, 60, 80`

In Bezug auf die obige Feststellung kann also ergänzt werden: Es gehen starke Mode Modell-Annahmen ins BRG-Verfahren ein, die von der gegebenen Situation auf *goett{6}* nicht erfüllt werden. Bemerkenswert daran ist nicht das schlechte Segmentierungsergebnis, sondern dass aufgezeigt werden konnte, warum das Verfahren nicht auf die Situation übertragbar ist. Weiterhin waren die Mängel schon an Tab. 6 abzulesen und könnten somit in Kenntnis der entsprechenden Voraussetzungen auch von einem automatischen Inferenzmechanismus ausgewertet werden.

Das Ineinanderlaufen der Regionen wird bei ERIKSON teilweise dadurch umgangen, dass weitere Saatpunkte innerhalb extrahierter Regionen gelöscht werden. Dadurch wird das grundsätzliche Problem, dass (im Widerspruch zu AN 22) radiometrisch ähnliche Objekte schlecht voneinander abgegrenzt werden, aber nicht gelöst. Teilweise könnte dem abgeholfen werden, indem zunächst die Ähnlichkeiten aller Pixel zu allen Saatpunkten berechnet und anschließend in einem eigenen Expansionschritt wie bei BROWNMO (mit `similarity` statt `numPass`) segmentiert würde. Dadurch würde das Merkmal Symmetrie stärker ins Verfahren eingehen und der Abstand vom Saatpunkt zum ausschlaggebenden Kriterium.

Weiteres Optimierungspotential könnte in der Nutzung von Texturinformationen liegen. So könnte das Ähnlichkeitsmaß in (5.1) einfach um eine Texturmaß erweitert werden. Da nur der relative Abstand zum Saatpunkt im entsprechenden Maß einginge, würden auch keine Vorab-Informationen über die Textur der gesuchten Objekte benötigt. Möglicherweise würden sich hierfür die Bandpassfilter nach KÖRBER *et al.* (2005) und MÖLLER *et al.* (2005) eignen.

Abb. 6.25 fasst die Auswirkungen der gewählten Parameter zusammen. Da das Verfahren weitgehend parameterfrei ist, fällt sie etwas kürzer aus.

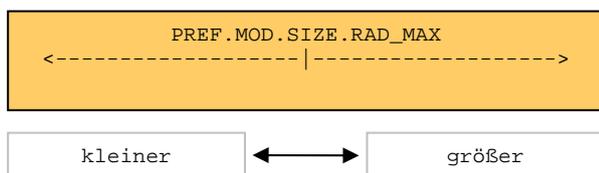


Abb. 6.25: Auswirkungen des Parameters `RAD_MAX` bei BRG

## 6.5 Brownsche Bewegung (BROWNM0)

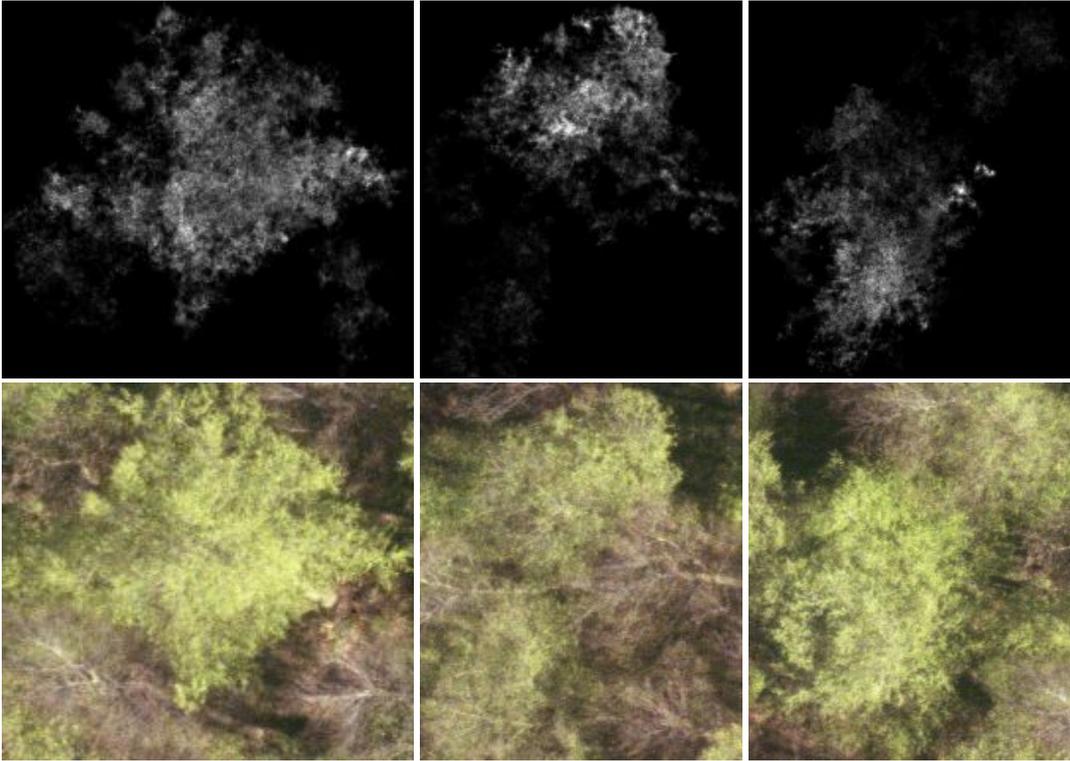
Im Folgenden werden exemplarisch repräsentative Ergebnisse des BROWNM0-Algorithmus vorgestellt. In Tab. 7 finden sich wieder wichtige Voraussetzungen des Verfahrens und Informationen über die Eignung der verwendeten Datensätze.

**Tab. 7: Eignung der verwendeten Datensätze für BROWNM0**

Voraussetzung	AN	BROWN_MO	<i>lamotte1</i>	<i>lamotte2</i>	<i>goett</i>	<i>namb1</i>
		<i>nötig?</i>	<i>erfüllt?</i>			
Kreisförmigkeit	(3)	schwach	sehr gut	befriedigend	ausreichend	nein
3D-Hülle konvex	(4)					
Symmetrie	(7)		ja	nein		nein
Kompaktheit	(11)	ja	gut	gut	überwiegend	nein
Größe	(13)	schwach	ja	ja	ausreichend	
Abstand	(14)	indirekt	ja	ja	ausreichend	nein
Blobs	(15)	ja	ja	ja	ja	nein
dunkle Täler	(21)	ja, stark	überwiegend	überwiegend	teilweise	nein
Homogenität	(22)	indirekt	gut	gut	mangelhaft	ja
keine dunklen Pixel	(23)	schwach	sehr gut	sehr gut	nein	nein
Intensitätsmaximum	(25)		gut	nein	teilweise	teilweise
radiom. konvex	(27)		sehr gut	gut		nein

Die entscheidende AN 11 (Kompaktheit) ist sowohl für die *lamotte*-Bilder als auch für die *goett*-Bilder im Wesentlichen zutreffend, von wenigen Kronen in *goett*, die in Teilkronen zerfallen kann abgesehen werden. Auch AN 15 (Kronen heller als Umgebung) wird von beiden Datensätzen erfüllt. Entscheidend dürfte dagegen AN 21 (Dunkle Täler zwischen Kronen) sein, da BROWNM0 ganz merklich auf diesem Charakteristikum basiert. Hier sind die Voraussetzungen recht unterschiedlich. In *goett{n}* sind teils mehrere Kronen ineinander gewachsen. Es ist zu erwarten, dass dadurch möglicherweise die Abgrenzung zwischen den Kronen nicht exakt ist (bei gleichmäßiger Helligkeit wird doch der Abstand zum Saatpunkt ausschlaggebend, wie in 5.4 gezeigt wurde), die Abgrenzung zwischen Kronen und Hintergrund müsste aber funktionieren.

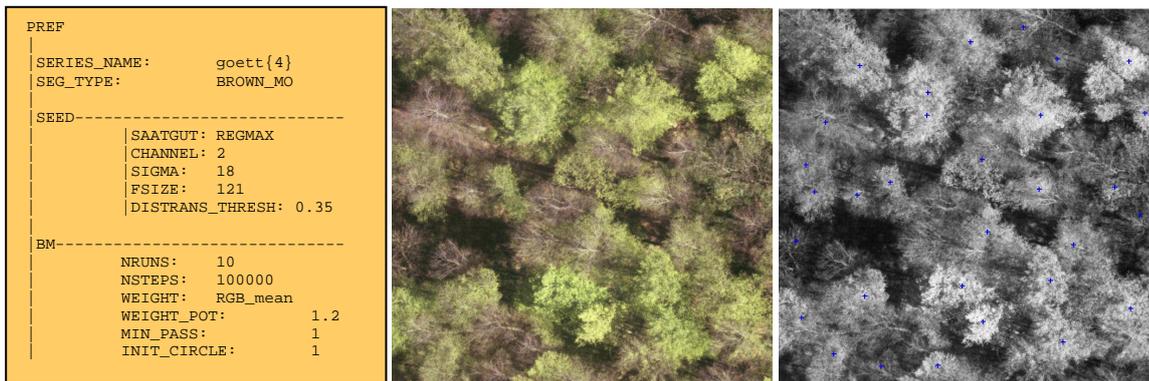
In den *lamotte*-Datensätzen grenzen dagegen teilweise helle Bodenflecken an die Kronen. Es ist zu erwarten, dass der Algorithmus die Kronen in diesem Fall nicht scharf abgrenzen kann. Es soll untersucht werden, ob dies durch Variation der Gewichtsfunktion  $w(\mathbf{I}(\mathbf{r}))$  verhindert werden kann. Außerdem sind die Schatten, wie in Abb. 6.9 und Abb. 6.13 deutlich zu erkennen ist, nicht völlig dunkel und das Bild also vergleichsweise kontrastarm. Es soll also ebenfalls untersucht werden, ob diesem Problem durch Wahl einer geeigneten Gewichtsfunktion begegnet werden kann.



**Abb. 6.26: NumPass-Ergebnisse von BROWNMO auf *goett{4}* von 3 Saatpunkten. Oben: numPass-Werte (BM.WEIGHT = 'RGB\_mean', WEIGHT\_POT = 1.5, NUM\_RUNS = 10, NUM\_STEPS = 100000); Unten: Korrespondierende Ausschnitte im Bild**

Abb. 6.26 zeigt die Ergebnisse einiger BROWNMO Zufallsläufe auf *goett{4}*. Die Helligkeit im Bild entspricht der Anzahl der Besuche in numPass (linear skaliert). Vor allem die helleren Teile der Kronen werden besucht, in Einzelfällen „verirrt“ das virtuelle Teilchen sich aber auch in angrenzende Bereiche. Der Intensitätsabfall nach außen und die Bestimmtheit werden durch die Parameter NUM\_STEPS und NUM\_RUNS gesteuert, wie in 5.4 hergeleitet wurde.

Abb. 6.27 zeigt den Ablauf von BROWNMO auf *goett{4}* als Sequenz.



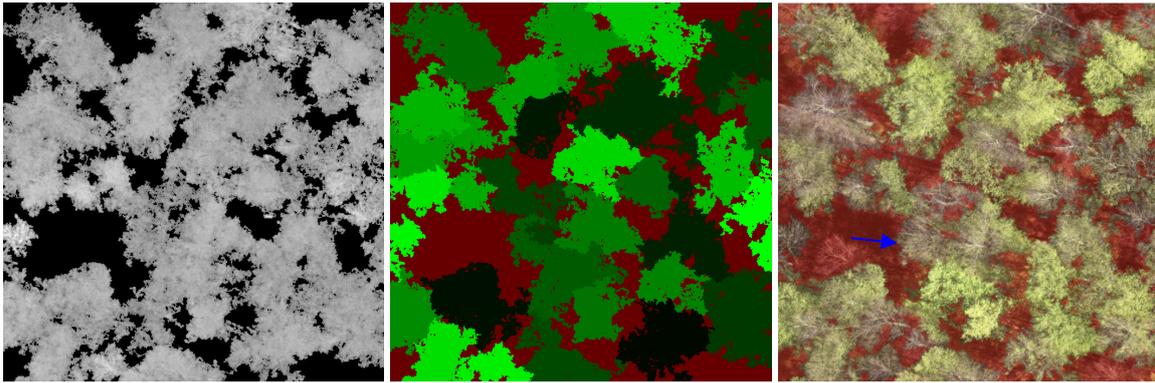


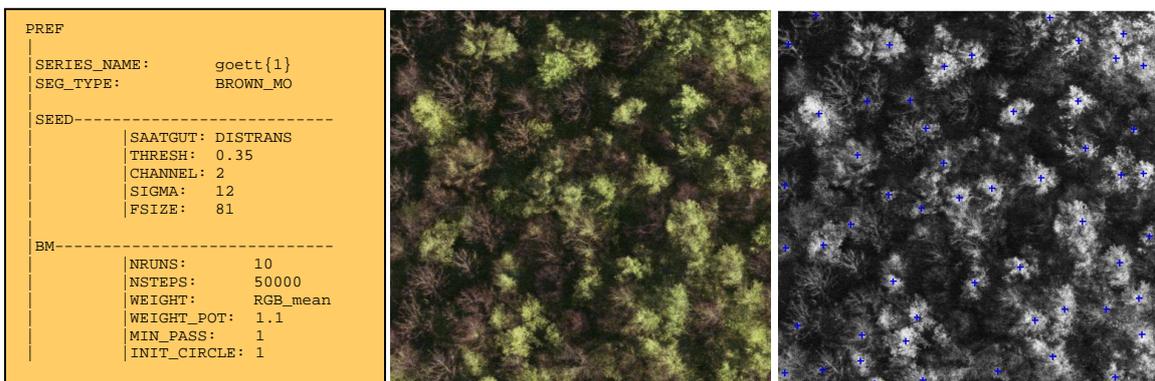
Abb. 6.27: Ergebnisse von BROWNMO auf *goett{4}*; Oben: Parameter, Bild, Saatpunkte; Unten: numPass, Segmentierungsergebnis (Regionen und Maske)

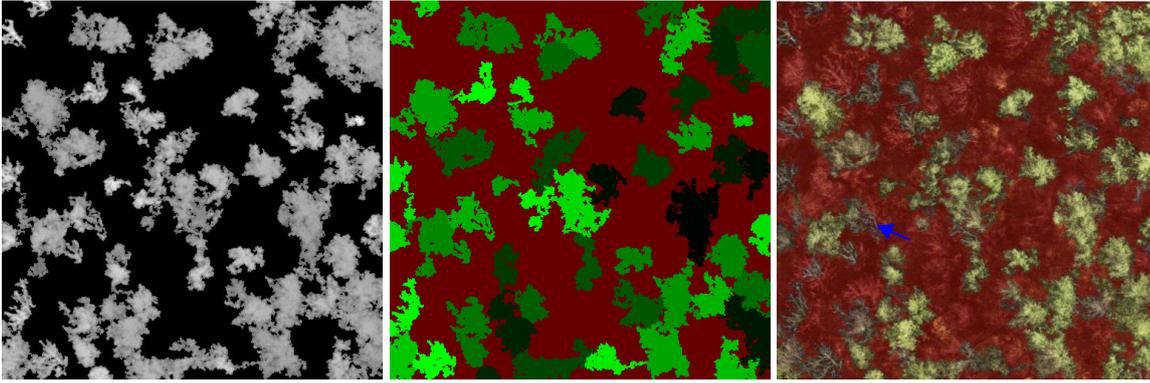
Oben sind die Liste der Steuerungsparameter, das Bild sowie die Saatpunkte zu sehen. Unten links ist die Besuchshäufigkeit der Pixel insgesamt dargestellt, also die Summe aller numPass-Ebenen. Im Gegensatz zu Abb. 6.26 wurde die Besuchshäufigkeit gamma-korrigiert, damit auch die äußeren Bereiche (mit geringeren Besuchshäufigkeiten) noch zu erkennen sind. Unten in der Mitte ist das Ergebnis des Expansionsschritts (also des RG mit numPass als Ähnlichkeitsmaß) dargestellt. Rechts ist das Segmentierungsergebnis vor dem Hintergrund des Eingangsbildes zu sehen.

Die Abgrenzung der Kronen von dunklen Bereichen ist sehr genau, die Abgrenzung zwischen den Kronen aber teilweise schlecht. Dies liegt aber vor allem daran, dass einige Kronen mehrere, andere dagegen gar keine Saatpunkte enthalten. Das REGMAX-Verfahren ist hier also nicht in der Lage, die anspruchsvolle Aufgabe zu lösen. Daher sollte darüber nachgedacht werden, ob Saatpunkte nicht auch auf verschiedenen Maßstabsebenen gewonnen werden müssen.

Erwartungsgemäß bereiten die unbelaubten Bäume dem Verfahren gewisse Schwierigkeiten, da sie ja die radiometrischen Annahmen nicht erfüllen. Umso beachtenswerter ist es, dass auch hier teilweise richtige Abgrenzungen geliefert werden (blauer Pfeil).

Abb. 6.28 zeigt den gleichen Ablauf für *goett{1}*.





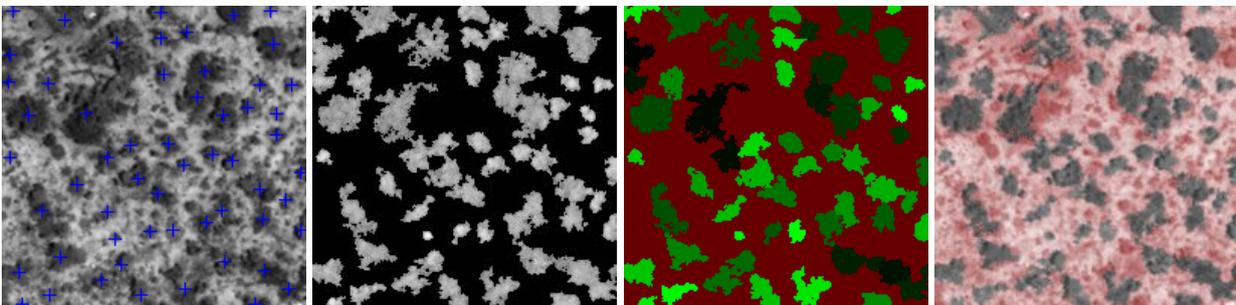
**Abb. 6.28:** Ergebnisse von BROWNMO auf *goett1*; Oben: Parameter, Bild, Saatpunkte; Unten: numPass (gamma-korrigiert), Segmentierungsergebnis (Regionen und Maske)

Auch hier sind die grünen Kronen gut vom Hintergrund abgegrenzt, die unbelaubten bereiten dagegen Schwierigkeiten. Da es wenige ineinander gewachsenen Kronen gibt (diese Kronen also meist durch klare Schatten getrennt sind) sind sie hier auch untereinander gut abgegrenzt. Die Regionen laufen doch vereinzelt in unbelaubte Kronen über, die keinen eigenen Saatpunkt bekommen haben (blauer Pfeil). Zur Verdeutlichung ist in Abb. 6.29 die rechte obere Ecke des Bildes und des Segmentierungsergebnisses vergrößert dargestellt.



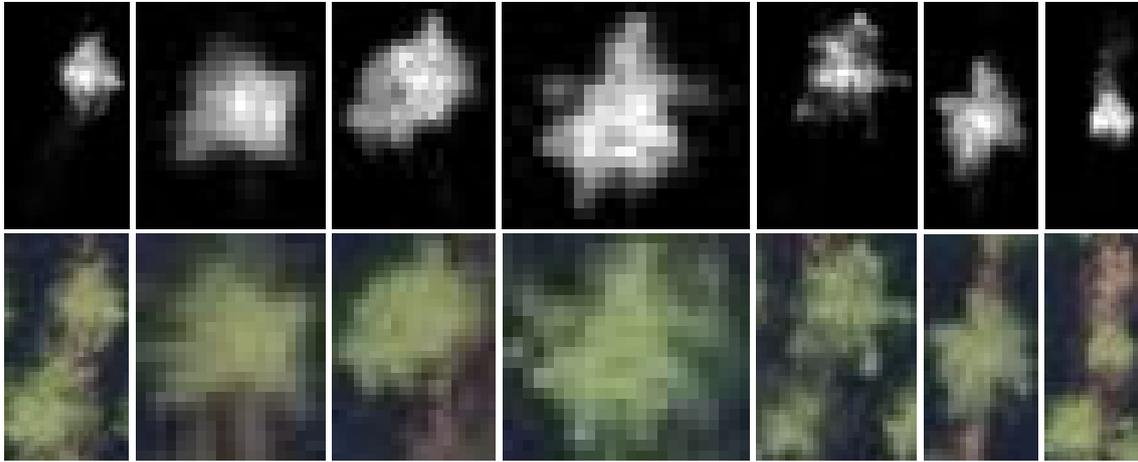
**Abb. 6.29:** Ergebnisse von BROWNMO auf *goett1*, vergrößerter Ausschnitt

Die Krone in der rechten oberen Ecke hat keinen Saatpunkt bekommen, möglicherweise da sie zu nah am Bildrand liegt. Das hellrote Segment ist in einen dunkleren, wahrscheinlich noch nicht voll ausgetriebenen, Baum „gelaufen“. Die übrigen Regionen treffen die Situation gut.



**Abb. 6.30:** Ergebnisse von BROWNMO auf *namb1*: Saatpunkte, numPass, Regionen, maskierter panchromatischer Kanal

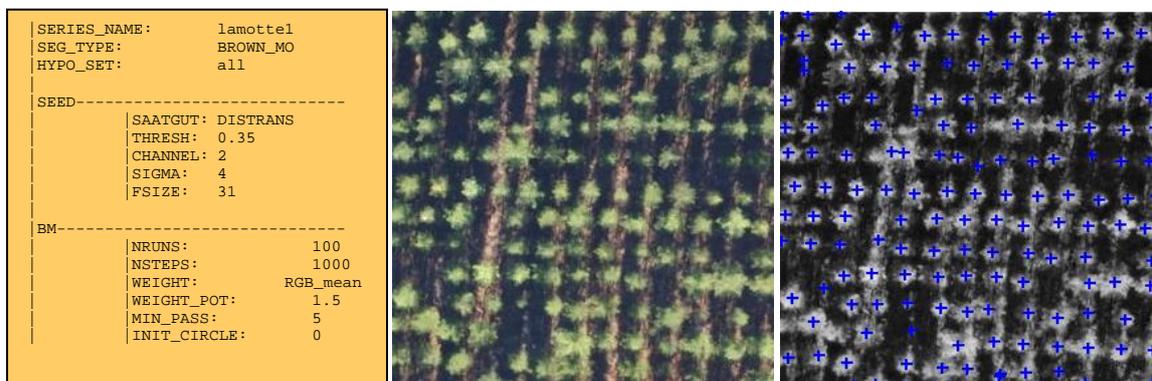
Abb. 6.30 zeigt Ergebnisse von BROWNMO auf *namb1*. Zur Bestimmung der Saatpunkte und der Gewichtsfunktionen wurde jeweils der invertierte panchromatische Kanal verwendet. Damit ist es zwar nicht möglich, einzelne Kronen zu segmentieren, die Baumgruppen und Vegetationsbereiche wurden aber relativ gut abgegrenzt, was für landschaftsökologische Anwendungen von Interesse sein könnte.



**Abb. 6.31: NumPass-Ergebnisse von BROWNMO auf *lamotte1*. Oben: numPass-Werte (BM.WEIGHT = 'RGB\_mean', WEIGHT\_POT = 1.5, NUM\_RUNS = 100, NUM\_STEPS = 1000); Unten: Korrespondierende Ausschnitte im Bild**

Abb. 6.31 zeigt Ergebnisse von Zufallsläufen auf *lamotte1*. Der gewählte Ausschnitt entspricht jeweils allen besuchten Spalten und Zeilen. Da die Helligkeit wie in Abb. 6.26 linear skaliert ist, sind die äußeren Bereiche aber eher dunkel und Kronen scheinen gut getroffen, was auch an der hohen Anzahl von Durchgängen (NUM\_RUNS) liegt. Besonders in den beiden rechten Bildern ist aber ebenfalls deutlich zu erkennen, dass das virtuelle Teilchen wie vermutet in angrenzende Bodenflecken gelaufen ist. Die gewählte Gewichtsfunktion hat eine vergleichsweise hohe Potenz von 1.5, wodurch der schlechte Kontrast zum Schatten ausgeglichen wird.

Abb. 6.32 zeigt die Ergebnisse auf dem gesamten Bild. Die Darstellung ist wie bisher, nur im letzten Bild wurde abweichend eine blaue Maske verwendet, um die segmentierten Bodenbereiche zu verdeutlichen. An Stellen, an denen die Bodenflecken eigene Saatpunkte bekommen haben (z.B. blauer Pfeil), ist die Abgrenzung zwischen Krone und Boden gut, in den anderen Fällen ist das RG aber direkt in die Bodenbereich „übergelaufen“.



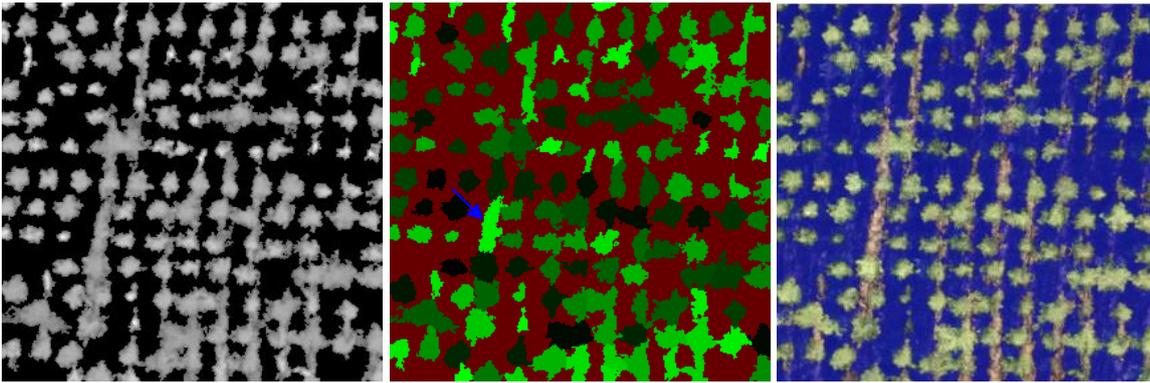


Abb. 6.32: Ergebnisse von BROWNM0 auf *lamotte1*; Oben: Parameter, Bild, Saatpunkte; Unten: numPass, Segmentierungsergebnis (Regionen und Maske)

Im Folgenden sollen nun wieder die Auswirkungen der eingeführten Steuerungsparameter untersucht werden. Da in 5.4 bereits hergeleitet wurde, dass NUM\_STEPS die Größe bestimmt, und dass durch NUM\_RUNS und MIN\_PASS die Bestimmtheit erhöht werden kann, liegt der Fokus dabei hauptsächlich auf den unterschiedlichen Gewichtsfunktionen. Abb. 6.33 und Abb. 6.34 zeigen die Auswirkungen des Parameters WEIGHT\_POT auf *goett{4}* bzw. *goett{2}*.

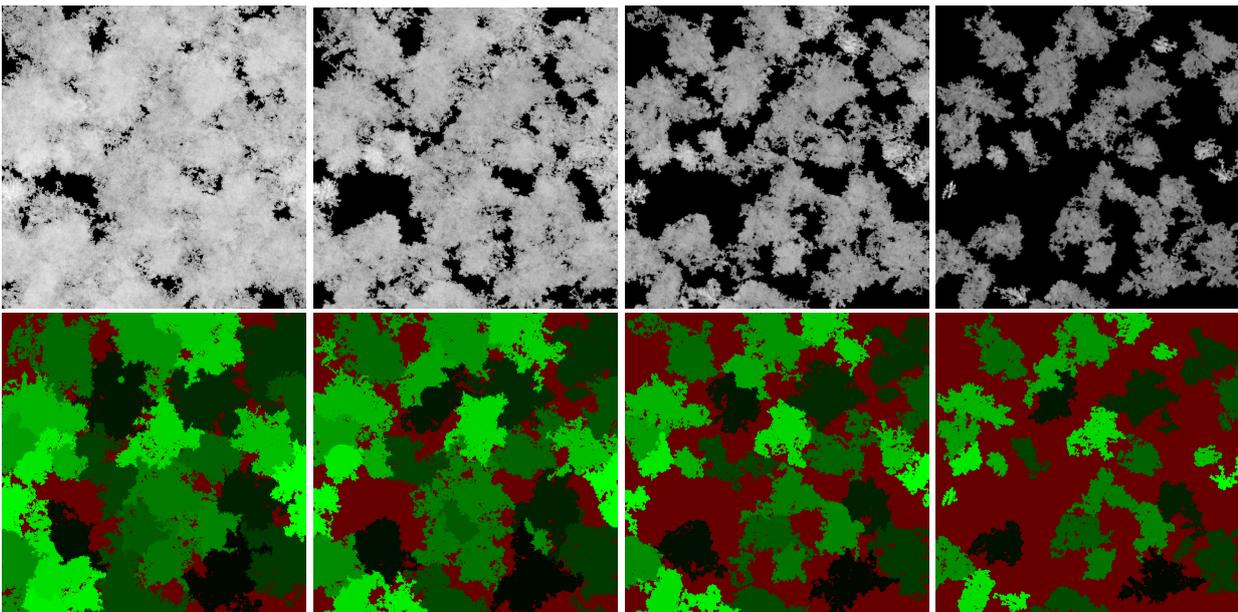
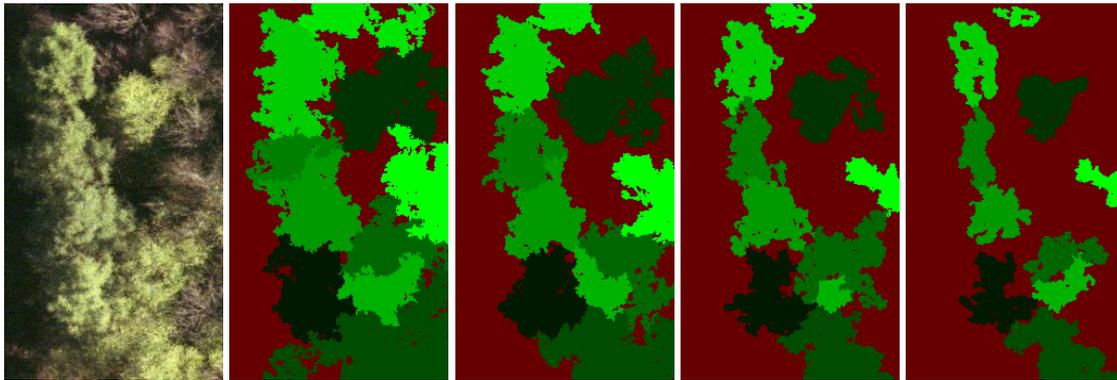
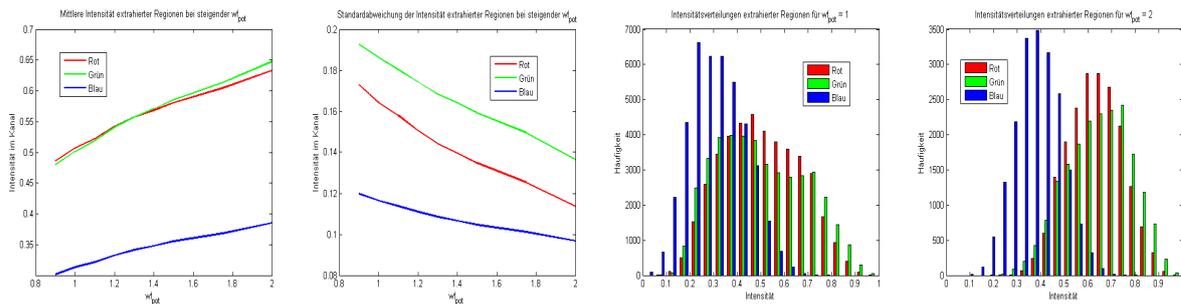


Abb. 6.33: Auswirkungen des Parameters BM.WEIGHT\_POT auf *goett{4}*, WEIGHT\_POT = 0.8, 1.1, 1.5 und 2.0; Oben: numPass; Unten: Segmentierungsergebnis



**Abb. 6.34: Auswirkungen des Parameters BM.WEIGHT\_POT auf *goett{2}*: Bild, Segmentierungsergebnisse für WEIGHT\_POT = 0,8, 1,1, 1,5 und 2,0,**

Es wird deutlich, dass die Größe wie in 5.4 vorhergesagt mit steigender Potenz abnimmt, da die durchschnittliche Schrittweite durch die Gewichtsfunktion reduziert wird (dies kann gegebenenfalls durch eine höhere Anzahl von Schritten ausgeglichen werden). Die entscheidende Frage ist, ob der Effekt nach Intensitäten differenziert auftritt, ob sich also auch die radiometrische Charakteristik durch den Parameter steuern lässt. Dies ist in Abb. 6.35 dargestellt.

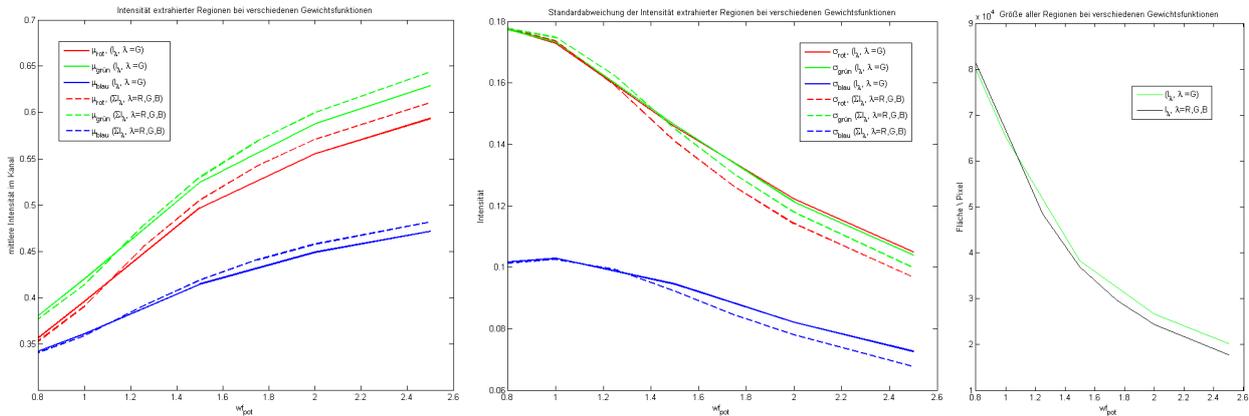


**Abb. 6.35: Veränderung der radiometrischen Charakteristik aller extrahierten Regionen bei Variation von WEIGHT\_POT auf *goett{2}***

Dafür wurden die Intensitätsverteilungen der Pixel aller extrahierten Regionen untersucht. In der linken Abbildung ist der Mittelwert der Intensitäten in den Kanälen Rot, Grün und Blau zu sehen, also der Mittelwert aller Pixel, die bei entsprechender Gewichtspotenz einer Region zugeordnet wurden. Die Kurve ist deutlich ansteigend. Somit kann festgehalten werden, dass die extrahierten Regionen bei steigender WEIGHT\_POT heller werden.

Die zweite Kurve zeigt die Veränderungen der entsprechenden Standardabweichungen. Die Standardabweichung fällt monoton mit steigender Gewichtspotenz. Dies liegt natürlich zum Teil daran, dass die Regionen kleiner werden, und somit auf Grund der texturspezifischen Semivarianz auch die Standardabweichung geringer wird. Dennoch bleibt festzuhalten, dass die Regionen mit steigender WEIGHT\_POT homogener werden, und dieser Parameter somit prinzipiell zur Anpassung an die radiometrische Beschaffenheit des Bildmaterials (etwa schwachen Kontrast) geeignet ist. Zur Verdeutlichung sind rechts noch einmal die kompletten Intensitätsverteilungen für WEIGHT\_POT = 1 und 2 dargestellt: Wie erwartet ist die rechte Verteilung gegenüber der linken deutlich schlanker und nach rechts verschoben.

Weiterhin soll untersucht werden, wie sich die Wahl der Farbkanäle, auf denen die Gewichtsfunktion arbeitet, auswirkt. Dies ist in Abb. 6.36 dargestellt. Alle drei Kurven beziehen sich auf *lamotte1*.



**Abb. 6.36: Radiometrische und geometrische Charakteristik extrahierter Regionen bei verschiedenen Gewichtsfunktionen: Intensität (links), Standardabweichung (Mitte) und Gesamtgröße (rechts)**

In den linken beiden Kurven ist wie oben die radiometrische Charakteristik aller extrahierten Regionen dargestellt. Dabei beziehen sich die durchgezogenen Linien auf Regionen, die mit Gewichtsfunktionen des Typs  $w_p(\mathbf{r}) = (I_\lambda(\mathbf{r}))^p$ ,  $\lambda = G$  segmentiert wurden. Die gestrichelten Linien

beziehen sich auf Gewichtsfunktionen des Typs  $w_p(\mathbf{r}) = \frac{1}{n} \sum_{\lambda=1}^n (I_\lambda(\mathbf{r}))^p$  für die Kanäle  $\lambda = R, G, B$ .

Die rote, grüne und blaue Linie geben die mittlere Intensität aller extrahierten Regionen im jeweiligen Farbkanal des Originalbildes an. Auf der x-Achse ist die Potenz  $p$  (WEIGHT\_POT, hier als  $wf_{pot}$  für Potenz der *weight function* bezeichnet) aufgetragen, auf der y-Achse die mittlere Intensität aller mit diesen Parametern extrahierten Regionen im jeweiligen Kanal. Um die Graphik zu verdeutlichen, sei ergänzt, dass sich diese Werte zwischen zwei Extremen bewegen: Für  $p \propto \infty$  geht die Intensität gegen die mittlere Intensität der Saatpunkte, für  $p \propto 0$  gegen die mittlere Intensität des gesamten Kanals.

Die Intensität wächst mit  $p$ , die Regionen werden also heller (was allerdings auch mit AN 25ff einhergeht). Gleichzeitig sinkt die Standardabweichung, die Regionen werden homogener. Die Potenz der Gewichtsfunktion kann daher als wirksamer Steuerungsparameter betrachtet werden.

Beide Effekte sind etwas stärker für die Gewichtsfunktionen, die alle drei Farbkanäle benutzen, eine grundlegend unterschiedliche Charakteristik ist aber nicht zu erkennen. Auch die geometrische Charakteristik (links: Größe der extrahierten Regionen) unterscheidet sich kaum zwischen den beiden Typen. Unter Berücksichtigung des erheblich höheren Rechenaufwands (die Berechnung der Gewichtsfunktion nimmt mit Abstand die meiste Zeit in jedem Schritt in Anspruch, so dass sich beinahe ein Faktor 3 ergibt) scheint der Mehrwert zu gering.

Ergänzend wurden nachträglich noch zwei weitere Gewichtsfunktionen namens 'p-norm' und 'min' implementiert. Die erste misst den euklidischen Abstand im Farbraum bzw. allgemeiner den

Abstand in einer  $p$ -Norm. Sie hat die Form:  $w_p(\mathbf{r}) = \|\mathbf{I}(\mathbf{r})\|_p = \left( \sum_{\lambda=1}^n (I_\lambda(\mathbf{r}))^p \right)^{\frac{1}{p}}$ . Die zweite nimmt

jeweils den kleinsten Farbwert als Basis der Gewichtsfunktion:  $w_p(\mathbf{I}(\mathbf{r})) = \left(\min_{\forall \lambda} (I_\lambda(\mathbf{r}))\right)^p$ . Auch hier konnten zunächst keine entscheidenden Unterschiede in der Charakteristik festgestellt werden, weitere Untersuchungen könnten aber lohnen.

Abb. 6.37 fasst die Auswirkungen der Steuerungsparameter bei BROWNMO zusammen.

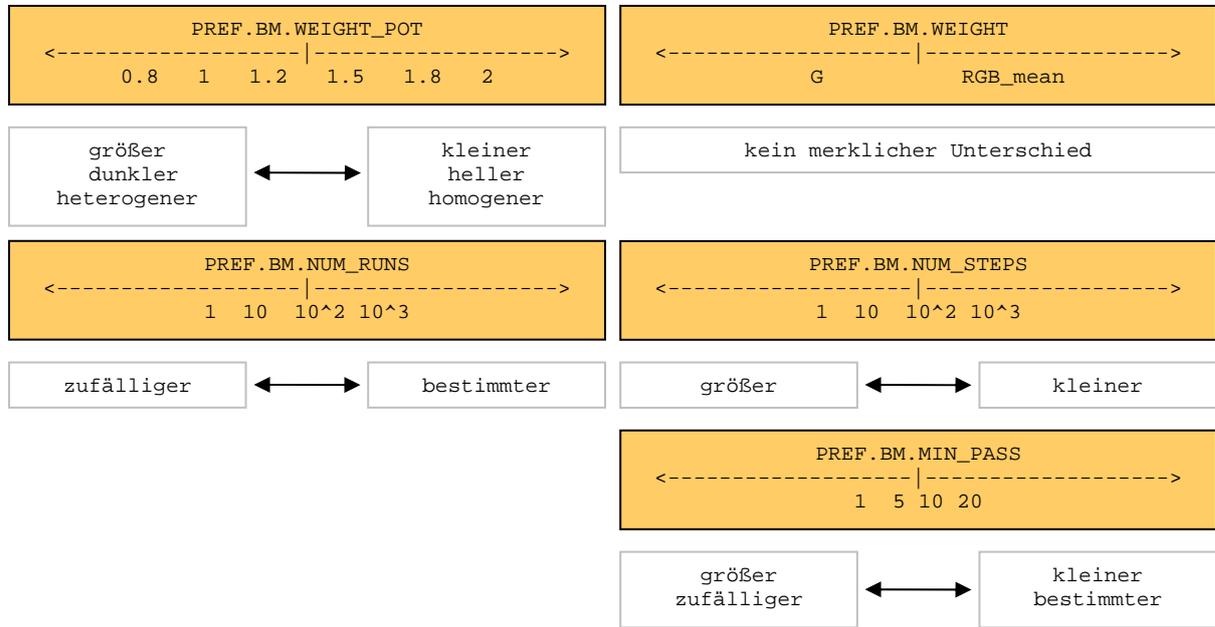


Abb. 6.37: Auswirkungen der Steuerungsparameter bei BROWNMO

Zum Abschluss soll kurz auf die mit Verfahren verbundene Unsicherheit in Bezug auf das Segmentierungsergebnis eingegangen werden. Auch wenn die vorgestellten Ergebnisse sehr vielversprechend sind, bleibt der Nachteil, dass das Verfahren nicht deterministisch ist, die Ergebnisse also nicht reproduzierbar sind. Zur Überprüfung wurde mehrfach mit der gleichen Gewichtsfunktion ( $w_p = (I_G(\mathbf{r}))^{1.3}$ ) und den gleichen Parametern (NUM\_RUNS: 100, NUM\_STEPS: 1000, MIN\_PASS: 5) auf *lamotte1* segmentiert. Die Ergebnisse sind in Abb. 6.38 dargestellt.

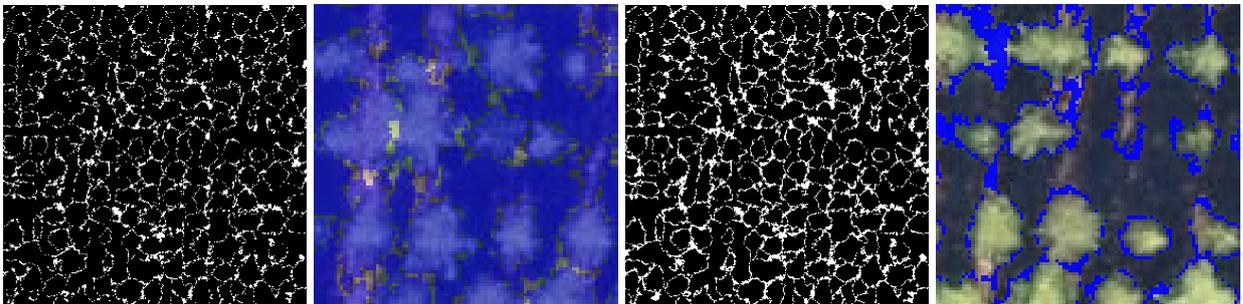


Abb. 6.38: Unterschiedlich klassifizierte Bereiche bei mehreren Durchgängen

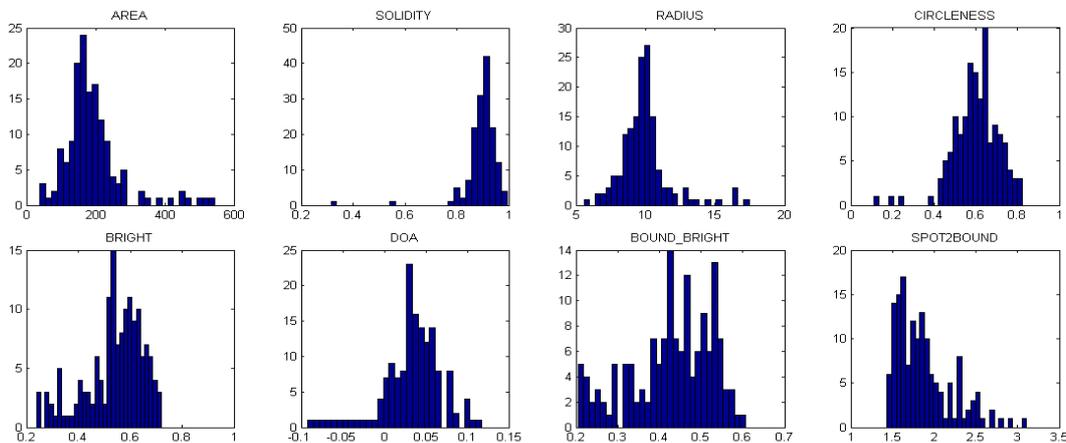
Die beiden linken Bilder zeigen Pixel, die bei 2 Durchgängen unterschiedlichen Segmenten zugeordnet wurden, die beiden rechten die Abweichungen bei 3 Durchgängen. Die zu beobachtende Unsicherheit ist nicht unerheblich. So werden knapp 10% der Pixel in zwei Durchgängen unterschiedlich zugeordnet, wobei wieder keine große Abweichung zwischen Gewichtsfunktionen verschiedenen Typs

festgestellt wurde. Die Bilder zeigen aber auch, dass die unsicheren Pixel, die ja in etwa mit der Grenzwahrscheinlichkeit zum Segment gehören, eine sehr gute Abgrenzung der Kronen liefern.

Die Unsicherheit könnte natürlich durch eine Erhöhung der Anzahl der Durchläufe reduziert werden. Dadurch würde sich aber auch der Rechenaufwand erheblich vergrößern, und es kann nur die in 5.4 aufgeworfene Überlegung unterstrichen werden, ob sich der erwartete Ausgang dieses aufwendigen Zufallsexperiments nicht auch analytisch berechnen ließe. Dies sollte Gegenstand weiterer Untersuchungen werden.

## 6.6 Ein einfaches Modell zur Hypothesenauswahl

Zur Merkmalsextraktion werden mit dem Modul `bbFeatureExtraction` die in 4.4.2 vorgeschlagenen Merkmale extrahiert. Die geometrischen Merkmale `AREA`, `SOLIDITY`, `CIRCLENESS` und `RADIUS` sind dabei Eigenschaften des gesamten Segments, die radiometrischen sind Mittelwerte aller Pixel (`BRIGHT` und `DOA`) bzw. einer Auswahl (`BOUND_BRIGHT`) von Pixeln. Abb. 6.39 zeigt die Merkmale für die Segmente aus Abb. 6.21 (unten).



**Abb. 6.39:** Merkmale aller Regionen in Abb. 6.21

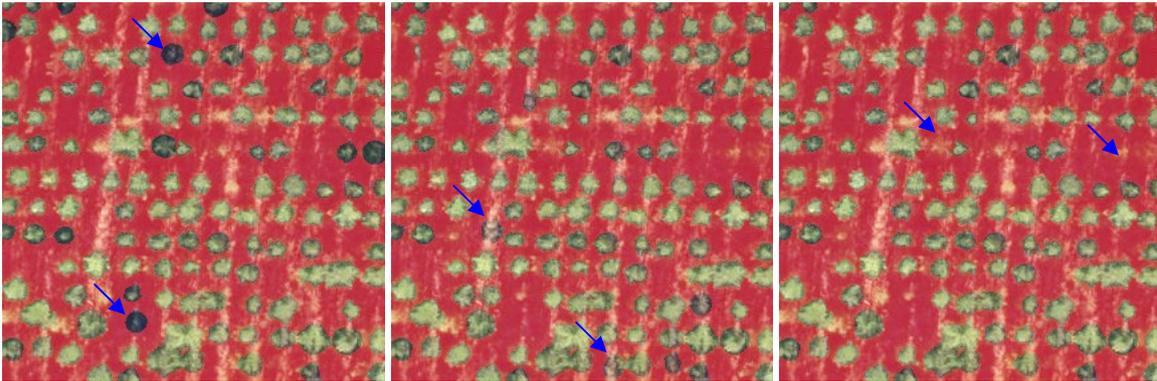
Erwartungsgemäß sind die extrahierten Segmente relativ kreisförmig, hell, haben einen `DOA`  $> 0$  und eine `SOLIDITY`  $> 0.6$ .

Da eine *top-down* Steuerung angestrebt wird (siehe 4.2), soll das Modell nicht aus den Merkmalen der Regionen abgeleitet werden, sondern möglichst selbsterklärend sein. Daher soll darauf verzichtet werden, die einzelnen Verteilungen eingehend zu analysieren.

Gleichzeitig soll das Modell eine der Situation angemessene Komplexität haben. Es wurde bereits festgestellt, dass in diesem Fall bei der Hypothesenauswahl im Wesentlichen Bäume von hellen Bodenflecken und dunklen Schattenbereichen abgegrenzt werden müssen. Dafür kommen vor allem radiometrische Merkmale in Frage, auch da die geometrischen schon stark in die Segmentierung eingingen (darum sind `SOLIDITY` und `AREA` relativ einheitlich).

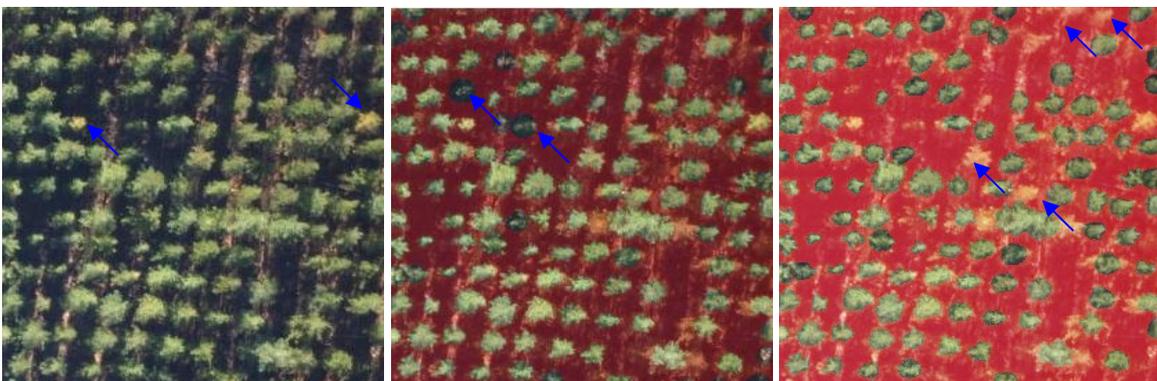
Um helle Bodenflecken zu erkennen bietet sich der `DOA` an, der beim roten Boden unter Null liegen sollte, bei grüner Vegetation sollte er dagegen merklich höher liegen. Als Entscheidungsgrenze soll daher der Wert Null festgelegt werden. In Abb. 6.40 (links) ist das Ergebnis einer Hypothesenauswahl anhand dieses einzelnen Merkmals zu sehen. Die Hypothesenauswahl erfolgt mit dem Modul

bbHypoSelect, bei dem über den Parameter `PREF.SET_NAME` verschiedene einfache Modelle ausgewählt werden können.



**Abb. 6.40: Hypothesenauswahl mit verschiedenen Modellen auf *lamotte1***

Gemäß der Zielsetzung wurden die beleuchteten Bodensegmente als Nicht-Baum klassifiziert, einige der schattigen (blaue Pfeile) aber als Bäume. Um diese zu finden bietet sich das Merkmal `BRIGHT` an. So wird im Modell bei der mittleren Abbildung gefordert, dass die Helligkeit des Segments über dem Mittelwert im Bild liegt, die Entscheidungsgrenze wird also aus dem Histogramm gewonnen. Das Auswahlergebnis in der Mitte zeigt die Eignung dieses Merkmals, schattige Bodensegmente auszusortieren, helle bleiben allerdings erhalten (Pfeile). Daher bieten sich beide Merkmale gemeinsam als Merkmalsraum an und mit den *a priori* festgesetzten Entscheidungsgrenzen ist bereits ein einfaches Modell generiert. Das Klassifikationsergebnis mit beiden Merkmalen (rechts) hat bereits eine hohe Qualität, obwohl auch einige richtige Bäume verworfen (blaue Pfeile rechts) wurden.



**Abb. 6.41: *lamotte2*, extrahierte Regionen und Hypothesenauswahl**

Bei *lamotte2* ist dies noch stärker der Fall (siehe Abb. 6.41 rechts). Dies hat verschiedene Gründe: Zunächst enthält das Bild verfärbte Bäume (Pfeile links), die wahrscheinlich von Holzwespen befallen sind. Dass diese nicht als grüne Vegetation klassifiziert werden, ist im Sinne des Modells und birgt sogar potentielle Anwendungen in der Schädlingsbekämpfung. Weitere Bäume, die ebenfalls wegen des DOA aussortiert werden (Pfeile rechts), sind zwar nicht offensichtlich krank, es könnte sich aber um eine frühere Stufe von Befall handeln. Andererseits wurde bereits mehrfach auf die schlechte radiometrische Qualität des Bildes hingewiesen, so dass ein solcher Befund ohne weitere Untersuchungen voreilig wäre. Es zeigt also auch, dass sich eine Anpassung des Modells an die radiometrische Charakteristik nicht ganz vermeiden lässt.

Schließlich gibt es Bäume, die größtenteils im Schatten liegen und daher als zu dunkel aussortiert werden. Hierin zeigt sich die Schwäche von BRG, kreisförmige Segmente einer bestimmten Größe zu produzieren.

## 6.7 Genauigkeit

Da die Segmentierungsergebnisse bislang vorwiegend qualitativ bewertet wurden, soll an dieser Stelle noch auf die Genauigkeit der extrahierten Objekte eingegangen werden. Dafür wurden für die beiden Bilder *lamotte1* und *lamotte2* von Hand Referenzpolygone der Kronen abgegrenzt (Abb. 6.42), die mit den extrahierten Regionen verglichen werden.

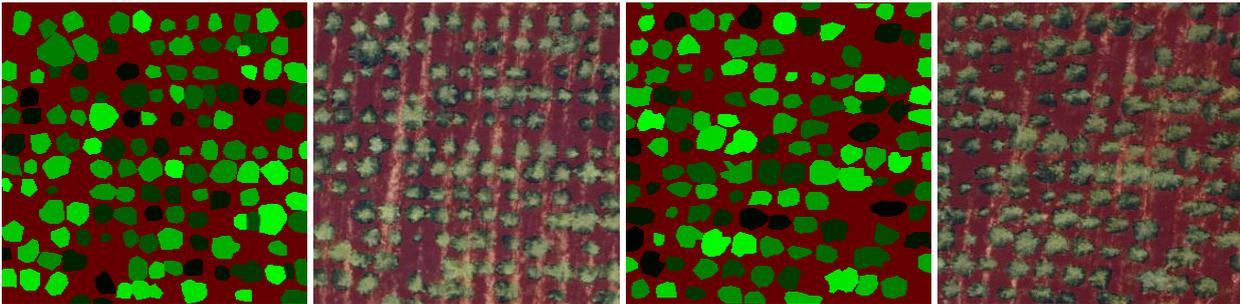


Abb. 6.42: Referenzpolygone für *lamotte1* und *lamotte2*, Regionen und Maske

Um übereinstimmende Regionen in Extraktionsergebnis und Referenz finden zu können, wird zunächst die Überlappung (*overlap*) nach der Definition von STRAUB & HEIPKE (2004, siehe (3.1)) zwischen allen extrahierten und allen Referenzregionen berechnet. Als Übereinstimmung wird ein *overlap* von mindestens 0.5 gewertet, gibt es mehrere Übereinstimmungen mit demselben Segment, wird dasjenige mit dem größten *overlap* gewählt und in einer Liste gespeichert.

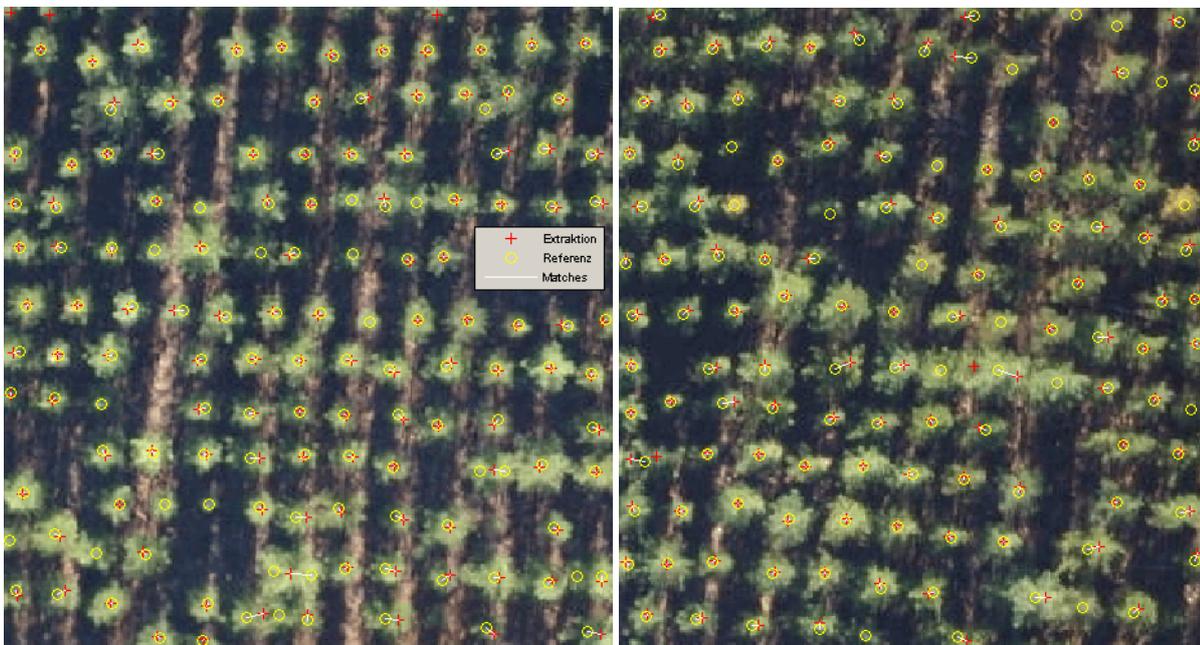
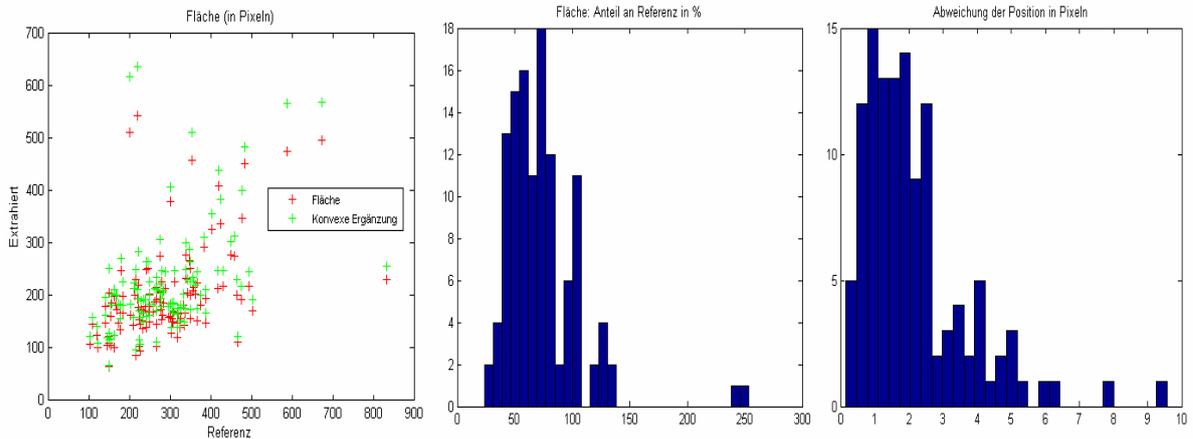


Abb. 6.43: Matching-Ergebnisse auf *lamotte1* (links) und *lamotte2*(rechts)

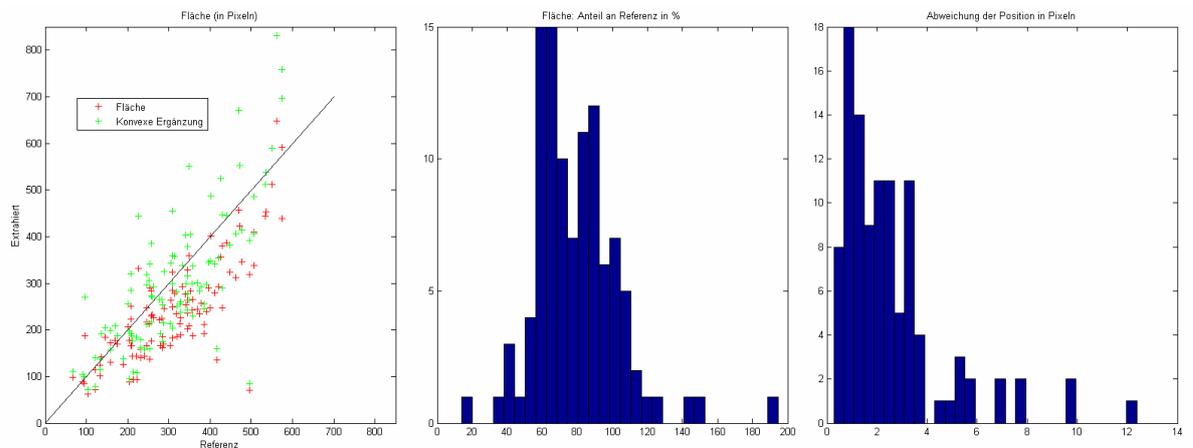
Abb. 6.43 zeigt die Ergebnisse des *matchings*. Die Schwerpunkte der extrahierten Bäume (BRG mit Modell aus dem letzten Abschnitt) sind als rote Kreuze dargestellt, die der Referenz-Bäume als gelbe Kreise und als Übereinstimmung gewertete Regionen sind durch eine weiße Linie verbunden.

Für *lamotte1* liegen die Vollständigkeit bei 86.3%, und die Korrektheit bei 97.5%. Die „fälschlicherweise“ als Baum erkannten Objekte sind aber sämtlich abgeschnittene Bäume am Bildrand, die daher nicht in die Referenz aufgenommen wurden.



**Abb. 6.44: Genauigkeit der extrahierten Bäume bei BRG auf *lamotte1***

In Abb. 6.44 sind Kennziffern zur Genauigkeit der einzelnen Segmente dargestellt. Im linken Diagramm ist die Fläche der extrahierten Kronen gegen die der Referenz-Kronen aufgetragen. Es ist zu erkennen, dass die Fläche von den extrahierten Regionen (rote Kreuze) tendenziell unterschätzt wurde, im Schnitt liegt sie bei 74% der Referenzfläche. Da schattige Bereiche am Kronenrand teilweise nicht mit segmentiert wurden, ist zum Vergleich auch die Fläche der konvexen Ergänzungen der Segmente aufgetragen (grüne Kreuze), die im Schnitt 83% der Referenzfläche beträgt. Das mittlere Diagramm zeigt die Verteilung der Flächenanteile, das rechte die Verteilung der Positionsabweichungen. Der mittlere Positionsfehler liegt bei 2.2 Pixeln bzw. 32cm, was im Vergleich zu den bekannten Verfahren (siehe Kapitel 3) als sehr gut bewertet werden kann.



**Abb. 6.45: Genauigkeit der extrahierten Bäume bei BROWNMO auf *lamotte2*, *rgb\_mean*, 1.5**

Abb. 6.45 zeigt die gleichen Werte für BROWNMO (*RGB\_mean*,  $w_p=1.5$ ) auf *lamotte2*. Hier liegen die extrahierten Flächen noch deutlich näher an der Referenz. Tab. 8 gibt schließlich einen Überblick

über die erzielten Genauigkeiten bei verschiedenen Datensätzen. Alle Angaben bis auf den Positionsfehler sind in Prozent. Die Vollständigkeitswerte liegen bis auf WST (was gesondert bewertet werden muss, da hier nur eine Maßstabsebene betrachtet wurde) sämtlich zwischen 80 und 90%, die Korrektheit liegt durchgehend über 95%. Diese Werte liegen im oberen Bereich der Vergleichswerte in der Literatur für einfache Waldsituationen (siehe etwa STRAUB 2003A: 38), da sie mit sehr einfachen Modellen erzielt wurden, können sie sogar als sehr gut bewertet werden.

**Tab. 8: Genauigkeit der extrahierten Ergebnisse für verschiedene Verfahren und Datensätze**

Verfahren	WST		BRG		BROWNMO			
	<i>lamotte1</i>	<i>lamotte2</i>	<i>lamotte1</i>	<i>lamotte2</i>	<i>lamotte1</i>	<i>lamotte1</i>	<i>lamotte2</i>	<i>lamotte2</i>
Datensatz								
Parameter	SIGMA = 4.5	SIGMA = 4.5	AUTO_ SIZE	AUTO_ SIZE	wf:'G'; wp: 1.5	wf:'RGB_ mean'; wp: 1.5	wf:'RGB_ mean'; wp: 1.5	wf:'RGB_ mean'; wp: 2
Modell	convex & SF_R	convex & SF_R	DOA & Bright	DOA & Bright	DOA & Bright	DOA & Bright	DOA & Bright	DOA & Bright
Vollständigkeit	73	72	86	87	83	81	84	87
Korrektheit	98	89	98	98	96	95	96	96
Mittlerer Positionsfehler (m)	0.28	0.32	0.32	0.35	0.34	0.35	0.38	0.41
Anteil an Referenzfläche	73	77	74	72	100	95	80	51
Anteil Konvexe Ergänzung	77	83	83	79	121	117	99	63

Die Genauigkeit der extrahierten Parameter muss differenziert betrachtet werden. Die Fläche der Kronen wird tendenziell bei allen Verfahren unterschätzt, da schattige Teile der Krone nicht mit segmentiert werden. Dies ist aber nicht unbedingt als Nachteil zu bewerten, da nur so der genaue Umriss zur Klassifikation herangezogen werden kann. Außerdem sei darauf hingewiesen, dass für waldkundliche Messungen eher der Kronendurchmesser als relevanter Parameter benötigt wird, der wesentlich genauer sein dürfte, da die schattigen Einbuchtungen den maximalen Durchmesser kaum beeinflussen.

Die Ergebnisse von WST sind wieder unter Vorbehalt zu betrachten, da die Hypothesenauswahl im LMR fehlt. Außerdem ist das Verfahren ja so konzipiert, dass entlang der Nullstellen der zweiten Ableitung segmentiert wird, wodurch systematisch Teile der Kronen abgeschnitten werden. Bei BROWNMO scheinen die extrahierten Flächen zwar auf den ersten Blick besser zu sein, zum Teil liegt dies aber daran, dass fälschlicherweise Bodenteile mit als Kronen segmentiert werden. Dafür spricht auch, dass der durchschnittliche *overlap* der übereinstimmenden Segmente bei BROWNMO (alle *lamotte1*, wf: 'G', wp: 1.5) mit 89% niedriger liegt als bei BRG mit 93% und WST ( $\sigma$ : 4.5) mit 94%. Das letzte Ergebnis wurde in die Tabelle aufgenommen um zu verdeutlichen, dass eine ungeeignete Wahl des Parameters `WEIGHT_POT` kaum Auswirkungen auf Vollständigkeit und Korrektheit hat, wohl aber auf die Größe der extrahierten Segmente.

Die Position der extrahierten Segmente weicht im Schnitt nur zwischen 28cm und 41cm ab, was auch an den relativ strengen *matching*-Regeln liegt. Würde der Begriff der Übereinstimmung wie bei vielen Studien etwas großzügiger definiert, würden die Positionsfehler zwar zunehmen, der *error of omission* aber abnehmen.

## 7 Zusammenfassung und Ausblick

Im Rahmen dieser Arbeit wurden verschiedene vorbereitende Studien im Hinblick auf ein Expertensystem zur automatischen Extraktion von Bäumen aus Fernerkundungsdaten geleistet. Ein solches System sollte für jede gegebene Problemstellung das beste verfügbare Verfahren auswählen und steuern, also mit geeigneten Parameter-Einstellungen ausführen. In Kapitel 3 wurde zunächst eine kategorisierende Übersicht relevanter Verfahren erstellt, und diese einer vergleichenden Evaluation im Hinblick auf ihre Verwendbarkeit in einem allgemeinen Expertensystem unterzogen. Dabei wurden Stärken und Schwächen der einzelnen Verfahren analysiert und die prinzipielle Eignung für erweiterte Problemstellungen diskutiert.

Es wurde festgestellt, dass die Verfahren oftmals für einzelne Datensätze optimiert wurden und Informationen zur Übertragbarkeit rar sind. Insbesondere gehen häufig Modell-Annahmen implizit in die Verfahren ein, ohne als solche gekennzeichnet zu sein. Daher wird nach Überzeugung des Autors neben der empirischen Untersuchung der Eignung einzelner Verfahren für weitere Problemstellungen (zu der es bereits erste Studien gibt, vgl. ERIKSON *et al.* 2006, ERIKSON & OLOFSSON 2005), ein systematischerer, konzeptionell-analytischer Zugriff auf die Frage der Übertragbarkeit von Verfahren (bzw. deren Eignung für bestimmte Datensätze) benötigt, insbesondere dann, wenn die Wahl des Verfahren perspektivisch innerhalb eines Expertensystems mit automatisierter Inferenz (etwa durch ein Bayes'sches Netz) stattfinden soll.

Dazu sollte diese Arbeit einen Beitrag leisten. Zur Analyse der Übertragbarkeit wurde ein Schema von Modell-Annahmen entwickelt, die aus bestehenden Verfahren entnommen bzw. abgeleitet wurden. Anschließend wurde exemplarisch für eine Reihe implementierter Segmentierungsverfahren detailliert aufgezeigt, welche Annahmen eingehen und welche Anforderungen damit an das eingehende Datenmaterial gestellt werden. Weiterhin wurde in 6.1 anhand von Beispielen verdeutlicht, wie vielfältig die Repräsentation von Bäumen in Luftbildern sein kann, und dass einzelne Datensätze unterschiedliche Voraussetzungen erfüllen. In 6.2 bis 6.5 wurde jeweils versucht, anhand des entwickelten Schemas vorab Aussagen über die Eignung einzelner Verfahren bezüglich der gegebenen Datensätze zu treffen.

Diese Aussagen können als Erfolg betrachtet werden. So ist es an verschiedenen Stellen gelungen, auftretende Probleme vorherzusagen und analytisch zu begründen. Es wurde gezeigt, dass ein solches Schema prinzipiell als Variable zur automatischen Inferenz von Verfahren geeignet sein könnte. Es wurde aber ebenfalls deutlich, dass das entwickelte Schema differenziert und vor allem quantifiziert werden muss. So wurden zwar qualitative Aussagen derart getroffen, dass eine bestimmte Modell-Annahme „stärker“ oder „schwächer“ in ein Verfahren eingeht, dass sie für einen Datensatz „gut“ oder „weniger gut“ erfüllt ist, und dass sich das Verfahren damit „tendenziell eignet“, oder „eher nicht eignet“. Damit konnten zwar die Grenzen der Übertragbarkeit anschaulich aufgezeigt werden, um eine automatische Entscheidung treffen zu können, müssten die Aussagen aber zu belastbaren und berechenbaren Quantitäten und objektivierbaren Merkmalen erweitert werden.

Als weiterer Schwerpunkt der Arbeit wurde in Kapitel 4 eine mögliche Architektur für ein Expertensystem zur Objektextraktion von Bäumen entwickelt. Entsprechend dem obigen Fokus wurde großer Wert darauf gelegt, die Relation zwischen einzelnen Verfahren und dem Modell aufzuzeigen. Zur Erweiterung der Modell-Annahmen wurden diese um quantifizierbare Merkmale erweitert. Es wurde

betont, dass im Sinne der Vielseitigkeit und Flexibilität des Systems eine *top-down* Steuerungsrichtung und ein selbsterklärendes Modell zu bevorzugen wären. Anhand der unterschiedlichen Voraussetzungen der einzelnen Datensätze wurde aber auch klar, dass eine gewisse Anpassung insbesondere der radiometrischen Komponenten im Einzelfall unumgänglich ist. Besonderes Augenmerk wurde der benötigten Modell-Komplexität in Abhängigkeit vom Szeneninhalte gewidmet. Dabei stellte sich heraus, dass insbesondere für Waldsituationen, in denen wenige weitere Objekte enthalten sind, oft bereits einfache Modelle genügen, was exemplarisch in 6.6 demonstriert wurde.

Weiterhin wurde ein schematisierter Ablauf entwickelt, der einer Spezialisierung des klassischen Ablaufs bei der ME entspricht. Der erste entscheidende Schritt ist hierbei die Segmentierung, unter der hier die Abgrenzung von Objekten im Bild verstanden wird, und die als Modul exemplarisch herausgegriffen und vertiefend getestet wurde.

Zu diesem Zweck wurden zwei Region Growing (BRG und BROWNMO) und ein morphologisches Segmentierungsverfahren (WST) sowie zwei Verfahren zur Gewinnung von Saatpunkten (REGMAX und DISTRANS) implementiert. Bei allen Segmentierungs-Algorithmen handelt es sich um eigenständige Weiterentwicklungen von in Kapitel 3 vorgestellten Verfahren aus der Literatur. Zusätzlich wurden weitere Steuerungsparameter eingeführt, um die Funktionalität zu erweitern und die Flexibilität und vielseitige Verwendbarkeit im Hinblick auf das Expertensystem zu erhöhen, die nach Möglichkeit derart beschaffen sein sollten, dass das Ergebnis eine stetige und monotone Funktion des Parameters ist. Die Eignung und Wirkungsweise der integrierten Parameter wurde in Kapitel 6 ausführlich getestet. Im Folgenden werden einige zentrale Ergebnisse zusammengefasst.

Mit dem WST-Verfahren ist es erfolgreich gelungen, ein Verfahren für DOM auf radiometrische Bilder zu übertragen. Es wurde bestätigt, dass die invertierte Summe des Quadrats der zweiten Ableitungen auch hier als Segmentierungsfunktion geeignet ist. Die beiden Saatgut-Verfahren REGMAX und DISTRANS lieferten vergleichbare Ergebnisse. Der Parameter  $\sigma$  reagierte unempfindlich auf kleine Veränderungen und ist geeignet, den Maßstab zu bestimmen. Es wird aber ein relativ einheitlicher Abstand der Bäume vorausgesetzt, bei komplexeren Situationen sinkt die Qualität der Saatpunkte. Bei radiometrisch homogenen Objekten kann mit dem DISTRANS-Verfahren die Größe des Objekts geschätzt und an weitere Schritte übergeben werden. In Verbindung damit liefert das BRG-Verfahren gute Ergebnisse für abgegrenzte, radiometrisch homogene Kronen, wobei die geometrischen Annahmen besonders stark ins Verfahren eingehen. Sind diese nicht erfüllt, ist das Verfahren dagegen weitgehend unbrauchbar. Beim BROWNMO-Verfahren wurden ausführlich die Auswirkungen verschiedener Gewichtsfunktionen getestet. Dabei wurde festgestellt, dass die Potenzierung der Gewichtsfunktion ein geeigneter Weg ist, die radiometrische Charakteristik der erzeugten Regionen zu bestimmen und das Verfahren etwa an kontrastarme Bilder anzupassen. Gleichzeitig beeinflusst sie aber auch die Größe der erzeugten Regionen, was durch die anderen eingeführten Parameter ausgeglichen werden kann. Weiterhin wurden verschiedene Typen von Gewichtsfunktionen getestet, deren Ergebnisse bei den vorliegenden Datensätzen aber alle vergleichbar waren. Das Verfahren liefert insgesamt sehr vielversprechende Segmentierungsergebnisse bei komplexen Situation und deutlich schwächeren geometrischen Voraussetzungen, hat aber den großen Nachteil nicht deterministisch zu sein. So wurden bei verschiedenen Ausführungen bis zu 10% der Pixel unterschiedlich klassifiziert, wobei die unsicheren Bereiche sehr gute Begrenzungen der Kronen darstellten. Die Bestimmtheit des Ergebnisses ließe sich durch eine höhere Anzahl von Durchläufen

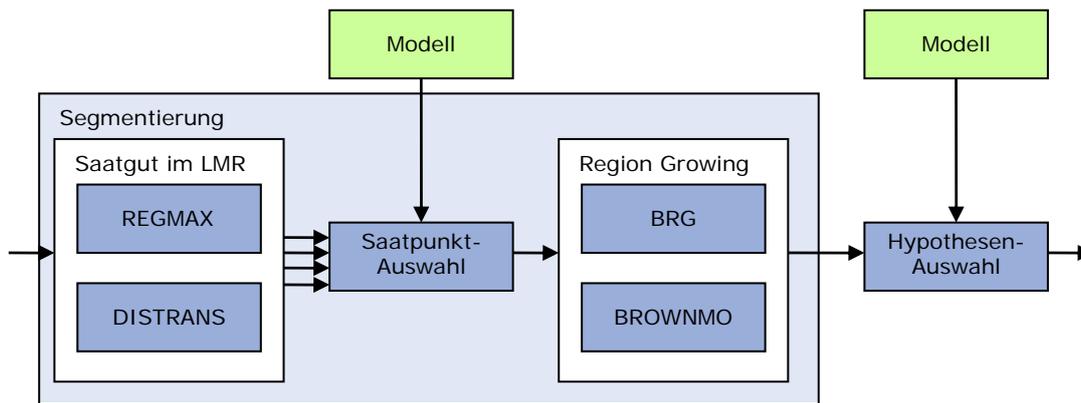
steigern, auf Grund der ohnehin sehr hohen Rechenaufwands wurde aber bereits in Abschnitt 5.4 anhand eines vereinfachten Experiments darüber nachgedacht, ob das Bild nicht als Potential begriffen werden könnte, mit dessen Hilfe sich die Aufenthaltswahrscheinlichkeiten direkt berechnen ließen. Dies sollte Gegenstand weiterer Untersuchungen werden.

Zusammenfassend kann festgehalten werden, dass der Funktionsumfang einzelner Algorithmen durch die Einführung geeigneter Steuerungsparameter erweitert und angepasst werden konnte. Andererseits erschweren zusätzliche Parameter die Steuerung der Verfahren. Nach der Festlegung auf bestimmte Parameter sollte also untersucht werden, ob diese nicht ihrerseits automatisch gewonnen werden können. So könnte der mittlere Baumabstand möglicher Weise aus den Frequenzen im Bild bestimmt werden (vgl. 3.5). Erste Versuche ergaben zumindest für die einfachen *lamotte*-Bilder vielversprechende Werte, weitere Untersuchungen müssten aber folgen. Besonders aufwendig könnte die Bestimmung einer optimalen Gewichtsfunktion beim BROWNMO Verfahren sein, wobei möglicherweise eine Analyse des Histogramms helfen könnte.

Auch wenn es ausdrücklich nicht darum ging, die Algorithmen für eine bestimmte Situation zu optimieren, wurden wichtige Ergebnisse in 6.7 quantifiziert. Sie liegen im Rahmen der Vergleichswerte für einfache Waldsituationen aus der Literatur, wobei ein direkter Vergleich der Werte unter Vorbehalt gesehen werden muss (vgl. 3.1.4). Im Hinblick auf die äußerst einfachen Modelle die dabei zugrunde gelegt wurden, können die Ergebnisse jedoch als sehr gut bewertet werden.

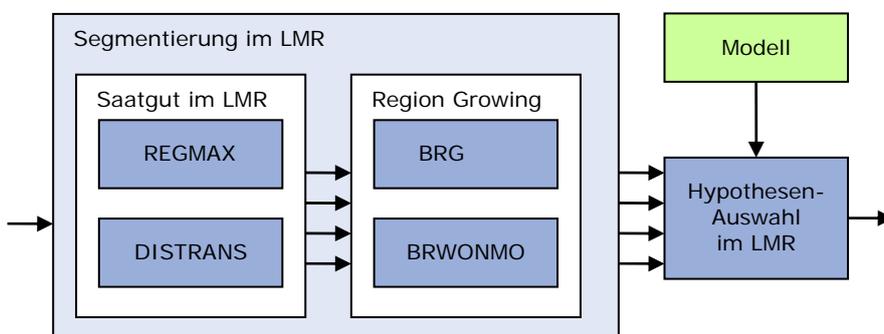
Eine zentrale Herausforderung für weitere Studien bleibt das Maßstabsproblem, das auch mit den vorgestellten Verfahren nicht befriedigend gelöst werden konnte. Es wurde aber vorgeschlagen, bei diesem Problem zwei Teile zu unterscheiden: Einerseits den Abstand zwischen den Kronen und andererseits die Größe der Kronen selbst, wobei beide Teilprobleme in unterschiedlichen Schritten adressiert werden können. So ist es gelungen beim BRG mit `AUTO_SIZE` Option Segmente unterschiedlicher Größe zu erzeugen und damit dem zweiten Teil des Problems zu begegnen. Für BROWNMO wurde zumindest aufgezeigt, dass die gewählten Parameter geeignet wären, Regionen verschiedener Größe zu generieren, wenn diese in Abhängigkeit von der erwarteten Größe unterschiedlich gesetzt würden.

Schwieriger ist das Maßstabsproblem des Abstandes, das bereits zentral in die Bestimmung der Saatpunkte eingeht. Sind die Abstände zu unterschiedlich (etwa mehr als eine Oktave) so versagen sämtliche Verfahren, die auf Tiefpass-Filterung einer Ausgangsfunktion beruhen. Daher wird es für RG-Verfahren in Ermangelung alternativer Methoden bei komplexen Situationen mit sehr unterschiedlichen Abständen zunächst unvermeidlich sein, auch die Saatpunkte auf verschiedenen Maßstabsebenen zu gewinnen. Abb. 7.1 zeigt einen Vorschlag für einen erweiterten Ablauf für RG-Verfahren.



**Abb. 7.1: Erweiterter Ablauf für Region Growing**

Da in mehreren Ebenen zu viele Saatpunkte erzeugt würden und falsche Saatpunkte das Segmentierungsergebnis verfälschten, müssten zunächst in einem Zwischenschritt die besten Saatpunkte ausgewählt werden. Dabei stellt sich das zusätzliche Problem, dass die Saatpunkte zur Auswahl deutlich weniger Merkmale zur Verfügung stellen als Segmente (da es sich um Punkte handelt, sind alle geometrischen Eigenschaften bereits abschließend festgelegt). Da die Klassifikation von extrahierten Regionen damit erheblich einfacher sein dürfte als die von Saatpunkten, müsste alternativ mit jeweils eigenen Saatpunkten auf mehreren Maßstabsebenen segmentiert werden und die Hypothesenauswahl im LMR stattfinden. Dieser alternative Ablauf ist in Abb. 7.2 dargestellt.



**Abb. 7.2: Alternativer Ablauf: Segmentierung im Maßstabsraum**

Der Preis dafür ist eine wesentlich komplexere Hypothesenauswahl, da die topologische Relationen der Segmente auf den verschiedenen Ebene berücksichtigt werden müssen (insbesondere muss entschieden werden, welche Form von Überlappungen toleriert werden). Eine Hypothesenauswahl im Maßstabsraum wurde für WST bereits von Straub (vgl. 3.7) vorgeschlagen, statt seiner Fuzzy-Zugehörigkeitsfunktionen wäre es aber sinnvoller, gleich PDF zu ermitteln. Eine solche Segmentierung im LMR und die entsprechende Hypothesenauswahl sollten Gegenstand weiterer Untersuchungen sein. Insbesondere müssen dafür auch bessere Modelle entwickelt werden, da die hier verwendeten insbesondere für komplexere Situationen unzureichend sein dürften, wobei gesonderte Objekt- und Bildmodelle wünschenswert wären. Weiterer Untersuchungsbedarf besteht auch bezüglich der Integration des Maßstabs ins Modell.

Ein bislang weitgehend ungelöstes Problem ist die Klassifikation von Baumarten, wobei die Verwendung von Hyperspektraldaten hier großes Potential für Fortschritte birgt. Der Klassifikationsschritt

---

könnte allerdings gegebenenfalls auch nachträglich in ein Expertensystem integriert werden. Selbiges gilt für die Vorsegmentierung, für die die Eignung von Texturmerkmalen weiter untersucht werden sollte.

Neben den bereits genannten Aufgaben wäre es daher zunächst nötig, die Fragen der Inferenz und der Wissensrepräsentation zu lösen. Bis zu einem vollautomatischen Expertensystem zur Extraktion von Bäumen aus Fernerkundungsdaten ist es also noch ein weiter Weg.

Andererseits wurde auch darauf hingewiesen, dass die benötigte Leistungsfähigkeit und damit die Spezifikation des Systems stark von der Anwendung abhängen. Da neben forstwirtschaftlichen auch weitere Anwendungen wie Umwelt-monitoring von Ökosystemen und Waldbeständen in Frage kommen, könnten sich abweichende Anforderungsprofile ergeben. Da zu erwarten ist, dass der Bedarf an solchen Anwendungen im Kontext des jüngeren Diskurses um Senken im globalen Kohlenstoffhaushalt und den Handel mit Verschmutzungsrechten stark zunehmen wird, könnte ein weniger umfassendes System für eine begrenzte Anwendung unter Umständen auch schneller implementiert werden.



## Literaturverzeichnis

- ANDERSEN, H.-E. REUTEBUCH, S.E. & G.F. SCHREUDER (2001): Automated Individual Tree Measurement through Morphological Analysis of LIDAR-Based Canopy Surface Model. In: Proceedings of the first International Precision Forestry Cooperative Symposium, S. 11-22. Seattle.
- ASRAR, G. (1989) (Hg): Theory and Applications of Optical Remote Sensing. New York.
- ATZBERGER C. & M. SCHLERF (2002): Automatisierte Bestimmung der Bestockungsdichte in Nadelwäldern aus räumlich hochauflösenden Ortholufbildern. In: Photogrammetrie Fernerkundung Geoinformation 3, S. 171-180.
- BAUMGARTNER, A., EKSTEIN, W., MAYER, H., HEIPKE, C. & H. EBNER (1997): Context-Supported Road Extraction. In: GRUEN, A., BALTSAVIAS, E.P. & O. HENRICSSON (Hg): Automatic Extraction of Man-Made Objects from Aerial and Space Images (II), S. 299-308. Basel.
- BISCHOF H. & A. PINZ (1989): Verwendung von neuronalen Netzen zur Bestimmung der Baumart aus digitalen Rasterbildern. In: PINZ A. (Hg): Wissensbasierte Mustererkennung, S. 149-161. Wien.
- BÖHM, J. (2005): Modellbasierte Segmentierung und Objekterkennung aus Distanzbildern. In: Deutsche Geodätische Kommission bei der Bayerischen Akademie der Wissenschaften: Reihe C, Dissertationen 583. München.
- BÖHNER, J. SELIGE, T. & A. RINGELER (2006): Image segmentation using representativeness analysis and region growing. In: BÖHNER, J. MCCLOY K.R. & J. STROBL (Hg): Göttinger Geographische Abhandlungen 115, S. 29-38.
- BORGEFORS, G. (1986): Distance Transform in Digital Images. In: Computer Vision, Graphics and Image Processing 34, S. 344-371.
- BORGEFORS, G., BRANDTBERG, T & F. WALTER (1999): Forest parameter extraction from airborne sensors. In: International Archives of Photogrammetry and Remote Sensing Vol. 32, Part 3-2W5, S. 151-157.
- BRANDBERG, T. & F. WALTER (1998): Automated delineation of individual tree crowns in high spatial resolution aerial images by multiple-scale analysis. In: Machine Vision and Applications 11, S. 64-73.
- BUSCHMANN, C. (1993): Fernerkundung von Pflanzen; Ausbreitung, Gesundheitszustand und Produktivität. In: Naturwissenschaften 80(10), S. 439-453.
- BUTENUTH, M. (2003): Steuerung bei der Qualitätskontrolle linienhafter Geo-Daten auf Basis von Orthofotos unter Nutzung eines GIS. Hannover.
- BUTENUTH, M., STRAUB, B.-M., HEIPKE, C. & F. WILLRICH (2003): Tree Supported Road Extraction from Aerial Images: Using Global and Local Context Knowledge. In: Lecture Notes in Computer Sciences 2626, S. 162-171.

- CAMPBELL, B. (2002): Introduction to Remote Sensing. London.
- CHARNIAK, E. (1991): Bayesian Networks without Tears. In: AI Magazine WINTER, S. 50-63.
- CHI, Z., YAN H. & T. PHAM (1996): Fuzzy algorithms: with applications to image processing and pattern recognition. Singapore.
- COLLINS, W. (1978): Remote Sensing of Crop Type and Maturity. In: Photogrammetric Engineering and Remote Sensing 44, S. 43-55.
- CONRAD, O. (2006): SAGA. Aufbau, Funktionsweise und Anwendung eines Systems für geowissenschaftliche Analysen. Göttingen.
- CURRAN, P.J. (1985): Principles of Remote Sensing. New York.
- CURTIS B. & C. MANNHEIMER (2005): Tree Atlas of Namibia. Windhoek.
- DIEDERSHAGEN, O., KOCH B. & H. WEINACKER (2003A): Automatic Estimation Of Forest Inventory Parameters Based On Lidar, Multi-Spectral And Fogis Data. Proceedings Optical 3-D Measurement Techniques VI II. S. 259-269.
- DIEDERSHAGEN, O., KOCH B., WEINACKER, H. & C. SCHÜTT (2003B): Combining LIDAR- and GIS data for the extraction of forest inventory parameters. In: Proceedings of the ScandLaser Scientific Workshop on Airborne Laser Scanning of Forests, Umeå, 2-4 Sep, S. 157-165.
- DIEDERSHAGEN, O., KOCH B. & H. WEINACKER (2004): Automatic segmentation and characterisation of forest stand parameters using airborne lidar data, multispectral and FOGIS data. In: THIES, M., KOCH, B., SPIEKER, H. & H. WEINACKER (Hg): International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences XXXVI 8/W2, S. 208-212.
- DRALLE, K. & M. RUDEMO (1997): Automatic estimation of individual tree positions from aerial photos. In: Canadian Journal of Forest Research 27, S. 1728-1736.
- DUDA, R.O. & P.E. HART (1973): Pattern Classification and Scene Analysis. New York.
- DUDA, R.O., HART P.E. & D.G. STORK (2001): Pattern Classification. New York.
- ELACHI, C. & J. VAN ZYL (2006): Introduction to the physics and techniques of Remote Sensing. Hoboken, NJ.
- ERIKSON, M. (2003): Segmentation of individual tree crowns in colour aerial photographs using region growing supported by fuzzy rules. In: Canadian Journal of Forest Research 33, S. 1557-1563.
- ERIKSON, M. (2006): Two preprocessing techniques based on grey level and geometric thickness to improve segmentation results. In: Pattern Recognition Letters 27(3), S. 160-166.
- ERIKSON, M. (2004A): Segmentation and Classification of Individual Tree Crowns in High Spatial Resolution Aerial Images. Uppsala.

- ERIKSON, M. (2004B): Species classification of individually segmented tree crowns in high-resolution aerial images using radiometric and morphologic image measures. In: *Remote Sensing of Environment* 91, S. 469-477.
- ERIKSON, M. & K. OLOFSSON (2005): Comparison of three individual tree crown detection methods. In: *Machine Vision and Applications* 16(4), S. 258-265.
- ERIKSON, M., PERRIN, G., DESCOMBES, X. & J. ZERUBIA (2006): A comparative study of three methods for identifying individual tree crowns in aerial images covering different types of forests. In: *Proc. International Society for Photogrammetry and Remote Sensing (ISPRS)*, Marne La Vallee, France, July 2006.
- GONG, P., SHENG, Y. & G. BIGING (2002): 3D Model Based Tree Measurement from High-Resolution Aerial Imagery. In: *Photogrammetric Engineering & Remote Sensing* 11 (68), S. 1203-1212.
- GONZALES, R.C. & R.E. WOODS (1993): *Digital Image Processing*. Reading.
- GONZALEZ, R.C., WOODS, R.E. & S.L. EDDINS (2004): *Digital Image Processing using MATLAB*. Pearson.
- GOUGEON, F.A. (1995): A crown-following approach to the automatic delineation of individual tree crowns in high resolution aerial images. In: *Canadian Journal of Remote Sensing* 21(3), S. 274-284.
- GOUGEON, F.A. (1998): Automatic individual tree crown delineation using a valley-following algorithm and a rule based system. In: HILL, LECKIE (Hg): *Proceedings of the International Forum on Automated Interpretation of High Spatial Resolution Digital Imagery for Forestry*. Victoria.
- GOUGEON, F.A. & T. MOORE (1988): Individual Tree Classification Using Meis-II Imagery. In: *Geoscience and Remote Sensing Symposium, 1988. IGARSS '88. Remote Sensing: Moving Toward the 21st Century.*, International 2, S. 927.
- HABERÄCKER, P. (1989): *Digitale Bildverarbeitung; Grundlagen und Anwendungen*. München.
- HERRERA, B., KLEINN, C., KOCH, B. & M. DEES (2004): Automatic Classification of Trees outside Forest using an object-driven Approach: an Application in a Costa Rican Landscape. In: *Photogrammetrie, Fernerkundung, Geoinformation* 2.
- HEURICH, M. (2005): Erfassung vertikaler Waldstrukturen mit flugzeuggetragenen Laserscannern. In: *Allg. Forst Z. Waldwirtsch. Umweltvorsorge* 60, S. 242-245.
- HEURICH, M., PERSSON, Å., HOLMGREN J. & E. KENNEL (2004): Detecting and measuring individual trees with laser scanning in mixed mountain forest of central Europe using an algorithm developed for Swedish boreal forest conditions. In: *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 36 - 8/W2.
- HEURICH, M. & E. KENNEL (2007): Projekt Fernerkundung für Waldinventuren erfolgreich. In: *Allg. Forst Z. Waldwirtsch. Umweltvorsorge* 62, S. 70.

- HINZ, S. (2003): Automatische Extraktion urbaner Straßennetze aus Luftbildern. München.
- HOOVER, A., JEAN-BAPTISTE, G., JIANG, X.Y., FLYNN, P.J., BUNKE, H., GOLDFOG, D.B., BOWYER, K. & D.W. EGGERT (1996): An Experimental Comparison of Range Image Segmentation Algorithms. In: IEEE Transactions on Pattern Analysis and Machine Intelligence 18(7), S. 673-689.
- HORN, W. (1984): Expertensysteme: Wissensrepräsentation und Inferenzprozesse. In: KROPATSCH, W. (Hg): Mustererkennung. Berlin.
- HYYPÄ, J. & M. INKINEN (1999): Detecting and estimating attributes for single trees using laser scanner. In: The Photogrammetric Journal of Finland 16, S. 27-42.
- HYYPÄ, J., HYYPÄ H., LITKEY P., YU, X., HAGGRÉN H., RÖNNHOLM P., PYYSALO U., PITKÄNEN J. & M. MALTAMO (2004): Algorithms and methods of airborne laser scanning for forest measurements. In: International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 36 - 8/W2, S. 82-89.
- HYYPÄ J., SCHARDT, M., HAGGRÉN, H., KOCH, B., LOHR, U., PAANANEN, R., SCHERRER, H.U., LUUKKONEN, H., ZIEGLER, M., HYYPÄ, H., PYYSALO, U., FRIEDLÄNDER, H., UUTTERA, J., WAGNER, S., INKINEN, M., WIMMER, A., KUKKO, A., AHOKAS, E. & M. KARJALAINEN (2001): HIGH SCAN: The first European-wide Attempt to Derive Single-Tree information from Laserscanner Data. In: The Photogrammetric Journal of Finland 2(17), S. 58-68.
- HYYPÄ, J., KELLE, O., LEHIKONEN, M. & M. INKINEN (2001): A segmentation-based method to retrieve stem volume estimates from 3-D tree height models produced by laser scanners. In: IEEE Transaction on Geoscience and Remote Sensing 39(5), S. 969-975.
- HYYPÄ, J., MIELONEN, T., HYYPÄ, H., MALTAMO, M., YU, X., HONKAVAARA, E. & H. KAARTINEN (2005): Using individual tree crown approach for forest volume extraction with aerial images and laser point clouds. In: ISPRS WG III/3, III/4, V/3 Workshop "Laser scanning 2005", Enschede, the Netherlands, September 12-14. Enschede.
- JÄHNE, B. (2005): Digitale Bildverarbeitung. Berlin.
- KAMWI, J.M. (2007): The use of high-resolution satellite imagery in forest inventory: a case of Hans Kanyinga Community Forest - Namibia. Stellenbosch.
- KASS, M., WITKIN, A. & D. TERZOPOULOS (1988): Snakes: active contour models. In: International Journal for Computer Vision 1, S. 321-331.
- KÄTSCH, C. (2002): Towards an Automatic Forest Inventory System Based on Remotely Sensed Data and Digital Image Processing Techniques. In: NOLLMANN, I. & R. COLOMB (Hg): Proceedings of 29th ISRE Conference, Argentina, April 2002, S. 207-212.

- KÄTSCH, C. & A. KUNNEKE (2006): Forest Inventory in the Digital Remote Sensing Age: New approaches to measure trees and forest stands in forest inventories. In: South African Forestry Journal 206, S. 43-49.
- KÖNIGSBERGER, K. (1999): Analysis 1. Berlin.
- KÖRBER, C., MÖLLER, D.P.F. & C. KÄTSCH (2005): Multi-Channel Texture Classification applied for Feature Extraction in Forestry. In: VKILZADIAN, H., SHARIF, H., SETH, S. & K. SAYOOD (Hg): Proceedings 2005 International IEEE Conference on Electro InformationTechnology, Lincoln, Nebraska, May 22-25, 2005, S. 433-438.
- KRUMMHEUER, F. (1998): Auswertung von Luftbildern ausgewählter Buchenbestände mit Fouriertechniken. In: Photogrammetrie, Fernerkundung, Geoinformation 3, S. 165-176.
- LARSEN, M. (1999): Individual Tree Top Position Estimation by Template Voting. In: 21. Canadian Symposium on Remote Sensing. Ottawa.
- LARSEN, M. & M. RUDEMO (1997): Estimation of Tree Positions from Aerial Photos. In: LINDEBERG, T. (Hg): Proceedings of the 1997 Swedish Symposium on Image Analysis, S. 130-134.
- LEHMANN, T., OBERSCHELP, W., PELIKA, E. & R. REPGES (1997): Bildverarbeitung für die Medizin; Grundlagen, Modell, Methode, Anwendungen. Berlin.
- LILLESAND, T. M. & R. W. KIEFER (1979): Remote Sensing und Image Interpretation. New York.
- LINDEBERG, T. (1994): Scale Space Theory in Computer Vision. Norwell.
- MARR, D. (1982): Vision: A Computational Investigation into the Human Representation and Processing of Visual Information. New York.
- MATLAB (1984-2007): Image Processing Toolbox. MA.
- MEI C. & S. DURRIEU (2004): Tree Crown Delineation from Digital Elevation Models and High Resolution Imagery. Route des Lucioles.
- MÖLLER, D.P.F., KÖRBER, C. & C. KÄTSCH (2005): Texture Classification applied on Aerial Imagery in Forestry. In: 19th European Conference on Modeling and Simulation, Riga, Latvia, June 1-4, S. 653-657. Nothingham, UK.
- MORS DORF, F., MEIER, E., ALLGÖWER, B. & D. NÜESCH (2004): Cluster-Analyse von LIDAR-Vektordaten zur Erfassung geometrischer Eigenschaften von einzelnen Bäumen. In: Photogrammetrie, Fernerkundung, Geoinformation 4 / 2004, S. 259-268.
- NÆSSET, E. (1997): Estimating timber volume of forest stands using airborne laser scanner data. In: Remote Sensing of Environment 52, S. 49-56.
- NÆSSET, E., GOBAKKEN, T., HOLMGREN, J., HYYPPÄ, H., HYYPPÄ, J., MALTAMO, M., NILSSON, M., OLSSON, H., PERSSON, Å. & U. SÖDERMAN (2004): Laser scanning of

forest resources: the nordic experience. In: Scandinavian Journal of Forest Research 19 (6), S. 482-499.

NIEMANN, H. (1983): Klassifikation von Mustern. Berlin.

OLOFSSON, K. (2002): Detection of Single Trees in Aerial Images using Template Matching. In: ForestSat 2002, Proceedings.

OLOFSSON, K., WALLERMANN, J., HOLGREN, J. & H. OLSSON (2006): Tree species discrimination using Z/I DMC imagery and template matching of single trees. In: Scandinavian Journal of Forest Research 21 (S7), S. 106-110.

PAULUS, D.W.R. & J. HORNEGGER (2003): Applied pattern recognition: algorithms and implementation in C++. Braunschweig.

PERSSON, A. (2001): Extraction of Individual Trees using Laser Radar Data. Linköping.

PERSSON, A., HOLMGREN J. & U. SÖDERMAN (2002): Detecting and Measuring Individual Trees Using an Airborne Laser Scanner. In: Photogrammetric Engineering & Remote Sensing 9 (68), S. 925-932.

PINZ, A. (1989): Final results of the Vision Expert System VES: finding trees in aerial Photographs. In: PINZ A. (Hg): Wissensbasierte Mustererkennung. Wien.

PINZ, A. (1994): Bildverstehen. Wien.

POLIKAR, R. (2006): Ensemble Based Systems in Decision Making. In: IEEE CIRCUITS AND SYSTEMS MAGAZINE 3, S. 21-45.

POLLOCK, R. J. (1996): The Automatic recognition of Individual trees in Aerial Images of Forests based on a Synthetic Tree Crown Image Model. Vancouver.

RICHARDS, J.A. & X. JIA (2006): Remote Sensing Digital Image Analysis. Berlin.

SABINS, F. F. (1987): Remote sensing: Principles and interpretation. New York.

SCHARDT, M., ZIEGLER, M., WIMMER, A., WACK R. & R. HYYPPÄ (2002): Assessment of Forest Parameters by Means of Laser Scanning. In: International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 34, Nr. 3A, S. 302-309.

SOILLE, P. (1998): Morphologische Bildverarbeitung: Grundlagen, Methoden, Anwendungen. Berlin.

STITELER W. & P. HOPKINS (2000): Using genetic algorithms to select tree crown templates for finding trees in digital imagery. In: Proceedings of the ASPRS National Convention, Washington D.C. May 22.-26.

STRAUB, B.-M. (2002): Investigation of the mpeg-7 homogenous Texture Descriptor for the automatic extraction of trees. In: International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 34, Nr. 3A, S. 351-355.

- STRAUB, B.-M. (2004): A multiscale approach for the automatic extraction of tree tops from remote sensing data. In: International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 35, Nr. B3, S. 418-423. Istanbul.
- STRAUB, B.-M. (2003A): Automatische Extraktion von Bäumen aus Fernerkundungsdaten. In: Wiss. Arbeiten, Fachrichtung Vermessungswesen 249. Hannover.
- STRAUB, B.-M. (2003B): A Top-Down-Operator for the Automatic Extraction of Trees - Concept and Performance Evaluation. In: International Archives of Photogrammetry and Remote Sensing and Spatial Information Sciences XXXIV (3/W13), S. 34-39.
- STRAUB, B.-M. & C. HEIPKE (2001): Automatic Extraction of Trees for 3D-City Models from Images and Height Data. In: Gool, B.G. van & A.A. Balkema (Hg): Automatic Extraction of Man-Made Objects from Aerial and Space Images 3, S. 267-277.
- STRAUB, B.-M. & C. HEIPKE (2004): Concepts for Internal and External Evaluation of Automatically Delineated Tree Tops. In: IntArchPhRS XXXVI 8/W2, S. 62-65. Freiburg.
- TÓVÁRI, D. & T. VÖGTLE (2004): Object classification in Laserscanning data. In: International Archives of Photogrammetry and Remote Sensing and Spatial Information Sciences XXXVI 8/W2.
- VERLINDEN, A. & R. LAAMANEN (2006): Long term fire scar monitoring with remote sensing in northern Namibia: Relations between fire frequency, rainfall, land cover fire management und trees. In: Environmental Monitoring and Assessment 112, S. 231-253.
- VÖGTLE, T. & E. STEINLE (2004): Detektion und Modellierung von 3D-Objekten aus flugzeuggetragenen Laserscannerdaten. In: Photogrammetrie, Fernerkundung, Geoinformation, S. 315-322.
- WITTIG, F. (2002): Maschinelles Lernen Bayes'scher Netze für benutzeradaptive Systeme. Saarbrücken.
- ZIEGLER, M., WIMMER, A., SCHARDT, M., ECKMÜLLNER, O. & J. HOFRICHTER (2001): Hochauflösende Gelände- und Oberflächenmodelle aus Laserscannerdaten - ein Anwendungsbeispiel aus der Forstinventur. In: Vermessung und Geoinformation 1, S. 18-25.

## Danksagung

Ich bedanke mich herzlich bei allen, die zum Gelingen dieser Arbeit beigetragen haben, insbesondere (*in order of appearance*) bei:

Meinen Eltern, Brigitte Spies-Bechtel und Manfred Bechtel, die mir mein Studium ermöglicht und geduldig finanziert sowie mich immer ermutigt und tatkräftig unterstützt haben.

Herrn D.P.F. Möller für die Anregung zu dem Thema, die Ermöglichung meines Südafrika-Aufenthaltes und die Zweitbetreuung.

Herrn C. Kätsch für die freundliche Aufnahme in Stellenbosch, die vielen Datensätze (insbesondere die goett-Bilder), anregende Diskussionen und anhaltenden Austausch.

Jonathan Kamwi sowie dem Ministry of Agriculture, Water & Forestry of Namibia für die Organisation der Feldarbeit in Rundu und den Quickbird-Datensatz.

Anton Kunneke für das lamotte-Bild.

Herrn J. Böhner für die freundliche Aufnahme in die AG Geosystemanalyse, die Erstbetreuung sowie zahlreiche konzeptionelle Verbesserungsvorschläge.

Den übrigen Mitgliedern der Arbeitsgruppe, insbesondere Elke, Helge und Olaf für SAGA-Tipps und Korrekturhilfen.

Den übrigen Korrekturlesern, insbesondere Ole, Basti, Jan, Felix, Sascha, Sebi und nochmals meinen Eltern.

Allen die mein Leben auf andere Weise bereichert haben, danke ich auch. Da sich dafür aber möglicherweise noch ein besserer Ort findet (vielleicht schreibe ich ja mal ein Buch), möchte ich das an dieser Stelle nicht namentlich tun.

## Anhang A: Beispielcode

```

%% bbGenerateNumPass
%
% erzeugt numPass für BROWNM0
%
% gültige Gewichtsfunktionen (PREF.BM.WEIGHT):
%
% 'G':          weight = I(round(vypos), round(vxpos)).^pot;
% 'RGB_mean':  weight = sum(Icol(round(vypos), round(vxpos),:).^pot)/3;
% 'p-norm':    weight = (sum(Icol(round(vypos), round(vxpos),:).^pot)).^pot_inv;
% 'min':       weight = (min(Icol(round(vypos), round(vxpos),:)).^pot);

%% preparations
fliters.SE4N = strel('diamond',1);
pot         = PREF.BM.WEIGHT_POT;
pot_inv     = 1/pot;
[ymax, xmax] = size(I);

%% erzeuge numPass
tic;
display(['[bb]>> BM > beginne Segmentierung mit BrownMo, Series ', PREF.SERIES_NAME]);
display(['[bb]>> BM > WEIGHT-FUNCTION: ', PREF.BM.WEIGHT, ', POTENZ: ', num2str(PREF.BM.WEIGHT_POT)]);

% initialisiere Anzeige
if PREF.DISP.SHOW_STEPS
    iptsetpref('imshowBorder', 'tight');
    if isfield(fh, 'brmo') && ishghandle(fh.brmo); close(fh.brmo);end
    if isfield(fh, 'seed') && ishghandle(fh.seed); close(fh.seed);end
    fh.brmo = figure('Position', [0 500 1000 500]); set(fh.brmo, 'Name', 'Brownsche Bewegungen');
    fh.seed = figure('Position', [1401, 700 230 230]); set(fh.brmo, 'Name', 'aktueller Seed'); imshow(Icol,[]); hold
on;
    buff.im = cell (0);
    buff.seed = ones(0,2);
end

%imshow(Icol);

for seednum = 1: anzSeeds
    % setze werte für actualSeed (index, row, col)
    actualSeed.ind = seed.list(seednum, 1);
    [actualSeed.r actualSeed.c] = ind2sub(size(I), actualSeed.ind);

    % anzeige des aktuellen seeds
    if PREF.DISP.SHOW_STEPS; plot(actualSeed.c, actualSeed.r, 'r+', 'LineWidth', 2); end

    % initialize
    actualNumPass = zeros(size(I)); % etwas schneller als gleich in numPass schreiben,
    % double deutlich schneller als uint!

    switch PREF.BM.WEIGHT
        case 'G'
            % neuen run beginnen
            for iruns = 1: PREF.BM.NUM_RUNS
                [ypos, xpos] = ind2sub(size(I), actualSeed.ind);
                % steps beginnen
                for isteps = 1:PREF.BM.NUM_STEPS
                    randvek = randn(2,1);
                    vxpos = xpos + randvek(1);
                    vypos = ypos + randvek(2);

                    % darf nicht aus dem bild laufen ...
                    if vypos < 1; vypos = 1; elseif vypos > ymax; vypos = ymax; end;
                    if vxpos > xmax; vxpos = xmax; elseif vxpos < 1; vxpos = 1; end;

                    % WEIGHT = G pot
                    weight = I(round(vypos), round(vxpos)).^pot;

                    % neue Position bestimmen
                    xpos = xpos + weight * randvek(1);
                    ypos = ypos + weight * randvek(2);

                    % sicherung, damit nicht aus bild läuft
                    if ypos < 1; ypos = 1; elseif ypos > ymax; ypos = ymax; end;
                    if xpos > xmax; xpos = xmax; elseif xpos < 1; xpos = 1; end;

                    iypos = round(ypos); ixpos = round(xpos); % etwas schneller, als 2*runden
                    actualNumPass(iypos, ixpos) = actualNumPass(iypos, ixpos) + 1;
                end % isteps
            end % iruns
        case 'RGB_mean'
            % neuen run beginnen
            for iruns = 1: PREF.BM.NUM_RUNS
                [ypos, xpos] = ind2sub(size(I), actualSeed.ind);
                % steps beginnen
                for isteps = 1:PREF.BM.NUM_STEPS
                    randvek = randn(2,1);
                    vxpos = xpos + randvek(1);
                    vypos = ypos + randvek(2);

                    % darf nicht aus dem bild laufen ...
                    if vypos < 1; vypos = 1; elseif vypos > ymax; vypos = ymax; end;
                    if vxpos > xmax; vxpos = xmax; elseif vxpos < 1; vxpos = 1; end;

                    % WEIGHT = RGB_mean pot
                    weight = sum(Icol(round(vypos), round(vxpos),:).^pot)/3;

                    % neue Position bestimmen
                    xpos = xpos + weight * randvek(1);
                    ypos = ypos + weight * randvek(2);
                end % isteps
            end % iruns
    end
end

```

```

        % sicherung, damit nicht aus bild läuft
        if ypos < 1;      ypos = 1;      elseif ypos > ymax; ypos = ymax;      end;
        if xpos > xmax;  xpos = xmax;  elseif xpos < 1;  xpos = 1;      end;

        iypos = round(ypos); ixpos = round(xpos); % etwas schneller, als 2*runden
        actualNumPass(iypos, ixpos) = actualNumPass(iypos, ixpos) + 1;
    end %isteps
end % iruns
case 'p_norm'
% neuen run beginnen
for iruns = 1: PREF.BM.NUM_RUNS
    [ypos, xpos] = ind2sub(size(I), actualSeed.ind);
    % steps beginnen
    for isteps = 1:PREF.BM.NUM_STEPS
        randvek = randn(2,1);
        vxpos = xpos + randvek(1);
        vypos = ypos + randvek(2);

        % darf nicht aus dem bild laufen ...
        if vypos < 1;      vypos = 1;      elseif vypos > ymax; vypos = ymax;      end;
        if vxpos > xmax;  vxpos = xmax;  elseif vxpos < 1;  vxpos = 1;      end;

        % WEIGHT = 'p-norm'
        weight = (sum(Icol(round(vypos), round(vxpos),:).^pot).^pot_inv);

        % neue Position bestimmen
        xpos = xpos + weight * randvek(1);
        ypos = ypos + weight * randvek(2);

        % sicherung, damit nicht aus bild läuft
        if ypos < 1;      ypos = 1;      elseif ypos > ymax; ypos = ymax;      end;
        if xpos > xmax;  xpos = xmax;  elseif xpos < 1;  xpos = 1;      end;

        iypos = round(ypos); ixpos = round(xpos); % etwas schneller, als 2*runden
        actualNumPass(iypos, ixpos) = actualNumPass(iypos, ixpos) + 1;
    end % isteps
end % iruns
case 'min'
% neuen run beginnen
for iruns = 1: PREF.BM.NUM_RUNS
    [ypos, xpos] = ind2sub(size(I), actualSeed.ind);
    % steps beginnen
    for isteps = 1:PREF.BM.NUM_STEPS
        randvek = randn(2,1);
        vxpos = xpos + randvek(1);
        vypos = ypos + randvek(2);

        % darf nicht aus dem bild laufen ...
        if vypos < 1;      vypos = 1;      elseif vypos > ymax; vypos = ymax;      end;
        if vxpos > xmax;  vxpos = xmax;  elseif vxpos < 1;  vxpos = 1;      end;

        % WEIGHT = 'min'
        weight = (min(Icol(round(vypos), round(vxpos),:).^pot).^pot);

        % neue Position bestimmen
        xpos = xpos + weight * randvek(1);
        ypos = ypos + weight * randvek(2);

        % sicherung, damit nicht aus bild läuft
        if ypos < 1;      ypos = 1;      elseif ypos > ymax; ypos = ymax;      end;
        if xpos > xmax;  xpos = xmax;  elseif xpos < 1;  xpos = 1;      end;

        iypos = round(ypos); ixpos = round(xpos); % etwas schneller, als 2*runden
        actualNumPass(iypos, ixpos) = actualNumPass(iypos, ixpos) + 1;
    end % isteps
end % iruns
end % switch

numPass(:, :, seednum) = actualNumPass;

%% anzeigen
if PREF.DISP.SHOW_STEPS
    buff.seed(end+1,:) = [actualSeed.r actualSeed.c];
    buff.im(end+1) = {actualNumPass(sum(actualNumPass,2) > 0, sum(actualNumPass,1) > 0)};
    if length(buff.seed) > 4;
        plot(buff.seed(1,2), buff.seed(1,1), 'b+', 'LineWidth', 2);
        buff.seed(1,:) = [];
    end
    if length(buff.im) > 4;    buff.im(1) = [];    end

    figure(fh.brmo);
    for i = 1:length(buff.im)
        subplot ('Position', [(i-1)* 0.25  0  0.24  1]);
        imshow(buff.im{i}, []);
        title(['BM für Seed ', int2str(buff.seed(i))]);
    end
    figure(fh.seed);
    plot(buff.seed(:,2), buff.seed(:,1), 'y+', 'LineWidth', 2);
end
display(['[bb]>> BM > numPass für Saatpunkt ' int2str(seednum), ' generiert, benötigte Zeit(s): ', num2str(toc)]);
end %for seednum

display(['[bb]>> SUCCESS: benötigte Zeit für numPass (s): ', num2str(toc)]);

% speicher sparen
if max3(numPass) < 2^16
    numPass = uint16(numPass);
end
clear buff actualNumPass actualSeed xmax ymax pot pot_inv;

% figures schließen
if PREF.DISP.SHOW_STEPS; close(fh.seed), close(fh.brmo);end

```

## Anhang B: Programmausgaben

```
[bb]>> Call > starte Sequenz für RGB auf lamotte 1
[bb]>> Saatgut > generate seed.list with DISTRANS; PREF.SEED.SIGMA: 4
[bb]>> Saatgut > SUCCESS. number of seeds: 150
[bb]>> RGB > füge region hinzu: 150 gröÙe 275
[bb]>> RGB > füge region hinzu: 149 gröÙe 197
[bb]>> RGB > füge region hinzu: 148 gröÙe 184
[bb]>> RGB > füge region hinzu: 147 gröÙe 232
[bb]>> RGB > füge region hinzu: 146 gröÙe 224
[bb]>> RGB > füge region hinzu: 145 gröÙe 281
[bb]>> RGB > füge region hinzu: 144 gröÙe 197

...

[bb]>> RGB > füge region hinzu: 7 gröÙe 93
[bb]>> RGB > füge region hinzu: 6 gröÙe 543
[bb]>> RGB > füge region hinzu: 5 gröÙe 205
[bb]>> RGB > füge region hinzu: 4 gröÙe 109
[bb]>> RGB > füge region hinzu: 3 gröÙe 199
[bb]>> RGB > füge region hinzu: 2 gröÙe 120
[bb]>> RGB > füge region hinzu: 1 gröÙe 110
[bb]>> RGB > Elapsed time is 10.989688 seconds.
[bb]>> FeatureExtract > beginne Merkmalsextraktion
[bb]>> FeatureExtract > SUCCESS; benötigte Zeit: 2.5801 s
[bb]>> HypoSelect > wähle beste Hypothesen, Modell: all
[bb]>> HypoSelect > schreibe ausgewählte Hypothesen in extract
[bb]>> HypoSelect > 150 von 150 Hypothesen gewählt
[bb]>> Document > Dokumentiere Durchlauf in Bildern...
[bb]>> Document > schreibe Regionen als Bilder
[bb]>> Call > sichere Ergebnisse als mat-files
[bb]>> Call > HABE FERTIG! benötigte Gesamtzeit in Sekunden: 18 (entspricht0.0048742 h)

-----

[bb]>> FeatureExtract > beginne Merkmalsextraktion
[bb]>> FeatureExtract > SUCCESS; benötigte Zeit: 2.2503
[bb]>> HypoSelect > wähle beste Hypothesen, Modell: DOA & Bright
[bb]>> HypoSelect > 111 von 135 Hypothesen gewählt
[bb]>> HypoSelect > schreibe ausgewählte Hypothesen in extract
[bb]>> ValiData > berechne Overlaps: Anzahl extrahierte Regionen: 111, Anzahl ReferenzRegionen: 125
[bb]>> ValiData > SUCCESS; benötigte Zeit: 5.106
[bb]>> ValiData > generate matchList
[bb]>> ValiData > Anzahl der Übereinstimmungen 110
[bb]>> ValiData > bereinige extrahierte Bäume, die mit mehreren Referenzbäumen übereinstimmen
[bb]>> ValiData > Anzahl der Übereinstimmungen: 110
[bb]>> ValiData > bereinige mehrfach extrahierte Bäume
[bb]>> ValiData > Anzahl der Übereinstimmungen: 109
[bb]>> ValiData > berechne Parameter für übereinstimmende Segmente
[bb]>> ValiData > Vollständigkeit: 0.872
[bb]>> ValiData > Korrektheit: 0.98198
[bb]>> ValiData > mean(delta r): 2.3187
[bb]>> ValiData > % Referenz-Fläche: 0.71692
[bb]>> ValiData > % " konvex ergänzt.: 0.79338
[bb]>> SUCCESS
```

## **Anhang C: Erklärung**

Hiermit erkläre ich, dass ich die vorliegende Arbeit ohne unzulässige Hilfe Dritter und nur unter Verwendung der angegebenen Hilfsmittel angefertigt habe. Die aus anderen Quellen oder indirekt übernommenen Daten und Gedanken sind unter Angabe der Quelle gekennzeichnet.

Diese Arbeit wurde bisher weder im In- noch im Ausland in gleicher oder ähnlicher Form in anderen Prüfungsverfahren vorgelegt. Mit einer zukünftigen Ausleihe der Arbeit bin ich einverstanden.

Hamburg, den 18. Dezember 2007,

---

Benjamin Bechtel