

**Ein Hochdurchsatz-Screeningsystem zur
Objekterkennung in Mikroskop-Farbbildern im
Rahmen der Analyse pflanzlicher Pathogenresistenz**

Dissertation

zur Erlangung des akademischen Grades

Doktoringenieur

(Dr.-Ing.)

von

Dipl.-Ing. Alexander Ihlow

geboren am 08.04.1976 in Magdeburg

genehmigt durch

die Fakultät für Elektrotechnik und Informationstechnik

der Otto-von-Guericke-Universität Magdeburg

Gutachter:

Prof. Dr.-Ing. habil. Bernd Michaelis

Prof. Dr.-Ing. Klaus Tönnies

Dr.-Ing. Udo Seiffert

Promotionskolloquium am 23. Juni 2006

Kurzfassung

Die vorliegende Dissertation beschreibt die ingenieurtechnische Entwicklung eines Hochdurchsatz-Screeningsystems zur bildbasierten Analyse von Mikroskoppräparaten als Automatisierung bis dato manuell durchgeführter Auswertungen. Technische Voraussetzung ist ein computergesteuertes motorisiertes Mikroskop mit Digitalkamera, welches hochaufgelöste und rauscharme Farbbilder des zu analysierenden Probenmaterials bereitstellt. In den Bildern finden sich grünbläulich gefärbte Zellen der oberen Epidermis von Gerstenblättern, welche mittels labortechnischer Präparation einer genetischen Transformation unterzogen wurden. Biologischer Hintergrund der vorgelegten Arbeit ist die Erlangung von Erkenntnissen zur Rolle bestimmter Gene der Pflanze bezüglich einer Resistenz gegen Mehltau. Dazu ist in den zu untersuchenden Zellen die Existenz des Mehltaupilzes in Form so genannter Haustorien zu detektieren und die Zelle den Klassen „*infiziert*“ (ein Haustorium oder mehrere Haustorien vorhanden) bzw. „*nicht infiziert*“ (kein Haustorium vorhanden) zuzuordnen.

Auf Basis des Bunttons werden die gefärbten, genetisch transformierten Zellen mittels des Canny-Kantendetektors zunächst grob segmentiert (Finden relevanter Bildausschnitte), anschließend erfolgt eine Verfeinerung der Segmentierung im betrachteten Bildausschnitt unter Zuhilfenahme eines aktiven Konturmodells, welches sich anhand der mittels eines adaptiven Farbraums bestimmten Bildmerkmale um die zu untersuchende Zelle legt.

Die anschließende Kontrastverstärkung mittels morphologischer Zylinderhuttransformationen ermöglicht eine robuste Segmentierung potenzieller Haustorienregionen innerhalb der Zelle. Zur Zurückweisung von Störungen, welche aufgrund des schwachen Farbkontrasts zwischen Zellgewebe und Haustorium sowie potenziell vorhandener Störobjekte innerhalb der Zelle auftreten, wird anschließend eine objektbasierte Klassifikation durchgeführt, die das segmentierte Objekt den Klassen „*Haustorium*“ bzw. „*kein Haustorium*“ zuweist. Mittels der verwendeten Formmerkmale wird eine Korrekturklassifikationsrate von mehr als 90 % erreicht, was eine hinreichend exakte Erstellung einer Statistik des haustoriellen Anfälligkeitsindex der Zellen und damit die Beantwortung der biologischen Fragestellung ermöglicht.

Die besondere Herausforderung der Arbeit bestand darin, sowohl allgemeine menschliche Erkennungs- und Kognitionsleistung als auch biologisches Expertenwissen in das System zu integrieren und dabei eine Flexibilität zu gewährleisten, die eine Adaption an ähnliche Fragestellungen ermöglicht.

Zusammen mit der implementierten vollautomatischen Ablaufsteuerung für die mikroskopseitige Bildaufnahme ergibt dies ein Laborassistenzsystem, welches dem Experimentator sehr zeitaufwändige Routineaufgaben abnimmt und ihm ein objektives Werkzeug zur Verfügung stellt, das u.a. durch Anbindung an eine Datenbank eine Dokumentation und Reproduzierbarkeit der Ergebnisse ermöglicht.

Diese Dissertation ist in elektronischer Form bei der Universitätsbibliothek Magdeburg unter der Internetadresse <http://diglib.uni-magdeburg.de> erhältlich.

Abstract

This thesis describes the development of a high-throughput screening system for image-based analysis of LM-micrographs as an automation of so far manual screening tasks. Hardware basis is a computer-operated, motorized microscope, equipped with a digital camera, which produces high-resolution low-noise color images of the specimens. These images contain greenish-blue dyed cells of the upper epidermis of barley leaves which were genetically transformed by biotechnological preparation. The biological background is the desire to understand the influence of certain plant genes on a resistance against powdery mildew. Therefore, the powdery mildew fungus has to be detected in terms of a so-called haustorium (which actually acts as a sucker and feeds the fungus by leaching the cell). The considered cells have to be classified into „*infected*“ (one haustorium or several haustoria present) or „*not infected*“ (no haustorium present).

Based on the hue feature, the dyed genetically transformed cells are first segmented in a coarse manner (finding regions of interest). Subsequently, the cell segmentation is improved by using an active contour model in connection with an adaptive color space. As an intermediate result, each dyed cell is accurately segmented according to its cell boundary.

Afterwards, potential haustoria regions are segmented inside these cells by a contrast enhancement via morphological top-hats. Due to the low contrast between haustoria and the remaining cell tissue the segmentation also catches other objects, such as the cell nucleus or several discolorations. Therefore, a succeeding object classification based on shape features assigns the segmented objects to the classes „*haustorium*“ or „*no haustorium*“. With this processing pipeline a classification accuracy of more than 90 % is obtained which provides sufficiently exact results to build a statistical analysis of the haustoria susceptibility index thereon and to finally answer the relevant biological questions.

The special challenge of this work was both to build a robot system with special human recognition abilities and to ensure a flexibility which offers an adaptation of this solution to similar problems.

In connection with the fully automatic microscope control, i.e., the automatic image acquisition, the developed solution helps the biological experimenter with very time-consuming routine jobs and provides a much less subjective analysis tool which enables a high-quality documentation and traceability of the results via database access.

This thesis is available online at the library of the Magdeburg University via the URL <http://diglib.uni-magdeburg.de>.

Vorwort

Die vorliegende Arbeit entstand während meiner Tätigkeit am Leibniz-Institut für Pflanzengenetik und Kulturpflanzenforschung (IPK) Gatersleben im Rahmen des vom BMBF geförderten Projekts „Erkennung und Analyse räumlich-zeitlicher Entwicklungsmuster“.

Für die Betreuung der Arbeit und die freundliche Unterstützung bedanke ich mich ganz herzlich bei Herrn Prof. Dr.-Ing. habil. Bernd Michaelis vom Institut für Elektronik, Signalverarbeitung und Kommunikationstechnik (IESK) der Otto-von-Guericke-Universität Magdeburg und Herrn Prof. Dr.-Ing. Klaus Tönnies vom Institut für Simulation und Graphik (ISG) der Otto-von-Guericke-Universität Magdeburg für die Bereitschaft, als Gutachter zur Verfügung zu stehen.

Ein ganz besonderes Dankeschön geht an dieser Stelle an Herrn Dr. Udo Seiffert, Leiter der Arbeitsgruppe Mustererkennung am IPK Gatersleben. Er brachte mir das entsprechende Vertrauen entgegen und ermöglichte es mir, dieses Thema zu bearbeiten. Dabei unterstützte und ermunterte er mich fach- und sachkundig in jeder Hinsicht. Er und meine Kollegen aus der Arbeitsgruppe Mustererkennung, Cornelia Brüß, Tobias Czauderna, Dr. Marc Strickert und Mathias Rossa sowie die Kollegen der Arbeitsgruppe Netzwerkanalyse unter Leitung von Dr. Falk Schreiber sorgten für eine angenehme und konstruktive Arbeitsatmosphäre und beflügelten so manchen Gedankengang. Unterstützt wurde dies auch durch die hervorragenden Arbeitsbedingungen und technischen Voraussetzungen am IPK.

Mein Dank gilt weiterhin den Herren Ralf Tautenhahn, Christian Schulze, Tobias Senst, Martin Kalev und Burkhard Steuernagel, die maßgeblich an der Entwicklung des Screeningsystems beteiligt waren. Ebenso danke ich Herrn Jörg Steinbach von der Firma Carl Zeiss, der mir bei technischen Fragen und Problemen zur Mikroskop-Hard- und -Software eine große Hilfe war.

Frau Simone Frintrop stellte dankenswerterweise die Grafiken zur Detektion der *Foci of Attention* mittels des Aufmerksamkeitssystems VOCUS bereit.

Mit der Biologie ist es für einen Ingenieur manchmal nicht so einfach. Zum Glück standen mir Stefanie Lück, Manuela Knauff, Gabriele Brantin, Grit Zimmermann, Daniela Nowara und natürlich Dr. Patrick Schweizer, Leiter der Arbeitsgruppe Transkriptomanalyse am IPK, jederzeit bei allen erdenklichen Fragen und Problemen bezüglich des biologischen Hintergrundes zur Seite und sorgten dafür, dass immer frisches Probenmaterial zur Verfügung stand.

Die Vertiefung weiterer Aspekte zur Pflanzen- und „Hamsterbiologie“ verdanke ich fachübergreifenden abendlichen Diskussionen in freier Natur, bei kulturellen Höhepunkten oder/und bei einem Glas Rotwein mit Sabine Jakob.

Nicht vergessen möchte ich an dieser Stelle die schöne Zeit „davor“. Mein herzlicher Dank gilt allen ehemaligen Kollegen des Fachgebiets Nachrichtentechnik der Technischen Universität Ilmenau, die mich während meiner „ersten Schritte“ im wissenschaftlichen Umfeld begleiteten: Prof. Dieter Kreß, Brigitte Epler, Dr. Ralf Irmer, Dr. Martin Huneck, Dr. Martin Janke, Dr. Wolfgang Reimer, Wolfgang Erdtmann, Dr. Wolfgang Röhr, Dr. Mike Wolf, Dr. Thomas Kuhwald, Dr. Enrico Jugl, Prof. Martin Haardt, Giovanni Del Galdo, Marko Hennhöfer, Marko Milojevic sowie den Mitarbeitern des Fachgebiets Elektronische Messtechnik, insbesondere Prof. Reiner Thomä und Gerd Sommerkorn sowie allen Studenten, die ich mit Telekommunikationstechnik, Informationstechnik und Systemtheorie begeistern konnte.

Danke an meine Freunde, ganz besonders an Claudia Heyer, den strahlenden Sonnenschein, der mich an den dunklen und hellen Tagen der letzten Jahre erleuchtete.

Ich danke meiner Familie für die liebevolle Unterstützung in allen Lebenslagen, zuletzt ganz besonders für den Rückhalt in der „heißen Phase“ der Entstehung dieser Arbeit. Insbesondere danke ich meiner Mutter für die akribische Durchsicht des Manuskripts in Bezug auf die Tücken der deutschen Grammatik und Orthograph(f)ie.

Ewiger Dank sei dem SCHÖPFER der Wissenschaft.

Für meine Eltern, Hans-Georg und Annerose Ihlow

Inhaltsverzeichnis

1	Einleitung	1
1.1	Biologischer Hintergrund und Motivation	1
1.2	Ziele und Herausforderungen	3
1.3	Technische Voraussetzungen und Vorüberlegungen	4
1.3.1	Mikroskophardware und Steuersoftware	4
1.3.2	Vorüberlegungen bezüglich des Zeitaufwandes	4
1.3.3	Objektivauswahl und generelle Strategie der Bildaufnahme	4
1.4	Weiteres Vorgehen	9
I	Farbbildverarbeitung und Segmentierung	11
2	Grundlagen der Farbwahrnehmung, Farbräume	13
2.1	Farbwahrnehmung	13
2.1.1	Metamerie und RGB-Spektralwertkurven	15
2.1.2	Normspektralwertkurven nach CIE	16
2.1.3	Normfarbtafel	16
2.2	Farbräume	16
2.2.1	Linearer <i>RGB</i> -Farbraum und gamma-korrigierter <i>R'G'B'</i> -Farbraum	16
2.2.2	Lineare Transformationen von <i>R'G'B'</i>	18
2.2.3	Wahrnehmungsorientierte Farbräume	19
2.2.4	Wahrnehmungsgetreue Farbräume mit visueller Gleichabständigkeit	19
3	Bildsegmentierung	21
3.1	Ein adaptiver Farbraum für die Zellsegmentierung	22
3.1.1	Parametrisches Datenmodell	22
3.1.2	Validierung des parametrischen Modells	24
3.1.3	Parameterschätzung, Expectation Maximization (EM)-Algorithmus	25
3.1.4	Herleitung des adaptiven Farbraumes	27
3.1.5	Leistungsfähigkeit und Diskussion des adaptiven Farbraumes	31
3.2	Kantenbasierte Segmentierung – Kantendetektion	34
3.2.1	Skalenraum und Multiskalen-Bildverarbeitung	34
3.2.2	Bildableitungen und Kantenfilter	36
3.2.3	Zellfindung mittels Canny-Kantendetektor im Buntton	37
3.3	Verbinden von Kanten mit aktiven Konturmodellen	38
3.3.1	Grundlagen aktiver Konturmodelle (Snakes)	41
3.3.2	Klassische Lösung im Ortsbereich	41
3.3.3	Schnelle Lösung im Spektralbereich mittels FFT	43
3.3.4	Lösung durch zyklische Faltung	44
3.3.5	Diskussion und Vergleich der Lösungsschemata	45

3.3.6	Anwendung zur genauen Bestimmung der Zellberandung	46
3.4	Segmentierung der Haustorien	47
3.4.1	Selektion der Fokusebene	48
3.4.2	Farbbasierte Segmentierung mittels EM-Clusterung	48
3.4.3	Segmentierung mittels morphologischer Kontrastverstärkung	52
3.5	Zusammenfassung der Ergebnisse	57
II	Merkmalsextraktion und Klassifikation	59
4	Formanalyse und Formmerkmale	61
4.1	Einfache, nicht informationserhaltende Formmerkmale	62
4.2	Regionenbasierte Formanalyse mittels Momentinvarianzen	63
4.2.1	Zentralmomente	63
4.2.2	Momenten-Tensor	63
4.2.3	Normalisierte Zentralmomente	64
4.2.4	Hu-Momente	64
4.2.5	Momentinvarianzen nach Flusser	64
4.3	Konturbasierte Formanalyse mittels Fourierdeskriptoren	65
4.3.1	Signal-Rausch-Verhältnis (SNR) von Konturen	67
4.3.2	Invarianzeigenschaften der Spektraldarstellung	68
4.3.3	Bedeutung der Betragsinformation des Spektrums	69
4.3.4	Spektralinvanzen nach Granlund	70
4.3.5	Oberwellenhaltigkeit, „Klirrfaktor“	70
4.3.6	Merkmale aus den Differentialen der Kontur	71
4.3.7	Kontur-Skalenraum	72
4.4	Weiteres Vorgehen	73
5	Merkmalsauswahl und Klassifikation	75
5.1	Anforderungen an Merkmale und Trainingsdaten	75
5.1.1	Datensatz zur Beschreibung der Haustorien	76
5.1.2	Aufteilung des Datensatzes in Trainings- und Testmenge	76
5.2	Merkmalsnormalisierung	77
5.3	Merkmalsnormalisierung	77
5.4	Transformation der Merkmale in den Eingaberaum	78
5.4.1	Bhattacharyya-Distanz	79
5.4.2	Hauptachsentransformation (PCA)	80
5.4.3	Lineare Diskriminanzanalyse (LDA)	80
5.5	Untersuchung des Merkmalsraums	81
5.5.1	Trennschärfe einzelner Merkmale	81
5.5.2	Trennschärfe von Merkmalskombinationen	84
5.5.3	Streudiagramme der Merkmale	84
5.5.4	Schlussfolgerungen	85
5.6	Klassifikationsergebnisse	86
5.6.1	Confusion Matrix	86
5.6.2	Ergebnisse	87
5.7	Fazit	88

III	Praktische Realisierung des Screeningsystems	93
6	Bildaufnahmesteuerung und Datenverwaltung	95
6.1	Auswertung der Vorschaubilder	96
6.2	Festlegung der Aufnahmepositionen für die Detailbilder	98
6.2.1	Minimierung der Anzahl der nötigen Aufnahmepositionen	99
6.2.2	Minimierung der vom X-Y-Tisch zurückzulegenden Wegstrecke	100
6.3	Austausch der Steuer- und Nutzinformation	100
6.4	Informationsspeicherung und -zugriff	101
7	Zusammenfassung	103
	Anhang	107
A	Farbräume – Eine Zusammenstellung	107
A.1	Lineare Transformationen von $R'G'B'$	107
A.1.1	$Y'UV$ -Farbraum (PAL-Fernsehnorm)	107
A.1.2	$Y'IQ$ -Farbraum (NTSC-Fernsehnorm)	107
A.1.3	$Y'D_bD_r$ -Farbraum (SECAM-Fernsehnorm)	107
A.1.4	$Y'P_bP_r$ -Farbraum (Komponentensignal bei analogem Video)	108
A.1.5	$Y'C_bC_r$ -Farbraum (Komponentensignal bei digitalem Video)	108
A.1.6	$I_1I_2I_3$ -Ohtas Farbraum	108
A.2	Wahrnehmungsorientierte Farbräume – HSV	108
A.3	Wahrnehmungsgetreue Farbräume	109
A.3.1	CIE $L^*a^*b^*$	109
A.3.2	CIE $L^*u^*v^*$	109
B	Zellfindung mit anderen Werkzeugen	111
B.1	VOCUS – Selbständiges Finden interessanter Regionen	111
B.2	SIOX – Selbständiges Freistellen von Objekten	113
C	Einige Kantendetektoren erster Ordnung	115
D	Laufzeitverhalten aktiver Konturmodelle	117
D.1	Lösungsschemata für aktive Konturmodelle	117
D.2	Tatsächliches Laufzeitverhalten	117
E	Reelle vs. komplexe Darstellung Kartesischer Fourierdeskriptoren	121
F	Streudiagramme der Merkmale	123
G	Entity-Relationship-Diagramm der Datenbank	127
	Literaturverzeichnis	129

Abbildungsverzeichnis

1.1	Lichtmikroskopische Aufnahmen gefärbter Epidermiszellen von Gerstenblättern	2
1.2	Schema der Entwicklung des Mehлтаupilzes auf einem Gerstenblatt	2
1.3	Computergesteuertes Mikroskop mit X-Y-Tisch und Digitalkamera	5
1.4	X-Y-Tisch mit Objektträgern	5
1.5	Kachelung der Vorschauaufnahmen bei kleiner Vergrößerung (5×-Objektiv)	7
1.6	Vorschauaufnahmen transformierter Zellen (5×-Objektiv)	8
1.7	Detailaufnahmen transformierter Zellen (10×-Objektiv)	8
2.1	Für den Menschen sichtbarer Teil des elektromagnetischen Spektrums	14
2.2	Normierte spektrale Empfindlichkeiten der drei Zapfentypen des menschlichen Auges	17
2.3	RGB-Spektralwertkurven	17
2.4	CIE-1931-Normspektralwertkurven	17
2.5	CIE-Normfarbtafel	17
3.1	Ausschnitt eines Beispielbildes und repräsentative Ausschnitte der Klassen „Hintergrund“ und „transformierte Zelle“	22
3.2	Farbcluster der Klassen „Hintergrund“ und „transformierte Zelle“ im RGB-Farbraum	23
3.3	Validierung der Gaußverteilung durch Analyse der Randverteilungen	25
3.4	Validierung der Gaußverteilung mittels Quantil-Quantil-Plot	26
3.5	Separierung der Klassen „Hintergrund“ und „transformierte Zelle“: quadratische Trennfläche in RGB, lineare Trennfläche in RGB, Schwellwert im Buntton	32
3.6	Verteilungsdichtefunktionen: lotrechte Distanz zur quadratischen Entscheidungsfläche, Euklidische Distanz zur linearen Entscheidungsfläche, Verteilungsdichte im Buntton	32
3.7	Abstandsmaße in Pseudofarbdarstellung	32
3.8	Resultierende Binärbilder nach Schwellwerthaltung	32
3.9	Einschaliges und zweischaliges Hyperboloid	33
3.10	Algebraische Distanz zur quadratischen Entscheidungsfläche	33
3.11	Ausschnitte aus typischen Mikroskopbildern	39
3.12	Pseudofarbdarstellung des Bunttons dieser Bildausschnitte	39
3.13	Betragsquadrat des Bildgradienten bei der Skala $\sigma = 5$	39
3.14	Betragsquadrat des Bildgradienten bei der Skala $\sigma = 12$	39
3.15	Detektierte Kanten bei der Skala $\sigma = 5$	40
3.16	Detektierte Kanten bei der Skala $\sigma = 12$	40
3.17	Zum Vergleich: Versuch einer mehrdimensionalen Kantendetektion in RGB	40
3.18	Steifigkeitsspektrum $\hat{\mathbf{k}}$ in logarithmischer Darstellung	46
3.19	Ausschnitt des Stabilisierungskerns $\text{IFFT}\{1/(1 + \tau \hat{\mathbf{k}})\}$	46
3.20	Einfluss des Farbraumes auf das Segmentierungsergebnis aktiver Konturmodelle	47

3.21	Haustoriensegmentierung durch Farbclustering: Segmentierungsergebnisse des <i>klassischen</i> EM-Algorithmus	49
3.22	Haustoriensegmentierung durch Farbclustering: Segmentierungsergebnisse des <i>modifizierten</i> EM-Algorithmus	49
3.23	Validierung der Gaußverteilung durch Analyse der Randverteilungen	51
3.24	Validierung der Gaußverteilung mittels Quantil-Quantil-Plot	52
3.25	Haustoriensegmentierung mittels morphologischer Kontrastverstärkung und anschließender Hysterese-Schwellwerthaltung der Farbsättigung	55
3.26	Veranschaulichung der morphologischen Kontrastverstärkung	56
3.27	Veranschaulichung der Hysterese-Schwellwerthaltung	56
4.1	4er und 8er Pixelnachbarschaft	62
5.1	Einfluss einer nichtlinearen Transformation auf die Verteilung im Merkmalsraum.	78
5.2	Trennschärfe der einzelnen Merkmale (Bhattacharyya-Distanz)	83
5.3	Projektion des Merkmalsraums in 2D mittels LDA	86
5.4	Bhattacharyya-Distanz bei wachsendem Eingaberaum	86
5.5	Lage der Ergebnisse im ROC-Diagramm	89
5.6	ROC-Diagramme der Objekt- und Zellklassifikation auf dem Trainingsdatensatz .	89
5.7	Korrektklassifikationsrate auf dem Trainingsdatensatz	89
5.8	Streuung der ROC-Kurve	90
5.9	ROC-Diagramme der Objekt- und Zellklassifikation bei Kreuzvalidierung	90
5.10	Mittlere Korrektklassifikationsrate bei Kreuzvalidierung	90
5.11	Bildverarbeitungskette zur Analyse der Mikroskopbilder	91
6.1	Zeitlich-örtlicher Ablauf des Bildaufnahmeprozesses	96
6.2	Segmentierung der Vorschaubilder mittels zwei Schwellwerten im Buntton	97
6.3	Segmentierte Vorschaubilder	98
6.4	Voraussetzungen zum gemeinsamen Fotografieren mehrerer Objekte	100
6.5	Grafische Nutzerschnittstelle (<i>Graphical User Interface, GUI</i>)	101
B.1	Mittels VOCUS detektierte interessante Regionen (<i>Foci of Attention, FOA</i>)	112
B.2	Interaktive Segmentierung transformierter Zellen mittels SIOX	113
D.1	Vergleich der Laufzeit aktiver Konturmodelle bei verschiedenen Lösungsschemata	119
D.2	Detailausschnitt von Abbildung D.1	119
D.3	Normierte Variante von Abbildung D.1	119
D.4	Detailausschnitt von Abbildung D.3	119
F.1	Streudiagramme der einfachen Merkmale.	123
F.2	Streudiagramme der Flusser-Momentinvarianzen	124
F.3	Streudiagramme der Spektralmerkmale	125

Tabellenverzeichnis

1.1	Unterstützte Bildauflösungen der Kamera und benötigte Zeiten für den Bildeinzug	4
1.2	Abschätzung des Zeitbedarfs für die Bildaufnahme eines Objektträgers	7
1.3	Speicherplatzbedarf der anfallenden Rohbilddaten in Megabyte (MB)	9
3.1	Segmentierungsergebnisse des modifizierten EM-Algorithmus bei Variation der minimal geforderten Entscheidungssicherheit R_{min}	51
4.1	Invarianzeigenschaften von Fourierdeskriptoren	68
4.2	Signifikante und insignifikante Beträge der Granlund-Invarianzen bei m -facher Rotationssymmetrie des Objekts	71
5.1	Merkmale zur Haustorienerkennung und angewendete nichtlineare Transformationen zur Entzerrung der Verteilung im Merkmalsraum	82
5.2	Bhattacharyya-Distanz verschiedener Merkmalskombinationen auf der Achse maximaler Diskriminanz (1. Diskriminanzachse)	85
6.1	Parameter zur Klassifikation der segmentierten Zellen	98
B.1	VOCUS' Merkmalswichtung nach dem Training auf die gefärbten Zellen	112
D.1	Drei verschiedene Lösungsschemata für aktive Konturmodelle	117

Symbolverzeichnis

Grundlegende Vereinbarungen zur Schreibweise von Formelzeichen

a, A	Skalare sind <i>kursiv</i> (in Klein- oder Großbuchstaben) gesetzt.
\mathbf{a}	Vektoren sind in fetten Kleinbuchstaben gesetzt.
\mathbf{A}	Matrizen sind in fetten Großbuchstaben gesetzt.

Mathematische Operationen

$\arg\{\cdot\}$	Argument (Phasenwinkel) einer komplexen Zahl
$\text{DFT}\{\cdot\}$	diskrete Fouriertransformation
$E[\cdot]$	Erwartungswert
$\text{FFT}\{\cdot\}$	schnelle Fouriertransformation
$\text{IFFT}\{\cdot\}$	inverse schnelle Fouriertransformation
$\text{Im}\{\cdot\}$	Imaginärteil einer komplexen Zahl
$\text{Re}\{\cdot\}$	Realteil einer komplexen Zahl
$\circ\bullet$	Fouriertransformation vom Orts- in den Spektralbereich
$\bullet\circ$	inverse Fouriertransformation vom Spektral- in den Ortsbereich
$*$	Faltung
\circledast	zyklische Faltung
\odot	elementweise Multiplikation (Hadamard-Produkt)
∇	Gradient-Operator
Δ	Laplace-Operator
$\lceil \cdot \rceil$	Aufrunden auf den nächsten ganzzahligen Wert (<i>ceiling</i>)
$\binom{n}{k}$	Binomialkoeffizient
$\delta_B(\cdot)$	morphologische Dilation
$\epsilon_B(\cdot)$	morphologische Erosion
$\gamma_B(\cdot)$	morphologisches Öffnen
$\phi_B(\cdot)$	morphologisches Schließen
$WTH_B(\cdot)$	morphologische weiße Zylinderhutoperation
$BTH_B(\cdot)$	morphologische schwarze Zylinderhutoperation
$\kappa(\cdot)$	morphologische Kontrastverstärkung

Matrixalgebra und spezielle Matrizen

\mathbf{A}^T	Transponierte der Matrix \mathbf{A}
\mathbf{A}^{-1}	Inverse der Matrix \mathbf{A}
$ \mathbf{A} $	Determinante der Matrix \mathbf{A}
$\mathbf{A} = \mathbf{V} \mathbf{\Lambda} \mathbf{V}^{-1}$	Eigenwertzerlegung der Matrix \mathbf{A}
\mathbf{C}	Zyklische (zirkulante) Matrix
\mathbf{I}	Einheitsmatrix
$\mathbf{\Lambda} = \text{diag}\{\lambda_1 \dots \lambda_i \dots\}$	Eigenwerte, angeordnet in einer Diagonalmatrix
\mathbf{V}	Matrix der zu $\mathbf{\Lambda}$ gehörenden Eigenvektoren
\mathbf{W}	Fouriermatrix

Formelzeichen

$2a$	$\in \mathbb{R}$	Länge der langen Hauptachse einer Ellipse
$2b$	$\in \mathbb{R}$	Länge der kurzen Hauptachse einer Ellipse
A	$\in \mathbb{R}$	Objektfläche
\tilde{A}	$\in \mathbb{R}$	nichtlineare Transformation von A (vgl. Tabelle 5.1, S. 82)
AC	$\in \mathbb{R}$	Korrektklassifikationsrate (<i>accuracy</i>)
b_n	$\in \mathbb{C}$	Granlund-Invarianzen ($n = 2, 3, \dots$)
\tilde{b}_n	$\in \mathbb{R}$	nichtlineare Transformation von b_n (vgl. Tabelle 5.1, S. 82)
c_{pq}	$\in \mathbb{C}$	komplexes Moment der Ordnung ($p + q$)
d	$\in \mathbb{N}$	Anzahl der Dimensionen (des Merkmalsraums)
d'	$\in \mathbb{N}$	Anzahl der Dimensionen des Eingaberaums
d_a	$\in \mathbb{R}$	algebraische Distanz zu einer quadratischen Entscheidungsfläche
d_l	$\in \mathbb{R}$	Euklidische Distanz zu einer linearen Entscheidungsfläche
d_p	$\in \mathbb{R}$	lotrechte Distanz zu einer quadratischen Entscheidungsfläche
E	$\in \mathbb{R}$	Verbiegungsenergie
$\mathbf{F}_{\text{ext}}(\mathbf{v})$	$\in \mathbb{R}^{N \times 2}$	Externe Kräfte, die auf die aktive Kontur wirken
$\hat{\mathbf{F}}_{\text{ext}}(\mathbf{v})$	$\in \mathbb{R}^{N \times 2}$	Fouriertransformierte von $\mathbf{F}_{\text{ext}}(\mathbf{v})$
$g_\sigma(\mathbf{x}; \sigma)$	$\mathbb{R}^d \times \mathbb{R} \rightarrow \mathbb{R}$	multivariate Gaußfunktion
h	$\in \mathbb{R}$	Schrittweite der Ortsdiskretisierung (akt. Konturmodelle)
$I(x, y) \rightsquigarrow F(u, v)$		Bildgrauwertfunktion und zugehörige Fouriertransformierte
$I_\sigma(x, y; \sigma)$		Skalenraumdarstellung eines Bildes
i, j, k	$\in \mathbb{N}$	Klassenbezeichnungen
j		imaginäre Einheit
K	$\in \mathbb{N}$	Anzahl Klassen
k	$\in \mathbb{R}$	Klirrfaktor
\mathbf{K}	$\in \mathbb{R}^{N \times N}$	Steifigkeitsmatrix eines aktiven Konturmodells
\mathbf{k}	$\in \mathbb{R}^{N \times 1}$	Steifigkeitskern eines aktiven Konturmodells
$\hat{\mathbf{k}}$	$\in \mathbb{R}^{N \times 1}$	Steifigkeitsspektrum eines aktiven Konturmodells

L	$\in \mathbb{R}$	Länge einer Kontur
m_{pq}	$\in \mathbb{R}$	Moment der Ordnung $(p + q)$
N	$\in \mathbb{N}$	Anzahl Samples
P	$\in \mathbb{R}$	Objektumfang
P_k	$\in \mathbb{R}$	A-priori-Wahrscheinlichkeit der Klasse k
P_q	$\in \mathbb{R}$	Quantisierungsgeräuschleistung
\tilde{P}_s	$\in \mathbb{R}$	Wechselleistung eines Signals
$p(\mathbf{x})$	$\mathbb{R}^d \rightarrow \mathbb{R}$	multivariate Wahrscheinlichkeitsdichtefunktion
R	$\in \mathbb{R}$	Entscheidungssicherheit der Klassenzugeordnung eines Samples
R	$\in \mathbb{R}$	Rundheit
\mathbf{S}_w	$\in \mathbb{R}^{d \times d}$	Intraklassen-Kovarianzmatrix
\mathbf{S}_b	$\in \mathbb{R}^{d \times d}$	Interklassen-Kovarianzmatrix
\mathbf{S}_m	$\in \mathbb{R}^{d \times d}$	Misch-Kovarianzmatrix
T	$\in \mathbb{R}$	Primitivperiode einer Kontur
TN	$\in \mathbb{R}$	Spezifität (<i>true negative rate</i>)
TP	$\in \mathbb{R}$	Sensitivität (<i>true positive rate</i>)
$v(s) = (x(s), y(s))$		Parameterkurve eines aktiven Konturmodells
\mathbf{v}	$\in \mathbb{R}^{N \times 2}$	abgetastete Parameterkurve eines aktiven Konturmodells
$\hat{\mathbf{v}}$	$\in \mathbb{R}^{N \times 2}$	Fouriertransformierte von \mathbf{v}
x, y	$\in \mathbb{R}$	Kartesische Koordinaten
\mathbf{X}	$\in \mathbb{R}^{d \times N}$	Menge multivariater Samples
\mathbf{x}	$\in \mathbb{R}^{d \times 1}$	multivariates Sample
z	$\in \mathbb{C}$	komplexe Repräsentation Kartesischer Koordinaten ($z = x + jy$)
$z(t) = x(t) + jy(t)$		komplexe Parameterdarstellung einer Kontur
\hat{Z}_μ	$\in \mathbb{C}$	Fourierdeskriptoren einer kontinuierlichen Kontur
\hat{Z}_ν	$\in \mathbb{C}$	Fourierdeskriptoren einer abgetasteten Kontur
\tilde{Z}_ν	$\in \mathbb{R}$	nichtlineare Transformation von \hat{Z}_ν (vgl. Tabelle 5.1, S. 82)
α	$\in \mathbb{R}$	Skalierungsparameter
α	$\in \mathbb{R}$	Elastizitätsparameter eines aktiven Konturmodells
β	$\in \mathbb{R}$	Verbiegungssteifigkeitsparameter eines aktiven Konturmodells
Δ	$\in \mathbb{R}$	Quantisierungsintervall
ϵ	$\in \mathbb{R}$	Exzentrizität
ϵ	$\in \mathbb{R}$	Bayes-Klassifikationsfehler
η_{pq}	$\in \mathbb{R}$	Normiertes Zentralmoment der Ordnung $(p + q)$
κ	$\in \mathbb{R}$	Krümmung (<i>Curvature</i>) einer Kontur
μ_{pq}	$\in \mathbb{R}$	Zentralmoment der Ordnung $(p + q)$
$\mu(s)$	$\in \mathbb{R}$	Chernoff-Distanz
$\mu(0.5)$	$\in \mathbb{R}$	Bhattacharyya-Distanz
μ_1	$\in \mathbb{R}$	Linearterm der Bhattacharyya-Distanz
μ_2	$\in \mathbb{R}$	Kovarianzterm der Bhattacharyya-Distanz
$\boldsymbol{\mu}$	$\in \mathbb{R}^{d \times 1}$	Mittelwertvektor von \mathbf{x}

ϱ	$\in \mathbb{R}$	Signal- zu Rauschleistungsverhältnis, angegeben in Dezibel (dB)
Σ	$\in \mathbb{R}^{d \times d}$	Kovarianzmatrix von \mathbf{x}
σ	$\in \mathbb{R}$	Skalenparameter, Glättungsgrad
$\phi_1 \dots \phi_7$	$\in \mathbb{R}$	Hu-Momentinvarianzen
$\psi_1 \dots \psi_{11}$	$\in \mathbb{R}$	Flusser-Momentinvarianzen
$\tilde{\psi}_1 \dots \tilde{\psi}_{11}$	$\in \mathbb{R}$	nichtlineare Transformation von $\psi_1 \dots \psi_{11}$ (vgl. Tabelle 5.1, S. 82)
Ψ	$\in \mathbb{R}$	<i>Normalized Multiscale Bending Energy, NMBE</i>
$\tilde{\Psi}$	$\in \mathbb{R}$	nichtlineare Transformation von Ψ (vgl. Tabelle 5.1, S. 82)
τ	$\in \mathbb{R}$	Schrittweite der Zeitdiskretisierung (akt. Konturmodelle)
ϕ	$\in \mathbb{R}$	Objektorientierung
φ	$\in \mathbb{R}$	Tangentenwinkel einer Kontur
χ_f^2		Zufallsprozess mit einer χ^2 -Verteilung mit f Freiheitsgraden

Anmerkung: Als Dezimaltrennzeichen wird in dieser Arbeit der in der angelsächsischen Notation übliche *Dezimalpunkt* verwendet.

Kapitel 1

Einleitung

1.1 Biologischer Hintergrund und Motivation

Die Untersuchung pflanzlicher Pathogenresistenz-Mechanismen ist sowohl aus biologischer als auch aus landwirtschaftlicher Sicht von großem Interesse. Dabei besteht unter anderem Bedarf am Verständnis der Pilz-Pflanze-Interaktion von Kulturgerste (*Hordeum vulgare*) mit Gerstenmehltau (*Blumeria graminis f. sp. hordei*) – einer der bedeutendsten Krankheiten dieser Kulturpflanze. Mehлтаubefall schädigt die Pflanze durch eine Beeinträchtigung der Photosynthese bei gleichzeitiger Erhöhung der Transpiration (Wasserdampfabgabe) und Respiration (Atmung) der Pflanze. Bei Gerste existieren Landrassen und Sorten, die als Resultat von Auswahl (Züchtung) resistent gegen Mehлтаubefall sind, d.h., diese Sorten sind in der Lage, Penetrationsversuche des Mehлтаupilzes in die Pflanzenzelle erfolgreich abzuwehren. Ziel vieler Forschungsbemühungen ist es, die für diese Abwehr verantwortlichen Resistenzgene zu identifizieren. Durch vorhergehende Arrayanalysen, d.h. durch die Analyse der Expression tausender Gene der Gerste, sind bereits einige hundert Kandidatengene bekannt, welche nun auf ihren tatsächlichen Einfluss auf das Resistenzverhalten hin überprüft werden sollen. Von SCHWEIZER *et al.* [SPSLD00] wurde ein transientes¹ Testverfahren entwickelt, mit dem sich die Bedeutung einzelner Gene hinsichtlich der Abwehrmechanismen der Pflanzenzelle gegen den Pilz untersuchen lässt.

In diesem Verfahren werden junge Pflanzenblätter im Labor mittels einer Hochdruckkanone mit DNA-beschichteten Goldpartikeln beschossen [KK04]. In Zellen der oberen Epidermis, deren Zellkern zufällig von solchen Partikeln getroffen wird, exprimiert das entsprechende DNA-Konstrukt auf eine transiente Weise, d.h. während zwei bis drei Tagen. Mit solchen DNA-Konstrukten lassen sich Gene transient überexprimieren oder in ihrer Expression unterdrücken, was auch als *gene silencing* bekannt ist. Eine der aktuellen Methoden für *gene silencing* ist RNA-Interferenz (RNAi), die auf der Wirkung doppelsträngiger RNA beruht. Mit einem als TIGS (*Transient Induced Gene-Silencing*) bezeichneten Verfahren lassen sich auf diese Weise einzelne zu testende Gene „ausschalten“ [SPSLD00]. Um feststellen zu können, welche Zellen von solchen DNA-tragenden Partikeln getroffen wurden, wird ein Reportergen mit eingebracht, welches einen blaugrünen Indigofarbstoff in den transformierten Zellen produziert, so dass diese bei Betrachtung unter dem Lichtmikroskop leicht identifiziert werden können².

Um den Abwehrmechanismus der transformierten Zellen zu testen, werden nun Mehltausporen auf die beschossenen Blätter aufgebracht (Inokulation) und diese über eine Zeit von 48 Stunden ihrer Entwicklung überlassen. Anschließend wird die Inokulation durch Infiltration eines Trichloressigsäure-Methanolgemischs gestoppt und gleichzeitig die Blätter entfärbt, gefolgt von der

¹Bei einem transienten Verfahren werden nur einzelne Zellen genetisch verändert – keine ganzen Organismen.

²Das GUS-Reportergen führt in den getroffenen Zellen zunächst zur Herstellung des Enzyms β -Glukuronidase. Dieses Enzym setzt die hinzugefügte X-GlcA-Substratlösung (5-Brom-4-Chlor-3-idoly- β -D-Glucuronid) in den Indigofarbstoff um [KK04].

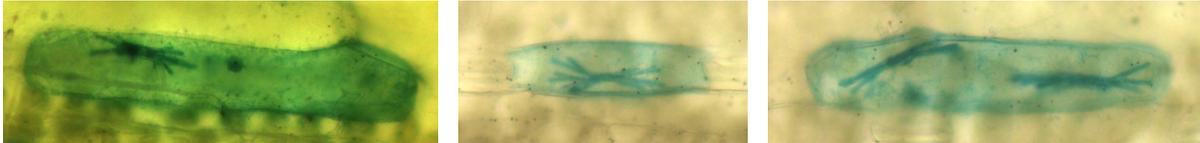


Abbildung 1.1: Lichtmikroskopische Aufnahmen gefärbter Epidermiszellen von Gerstenblättern (*Hordeum vulgare*). Eine blaugrünliche Färbung zeigt an, dass die Zelle genetisch transformiert ist. Die darin sichtbaren Objekte mit den typischen fingerartigen Fortsätzen sind **Haustorien** des Mehltaupilzes (*Blumeria graminis f. sp. hordei*). Über diese Saugorgane ernährt sich der Pilz, indem er der Zelle Nährstoffe entzieht.

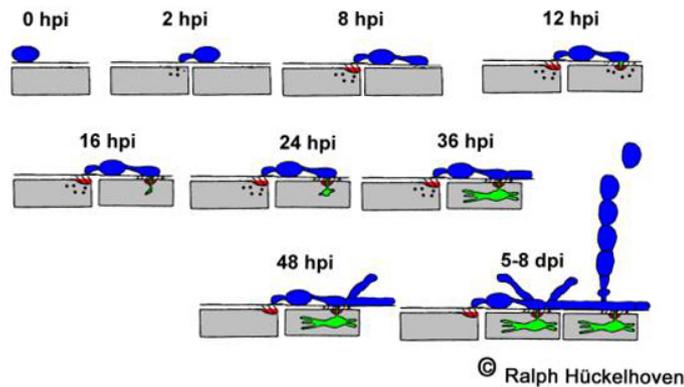


Abbildung 1.2: Schema der Entwicklung des Mehltaupilzes auf einem Gerstenblatt. (hpi/dpi = Stunden/Tage nach Inokulation) [HFPK00].

GUS-Färbung. Schließlich werden die präparierten Blätter auf Objektträger aufgebracht (zwei bis drei Blätter pro Objektträger).

Unter dem Lichtmikroskop lassen sich nun die von solchen DNA-tragenden Partikeln getroffenen, transformierten Zellen über ihre blaugrünliche Färbung identifizieren. Ob diese Zellen den Eindringversuch des Mehltaupilzes abwehren konnten, d.h., ob das Testgen möglicherweise einen Einfluss auf die Pathogenresistenz hat, lässt sich dabei über einen „Fußabdruck“ (*footprint*) des Mehltaupilzes überprüfen: Der Pilz heftet bei seinem Eindringversuch zunächst ein so genanntes Appressorium an die Zelloberfläche. Von diesem Haftorgan aus dringt eine Infektionshyph in die Wirtszellwand ein und bildet in der Zelle ein so genanntes **Haustorium**. Über dieses Saugorgan ernährt sich der Pilz, indem er der penetrierten Wirtszelle Nährstoffe entzieht. Ausgereifte Haustorien weisen typischerweise fingerartige Fortsätze (zur Vergrößerung ihrer Oberfläche) auf und sind ein eindeutiger Indikator für das erfolgreiche Eindringen des Pilzes in die Wirtszelle, stellen also den oben erwähnten „Fußabdruck“ des Pathogens dar. In Abbildung 1.1 sind exemplarisch einige transformierte Zellen mit Haustorien abgebildet. Abbildung 1.2 zeigt eine schematische Darstellung der Entwicklung des Mehltaupilzes und die Ausbildung von Haustorien in befallenen Wirtszellen.

Über eine große Versuchsreihe ist nun zu überprüfen, welche Gene die Pathogenresistenz der Gerste beeinflussen. Dazu ist für jedes Testgen zuerst die Erstellung eines RNAi-Konstruktes notwendig. Anschließend müssen pro Konstrukt einige hundert Zellen ausgewertet werden, um aus der statistischen Häufigkeit des Haustorienbefalls (Anfälligkeitsindex, *susceptibility index*) Rückschlüsse auf die Resistenz ziehen zu können. Zur Überprüfung von etwa tausend potenziellen Testgenen sind also insgesamt weit über 100 000 transformierte Zellen auf das Vorhandensein von Haustorien zu untersuchen.

Die manuelle Auswertung der Proben unter dem Mikroskop ist jedoch einerseits ein sehr zeitaufwändiges Verfahren, andererseits bringt der menschliche Betrachter eine gewisse subjektive Komponente in die Ergebnisse ein. Außerdem erlaubt diese Form der manuellen Auswertung keine Dokumentation der Rohdaten – es ist nicht möglich, die Ergebnisse der Experimente zu einem späteren Zeitpunkt nachzuvollziehen, weil nach dem Verderben der Frischpräparate keine Möglichkeit einer erneuten Betrachtung der Proben besteht.

Ziel ist daher die Entwicklung eines Screeningsystems, welches die aufwändige manuelle Arbeit durch eine automatische und nachhaltige Hochdurchsatz-Analyse ersetzt.

1.2 Ziele und Herausforderungen

Die Vorteile eines solchen automatischen Hochdurchsatz-Screeningsystems liegen auf der Hand:

- Die zeitaufwändige und für den Menschen eintönige und ermüdende Arbeit am Mikroskop entfällt völlig, und die dadurch entlasteten Laborassistenten können sich anderen Aufgaben widmen – beispielsweise der Herstellung neuer Proben.
- Der vollautomatische Screeningprozess (lediglich die Proben sind manuell einzulegen) ermöglicht einen 24-Stunden-Betrieb des Systems, was die teuren Laborgeräte besser auslastet.
- Durch die Speicherung der Rohdaten der Experimente in Form von Digitalbildern wird eine qualitativ hochwertige Dokumentation erreicht, so dass auch zu einem späteren Zeitpunkt (lange nachdem die Frischpräparate bereits verdorben sind) weitere Untersuchungen auf Basis der Ausgangsdaten erfolgen können.
- Die subjektive Komponente des menschlichen Auswerters wird durch deterministische Bildauswertungsalgorithmen ersetzt.

Die geschilderten Hintergründe lassen vermuten, dass für diese recht spezielle Problemstellung keine Standardlösung existiert. Vielmehr erfordert die Entwicklung eines solchen vollautomatischen Hochdurchsatz-Screeningsystems ein sehr breites Spektrum ingenieurtechnischer Eigenarbeit. Zentraler Punkt dabei ist die Übertragung der biologischen Expertise auf ein technisches System. Dabei sind folgende Kernaufgaben zu lösen:

- **Digitalisierung relevanter Bereiche des Probenmaterials**
Die bisher vom Experimentator manuell durchgeführte Suche nach transformierten Zellen auf dem Probenmaterial muss vollständig automatisiert ablaufen. Dazu gehören:
 - Navigation der Proben unter dem Mikroskop
 - Auffinden interessanter Regionen mit gefärbten, genetisch transformierten Zellen
 - korrekte Fokussierung der abzubildenden Objekte
 - strukturierte Speicherung und Archivierung von Bildmaterial.
- **Automatische Bildanalyse**
Das aufgenommene Bildmaterial muss vollautomatisch ausgewertet werden – der prüfende Blick des menschlichen Auswerters in das Okular des Lichtmikroskops ist mittels automatischer Bilderkennungsalgorithmen nachzubilden. Vor dem Hintergrund der konkreten Aufgabenstellung umfasst dies die folgenden Kernpunkte:
 - Segmentierung der gefärbten, genetisch transformierten Zellen
 - Analyse der transformierten Zellen auf Haustorien des Mehltaupilzes
 - Statistische Auswertung der Ergebnisse und Dokumentation sowohl der relevanten Rohdaten als auch der Ergebnisse.

1.3 Technische Voraussetzungen und Vorüberlegungen

1.3.1 Mikroskophardware und Steuersoftware

Die Hardwarebasis des Screeningsystems besteht aus einem vollmotorischen Mikroskop „Axio-plan Imaging 2“ der Firma Carl Zeiss. Aufgrund der Motorisierung lassen sich nahezu sämtliche Komponenten computergesteuert betätigen. Insbesondere ist das Mikroskop mit einem motorischen X-Y-Tisch ausgestattet, auf dem bis zu acht Objektträger Platz haben. Die Bildbetrachtung kann wahlweise im Okular und/oder auf dem Computermonitor erfolgen.

Die Steuersoftware „AxioVision“ bietet umfassende Möglichkeiten der Computer-Mikroskop-Interaktion, sie ist jedoch mehr für eine interaktive Bedienung durch den Nutzer und weniger für eine automatische Ablaufsteuerung ausgelegt. Durch die Scriptingmöglichkeiten über die optional erhältliche Visual Basic for Applications (VBA)-Schnittstelle bieten sich jedoch genügend Freiheiten, eine vollautomatische Ablaufsteuerung zu realisieren. Abbildung 1.3 vermittelt einen Eindruck der Hardware im praktischen Einsatz. Auffällig ist der große motorische Tisch, welcher in Abbildung 1.4 näher betrachtet werden kann. Auf ihm finden bis zu acht Objektträger Platz.

1.3.2 Vorüberlegungen bezüglich des Zeitaufwandes

Vor dem Hintergrund eines Hochdurchsatzbetriebs spielen insbesondere die *zeitlichen* Randbedingungen des Screenings eine wichtige Rolle. Der zügige Ablauf des Screeningprozesses wird vordergründig durch die *Bildaufnahme* und weniger durch die *Bildauswertung* bestimmt, da letztere durch parallele Datenverarbeitung nahezu beliebig beschleunigt werden kann. Als Flaschenhals tritt dabei besonders der Bildeinzug in Erscheinung. Tabelle 1.1 gibt dazu praktische Zahlenwerte an. Die native Auflösung des Kamerasensors beträgt 1300×1030 Pixel; eine Verdopplung bzw. Verdreifachung der Auflösung wird optional durch eine piezomechanische Verschiebung des Sensors und der Interpolation der entstehenden vier bzw. neun Einzelbilder realisiert. Diese virtuelle Erhöhung der Bildauflösung beeinträchtigt jedoch den zügigen Bildeinzug erheblich.

Bildauflösung der Kamera (Pixel)	432×342	1300×1030	2600×2060	3900×3090
Zeitaufwand pro Bildeinzug	< 1 s	≈ 2 s	≈ 7 s	≈ 9 s

Tabelle 1.1: Unterstützte Bildauflösungen der verwendeten Kamera und benötigte Zeiten für den Bildeinzug. Der hohe Zeitbedarf für Bildauflösungen $\geq 2600 \times 2060$ Pixel rührt daher, dass die Kamera eine native Auflösung von 1300×1030 Pixel hat, welche durch eine 4- bzw. 9-fache piezomechanische Verschiebung des Sensors vervielfacht werden kann.

1.3.3 Objektivauswahl und generelle Strategie der Bildaufnahme

Der Bildaufnahmeprozess soll in einer *zeiteffizienten* Weise qualitativ hochwertiges Bildmaterial bereitstellen, welches die Beantwortung der biologischen Fragestellung möglichst *zuverlässig* zulässt. Dabei sind folgende Anforderungen zu erfüllen:

- Es sind möglichst alle relevanten Bereiche auf dem Objektträger – nämlich die grünbläulich angefärbten transformierten Zellen – zu fotografieren.
- Die Abbildung muss korrekt fokussiert sein, damit der potenzielle Befall der transformierten Zellen mit Haustorien des Mehltaupilzes sicher erkannt werden kann.
- Die Vergrößerung soll derart gewählt sein, dass möglichst viel Information pro Bild untergebracht wird, ohne jedoch die Erkennungsrate zu beeinträchtigen.

1.3 TECHNISCHE VORAUSSETZUNGEN UND VORÜBERLEGUNGEN



Abbildung 1.3: Computergesteuertes Mikroskop mit X-Y-Tisch und Digitalkamera. Nahezu sämtliche Komponenten lassen sich softwaregesteuert betätigen.

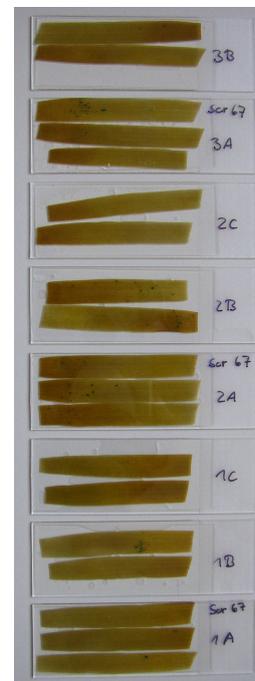
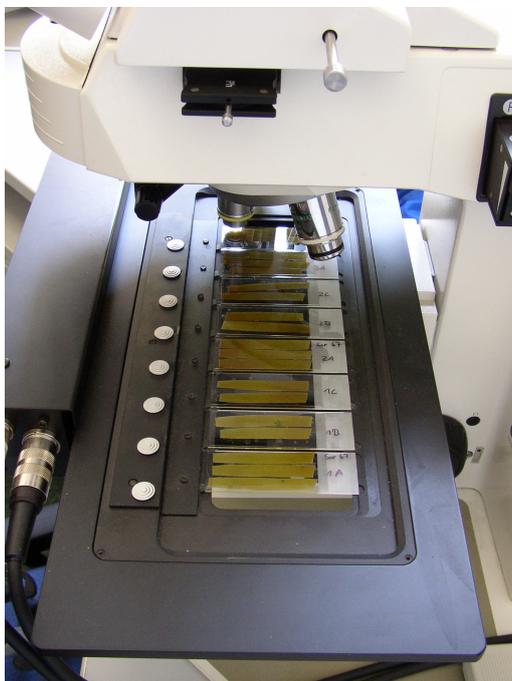


Abbildung 1.4: *Links*: Blick auf den X-Y-Tisch. Auf ihm finden bis zu acht Objektträger Platz. *Rechts*: Acht Objektträger mit jeweils zwei oder drei präparierten Gerstenblättern. In ähnlicher Weise kommen die Proben auch auf dem X-Y-Tisch (vgl. linkes Bild) zu liegen.

Wahl des Objektivs und der Bildauflösung

Der letztgenannte Aspekt zur Wahl eines geeigneten Objektivs lässt sich leicht durch einfaches Probieren beantworten: Aus den zur Verfügung stehenden Objektiven (mit Vergrößerungen zwischen $5\times$ und $100\times$) kommen nur die Vergrößerungen $5\times$ und $10\times$ in Betracht, da die Zellen bei größeren Werten sonst nicht mehr in voller Größe abgebildet werden können. Die $5\times$ -Vergrößerung zeichnet sich durch eine relativ große Schärfentiefe aus, bei der auch tiefer liegende Bereiche des Blattmaterials deutlich in Erscheinung treten. Dies erschwert die Erkennung der Haustorien erheblich. Des Weiteren sind die zu detektierenden Strukturen zu klein, um bei dieser Vergrößerung sicher erkannt zu werden. Die Zielvergrößerung für die Bildaufnahme ist daher das $10\times$ -Objektiv. Dieses Objektiv wird auch von den Laborassistenten bei der manuellen Auswertung benutzt.

Mit dem $10\times$ -Objektiv sind die Verzweigungen der Haustorien („Finger“) gerade noch bei einer Bildauflösung von 1300×1030 Pixel abbildbar – ein Haustorienfinger erscheint dann etwa 3–4 Pixel breit. Zwar ließe sich eine sicherere Abbildung mittels einer höheren Kameraauflösung erreichen, aufgrund der zeitlichen Restriktionen scheidet dies jedoch aus. Für die Auflösung 2600×2060 Pixel benötigt der Bildeinzug mehr als die dreifache Zeit (siehe Tabelle 1.1).

Aufnahmestrategie

Nur wenige Zellen auf dem Probenmaterial sind tatsächlich transformiert (einige -zig Zellen pro Gerstenblatt) und müssen fotografiert werden. Keinesfalls ist es unter zeitlichen und ökonomischen Aspekten möglich, den gesamten Objektträger bei $10\times$ -Vergrößerung lückenlos mit hoher Bildauflösung zu fotografieren. Statt dessen müssen in einem möglichst aufwandsgünstigen Verfahren zuerst relevante Bereiche mit transformierten Zellen aufgefunden werden, bevor die zeitaufwändige Bildaufnahme dieser Bereiche in Betracht kommt.

Für das *Auffinden* transformierter Zellen ist die Vergrößerung des $5\times$ -Objektivs ausreichend. Des Weiteren ist nur eine *annähernd* korrekte Fokussierung erforderlich, da zunächst lediglich die *Färbung* der transformierten Zellen detektiert werden muss.

Aus diesen Gründen findet die Bildaufnahme in einem zweistufigen Verfahren statt:

1. Vorschuaufnahmen

Die erste Stufe stellt zunächst Vorschaubilder bei kleiner Vergrößerung ($5\times$ -Objektiv) mit geringer Auflösung (432×342 Pixel) bereit. Dazu werden alle Probenbereiche mäanderförmig abgefahren und fotografiert. Auf diese Weise entstehen $15 \times 38 = 570$ Vorschaubilder pro Objektträger, wie in Abbildung 1.5 schematisch dargestellt. Abbildung 1.6 zeigt *links* einige typische Vorschaubilder mit gefärbten Bereichen. In Abbildung 1.6 *rechts* wurden alle Vorschuaufnahmen zu einer Gesamtansicht des Objektträgers zusammengesetzt.

2. Z-Stapel von Detailaufnahmen

In der zweiten Stufe werden die in den Vorschaubildern detektierten transformierten Zellen bei $10\times$ -Vergrößerung mit hoher Bildauflösung (1300×1030 Pixel) in Form eines Z-Stapels (5 Bilder) abfotografiert. Die Notwendigkeit mehrerer fokaler Ebenen in Form eines Z-Stapels ergibt sich durch die „hügelige“ Struktur des Probenmaterials: Es existieren verschiedene Schärfeebenen, bei denen unterschiedliche Strukturen hervortreten. Dies wird in Abbildung 1.7 verdeutlicht, wo zwei transformierte Zellen in unterschiedlichen Schärfeebenen zu sehen sind.

Zeitbedarf

In Tabelle 1.2 findet sich eine Abschätzung des Zeitaufwands für die zwei Stufen der Bildaufnahme. Der Aufwand für die Detailaufnahmen ist – im Gegensatz zu den Vorschuaufnahmen –

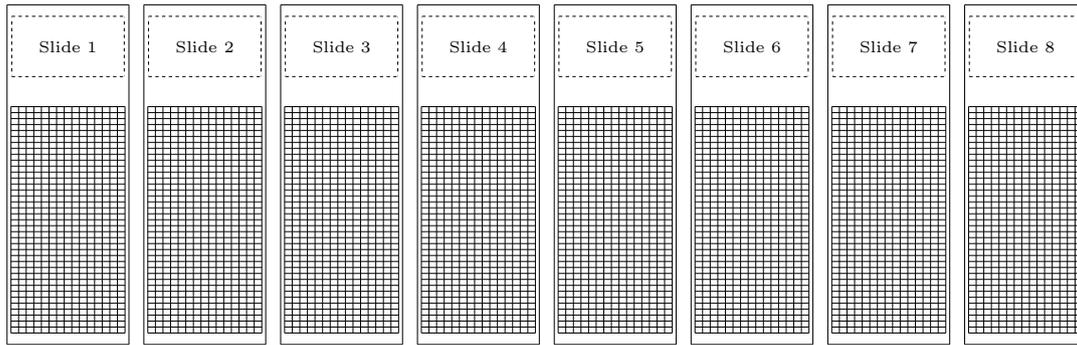


Abbildung 1.5: Kachelung der Vorschaufnahmen bei kleiner Vergrößerung (5×-Objektiv). Für jeden der acht Objektträger (*Slides*) sind $15 \times 38 = 570$ Einzelaufnahmen nötig, um eine vollständige Abdeckung zu gewährleisten. Die Aufnahme dieser Bilder bei der Zielvergrößerung mit dem 10×-Objektiv würde den Aufwand *vervierfachen!*

abhängig vom tatsächlichen Probeninhalt. Dabei stellt die Annahme von 100 Positionen pro Objektträger einen Richtwert dar, der auch deutlich über- bzw. unterschritten werden kann.

Speicherplatzbedarf

Tabelle 1.3 gibt den Speicherplatzbedarf für die anfallenden Bilder an. Auf eine Kompression sollte während der Verarbeitungsphase aus Zeit- und Effizienzgründen verzichtet werden. Für die Archivierung kann aus Gründen der Platzersparnis eine mit geringen Verlusten behaftete Komprimierung in Erwägung gezogen werden. Dabei ist zu beachten, dass verlustbehaftete Kompressionsverfahren in der Regel anhand der *menschlichen* Wahrnehmung optimiert sind. Eine maschinelle Bildauswertung solcher komprimierten Bilder kann jedoch zu unerwarteten Ergebnissen führen. So verursacht die 8×8-Blockbildung des etablierten JPEG-Formats unter Umständen Artefakte („Klötzchenbildung“), wodurch Bildauswertungsalgorithmen durch die künstlich eingebrachten Kanten erheblich gestört werden können. Im Gegensatz dazu äußern sich die Kompressionsverluste des auf Wavelets basierenden JPEG2000-Formats lediglich in leichten Unschärfen (Tiefpassfilterung), was aus Sicht der Bildverarbeitung weitaus weniger kritisch ist.

Notwendige Aktionen	Vorschaubilder (432×342 Pixel)	Detailbilder (1300×1300 Pixel)
Tischpositionierung	≈ 0.5 s	< 1 s
Autofokus	(1× pro Objektträger)	2...3 s
Einzelbildaufnahme	≈ 0.7 s	≈ 2 s
Aufnahme des Z-Stapels	—	5 Bilder je Position
Zeitaufwand pro Objektträger	≈ 10 min	≈ 25 min (bei 100 Positionen)

Tabelle 1.2: Abschätzung des Zeitbedarfs für die Bildaufnahme *eines* Objektträgers. Zunächst liefern die *Vorschauaufnahmen* eine vollständige, grobe Abbildung des Objektträgers mit dem 5×-Objektiv. Anschließend werden *Detailaufnahmen* von Regionen mit transformierten Zellen in mehreren fokalen Ebenen unter Nutzung des 10×-Objektivs bereitgestellt.

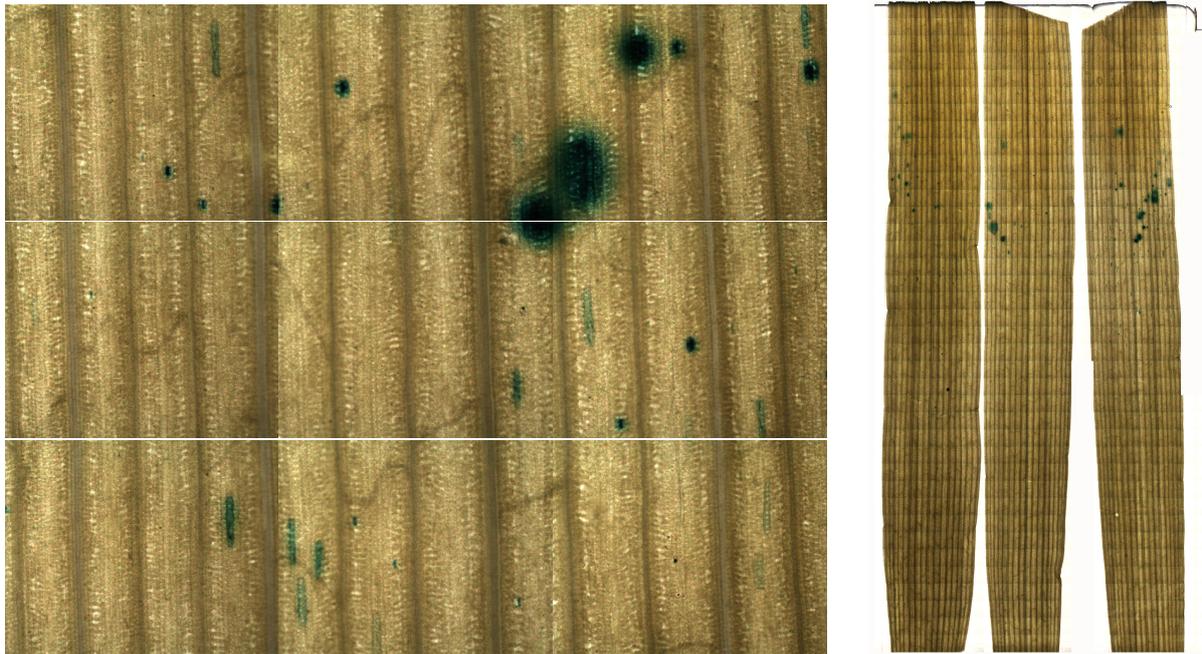


Abbildung 1.6: *Links*: Drei mal drei Vorschaubilder horizontal und vertikal aneinander gesetzt. Die Vorschaubilder werden mit dem 5 \times -Objektiv bei einer Bildauflösung von 432 \times 342 Pixel aufgenommen. *Rechts*: Gesamtansicht aller Vorschaubilder eines Objektträgers. Der *links* dargestellte Ausschnitt befindet sich auf dem rechten Blatt an der Stelle des untersten stark gefärbten Flecks (ca. $\frac{1}{3}$ von oben).

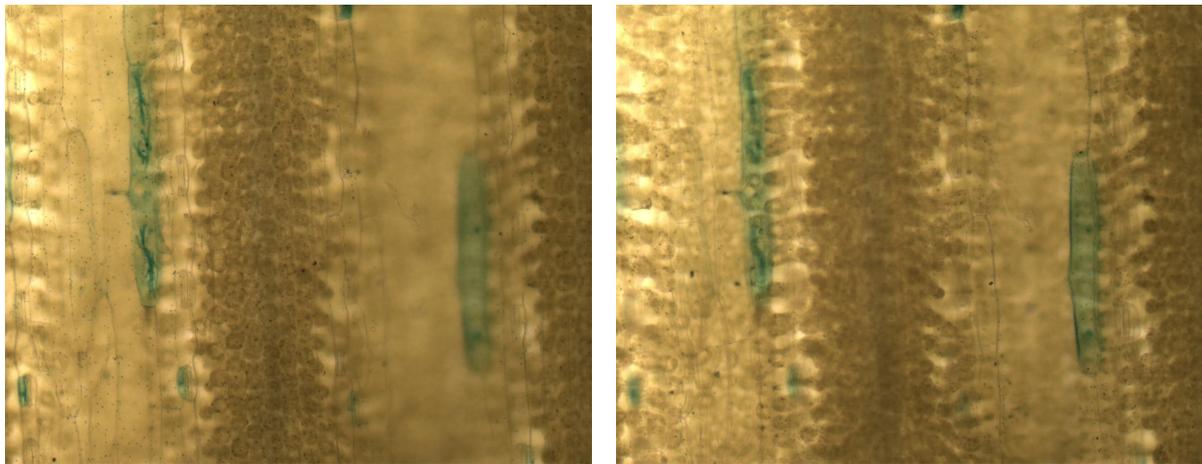


Abbildung 1.7: Detailaufnahmen transformierter Zellen. Die beiden transformierten Zellen liegen in unterschiedlichen fokalen Ebenen. Beide Bilder weisen eine Fokusdifferenz von $\Delta Z = 20\mu\text{m}$ auf. Die Detailbilder werden mit dem 10 \times -Objektiv mit einer Bildauflösung von 1300 \times 1030 Pixel aufgenommen. Sie sind die Eingangsdaten der zu entwickelnden automatischen Bildauswertungsalgorithmen.

Verfahren	Standard	Vorschauansicht eines Objektträgers	Detailbild (1300×1300 Pixel)
keine Kompression	TIF, BMP	≈ 300 MB	≈ 4.0 MB
verlustlose Kompression	PNG	≈ 150 MB	≈ 1.7 MB
geringe Verluste	JPEG2000	≈ 50 MB	≈ 0.3 MB

Tabelle 1.3: Speicherplatzbedarf der anfallenden Rohbilddaten in Megabyte (MB).

1.4 Weiteres Vorgehen

Mit diesen einleitenden Betrachtungen sind die Randbedingungen des zu realisierenden Systems definiert. Zentraler Punkt ist dabei die Entwicklung von Verfahren, welche die automatische Auswertung bisher manuell betrachteter Bilder ermöglichen. Dies betrifft im Wesentlichen Segmentierungsalgorithmen, welche die Sichtweise des menschlichen Betrachters für diese spezielle Klasse von Bildmaterial nachbilden. Da das Merkmal „Farbe“ dabei eine zentrale Rolle spielt, werden die beiden Aspekte *Farbbildverarbeitung und Segmentierung* an den Anfang der Arbeit gestellt.

Später wird sich zeigen, dass mit den entwickelten Algorithmen sowohl für die Zellsegmentierung als auch für die Haustoriensegmentierung zwar sehr robuste Ergebnisse erzielt werden können, dass jedoch in dem sehr komplexen Fall der Haustoriensegmentierung das Farbmerkmal allein nicht ausreicht, um diese Strukturen mit der nötigen Sicherheit zu erkennen. Daher muss im Anschluss an die Segmentierung eine *objektbasierte Klassifikation* erfolgen, welche fälschlicherweise als Haustorien segmentierte Objekte erkennt und zurückweist. Dies wird im zweiten Teil der Arbeit behandelt.

Im dritten und letzten Teil wird schließlich das Screeningsystem in seiner Gesamtheit vorgestellt. Teile der im ersten und zweiten Teil der Arbeit gewonnenen Erkenntnisse zur Bildsegmentierung und Objektklassifikation werden dort zur Realisierung der automatischen Vorauswahl interessanter Bereiche (Analyse der Vorschaubilder) wieder verwendet. Außerdem wird gezeigt, mit welchen Optimierungsmethoden eine möglichst zeitsparende Aufnahme des Bildmaterials erfolgen kann.

Ergebnis der Arbeit ist ein vollautomatisches Laborassistenzsystem, welches für den täglichen Laboreinsatz tauglich ist und wesentlich zu neuen Erkenntnissen im Bereich der Pflanzenbiologie beitragen wird.

Teil I

Farbbildverarbeitung und Segmentierung

Kapitel 2

Grundlagen der Farbwahrnehmung, Farbräume

Diese Arbeit beschäftigt sich mit dem umfangreichen Thema **Farbbildverarbeitung** mit dem konkreten Ziel, lichtmikroskopische, qualitative Untersuchungen an farblich markierten Proben zu automatisieren. Dabei ist eine bisher von Menschen durchgeführte Analyse nunmehr von einem technischen Sehsystem zu übernehmen. Weil dem Merkmal „Farbe“ vor diesem Hintergrund eine so große Bedeutung zukommt, erscheint es angemessen und wichtig, zunächst einige grundlegende Fragen zu dieser Thematik zu klären:

- Was genau ist Farbe?
- Worin unterscheidet sich die *menschliche* und *maschinelle* „Farbwahrnehmung“?
- Wie kann Farbe mittels eines technischen Sehsystems repräsentiert werden?

2.1 Farbwahrnehmung

Farbe ist ein Sinneseindruck, welcher durch Licht einer bestimmten spektralen Verteilung bei einem Beobachter hervorgerufen wird. Ein Farbeindruck existiert also rein virtuell als Sinnesempfindung bei einem Lebewesen, welches die zugrunde liegenden physikalischen Phänomene entsprechend wahrnehmen kann. Physikalischer Auslöser einer Farbwahrnehmung ist Energie in Form von elektromagnetischen Wellen, die eine bestimmte spektrale Leistungsdichte (*Power Spectral Density, PSD* oder *Spectral Power Distribution, SPD*) aufweisen. Elektromagnetische Wellen im Bereich von 380 bis 780 nm Wellenlänge werden vom Menschen als Spektralfarben wahrgenommen (vgl. Abbildung 2.1). Dies kann entweder direkt durch eine Lichtquelle oder indirekt durch Reflexionen, die von einem beleuchteten Körper ausgehen, geschehen. Das System „Farbwahrnehmung“ besteht damit aus drei wesentlichen Komponenten: der *Lichtquelle*, dem *Objekt* und dem *Betrachter*.

Lichtquelle

Die von einer Lichtquelle emittierte elektromagnetische Strahlung lässt sich mittels ihrer spektralen Leistungsdichte beschreiben. Häufig wird dabei mit normierten Werten gearbeitet, da der Absolutwert der Gesamtleistung an dieser Stelle keine Rolle spielt, sondern allein die spektrale *Verteilung* der Leistung interessiert. Als Referenzwert für die Normierung der Spektralfunktion dient dabei der Wert bei $\lambda = 560$ nm. Daraus resultiert die relative spektrale Leistungsdichte $E(\lambda)$.

Die bedeutendste Lichtquelle ist die Sonne und Hersteller von Leuchtmitteln sind bemüht, das Emissionsspektrum ihrer Produkte dem natürlichen Ideal „Sonne“ möglichst nahe kommen zu

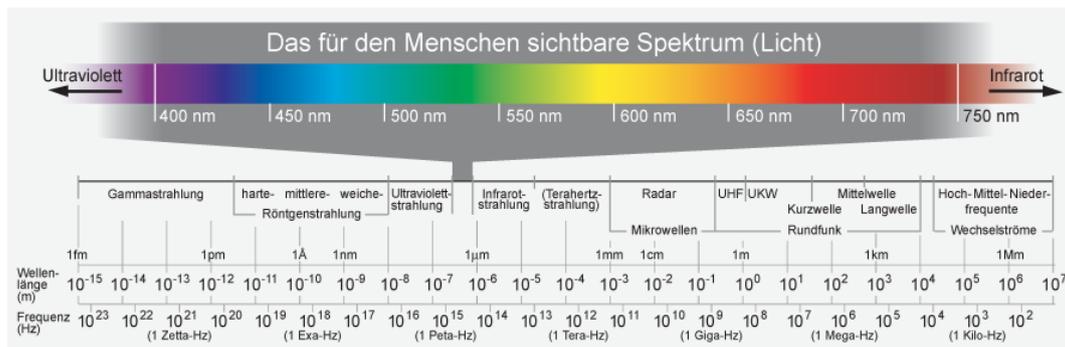


Abbildung 2.1: Für den Menschen sichtbarer Teil des elektromagnetischen Spektrums. (Quelle: http://de.wikipedia.org/wiki/Elektromagnetische_Strahlung)

lassen. Mit Hilfe von gefiltertem Xenonlicht ist dies bereits näherungsweise realisierbar; eine *preiswerte* Beleuchtung mit tageslichtähnlichen Eigenschaften in Form eines Massenproduktes ist derzeit aber noch nicht erhältlich [Lew01].

Das Lichtspektrum von Temperaturstrahlern lässt sich näherungsweise mittels der *Farbtemperatur* parametrieren – jener Temperatur, die ein idealer schwarzer Strahler aufweisen müsste, um einen ähnlichen Farbeindruck wie der Temperaturstrahler hervorzurufen. Die Farbe des schwarzen Strahlers variiert von *schwarz* bei 0 K nach *rot* bei ca. 1000 K über *weiß* bei ca. 4500 K nach *blau* bei ca. 6500 K. Eine Lampe mit weißglühendem Glühfaden liegt bei ca. 2900 K, Halogenlampen bei 3400 K, Xenonlampen bei 4500 bis 5000 K.

Von der Commission Internationale de l’Eclairage (CIE), der Internationalen Beleuchtungskommission¹, wurde für das durchschnittliche Tageslicht die Annahme einer Farbtemperatur von 6500 K vorgeschlagen, was zur Spezifizierung der Standardbeleuchtung D65 führte. Weitere an das Tageslicht angelehnte Standardbeleuchtungen sind D50, D55 und D75, welche jeweils Farbtemperaturen von 5000 K, 5500 K bzw. 7500 K entsprechen. Der Vollständigkeit halber seien an dieser Stelle noch die Standards A (schwarzer Strahler bei 2856 K), C (frühere Spezifikation einer tageslichtähnlichen Beleuchtung, abgelöst durch D65) sowie F1–F12 (verschiedene Fluoreszenzlampen) erwähnt.

Die spektralen Leistungsdichten einiger Referenz-Lichtquellen lassen sich als diskrete Werte auf diversen Internetseiten² finden.

Objekt

Bezogen auf ihre Lichtdurchlässigkeit können Objekte in die Klassen „transparent“, „durchlässig“ oder „undurchsichtig“ eingestuft werden, wobei letztgenannte Eigenschaft auf den größten Teil der uns umgebenden Objekte zutrifft. Das von Objekten reflektierte Lichtspektrum kann mit Hilfe eines Spektrofotometers gemessen und in Form des spektralen Reflexionsgrades $S(\lambda)$, einer dimensionslosen Größe zwischen 0 und 1, angegeben werden.

Die von einem beleuchteten Objekt ausgehende spektrale Leistungsdichte $P(\lambda)$ ergibt sich aus der spektralen Leistungsdichte der Lichtquelle, $E(\lambda)$, multipliziert mit dem spektralen Reflexionsgrad des Objekts, $S(\lambda)$.

$$P(\lambda) = E(\lambda) S(\lambda) \quad (2.1)$$

¹Die Anfang des 20. Jahrhunderts gegründete Internationale Beleuchtungskommission (CIE) ist eine gemeinnützige Organisation mit Hauptsitz in Wien, die die internationale Kooperation und den Informationsaustausch sowie die Standardisierung in allen wissenschaftlichen Belangen zum Thema Beleuchtung, Farbwahrnehmung, Fotometrie etc. unterstützt.

²beispielsweise auf der Seite der CIE unter <http://www.cie.co.at> oder in der CVRL Colour & Vision database unter <http://cvrl.ioo.ucl.ac.uk/index.htm>

Beobachter

Die Retina des menschlichen Auges enthält zwei verschiedene Fotorezeptoren – Stäbchen und Zapfen. Die *Stäbchen* sind sehr lichtempfindlich und für das Sehen bei Dämmerung und Nacht verantwortlich (skotopisches Sehen). Eine Farbwahrnehmung ist für den Menschen nur möglich, wenn genügend Licht das Auge erreicht, so dass die Empfindlichkeitsschwelle der *Zapfen* überschritten wird. Die Zapfen unterteilen sich in drei verschiedene Typen: den *Blaurezeptor*, den *Grünrezeptor* und den *Rot-(Gelb-)Rezeptor*. Mit Hilfe dieser drei Farbrezeptoren wird der sichtbare Bereich des Spektrums „abgetastet“. Signaltheoretisch handelt es sich aus Sicht des Betrachters bei einem Farbsignal also um drei Signalkomponenten, aus denen sich jede beliebige Farbe mischen lässt. Diese Gesetzmäßigkeiten wurden bereits im 19. Jahrhundert von YOUNG [You02] und VON HELMHOLTZ [Hel67] beschrieben und sind als die Young-Helmholtzsche Dreifarben-theorie bekannt. Die spektralen Empfindlichkeiten der drei Fotorezeptoren des menschlichen Auges wurden erstmals 1964 von BROWN *et al.* [BW64] und MARKS *et al.* [MDMJ64] mit mikrospektralphotometrischen Methoden gemessen. In Abbildung 2.2 sind die spektralen Empfindlichkeiten der Zapfen aus neuen Messungen nach STOCKMAN *et al.* dargestellt³.

2.1.1 Metamerie und RGB-Spektralwertkurven

Abbildung 2.3 zeigt die Ergebnisse folgenden Experiments: Ein menschlicher Beobachter soll zwei Lichtquellen vergleichen. Die eine besteht aus einer *Mischung* der drei Grundfarben (Primärvalenzen) rot, grün und blau, die andere emittiert monochromatisches Licht einer beliebig einstellbaren Wellenlänge mit konstanter Intensität. Die Aufgabe des Beobachters ist es, die Intensität der drei Grundfarben derart einzustellen, dass die entstehende Mischfarbe der des monochromatischen Lichts entspricht. Diese Anpassbarkeit beruht auf einem psychophysischen Phänomen namens *Metamerie*, welches bedeutet, dass zwei Lichtquellen visuell gleich erscheinen können, jedoch völlig verschiedene spektrale Leistungsdichten aufweisen.

Wird das ganze sichtbare Spektrum mit der monochromatischen Lichtquelle durchfahren, so wird es dem Beobachter in einigen Spektralbereichen nicht möglich sein, eine Übereinstimmung mit den Mischfarben zu erreichen. In solchen Spektralbereichen ist es jedoch möglich, mittels eines „Tricks“ trotzdem eine künstliche Übereinstimmung zu generieren – indem eine bestimmte Intensität einer Grundfarbe *anstatt zur Mischlichtquelle zur monochromatischen Lichtquelle* hinzugemischt wird. In diesem Fall ist mathematisch gesehen eine negative Intensität der Grundfarbe für die Mischlichtquelle nötig. Natürlich sind negative Intensitäten physikalisch nicht realisierbar. Deshalb lassen sich auch nicht alle Spektralfarben mittels einer Mischlichtquelle aus den drei genannten Grundfarben nachbilden.

Aus diesem Grund haben sowohl Aufnahmegeräte wie Farbkameras als auch Wiedergabegeräte wie Farbbildschirme, die auf dem Prinzip der Farbmischung aus den genannten Grundfarben rot, grün und blau (RGB) basieren, nur einen eingeschränkten Farbdarstellungsbereich. Jener Bereich von Farben, die von einem Gerät dargestellt werden können, wird als Farbskala (*Color Gamut*) bezeichnet.

Angewendet auf das Ziel einer möglichst *farbgetreuen* Abbildung der Mikroskopbilder lässt sich daraus Folgendes schlussfolgern:

- Die Farbdarstellung eines Mikroskopbildes auf dem (Kontroll)Bildschirm kann dem Farbumfang, wie er sich im Okular ergibt, höchstens *angenähert* werden – eine *exakte* Reproduktion ist jedoch nicht möglich.

³Datenquelle: <http://cvr1.iio.ucl.ac.uk/cones.htm>

2.1.2 Normspektralwertkurven nach CIE

Da sich nicht alle sichtbaren Farben mittels der RGB-Grundfarben darstellen lassen (weil dies teilweise negative Intensitäten erfordern würde), wurde von der CIE im Jahr 1931 ein System basierend auf drei virtuellen Spektralwertkurven $\bar{x}(\lambda)$, $\bar{y}(\lambda)$ und $\bar{z}(\lambda)$ definiert, welches folgenden beiden Bedingungen genügt (vgl. Abbildung 2.4):

- Alle sichtbaren Farben lassen sich aus drei *nichtnegativen* Spektralwertkurven darstellen.
- Die Spektralwertkurve $\bar{y}(\lambda)$ ist proportional zum menschlichen Helligkeitsempfinden, wogegen die Kurven $\bar{x}(\lambda)$ and $\bar{z}(\lambda)$ reine Farbkomponenten sind.

Mittels der Wichtung der spektralen Leistungsdichte $C(\lambda)$ mit diesen virtuellen Spektralwertkurven und der Integration über den interessierenden Spektralbereich (im Allgemeinen $\lambda = 380 \dots 780$ nm) ergeben sich die Normfarbwerte X , Y und Z .

$$X = \int \bar{x}(\lambda) C(\lambda) d\lambda \quad Y = \int \bar{y}(\lambda) C(\lambda) d\lambda \quad Z = \int \bar{z}(\lambda) C(\lambda) d\lambda \quad (2.2)$$

Diese künstlichen Primärvalenzen repräsentieren einen Helligkeitsanteil Y und zwei Farbanteile X und Z . Der Helligkeitsanteil Y ist die so genannte Luminanz (*luminance*), ein gemäß der spektralen Lichtempfindlichkeit des Menschen gewichtetes Intensitätsmaß. Das XYZ -Modell spielt eine wichtige Rolle in der Farbmessung und ist die Basis vieler anderer Farbmodelle.

2.1.3 Normfarbtafel

Um eine Referenzdarstellung aller möglichen Farben zu erreichen, wird durch eine Normierung des XYZ -Systems zur Normfarbtafel (*Chromaticity Diagram*) übergegangen.

$$x = \frac{X}{X + Y + Z} \quad y = \frac{Y}{X + Y + Z} \quad z = \frac{Z}{X + Y + Z} \quad (2.3)$$

Aufgrund der linearen Abhängigkeit der Gleichungen ($z = 1 - (x + y)$) wird dabei eine Komponente redundant. In der xy -Ebene ergibt sich eine hufeisenförmige Region, in der alle möglichen Farben lokalisiert sind (vgl. Abbildung 2.5). An der gebogenen äußeren Begrenzung des „Hufeisens“ sind die Spektralfarben zu finden; die gedachte Verbindungslinie zwischen kurzen und langen Wellenlängen enthält die Magentatöne (Purpurlinie). Magentatöne sind keine Spektralfarben, sondern lassen sich nur als Mischfarben erzeugen.

Anhand der Normfarbtafel werden Farben durch die Angabe ihrer x - und y -Koordinaten eindeutig festgelegt.

2.2 Farbräume

2.2.1 Linearer RGB-Farbraum und gamma-korrigierter $R'G'B'$ -Farbraum

Der RGB-Farbraum ist in der technischen Farbbildverarbeitung der Dreh- und Angelpunkt. Sowohl Aufnahmegeräte wie Kameras, Scanner, etc. als auch Wiedergabegeräte wie Bildschirme arbeiten mit den Grundfarben Rot, Grün und Blau als Primärvalenzen. In Abhängigkeit der exakten Farbe der Primärvalenzen und des Weißpunktes existieren verschiedene Varianten des RGB-Farbraumes [SBS99, Poy96], wie ISO RGB, sRGB, ROMM RGB, Adobe RGB, Apple RGB, NTSC RGB und EBU RGB. Für die moderne Farbbildverarbeitung ist der im Standard ITU-R BT.709 für das hochauflösende Fernsehen (HDTV) festgeschriebene RGB-Farbraum repräsentativ.

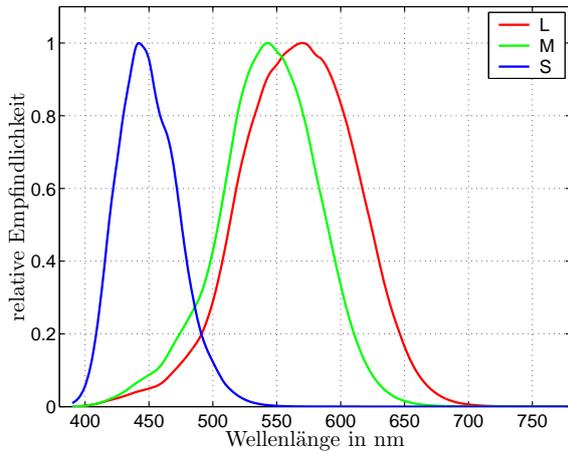


Abbildung 2.2: Normierte spektrale Empfindlichkeiten der drei Zapfentypen des menschlichen Auges bei einem Öffnungswinkel von 2° (2° -Betrachter).

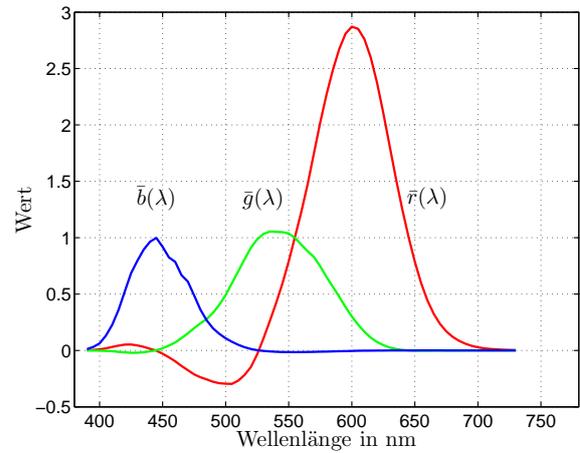


Abbildung 2.3: RGB-Spektralwertkurven nach STILES und BURCH (1959) [SB59]. Negative Werte bedeuten, dass die Primärvalenz zur Referenzfarbe hinzugesetzt werden muss, um eine Übereinstimmung zu erreichen (vgl. Abschnitt 2.1.1).

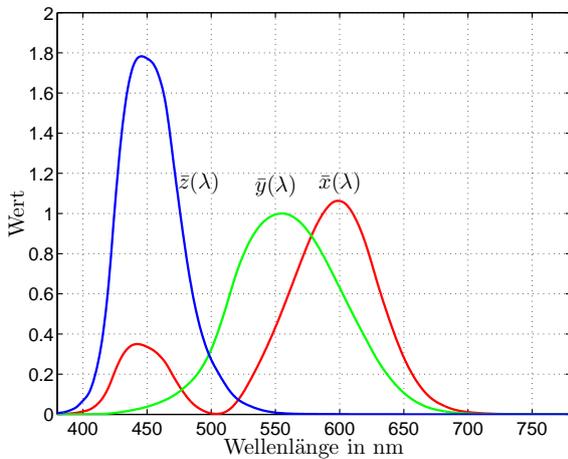


Abbildung 2.4: CIE-1931-Normspektralwertkurven (vgl. Abschnitt 2.1.2).

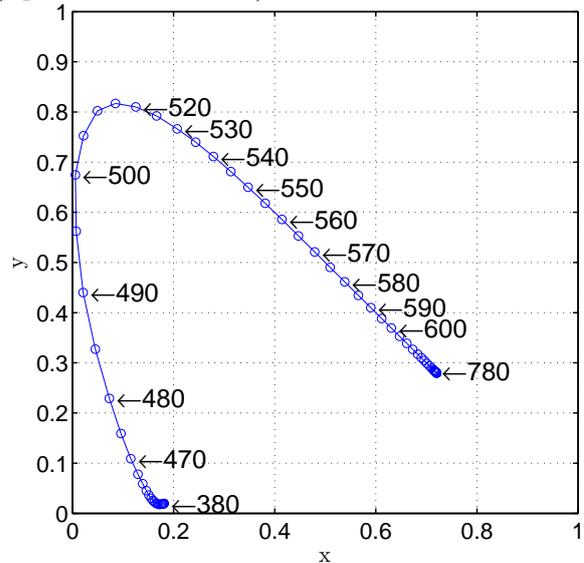


Abbildung 2.5: Innerhalb der hufeisenförmigen Region der CIE-Normfarbtafel sind alle möglichen Farbwerte zu finden. An der Begrenzung sind die Spektralfarben lokalisiert, die gedachte Verbindungslinie zwischen den Enden des Spektrums (380 nm \rightarrow 780 nm) enthält die Purpurtöne (vgl. Abschnitt 2.1.3).

Gammakorrektur

Die von einem Wiedergabegerät erzeugte Lichtintensität hängt nicht notwendigerweise *linear* mit den darzustellenden RGB-Werten zusammen. Die für die Bildwiedergabe seit Jahrzehnten bewährten Kathodenstrahlröhren (*cathode ray tube, CRT*) folgen beispielsweise einem Potenzgesetz (Exponent $\gamma \approx 2.2$) zwischen Steuerspannung und erzeugter Lichtintensität. Deshalb wird in der Regel bereits im Aufnahmegerät eine Gammakorrektur durchgeführt, die diese Nichtlinearität kompensiert. Im Standard ITU-R BT.709 wird die Abbildung der aufgenommenen linearen Intensität folgendermaßen festgeschrieben [Poya]:

$$R' = \begin{cases} 4.5 R & R \leq 0.018 \\ 1.099 R^{\frac{1}{\gamma}} - 0.099 & R > 0.018 \end{cases} \quad \text{mit} \quad \frac{1}{\gamma} = 0.45 \quad (2.4)$$

Die Gleichung gilt für die Transformation $G \rightarrow G'$ und $B \rightarrow B'$ entsprechend. Der Farbraum der gamma-korrigierten *RGB*-Werte soll als $R'G'B'$ -Farbraum bezeichnet werden.

In der Literatur wird häufig nicht klar zwischen linearen *RGB*- und gamma-korrigierten $R'G'B'$ -Werten unterschieden. Statt dessen werden die gamma-korrigierten Werte ebenfalls mit *RGB* bezeichnet, was leicht zu Verwechslungen führen kann. Für eine grobe Orientierung sei an dieser Stelle gesagt, dass technische Farbräume (wie sie z.B. zur Übertragung und Speicherung von analogen und digitalen Bildsignalen verwendet werden) aus den gamma-korrigierten $R'G'B'$ -Werten abgeleitet werden, während Normfarb-Farbräume (CIE-XYZ und daraus abgeleitete Farbräume) aus den intensitätslinearen *RGB*-Werten abzuleiten sind.

Zwischen dem linearen *RGB*-Farbraum nach ITU-R BT.709 und dem CIE-XYZ-Farbraum besteht der lineare Zusammenhang

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 3.240497 & -1.537150 & -0.498535 \\ -0.969256 & 1.875992 & 0.041556 \\ 0.055648 & -0.204043 & 1.057311 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}. \quad (2.5)$$

2.2.2 Lineare Transformationen von $R'G'B'$

Für eine effiziente Übertragung und Speicherung von Farbbildern ist der $R'G'B'$ -Farbraum schlecht geeignet. Die Komponenten weisen in der Regel eine sehr große Korrelation auf, beanspruchen also wesentlich mehr Bandbreite bzw. Speicherplatz als nötig, weil Helligkeits- und Farbinformation *gemeinsam* (in korrelierter Form) vorliegen.

Dieses Problem äußerte sich in der Geschichte erstmalig bei der Entwicklung des Farbfernsehens, wobei zwei wesentliche Kompatibilitätsprobleme zu lösen waren: Zum einen mussten bestehende Schwarz-Weiß-Geräte in Farbe ausgestrahlte Sendungen ebenso empfangen können wie die neu zu etablierenden Farbempfänger, zum anderen musste zusätzlich das Farbsignal im zur Verfügung stehenden Frequenzband untergebracht werden. Eine direkte Übertragung der $R'G'B'$ -Werte war dazu völlig ungeeignet.

Bereits im Jahr 1938 entwickelte und patentierte der französische Telekommunikationsingenieur VALENSI dazu die Idee, ein Farbsignal in seine Helligkeits- und Farbinformation aufzuspalten. Alle heutzutage etablierten Farbbildübertragungssysteme greifen auf diese Idee zurück. Dazu werden aus dem $R'G'B'$ -Signal die Helligkeitsinformation $Y' = Y_r R' + Y_g G' + Y_b B'$ (durch eine gewichtete Aufsummierung der Kanäle) sowie die Farbdifferenzen $B' - Y'$ und $R' - Y'$ extrahiert. Der in der Fachliteratur für die Helligkeitskomponente häufig verwendete Begriff „Luminanz“ ist an dieser Stelle nicht korrekt. Laut Definition der CIE ist die Luminanz proportional zur spektralen Leistungsdichte, gewichtet mit einer spektralen Empfindlichkeitsfunktion. Korrekterweise sollte der Begriff „Luminanz“ deshalb ausschließlich für die Größe Y aus dem XYZ-System (vgl. Abbildung 2.1.2) verwendet werden. POYNTON schlägt in seinem Appell [Poyb] für die aus den

gamma-korrigierten $R'G'B'$ -Werten abgeleitete Helligkeitsinformation Y' den Begriff „Luma“ vor. Des Weiteren warnt er vor dem unachtsamen Gebrauch des Begriff „YUV“ in der Literatur, welcher häufig für die im Anhang A.1 vorgestellten Farbräume verwendet wird, für die es jedoch korrekte Bezeichnungen gibt.

2.2.3 Wahrnehmungsorientierte Farbräume

Die menschliche Interpretation von Farbinformation orientiert sich (statt an Primärvalenzen) an den folgenden Eigenschaften:

- Buntton (*Hue*) (rot — gelb — grün — türkis — blau — violett — rot)
- Farbsättigung (*Saturation*) (satte Färbung — Pastelltöne — Grauton)
- Helligkeit (*Intensity, Lightness, Value*)

Die am Buntton orientierten technischen Farbräume HSI , HLS und HSV tragen diesem Ansatz Rechnung [PV00, Fre88].

2.2.4 Wahrnehmungsgetreue Farbräume mit visueller Gleichabständigkeit

Es gab in der Geschichte viele Versuche, Farben anhand ihrer Affinität (d.h. auf Grundlage der „empfindungsgemäßen Gleichabständigkeit“) in einem System anzuordnen. Dabei ist insbesondere der Ansatz des amerikanischen Malers Albert Henry Munsell (1858–1918) bis heute von Bedeutung. In seiner ursprünglichen Arbeit „A Color Notation“ [Mun05] schlug er eine kugelförmige Anordnung von Farbproben vor, jedoch stellte sich die Kugelsymmetrie als ungünstig für die empfindungsgemäße Gleichabständigkeit heraus. Seine Bemühungen, Farben derart in einem System anzuordnen, dass die Abstände jeder Farbprobe zur Nachbarprobe als gleich groß empfunden wird, mündeten schließlich in seinem 1915 veröffentlichten Farbatlas (*Color Atlas*) [Mun15], mit dem sich jede Farbe durch ein Tripel

- Buntton (*Hue, H*)
- Schwarz-Weiß-Wert (*Value, V*)
- Farbsättigung (*Chroma, C*)

beschreiben lässt.

CIE $L^*a^*b^*$ und CIE $L^*u^*v^*$

Für eine objektive Bewertung von Farbunterschieden (Farbmetrik) wurden von der CIE im Jahr 1976 zwei Farbräume definiert, in denen der euklidische Abstand zweier Farben annähernd mit dem empfindungsgemäßen „Farbabstand“ übereinstimmt. Beide Farbräume definieren eine Komponente für die empfundene Helligkeit (*Lightness L^**) sowie zwei Chrominanzkomponenten. Die Transformationsbeziehungen sind im Anhang A.3 angegeben.

Kapitel 3

Bildsegmentierung

Nach diesen Ausführungen zur Wahrnehmung und Repräsentation von Farbe wird in diesem Kapitel nun die Anwendung der Farbinformation zur Bildsegmentierung behandelt. Vor dem Hintergrund des zu realisierenden Bildauswertungssystems stehen dabei die folgenden zwei Teilprobleme im Mittelpunkt:

1. Segmentierung der gefärbten, genetisch transformierten Zellen
2. Segmentierung der Haustorien in den transformierten Zellen.

Dazu werden zunächst Untersuchungen zum Färbungsmerkmal der transformierten Zellen durchgeführt und ein adaptiver Farbraum hergeleitet, welcher eine (unter bestimmten Modellannahmen) bestmögliche Trennung zwischen den Klassen „Hintergrund“ und „transformierte Zelle“ bereitstellt. Es stellt sich heraus, dass dieser adaptive Farbraum sehr gut durch den Buntton (*Hue*) approximiert werden kann und dass daher eine aufwandsgünstige Vorsegmentierung der Bilder (Auffinden transformierter Zellen) mit Hilfe des Bunttons möglich ist. Um das Segmentierungsergebnis weiter zu verfeinern, wird die detektierte Region anschließend in den adaptiven Farbraum überführt. In dem entstehenden einkanaligen Bild kann nun die Berandung der transformierten Zelle sehr genau bestimmt werden. Dies ist für die später eingesetzten Algorithmen zur Haustoriensegmentierung notwendig.

Methodisch wird zur Zellsegmentierung zunächst eine Kantendetektion mit dem Canny-Kantendetektor auf Basis des Bunttons durchgeführt [IS03]. Die gefärbten, transformierten Zellen werden dabei mit einer geschlossenen Kante markiert. Anschließend werden auf diesen Bildausschnitt aktive Konturmodelle, so genannte Schlangen (*Snakes*) angewendet, um die Kontur der tatsächlichen Zellberandung möglichst gut anzunähern. Bei der Implementierung der aktiven Konturmodelle wurde ein neues Verfahren angewendet, welches die zyklische Struktur des zu lösenden Gleichungssystems ausnutzt und auf der schnellen Fouriertransformation (*Fast Fourier Transform, FFT*) basiert [IS05]. Dieses Verfahren ist allgemein gültig und reduziert die numerische Komplexität von $\mathcal{O}(N^2)$ auf $\mathcal{O}(N \log N)$.

Liegen die gefärbten Zellen schließlich räumlich isoliert vor, sind sie anschließend auf das eventuelle Vorhandensein von Haustorien des Mehltaupilzes zu untersuchen. Diese Saugorgane, über die sich der Pilz ernährt, deuten sich als fingerartige Strukturen innerhalb der Zelle an, heben sich jedoch bezüglich des Farbkontrasts nur sehr schwach vom Zellhintergrund ab. Zur Segmentierung der Haustorien werden zwei Verfahren vorgestellt: Eine Clusterung im RGB-Farbraum [IS04a] führt bei „gutmütigen“ Daten bereits zu zufrieden stellenden Ergebnissen. Als robuster erweist sich eine gezielte Hervorhebung der gesuchten Objekte mit Hilfe einer morphologischen Kontrastverstärkung [IS04b]. Um dieses Verfahren sinnvoll anwenden zu können, muss im Vorfeld eine qualitativ hochwertige Zellsegmentierung durchgeführt werden – die Berandung der Zelle muss möglichst genau bekannt sein. Dies ist darin begründet, dass die Zellwand häufig eine ähnliche Färbung aufweist wie Haustorien, so dass eine Kontrastverstärkung diese Bereiche ebenfalls



Abbildung 3.1: Ausschnitt eines Beispielbildes (*links*) und repräsentative Ausschnitte des Hintergrundes (*mittig*) und einer transformierten Zelle (*rechts*).

hervorheben würde. Durch ein exaktes „Ausschneiden“ der Zellen vor der Kontrastverstärkung wird dieses Problem vermieden.

Da bei dieser farbbasierten Segmentierung auch andere Objekte innerhalb der Zelle markiert werden können, ist es notwendig, anschließend eine Überprüfung der Ergebnisse durchzuführen. Selbst für den geschulten Blick eines Experten treten an dieser Stelle gelegentlich Zweifelsfälle auf, ob es sich bei dem betrachteten Objekt tatsächlich um ein Haustorium oder ein anderes Objekt (z.B. den Zellkern oder eine Verfärbung) handelt. Diese Entscheidung lässt sich nicht allein auf Basis der Farbinformation fällen. Vielmehr müssen die segmentierten Regionen einer *objektbasierten Klassifikation* unterzogen werden, in welcher fälschlicherweise als Haustorium segmentierte Objekte zurückgewiesen werden. Dies wird im zweiten Teil der Arbeit im Anschluss an die Thematik „Bildsegmentierung“ behandelt.

Im Vorfeld und während der Bearbeitung dieser Thematik wurden auch jüngst entwickelte Verfahren untersucht, welche versuchen, dem Ideal eines möglichst allgemein gültigen Segmentierungsverfahrens ein Stück näher zu kommen. Zwei Vertreter werden im Anhang B vorgestellt.

3.1 Ein adaptiver Farbraum für die Zellsegmentierung

Im folgenden Abschnitt wird die Beschreibung von Farben mit Hilfe eines parametrischen statistische Modells erörtert. Ziel ist ein adaptiver Farbraum, welcher das Farbbild in ein einkanaliges Bild überführt, in dem die zu separierenden Objekte bestmöglich getrennt werden können.

Abbildung 3.1 zeigt zwei repräsentative Ausschnitte für Bildbereiche aus den Klassen „Hintergrund“ bzw. „transformierte Zelle“. In Abbildung 3.2 ist links das zugehörige Streudiagramm (*Scatter Plot*) im RGB-Farbraum zu sehen, wobei die Klasse „Hintergrund“ bläulich und die Klasse „transformierte Zelle“ grünlich dargestellt ist. Die rechte Abbildung verdeutlicht die Modellbildung: Die Cluster in der linken Abbildung formen Punktwolken, welche sich ellipsoidal im Raum ausdünnen. Dementsprechend ist es möglich, diese Cluster mittels eines ellipsoidalen Wahrscheinlichkeitsdichtemodells parametrisch zu beschreiben. Dies führt auf ein sehr einfach zu handhabendes und weit verbreitetes Modell – die multivariate Gaußverteilung.

3.1.1 Parametrisches Datenmodell

Für die Beschreibung der Farbcluster über ein statistisches Modell werden die einzelnen Bildpunkte als statistisch unabhängige Proben (Samples) aufgefasst. Dies ignoriert natürlich die inhärente räumliche Korrelation der Farbinformation, die insbesondere im vorliegenden Bildmaterial gegeben ist. Es wird also zunächst keine Ortsinformation im Modell berücksichtigt – man könnte die Bildpunkte beliebig umsortieren, ohne dass sich an den Modellparametern etwas änderte.

Momente und multivariate Gaußverteilung

Folgt eine d -dimensionale Zufallsvariable einer multivariaten Gaußverteilung (wie dies für die Farbverteilung der Klassen „Hintergrund“ und „transformierte Zelle“ angenommen und im nächsten Abschnitt auch überprüft wird), so sind ihre statistischen Eigenschaften über den

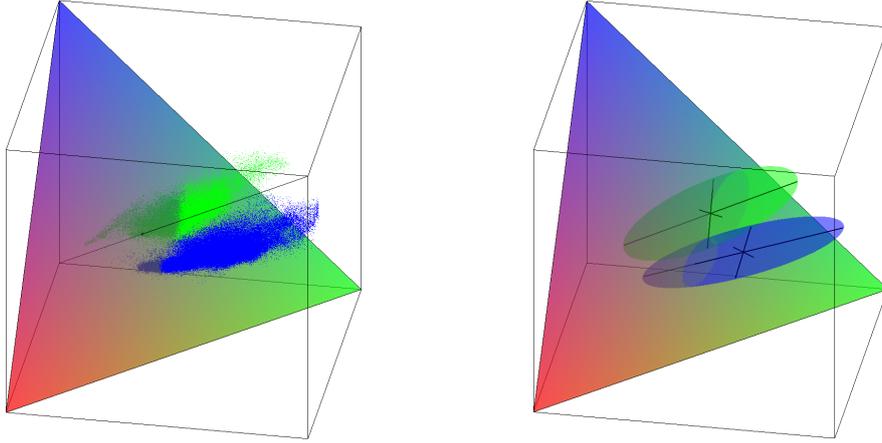


Abbildung 3.2: Farbcluster der Klassen „Hintergrund“ (blau) und „transformierte Zelle“ (grün) im RGB-Farbraum. *Links*: tatsächliche Cluster unter Verwendung der Bildausschnitte aus Abbildung 3.1. *Rechts*: Cluster approximiert mittels multivariater Gaußverteilung. Die Hülle der dargestellten Ellipsoiden beschreibt die dreifache Standardabweichung der Verteilung.

Mittelwertvektor $\boldsymbol{\mu}$ und die Kovarianzmatrix¹ $\boldsymbol{\Sigma}$ vollständig bestimmt.

$$\boldsymbol{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_d \end{bmatrix} \quad \boldsymbol{\Sigma} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \dots & \sigma_{1d} \\ \sigma_{21} & \sigma_{22} & \dots & \sigma_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{d1} & \sigma_{d2} & \dots & \sigma_{dd} \end{bmatrix} \quad (3.1)$$

Die entsprechende d -dimensionale Wahrscheinlichkeitsdichtefunktion (*Probability Density Function, PDF*) ist gegeben durch

$$p(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^d |\boldsymbol{\Sigma}|}} e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})}, \quad (3.2)$$

wobei $|\boldsymbol{\Sigma}|$ die Determinante der Kovarianzmatrix bezeichnet. In diesem Modell dünnt sich die Aufenthaltswahrscheinlichkeit der Samples ellipsoidal um den Mittelwert herum aus, wobei der Mittelwertvektor $\boldsymbol{\mu}$ den *Schwerpunkt* und die Kovarianzmatrix $\boldsymbol{\Sigma}$ die *Form und Orientierung* des Ellipsoiden im d -dimensionalen Raum beschreibt.

Parameterschätzung aus einer Stichprobe

Die Modellparameter

$$\boldsymbol{\mu} = \mathbb{E}[\mathbf{x}] \quad \text{und} \quad \boldsymbol{\Sigma} = \mathbb{E}[(\mathbf{x} - \boldsymbol{\mu})(\mathbf{x} - \boldsymbol{\mu})^T] \quad (3.3)$$

müssen aus zur Verfügung stehenden Stichproben der Zufallsvariablen geschätzt werden. Dies kann über die empirischen Momente der Stichprobe

$$\bar{\mathbf{x}} = \frac{1}{N} \sum_{n=1}^N \mathbf{x}_n \quad \text{und} \quad \mathbf{S} = \frac{1}{N} \sum_{n=1}^N (\mathbf{x}_n - \bar{\mathbf{x}})(\mathbf{x}_n - \bar{\mathbf{x}})^T \quad (3.4)$$

¹Die Kovarianzmatrix ist symmetrisch ($\sigma_{ij} = \sigma_{ji}$), die Kovarianz einer Zufallsvariablen mit sich selbst ist gleich ihrer Varianz ($\sigma_{ii} = \sigma_i^2$).

geschehen. Beide Schätzungen $\hat{\boldsymbol{\mu}} = \bar{\mathbf{x}}$ und $\hat{\boldsymbol{\Sigma}} = \mathbf{S}$ sind konsistent, d.h., sie konvergieren mit steigendem Stichprobenumfang gegen den wahren Parameter. Jedoch ist die Schätzung der Kovarianzmatrix nur asymptotisch (für einen sehr großen Stichprobenumfang N) erwartungstreu, da der Mittelwert aus derselben Stichprobe geschätzt wird und sich dadurch die Zahl der Freiheitsgrade um Eins verringert. Daher ist für eine erwartungstreue Schätzung einer relativ kleinen Stichprobe $\hat{\boldsymbol{\Sigma}} = \frac{N}{N-1}\mathbf{S}$ anzusetzen.

3.1.2 Validierung des parametrischen Modells

Zur Überprüfung der Zulässigkeit der Approximation der Farbverteilung mittels des multivariaten Gaußmodells sollen an dieser Stelle zwei Tests durchgeführt werden.

Grafischer Test I – Analyse der Hauptkomponenten

Ein erster Test auf das Vorliegen einer multivariaten Gaußverteilung ist die Analyse der eindimensionalen Randverteilungen (*Marginal Distributions*), da sich eine multivariate Gaußverteilung als Produkt eindimensionaler Gaußverteilungen ausdrücken lässt. Mit Hilfe der Eigenwertzerlegung der Inversen der Kovarianzmatrix $\boldsymbol{\Sigma}^{-1} = \mathbf{V}\boldsymbol{\Lambda}^{-1}\mathbf{V}^T$ lässt sich die multivariate Gaußverteilung schreiben als

$$p(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^d |\boldsymbol{\Sigma}|}} \exp \left[-\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})^T \mathbf{V} \boldsymbol{\Lambda}^{-1} \mathbf{V}^T (\mathbf{x} - \boldsymbol{\mu}) \right] \quad (3.5)$$

$$= \frac{1}{\sqrt{(2\pi)^d |\boldsymbol{\Sigma}|}} \exp \left[-\frac{1}{2} \sum_{i=1}^d (\mathbf{x} - \boldsymbol{\mu})^T \mathbf{v}_i \frac{1}{\lambda_i} \mathbf{v}_i^T (\mathbf{x} - \boldsymbol{\mu}) \right] \quad (3.6)$$

$$= \frac{1}{\sqrt{(2\pi)^d |\boldsymbol{\Sigma}|}} \prod_{i=1}^d \exp \left[-\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})^T \mathbf{v}_i \frac{1}{\lambda_i} \mathbf{v}_i^T (\mathbf{x} - \boldsymbol{\mu}) \right], \quad (3.7)$$

wobei der Vektor \mathbf{v}_i der i -te Eigenvektor (die i -te Spalte aus \mathbf{V}) ist. Mit Hilfe des folgenden Theorems, wonach die Determinante einer Matrix gleich dem Produkt ihrer Eigenwerte ist

$$|\boldsymbol{\Sigma}| = \prod_{i=1}^d \lambda_i, \quad (3.8)$$

lässt sich die multivariate Gaußverteilung letztendlich als Produkt eindimensionaler Gaußverteilungen entlang der Hauptkomponenten der Kovarianzmatrix schreiben.

$$p(\mathbf{x}) = \prod_{i=1}^d \frac{1}{\sqrt{2\pi\lambda_i}} \exp \left[-\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})^T \mathbf{v}_i \frac{1}{\lambda_i} \mathbf{v}_i^T (\mathbf{x} - \boldsymbol{\mu}) \right] \quad (3.9)$$

Dabei führt die Operation $(\mathbf{x} - \boldsymbol{\mu})^T \mathbf{v}_i$ bzw. $\mathbf{v}_i^T (\mathbf{x} - \boldsymbol{\mu})$ eine Projektion der zentrierten Daten auf die jeweilige Hauptachse (welche durch den Vektor \mathbf{v}_i beschrieben ist) durch.

Über eine Hauptachsentransformation lässt sich also zunächst die Frage beantworten, inwieweit die *Hauptkomponenten* der Daten jeweils einer *eindimensionalen* Gaußverteilung entsprechen. Nach Abbildung 3.3 ist die Normalverteilung der Hauptkomponenten der Klassen „Hintergrund“ und „transformierte Zelle“ in guter Näherung gewährleistet.

Die Normalverteilung der Randverteilungen ist zwar notwendig, aber nicht hinreichend für eine multivariate Gaußverteilung (siehe [Koz93] für ein Gegenbeispiel). Deshalb soll im Folgenden ein weiterer Test zu Rate gezogen werden.

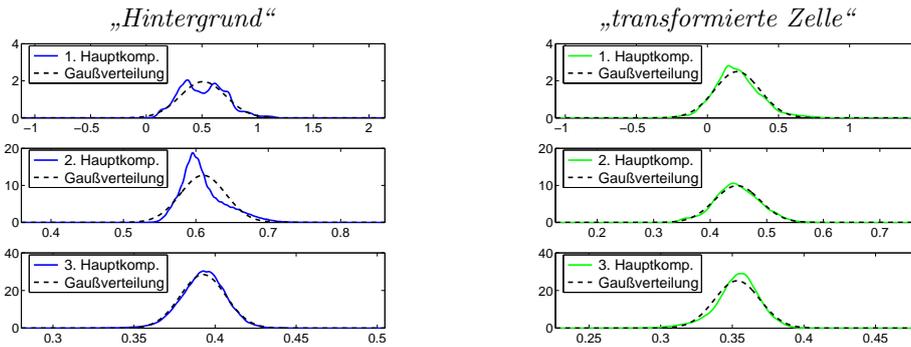


Abbildung 3.3: Wahrscheinlichkeitsdichtefunktion der Hauptkomponenten der in Abbildung 3.2 dargestellten Farbcluster von „Hintergrund“ (links) und „transformierte Zelle“ (rechts). Die Verteilungen spiegeln in guter Näherung Gaußverteilungen wider. Mit dieser Testmethode ergibt sich *kein* Anlass, die Modellannahme einer multivariaten Gaußverteilung abzulehnen.

Grafischer Test II – χ^2 -Test der Mahalanobis-Distanz

Werden Samples, welche aus d unkorrelierten, mittelwertfreien Gaußprozessen mit der Varianz $\sigma^2 = 1$ stammen, quadriert und aufsummiert, so entsteht eine χ^2 -Verteilung mit d Freiheitsgraden. Diese quadratische Aufsummierung geschieht genau bei der Berechnung der Mahalanobis-Distanz

$$d^2 = (\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}). \quad (3.10)$$

Mittels dieser Abbildung $\mathbb{R}^d \rightarrow \mathbb{R}^1$ lässt sich ein grafischer Test auf multivariate Normalverteilung auch im eindimensionalen Raum durchführen. Abbildung 3.4 zeigt links die Wahrscheinlichkeitsdichtefunktionen für die Mahalanobis-Distanzen der Klassen „Hintergrund“ und „transformierte Zelle“ und rechts die zugehörigen Quantil-Quantil-Plots (Q-Q-Plots) [NIS]. In einem Q-Q-Plot werden die Quantile der Stichprobe über den theoretischen Quantilen der angenommenen Verteilungsfunktion aufgetragen. Dazu sind die skalaren Prüfgrößen der Samples nach steigender Größe zu sortieren und der i -te Wert dieser Rangliste über dem Quantil der für die Prüfgröße angenommenen Verteilungsfunktion aufzutragen, für das die (kumulative) Verteilungsfunktion F (Cumulative Distribution Function, CDF) den Wert $(i - \frac{1}{2})/N$ annimmt. Im vorliegenden Fall wurden dazu die sortierten Mahalanobis-Distanzen $d_{(1)}^2 \leq \dots \leq d_{(i)}^2 \leq \dots \leq d_{(N)}^2$ der Samples über den Quantilen $x_q : F_{\chi_3^2}(x_q) = (i - \frac{1}{2})/N$ aufgetragen. Die Ergebnisse aus Abbildung 3.4 zeigen eine recht gute Übereinstimmung von theoretischer und praktischer Verteilung.

Fazit

Bei der Wahl eines Modells ist immer ein Kompromiss zwischen Anpassungsgüte und Komplexität zu finden. Insbesondere vor dem Hintergrund der praktischen Einsetzbarkeit ist dabei einem möglichst *einfachen* Modell der Vorzug zu geben. In der Wissenschaft wird dieses Prinzip häufig mit der Bezeichnung „Ockhams Rasiermesser“² (*Occam's razor*) umschrieben [For00]. Nach diesem Prinzip und bedingt durch den Fakt, dass die auf der multivariaten Gaußverteilung basierenden adaptiven Farbräume vor dem Hintergrund der geschilderten Anwendung *praxistauglich* sind, ist dieses Modell die erste Wahl.

3.1.3 Parameterschätzung, Expectation Maximization (EM)-Algorithmus

Eine Beschreibung von Farbe mittels eines parametrischen Modells erfordert die korrekte Bestimmung der Modellparameter. Bei Farbänderungen ist eine entsprechende Parameteranpas-

²Wilhelm von Ockham (1285–1349) hat dieses Prinzip in seinen Schriften gebraucht.

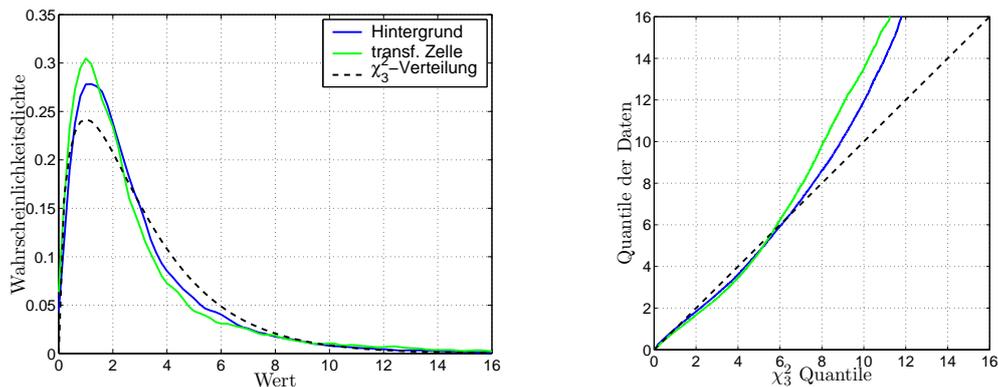


Abbildung 3.4: Wahrscheinlichkeitsdichte der Mahalanobis-Distanz (*links*) und Quantil-Quantil-Darstellung (Q-Q-Plot) (*rechts*). Dieser Test ist neben Abbildung 3.3 ein weiteres Indiz für die Zulässigkeit der Modellannahme einer multivariaten Gaußverteilung.

sung erforderlich. Für die Modellierung des Bildhintergrundes und der transformierten Zellen bedeutet dies, dass für jedes Bild eine Schätzung der aktuellen Modellparameter notwendig ist.

Mischmodell

Die bisherigen Überlegungen führen zu der Modellannahme, dass sich das Bild in $K = 2$ Klassen („Hintergrund“ und „transformierte Zelle“) zerlegen lässt, wobei jede Klasse mittels einer multivariaten Gaußverteilung beschrieben wird (*Mixture of Gaussians*). Damit lässt sich die Farbstatistik des Gesamtbildes durch eine Linearkombination von Wahrscheinlichkeitsdichtefunktionen beschreiben, wobei die Wichtungsfaktoren die A-priori-Wahrscheinlichkeiten der einzelnen Klassen darstellen. Allgemein setzt sich die Gesamtwahrscheinlichkeitsdichtefunktion $p(\mathbf{x})$ folgendermaßen aus K Komponenten zusammen:

$$p(\mathbf{x}) = \sum_{k=1}^K P_k p(\mathbf{x} | k) \quad \text{mit} \quad \sum_{k=1}^K P_k = 1 \quad \text{und} \quad \int_{\mathbf{x}} p(\mathbf{x} | k) d\mathbf{x} = 1. \quad (3.11)$$

Bayes' Theorem

BAYES' Theorem [Bay53] gibt die Wahrscheinlichkeit an, mit der ein Sample \mathbf{x}_n zur Klasse k gehört.

$$P(k | \mathbf{x}_n) = \frac{P_k p(\mathbf{x}_n | \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)}{\sum_{j=1}^K P_j p(\mathbf{x}_n | \boldsymbol{\mu}_j, \boldsymbol{\Sigma}_j)} \quad (3.12)$$

Dazu müssen die Parameter der Klassen (das sind im Fall des hier verwendeten multivariaten Gaußmodells $\boldsymbol{\mu}_k$ und $\boldsymbol{\Sigma}_k$) und die A-priori-Wahrscheinlichkeiten der Klassen, P_k , bekannt sein. Die Parameter $\boldsymbol{\mu}$ und $\boldsymbol{\Sigma}$ sind nicht konstant, sondern je nach Probenmaterial und Ausprägung der Zellfärbung unterschiedlich. Weiterhin ist die A-priori-Wahrscheinlichkeit der beiden Klassen schwierig festzulegen.

Wünschenswert wäre an dieser Stelle, diese Parameter aus dem jeweils vorliegenden Bild schätzen zu können. Dies leistet der Expectation Maximization (EM)-Algorithmus, der im Folgenden vorgestellt wird.

EM-Algorithmus

Der Expectation Maximization (EM)-Algorithmus [DLR77, RW84] löst die Aufgabe, die einzelnen Bestandteile eines Modells aus den vorliegenden Daten zu bestimmen. Es handelt sich dabei um ein iteratives zweistufiges Verfahren, welches für den Fall des multivariaten Gauß-Mischmodells folgendermaßen abläuft:

In jedem Iterationsdurchlauf t werden zunächst im E-Schritt (*Expectation Step*) die Klassenzugehörigkeitswahrscheinlichkeiten für jedes Sample $\mathbf{x}_n \in \mathbb{R}^{d \times 1}$ der zu entmischenden Daten $\mathbf{X} \in \mathbb{R}^{d \times N}$ berechnet.

$$P^t(k | \mathbf{x}_n) = \frac{P_k^t p(\mathbf{x}_n | \boldsymbol{\mu}_k^t, \boldsymbol{\Sigma}_k^t)}{\sum_{j=1}^K P_j^t p(\mathbf{x}_n | \boldsymbol{\mu}_j^t, \boldsymbol{\Sigma}_j^t)} \quad (3.13)$$

In der nachfolgenden Parameteraktualisierung (M-Schritt, *Maximization Step*) werden die A-priori-Wahrscheinlichkeiten, der Mittelwertvektor und die Kovarianzmatrix mit neuen Werten belegt.

$$P_k^{t+1} = \frac{1}{N} \sum_{n=1}^N P^t(k | \mathbf{x}_n) \quad (3.14)$$

$$\boldsymbol{\mu}_k^{t+1} = \frac{1}{NP_k^{t+1}} \sum_{n=1}^N P^t(k | \mathbf{x}_n) \mathbf{x}_n \quad (3.15)$$

$$\boldsymbol{\Sigma}_k^{t+1} = \frac{1}{NP_k^{t+1}} \sum_{n=1}^N P^t(k | \mathbf{x}_n) (\mathbf{x}_n - \boldsymbol{\mu}_k^{t+1})(\mathbf{x}_n - \boldsymbol{\mu}_k^{t+1})^T \quad (3.16)$$

In jeder Iteration ergibt sich eine verbesserte Schätzung der Parameter P_k , $\boldsymbol{\mu}_k$ und $\boldsymbol{\Sigma}_k$.

Der EM-Algorithmus erfordert geeignete Startwerte. Diese können jedoch aufgrund der Kenntnis des zu parametrierenden Bildmaterials leicht bereitgestellt werden.

3.1.4 Herleitung des adaptiven Farbraumes

Nachdem die Bereitstellung der Modellparameter mit Hilfe des EM-Algorithmus gesichert ist, soll nun konkret erörtert werden, wie ein Farbbild in ein äquivalentes *einkanaliges* Bild (Merkmalsbild) überführt wird, welches die Klassenzugehörigkeitsinformation in Form von Grauwerten repräsentiert. Auf dieses Merkmalsbild können dann die aus der Graubildverarbeitung bekannten Segmentierungsalgorithmen (z.B. Schwellwerthaltung, Kantendetektion, aktive Konturmodelle, etc.) angewendet werden.

Die Methodik ist aus der statistischen Klassifikation entlehnt, orientiert sich aber stark an dem Ziel, statt einer Binärentscheidung (Klasse 1 oder Klasse 2) vielmehr die *Unschärfe* der Klassenzugehörigkeit als Information bereitzustellen.

Distanz eines Samples zur Klasse k

Die Klasse, zu der ein Sample \mathbf{x} am wahrscheinlichsten gehört, kann – wie vorhin gezeigt – mit Hilfe von BAYES' Theorem bestimmt werden. Da der Nenner von Gleichung (3.12) für alle Klassen gleich ist, muss dieser dazu nicht explizit berechnet werden, sondern es genügt, die Klasse auszuwählen, welche den Zähler

$$P_k p(\mathbf{x} | \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) = \frac{P_k}{\sqrt{(2\pi)^d |\boldsymbol{\Sigma}_k|}} e^{-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_k)^T \boldsymbol{\Sigma}_k^{-1} (\mathbf{x} - \boldsymbol{\mu}_k)} \quad (3.17)$$

maximiert. Durch Logarithmierung und Multiplikation mit -2 ist zu sehen, dass dies äquivalent zu jener Klasse ist, welche

$$-2 \ln(P_k) + \ln(|\Sigma_k|) + d \ln(2\pi) + (\mathbf{x} - \boldsymbol{\mu}_k)^T \Sigma_k^{-1} (\mathbf{x} - \boldsymbol{\mu}_k) \quad (3.18)$$

bzw.

$$D_k^2 = -2 \ln(P_k) + \ln(|\Sigma_k|) + (\mathbf{x} - \boldsymbol{\mu}_k)^T \Sigma_k^{-1} (\mathbf{x} - \boldsymbol{\mu}_k) \quad (3.19)$$

minimiert. In der letzten Gleichung wurde der Ausdruck $d \ln(2\pi)$ weggelassen, da er für alle Klassen konstant ist. Der Ausdruck $D_k = \sqrt{D_k^2}$ gibt die Distanz eines Samples \mathbf{x} zum Klassenmittelpunkt $\boldsymbol{\mu}_k$ (gemessen im Hauptachsen-Koordinatensystem der Klasse k) an. Die Achsen dieses Koordinatensystems sind in die Richtungen maximaler Varianz der Klasse orientiert und entsprechend der Standardabweichungen der Klasse skaliert. Das heißt, der Ausdruck D_k gibt den Abstand zu $\boldsymbol{\mu}_k$ im Vergleich zur Standardabweichung in Richtung \mathbf{x} an. Flächen mit gleichem Abstand bilden – je nach Anzahl der Dimensionen d – Ellipsen ($d = 2$), Ellipsoiden ($d = 3$) bzw. Hyperellipsoiden ($d > 3$). Auf diesen (Hyper)Flächen ist die Wahrscheinlichkeitsdichte $p(\mathbf{x} | \boldsymbol{\mu}_k, \Sigma_k)$ konstant.

Quadratische Entscheidungsfläche für $K = 2$ Klassen

Im nächsten Schritt soll eine grafische Interpretation der Klassentrennung abgeleitet werden, welche letztlich auf adaptive, speziell auf das Problem der Zellsegmentierung angepasste Farbräume führt. Die hier gezeigte Herleitung ist in etwas weniger ausführlicher Form in [GJJ96] zu finden.

Mit Hilfe des Distanzmaßes aus Gleichung (3.19) lässt sich jene Trennfläche im Raum bestimmen, welche die zwei Klassen i und j – gegeben durch ihre Parameter $\boldsymbol{\mu}_i, \Sigma_i$ und $\boldsymbol{\mu}_j, \Sigma_j$ – separiert. Samples auf der einen bzw. anderen Seite dieser Fläche gehören entsprechend zur Klasse i bzw. Klasse j . Auf der Trennfläche muss das Distanzmaß für beide Klassen gleich sein: $D_i^2 = D_j^2$.

$$\begin{aligned} & -2 \ln(P_i) + \ln(|\Sigma_i|) + (\mathbf{x} - \boldsymbol{\mu}_i)^T \Sigma_i^{-1} (\mathbf{x} - \boldsymbol{\mu}_i) \\ = & -2 \ln(P_j) + \ln(|\Sigma_j|) + (\mathbf{x} - \boldsymbol{\mu}_j)^T \Sigma_j^{-1} (\mathbf{x} - \boldsymbol{\mu}_j) \end{aligned} \quad (3.20)$$

Da Kovarianzmatrizen immer symmetrisch sind, kann der Term $(\mathbf{x} - \boldsymbol{\mu}_i)^T \Sigma_i^{-1} (\mathbf{x} - \boldsymbol{\mu}_i)$ umgeformt werden zu

$$\begin{aligned} (\mathbf{x} - \boldsymbol{\mu}_i)^T \Sigma_i^{-1} (\mathbf{x} - \boldsymbol{\mu}_i) &= \mathbf{x}^T \Sigma_i^{-1} \mathbf{x} - \mathbf{x}^T \Sigma_i^{-1} \boldsymbol{\mu}_i - \boldsymbol{\mu}_i^T \Sigma_i^{-1} \mathbf{x} + \boldsymbol{\mu}_i^T \Sigma_i^{-1} \boldsymbol{\mu}_i \\ &= \mathbf{x}^T \Sigma_i^{-1} \mathbf{x} - 2\boldsymbol{\mu}_i^T \Sigma_i^{-1} \mathbf{x} + \boldsymbol{\mu}_i^T \Sigma_i^{-1} \boldsymbol{\mu}_i. \end{aligned} \quad (3.21)$$

Nun lassen sich die Konstanten aus Gleichung (3.20) zusammenfassen als

$$c_q = -2 \ln(P_j) + 2 \ln(P_i) + \ln(|\Sigma_j|) - \ln(|\Sigma_i|) + \boldsymbol{\mu}_j^T \Sigma_j^{-1} \boldsymbol{\mu}_j - \boldsymbol{\mu}_i^T \Sigma_i^{-1} \boldsymbol{\mu}_i, \quad (3.22)$$

so dass diese weiter vereinfacht werden kann zu

$$\mathbf{x}^T \Sigma_i^{-1} \mathbf{x} - 2\boldsymbol{\mu}_i^T \Sigma_i^{-1} \mathbf{x} = \mathbf{x}^T \Sigma_j^{-1} \mathbf{x} - 2\boldsymbol{\mu}_j^T \Sigma_j^{-1} \mathbf{x} + c_q \quad (3.23)$$

$$0 = \mathbf{x}^T \Sigma_j^{-1} \mathbf{x} - \mathbf{x}^T \Sigma_i^{-1} \mathbf{x} - 2\boldsymbol{\mu}_j^T \Sigma_j^{-1} \mathbf{x} + 2\boldsymbol{\mu}_i^T \Sigma_i^{-1} \mathbf{x} + c_q \quad (3.24)$$

$$0 = \mathbf{x}^T \underbrace{(\Sigma_j^{-1} - \Sigma_i^{-1})}_{\mathbf{A}_q} \mathbf{x} - 2 \underbrace{(\boldsymbol{\mu}_j^T \Sigma_j^{-1} - \boldsymbol{\mu}_i^T \Sigma_i^{-1})}_{\mathbf{b}_q^T} \mathbf{x} + c_q. \quad (3.25)$$

Entstanden ist eine quadratische Gleichung

$$\mathbf{x}^T \mathbf{A}_q \mathbf{x} + \mathbf{b}_q^T \mathbf{x} + c_q = 0 \quad (3.26)$$

mit den Parametern

$$\mathbf{A}_q = \boldsymbol{\Sigma}_j^{-1} - \boldsymbol{\Sigma}_i^{-1} \quad (3.27)$$

$$\mathbf{b}_q = -2\boldsymbol{\Sigma}_j^{-1}\boldsymbol{\mu}_j + 2\boldsymbol{\Sigma}_i^{-1}\boldsymbol{\mu}_i \quad (3.28)$$

$$c_q = -2\ln(P_j) + 2\ln(P_i) + \ln(|\boldsymbol{\Sigma}_j|) - \ln(|\boldsymbol{\Sigma}_i|) + \boldsymbol{\mu}_j^T \boldsymbol{\Sigma}_j^{-1} \boldsymbol{\mu}_j - \boldsymbol{\mu}_i^T \boldsymbol{\Sigma}_i^{-1} \boldsymbol{\mu}_i. \quad (3.29)$$

Im dreidimensionalen Raum ($d = 3$) nimmt diese quadratische Fläche eine von 17 möglichen Gestalten in Form von Ellipsoiden, Kegeln, einschaligen bzw. zweisechaligen Hyperboloiden, Paraboloiden, Zylindern oder Ebenenpaaren an. Diese Gestalteigenschaften lassen sich aus den Parametern der Gleichung ableiten (siehe dazu [BSMM01, Wei]).

Lineare Entscheidungsfläche für $K = 2$ Klassen

Für den Fall, dass die Kovarianzmatrizen beider Klassen gleich sind, wird die Entscheidungsfläche linear. Eine Linearisierung lässt sich auch erreichen, indem die Kovarianzmatrizen als gleich *angenommen* werden. Die allgemeine Gleichung für die Berechnung einer solchen Misch-Kovarianzmatrix für K Klassen ist

$$\boldsymbol{\Sigma}_l = \frac{1}{N - K} \sum_{i=1}^K (n_i - 1) \boldsymbol{\Sigma}_i, \quad (3.30)$$

wobei n_i die Anzahl der Samples in der Klasse i und N die Gesamtanzahl der Samples ist. Die lineare Entscheidungsfläche ist nun gegeben durch

$$\mathbf{b}_l^T \mathbf{x} + c_l = 0 \quad (3.31)$$

mit den Parametern

$$\mathbf{b}_l = -2\boldsymbol{\Sigma}_l^{-1}\boldsymbol{\mu}_j + 2\boldsymbol{\Sigma}_l^{-1}\boldsymbol{\mu}_i = 2\boldsymbol{\Sigma}_l^{-1}(\boldsymbol{\mu}_i - \boldsymbol{\mu}_j) \quad (3.32)$$

$$c_l = -2\ln(P_j) + 2\ln(P_i) + \boldsymbol{\mu}_j^T \boldsymbol{\Sigma}_l^{-1} \boldsymbol{\mu}_j - \boldsymbol{\mu}_i^T \boldsymbol{\Sigma}_l^{-1} \boldsymbol{\mu}_i. \quad (3.33)$$

Von der Entscheidungsfläche zum adaptiven Farbraum

Die Lage eines Samples im Raum führt zu einer Binärentscheidung, bei der die Samples einer der beiden Klassen zugeordnet wird (Maximum Likelihood Entscheidung). Aus Sicht der Bildverarbeitung ist jedoch an dieser Stelle kein binäres Ausgangssignal gewünscht, sondern ein Grauwertbild, welches den *Abstand* der Farbwerte des Eingangsbildes zu den vorgegebenen Klassen (bzw. der die Klassen separierenden Fläche) widerspiegelt. Ergebnis ist ein adaptiver Farbraum, der die Farbsamples $\mathbf{x} \in \mathbb{R}^3$ auf eine reelle Zahl $d \in \mathbb{R}$ abbildet.

Linearisierter Fall Für den adaptiven Farbraum auf Basis der *linearen* Entscheidungsfläche nach Gleichung (3.31) ist dazu der Abstand eines Punktes zu einer Ebene zu bestimmen.

$$d_l = \mathbf{b}_l^T \mathbf{x} + c_l \quad (3.34)$$

Dabei ist der Ebenen-Normalenvektor auf $\|\mathbf{b}_l\| = 1$ zu normieren (Normaleneinheitsvektor).

Quadratischer Fall – lotrechte Distanz Für quadratische Flächen existieren verschiedene Definitionsmöglichkeiten eines Abstandsmaßes [TK03]. Dabei stellt sich die *lotrechte Distanz* (*perpendicular distance*) aus intuitiver Sicht zunächst als die vernünftigste Wahl dar, da sie einer unmittelbaren Verallgemeinerung des Abstandes eines Punktes zu einer Ebene entspricht.

Die lotrechte Distanz ist definiert als der Euklidische Abstand zwischen dem Sample \mathbf{x} und dem Punkt \mathbf{z} auf der quadratischen Entscheidungsfläche, der \mathbf{x} am nächsten liegt.

$$d_p = \min_{\mathbf{z}} \|\mathbf{x} - \mathbf{z}\| \quad \text{mit} \quad \mathbf{z}^T \mathbf{A}_q \mathbf{z} + \mathbf{b}^T \mathbf{z} + c_q = 0 \quad (3.35)$$

Die Berechnung dieses Distanzmaßes ist jedoch recht aufwändig, weil dazu für jedes Sample die Nullstellen von Polynomen sechsten Grades bestimmt werden müssen [LS02, Len05, TK03]. Abbildung 3.5 *links* zeigt einen Teil der quadratischen Trennfläche, welche im vorliegenden Fall zu einem einschaligen Hyperboloiden gehört, der sich wie eine „Röhre“ um die Klasse „*Hintergrund*“ legt. In Abbildung 3.9 *links* ist ein größerer Ausschnitt der Trennfläche dargestellt, so dass ihre Form besser zu erkennen ist.

An dieser Stelle sei darauf hingewiesen, dass das Distanzmaß bei Flächenpaaren, wie z.B. bei zweischaligen Hyperboloiden (Abbildung 3.9 *rechts*) unter Umständen nicht monoton ist: In Abbildung 3.9 *rechts* liegt die Klasse „*Hintergrund*“ zwischen den beiden Schalen des Hyperboloiden, wobei die rechte Schale in der Abbildung die beiden Cluster trennt. Bewegt man sich von der rechten Schale weg hin zur linken Schale, so wird die lotrechte Distanz zunächst größer – dann jedoch wieder kleiner, weil der Abstand zur *linken* Schale zum Tragen kommt.

Das Vorzeichen von d_p muss aus der Lage des Samples zur Fläche bzw. mittels der algebraischen Distanz (siehe nächster Abschnitt) bestimmt werden.

Quadratischer Fall – algebraische Distanz Die *algebraische Distanz* (*algebraic distance*) d_a ergibt sich direkt durch Einsetzen des Samples in die Definitionsgleichung der quadratischen Entscheidungsfläche.

$$d_a = \mathbf{x}^T \mathbf{A}_q \mathbf{x} + \mathbf{b}_q^T \mathbf{x} + c_q \quad (3.36)$$

Eine anschauliche Deutung dieses Distanzmaßes ist im Kontext der verallgemeinerten Distanzen zu den Klassen (vgl. Gleichung (3.19)) möglich, denn d_a spiegelt das Distanzmaß $D_j^2 - D_i^2$ wider. So ergeben sich – je nach Klassenzugehörigkeit des Samples – positive oder negative Werte für d_a . Abbildung 3.10 stellt die Ergebnisse bei der Nutzung dieses Distanzmaßes als adaptiver Farbraum dar.

Radiale Distanzmaße Die radiale Distanz (*radial distance*) bzw. normierte radiale Distanz (*normalised radial distance*) [TK03] ist nur geeignet, wenn die quadratische Fläche einen Ellipsoiden beschreibt. In diesem Fall wird eine Klasse von der Trennfläche *komplett* (als Ellipsoid) umschlossen. Dies ist eine starke Einschränkung und tritt bei den vorliegenden Daten nur in sehr seltenen Fällen auf; deshalb soll dieses Distanzmaß nicht weiter betrachtet werden.

Vor- und Nachteile der Distanzmaße

Im Folgenden sollen die wesentlichen positiven (+) und negativen (–) Eigenschaften der vorgestellten Distanzmaße zusammengefasst werden.

Euklidische Distanz von der Trennebene (linearisierter Fall des quadratischen Modells)

- + sehr einfach und robust zu berechnen
- + intuitive und anschauliche Deutung problemlos möglich
- gilt nur dann exakt, wenn die Kovarianzen beider Klassen gleich sind ($\Sigma_i = \Sigma_j$)

Lotrechte Distanz zur quadratischen Trennfläche

- + intuitive und anschauliche Deutung möglich (kürzester Euklidischer Abstand zur quadratischen Fläche)
- + Ausnutzung aller Modellparameter (Berücksichtigung der Kovarianzen)
- relativ aufwändig zu berechnen (Bestimmung der Nullstellen eines Polynoms sechsten Grades für jeden Bildpunkt)
- bei Flächenpaaren ist das Distanzmaß u.U. nicht monoton (vgl. Abbildung 3.9 rechts)

Algebraische Distanz zur quadratischen Trennfläche

- + einfach zu berechnen (Sample \mathbf{x} in die Flächengleichung einsetzen)
- + Ausnutzung aller Modellparameter (Berücksichtigung der Kovarianzen)
- Deutung des Abstandsmaßes eher über die verallgemeinerte Distanz zu den Klassen ($d_a = D_j^2 - D_i^2$) als zur Trennfläche möglich

3.1.5 Leistungsfähigkeit und Diskussion des adaptiven Farbraumes

Unter Annahme der Gültigkeit des verwendeten parametrischen Modells, welches die Momente erster und zweiter Ordnung (Mittelwertvektor und Kovarianzmatrix) berücksichtigt, bietet eine quadratische Entscheidungsfläche eine bestmögliche Separation zwischen den zu unterscheidenden Klassen. Weisen beide Klassen gleiche Kovarianzen auf, so ergibt sich als Spezialfall eine lineare Trennebene.

Mit Hilfe eines geeigneten Abstandsmaßes zur Entscheidungsfläche lassen sich die Eingangsvektoren (hier Farbwerte) auf reelle Zahlen abbilden, welche ein Grauwertbild mit maximaler Diskriminanz der Helligkeitswerte bezüglich der Klassen formen.

Die Bereitstellung des adaptiven Farbraumes bzw. der Modellparameter erfordert jedoch einerseits einen relativ hohen Rechenaufwand, da die Klassenparameter für jedes Bild individuell mittels des EM-Algorithmus' geschätzt werden müssen. Andererseits ist eine gute Initialisierung der Modellparameter sowie unter Umständen sogar ein Constraining-Mechanismus nötig, damit der EM-Algorithmus auch tatsächlich die gewünschte Klassentrennung erreicht und sich nicht „verschätzt“. Dieses Problem wird in Abschnitt 3.4.2 vor dem Hintergrund der Haustoriensegmentierung genauer untersucht.

Ein Standardfarbraum für die Zell(vor)segmentierung

Aus Gründen des numerischen Aufwandes ist es völlig unangebracht, die hochaufgelösten Mikroskopbilder direkt in den adaptiven Farbraum zu überführen. Stattdessen stellt sich die Frage, inwieweit das adaptive Ideal aufwandsgünstig durch einen statischen Farbraum, möglichst einen Standardfarbraum, approximiert werden kann. Die Färbung der transformierten Zellen lässt sich sehr gut im Buntton erkennen. In Abbildung 3.5 – 3.8 (jeweils *rechte Spalte*) wird dieser dem adaptiven Farbraum gegenübergestellt. Es zeigt sich, dass der Buntton eine ähnlich gute Klassentrennung zwischen „Hintergrund“ und „transformierte Zelle“ wie der adaptive Farbraum aufweist. Daher eignet sich der Buntton sehr gut als aufwandsgünstige Approximation zum generellen Auffinden transformierter Zellen in den Mikroskopbildern. Dies gilt auch für den Fall, dass die transformierten Zellen nur eine sehr schwache Färbung aufweisen.

Die Motivation zum Einsatz des adaptiven Farbraums bei der Zellfindung ergibt sich jedoch spätestens bei der Aufgabe, die Berandung transformierter Zellen *möglichst genau* zu bestimmen.

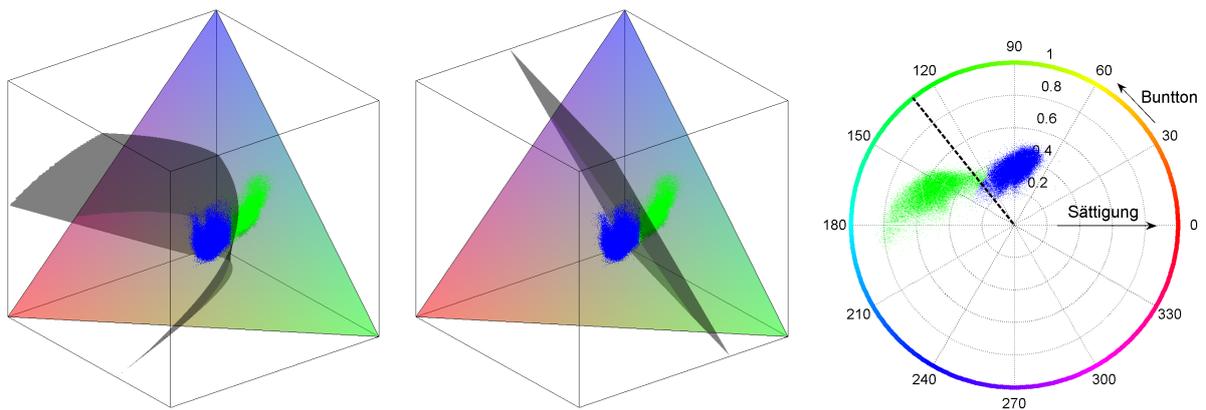


Abbildung 3.5: Separierung der Klassen „Hintergrund“ und „transformierte Zelle“ im $R'G'B'$ - und HSV -Farbraum am Beispielbild aus Abbildung 3.1.

Links: Quadratische Entscheidungsfäche (vorliegender Fall: einschaliger Hyperboloid).

Mitte: Unter Annahme gleicher Klassenkovarianz wird die Entscheidungsfäche linear (Ebene).

Rechts: Polardarstellung von Buntton (*Hue*) und Sättigung (*Saturation*) mit Buntton-Schwellwert.

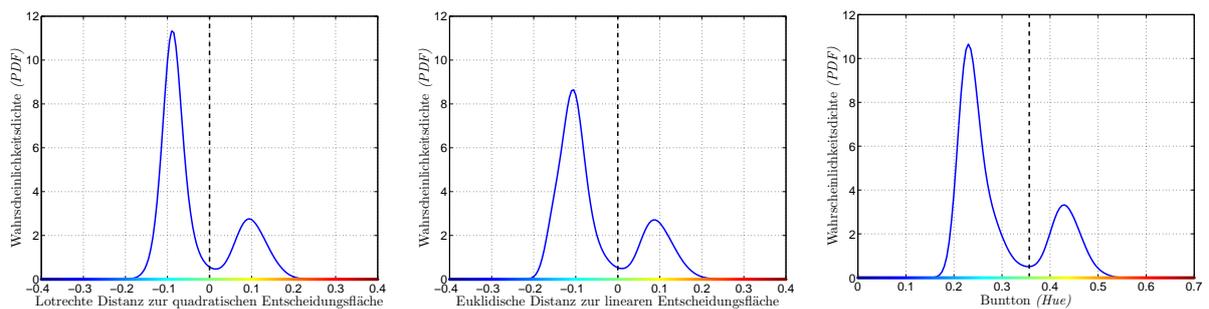


Abbildung 3.6: *Links:* Verteilungsdichte der lotrechten Distanz der Samples zur quadratischen Entscheidungsfäche. *Mitte:* Verteilungsdichte der Euklidischen Distanz der Samples zur linearen Entscheidungsfäche. *Rechts:* Verteilungsdichte des Bunttons (*Hue*).

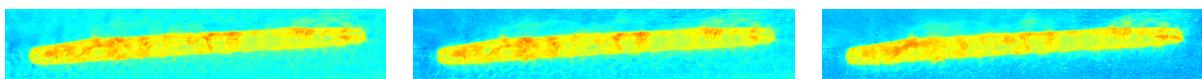


Abbildung 3.7: Abstandsmaße in Pseudofarbdarstellung. Die Farblegende ist Abbildung 3.6 zu entnehmen (Farbbalken an der Abszisse).

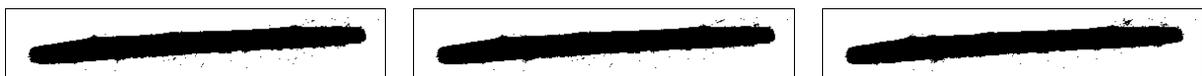


Abbildung 3.8: Resultierende Binärbilder nach Schwellwerthaltung.

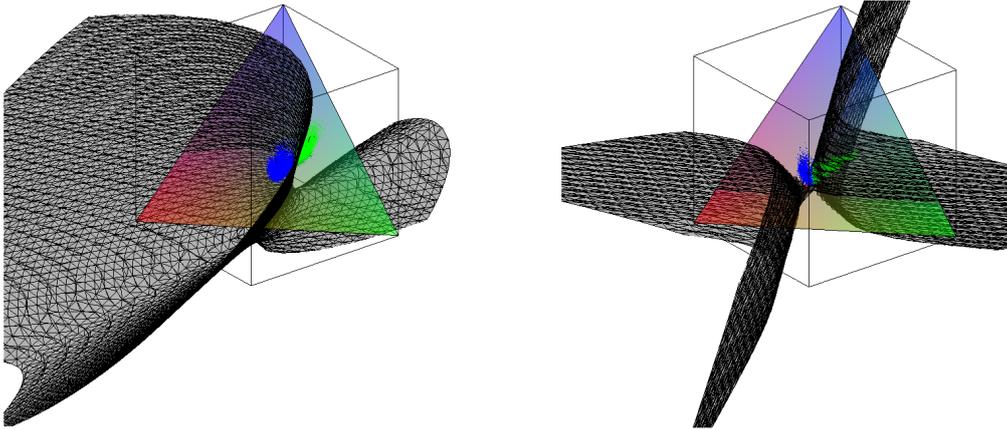
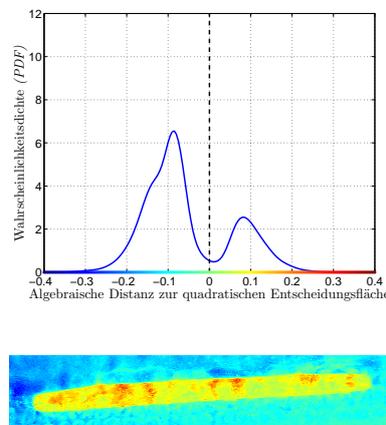
Abbildung 3.9: Einschaliges (*links*) und zweischaliges Hyperboloid (*rechts*).

Abbildung 3.10: Algebraische Distanz zur quadratischen Entscheidungsfläche.

Dabei können Strukturen (z.B. Haustorien) innerhalb der Zelle zu Störungen führen, weil diese Strukturen den Buntton ebenfalls beeinflussen und eine schwache Färbung damit „überstrahlen“ können. An dieser Stelle leistet der adaptive Farbraum Abhilfe, weil er sich konkret an das vorliegende Bildmaterial anpasst und potenzielle Störungen dadurch viel weniger in Erscheinung treten. Dies wird in Abschnitt 3.3 bei der genauen Bestimmung der Zellberandung mittels aktiver Konturmodelle diskutiert.

Weiteres Vorgehen

Nach der Überführung des Bildes in einen geeigneten Farbraum stellt sich nunmehr die Frage nach der eigentlichen Zellsegmentierung. In Abbildung 3.8 wird die einfachste Form der Bildsegmentierung bereits durchgeführt – die globale Schwellwerthaltung. Bei der Berechnung des adaptiven Farbraumes wird der optimale Schwellwert implizit mitberechnet, denn die Trennfläche selbst wirkt – wie ihr Name sagt – als bestmögliche Trennung der beiden Klassen. Der Schwellwert liegt daher bei Null. Wollte man hingegen eine Schwellwerthaltung im Buntton durchführen, so könnte beispielsweise das Minimum des geglätteten Histogramms zwischen den beiden Maxima als Schwellwert verwendet werden.

Eine einfache Schwellwerthaltung führt jedoch nur bei solchen Bildern zu zufriedenstellenden Ergebnissen, in denen sich die zu segmentierenden Objekte durch deutliche Helligkeits- bzw. Farbunterschiede vom Hintergrund abheben. Im vorliegenden Fall werden transformierte Zellen häufig aufgrund von schwacher bzw. ungleichmäßiger Färbung nicht vollständig oder gar nicht

segmentiert, d.h., es werden nicht alle Pixel, welche zu transformierten Zellen gehören, als solche markiert. Weiterhin werden viele kleine Bereiche des Hintergrundes aufgrund von Rauscheffekten fälschlicherweise markiert. Wie auch immer der Schwellwert gewählt wird – eine einfache Schwellwerthaltung führt für das vorliegende Problem der Erkennung transformierter Zellen nicht zu zufriedenstellenden Ergebnissen.

Kantendetektoren hingegen sind (als Vertreter der nächsthöheren Stufe von Segmentierungsverfahren) unabhängig von einem absoluten Grauschwellwert. Durch die Differenzierung des Bildes verschwinden Absolutwerte; stattdessen rückt die *Veränderung* des Grauwertes ins Zentrum der Aufmerksamkeit.

Bevor im nächsten Kapitel im Detail auf das Thema Kantendetektion eingegangen wird, ist zunächst ein kleiner Exkurs in die Skalenraumtheorie notwendig. Sie vereinigt das Verständnis für verschiedene Auflösungsstufen von Bildern unter einem soliden mathematischen Gerüst und ist für das Verständnis und die korrekte Anwendung von Kantendetektoren von fundamentaler Bedeutung.

3.2 Kantenbasierte Segmentierung – Kantendetektion

3.2.1 Skalenraum und Multiskalen-Bildverarbeitung

Jede Betrachtung, jedes „Sehen“ – egal ob menschlich oder maschinell – ist skalenabhängig. Versetzen wir uns als Gedankenexperiment in die Lage, Objekte aus beliebiger Entfernung bzw. aus beliebiger Nähe betrachten zu können. Der Grashalm, auf den unser Fokus aktuell gelenkt ist, entpuppt sich bei zunehmender Entfernung als zu einer großen Wiese gehörig, welche – bei zunehmender Entfernung – als grüne Lichtung eines Waldes erscheint; dieser Wald wiederum könnte Teil einer Gebirgslandschaft sein, welche zu einem bestimmten Teil des Festlandes der Erde gehört... Dieses Gedankenexperiment ließe sich in beide Richtungen weiter fortsetzen, z.B. bis zu einem Blick auf das gesamte Universum oder in die andere Richtung bis hin zum inneren Aufbau der Elementarteilchen, aus denen der ursprünglich betrachtete Grashalm besteht.

Genauso wie das menschliche Sehen verschiedene, dem Problem angepasste Detaillierungsgrade – bestimmte Skalen – erfordert, so benötigt maschinelles Sehen einen ähnlichen Mechanismus, um eine Szene erfassen und interpretieren zu können. Die Skalenraumtheorie (*Scale-Space Theory*) stellt das dafür notwendige theoretische Fundament bereit.

Linearer Skalenraum

Die Grundlagen für die Skalenraumtheorie wurden von WITKIN [Wit83] und KOENDERINK [Koe84] gelegt. WITKIN formulierte das Prinzip des Skalenraums zunächst für eindimensionale Signale, indem er forderte, dass bei wachsendem „Verallgemeinerungsgrad“ keine neuen Strukturen entstehen dürfen. Vielmehr sollten die *Zusammenhänge* der bei feinerer Skale betrachteten Details *untereinander* zutage treten; es sollte also eine strikte Vereinfachung stattfinden. Als Glättungsoperator schlug er dazu einen Gaußkern vor.

KOENDERINK erkannte, dass das von WITKIN geforderte Prinzip als Diffusionsprozess aufgefasst werden kann, und zeigte, dass der Gaußkern zugleich die *einzig*e Lösung zur Generierung eines Skalenraumes ist.

Für ein d -dimensionales Signal $f: \mathbb{R}^d \rightarrow \mathbb{R}$ ist die Skalenraumrepräsentation $f_\sigma: \mathbb{R}^d \times \mathbb{R}_+ \rightarrow \mathbb{R}$ gegeben durch

$$f_\sigma(\mathbf{x}; \sigma) = \int f(\mathbf{x} - \boldsymbol{\xi}) g_\sigma(\boldsymbol{\xi}) d\boldsymbol{\xi} \quad (3.37)$$

mit der Gaußfunktion $g_\sigma: \mathbb{R}^d \times \mathbb{R}_+ \rightarrow \mathbb{R}$

$$g_\sigma(\mathbf{x}; \sigma) = \frac{1}{(2\pi\sigma^2)^{d/2}} e^{-\frac{|\mathbf{x}|^2}{2\sigma^2}}. \quad (3.38)$$

Die Skalenraumdarstellung eines Bildes berechnet sich also aus der Faltung des Bildes mit einem 2D-Gaußkern.

$$I_\sigma(x, y; \sigma) = I(x, y) * g_\sigma(x, y; \sigma) \quad (3.39)$$

Bei der numerischen Implementation der Faltung sollte aus Effizienzgründen die *Separierbarkeit* des Gaußkerns ausgenutzt werden: In Kartesischen Koordinaten lässt sich ein mehrdimensionaler Gaußkern als Produkt von 1D-Gaußkernen schreiben. Dementsprechend sollten die einzelnen Komponenten des mehrdimensionalen Signals nacheinander mit 1D-Gaußkernen gefaltet werden. Die Komplexität verringert sich dadurch von $\mathcal{O}(N^d)$ auf $\mathcal{O}(dN)$.

Ableitungen der Gaußfunktion

Nicht nur die Gaußfunktion, sondern auch ihre Ableitungen sind in der Bildverarbeitung von zentraler Bedeutung. Der Gradient der Skalenraumdarstellung eines Bildes lässt sich mit Hilfe der Ableitungen der Gaußfunktion schreiben als

$$\nabla I_\sigma = (\nabla I) * g_\sigma = I * (\nabla g_\sigma). \quad (3.40)$$

Mittels der Hermite-Polynome $H_n(x)$ lässt sich die n te Ableitung des Gaußkerns analytisch explizit angeben.

$$\frac{\partial^n g_\sigma(x; \sigma)}{\partial x^n} = (-1)^n \frac{1}{(2\sigma^2)^{n/2}} H_n\left(\frac{x}{\sqrt{2\sigma^2}}\right) g_\sigma(x; \sigma) \quad (3.41)$$

Die Hermite-Polynome lassen sich anhand folgender Rekursionsvorschrift berechnen:

$$H_0(x) = 1 \quad (3.42)$$

$$H_1(x) = 2x \quad (3.43)$$

$$H_2(x) = 4x^2 - 2 \quad (3.44)$$

⋮

$$H_{n+1}(x) = 2x H_n(x) - 2n H_{n-1}(x). \quad (3.45)$$

Linearer und nichtlinearer Skalenraum auf Basis der Diffusionsgleichung

Alternativ kann die Erzeugung der Skalenraumdarstellung auch als Diffusionsprozess dargestellt werden.

$$\frac{\partial I_\sigma}{\partial t} = \nabla(D \nabla L) \quad (3.46)$$

Im Fall eines homogenen Diffusionsprozesses (Diffusionskoeffizient D ist konstant) vereinfacht sich diese Gleichung zu

$$\frac{\partial I_\sigma}{\partial t} = D \Delta I_\sigma \quad \text{mit} \quad \sigma = \sqrt{2Dt}. \quad (3.47)$$

Eine Verallgemeinerung dieses Ansatzes auf eine *ortsabhängige (non-linear diffusion)* bzw. *ortsabhängige und anisotrope (anisotropic diffusion)* Diffusion führt zur Theorie nichtlinearer Skalenräume. Die Grundlagen dazu gehen auf die Arbeiten von PERONA und MALIK [PM90] zurück; von WEICKERT *et al.* [Wei98a, Wei98b] wurden dazu weitere fundamentale Beiträge geliefert. Dazu ist die nichtlineare Differentialgleichung

$$\frac{\partial L}{\partial t} = \nabla(D(|\nabla I_\sigma|^2) \nabla L) \quad (3.48)$$

zu lösen. Die Diffusionsfähigkeit ist dabei eine Funktion des Gradienten der (linearen) Skalenraumdarstellung des Bildes. PERONA und MALIK [PM90] haben dazu die Funktion

$$D(|\nabla I_\sigma|^2) = \frac{1}{1 + |\nabla I_\sigma|^2 / \lambda^2} \quad (3.49)$$

vorgeschlagen. Nichtlineare Skalenräume bieten die Möglichkeit einer inhaltsadaptiven, kanten-erhaltenden Bildglättung.

3.2.2 Bildableitungen und Kantenfilter

Für die Detektion von Kanten wird innerhalb des Bildes nach Bereichen gesucht, welche eine *große örtliche Veränderung des Grauwertes* aufweisen. Diese Definition trifft zunächst nur auf Graubilder bzw. einkanalige Bilder zu. In der Tat ist die Behandlung von Farbbildern (allgemeiner: vektorwertige Bilder oder Mehrkanalbilder) sehr viel komplexer, deshalb werden für die folgenden Erklärungen zunächst Graubilder vorausgesetzt und später eine Erweiterung auf Farbbilder vorgenommen.

Die i -te Ortsableitung eines Bildes nach den Koordinaten x bzw. y korrespondiert im Spektrum mit einer Multiplikation der entsprechenden Spektralanteile mit einer Potenzfunktion i -ten Grades, sie führt also zu einer Anhebung hoher Spektralanteile.

$$I(x, y) \quad \circ\bullet \quad F(u, v) \quad (3.50)$$

$$\frac{\partial^i}{\partial x^i} I(x, y) \quad \circ\bullet \quad (j2\pi u)^i F(u, v) \quad (3.51)$$

$$\frac{\partial^i}{\partial y^i} I(x, y) \quad \circ\bullet \quad (j2\pi v)^i F(u, v) \quad (3.52)$$

Die Ordnung i der Bildableitung kann gleichzeitig für eine Klassifikation von Kantendetektoren herangezogen werden, wobei in der Praxis nur den Ableitungsoperatoren erster und zweiter Ordnung eine Bedeutung zukommt.

Die Frage, warum es überhaupt verschiedene Ableitungsoperatoren gibt, liegt in der diskreten Ortsrepräsentation von Bildern – in ihrer Darstellung in Form von Pixeln – begründet. Eine Repräsentation des Bildes in Form einer kontinuierlichen Funktion würde in der Tat eine eindeutige Definition einer Ableitung – der Bildableitung – zulassen. Durch die in der Praxis vorliegende Repräsentation auf einem zweidimensionalen diskreten Gitter ist dies jedoch nicht so einfach; hier kann die Ableitung höchstens näherungsweise anhand bestimmter Optimierungskriterien approximiert werden. In [Jäh02] werden dazu ausführliche Untersuchungen zur Approximationsgüte diskreter Ableitungsoperatoren vorgestellt.

Kantendetektoren erster Ordnung

Kantendetektoren erster Ordnung bilden den Gradienten der diskreten Bildfunktion nach.

$$\nabla I = \begin{bmatrix} \frac{\partial I}{\partial x} \\ \frac{\partial I}{\partial y} \end{bmatrix} \quad |\nabla I| = \sqrt{\left(\frac{\partial I}{\partial x}\right)^2 + \left(\frac{\partial I}{\partial y}\right)^2} \quad \varphi_{\nabla I} = \arctan\left(\frac{\partial I}{\partial y} / \frac{\partial I}{\partial x}\right) \quad (3.53)$$

Dies wird über Ableitungsoperatoren realisiert. Einige Vertreter sind z.B. der Roberts-Operator, der Prewitt-Operator und der Sobel-Operator. Die entsprechenden Filtermasken sind im Anhang C zu finden.

Kantendetektoren zweiter Ordnung – Laplaceoperator

Der auch als Marr-Hildreth-Kantendetektor [MH80] bekannte Ansatz basiert auf der zweiten Ableitung des Bildes. Der Laplaceoperator Δ , angewendet auf das geglättete Bild bei der Skala σ , liefert eine Funktion, deren Nullstellen potenzielle Kanten darstellen. Die Intensität der Kante wird dabei von der „Überhöhung“ der Umgebung der Nullstelle repräsentiert, welche über eine Schwellwerthaltung schließlich zu einem binären Kantenbild führt.

$$\Delta I = \frac{\partial^2 I_\sigma}{\partial x^2} + \frac{\partial^2 I_\sigma}{\partial y^2} \quad (3.54)$$

Die diskrete Approximation des Laplaceoperators ist gegeben durch die folgenden Filtermasken:

$$\frac{\partial^2}{\partial x^2} \approx \begin{bmatrix} 0 & 1 & -1 \end{bmatrix} * \begin{bmatrix} 1 & -1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & -2 & 1 \end{bmatrix} \quad (3.55)$$

$$\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \approx \begin{bmatrix} 1 & -2 & 1 \end{bmatrix} + \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad (3.56)$$

Kantendetektor nach Canny

CANNY [Can86] hat sich Gedanken bezüglich eines optimalen Kantendetektors gemacht und dazu formale Kriterien aufgestellt, anhand derer er ein optimales Kantentfilter hergeleitet hat. Dazu fasste er die drei Kriterien

- geringe Fehlerrate
- genaue Lokalisierung
- keine Mehrfacherkennungen

zu einer Kostenfunktion zusammen und bestimmte für den Fall einer idealen Stufenkante das Optimalfilter. Dieses Optimalfilter lässt sich sehr gut durch die erste Ableitung des Gaußfilters approximieren.

Die Glättung mittels eines Gaußfilters überführt das Bild in den Skalenraum. Das Canny-Kantentfilter kann also als ein Multiskalenansatz interpretiert werden, mit dem das Bild – je nach vorgegebener Standardabweichung σ des Gauß- bzw. abgeleiteten Gaußfilters – bei einer bestimmten Auflösungsstufe behandelt wird.

Der Gradientenvektor des Bildes wird im Fall des Canny-Kantendetektors approximiert durch

$$\frac{\partial I(x, y)}{\partial x} \approx I(x, y) * \left(-\frac{x}{\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \right), \quad \frac{\partial I(x, y)}{\partial y} \approx I(x, y) * \left(-\frac{y}{\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \right). \quad (3.57)$$

Aus Sicht des numerischen Berechnungsaufwandes sei an dieser Stelle nochmals auf die *Separierbarkeit* der Gaußfunktion in Kartesischen Koordinaten hingewiesen. Dadurch lassen sich die Richtungsableitungen durch je zwei Faltungen mit 1D-Filterkernen mit einem Gesamtaufwand von $\mathcal{O}(2N)$ bereitstellen, anstatt $\mathcal{O}(N^2)$ im allgemeinen Fall bei der Faltung mit einem 2D-Kern.

$$\frac{\partial I(x, y)}{\partial x} \approx \left[I(x, y) * \left(-\frac{x}{\sigma^2} e^{-\frac{x^2}{2\sigma^2}} \right) \right] * e^{-\frac{y^2}{2\sigma^2}}, \quad \frac{\partial I(x, y)}{\partial y} \approx \left[I(x, y) * \left(-\frac{y}{\sigma^2} e^{-\frac{y^2}{2\sigma^2}} \right) \right] * e^{-\frac{x^2}{2\sigma^2}} \quad (3.58)$$

Im zweiten Verarbeitungsschritt des Canny-Kantendetektors werden mittels der Betrags- und Richtungsinformation der Bildableitung und unter Anwendung einer Hysterese-Schwelwerhaltung jene Punkte entfernt, die kein lokales Maximum darstellen (*non-maximum supression*). Dieser Vorgang dünnt die Kanten auf eine Breite von einem Pixel aus und kann als ein intelligentes Skelettierungsverfahren gedeutet werden.

3.2.3 Zellfindung mittels Canny-Kantendetektor im Buntton

Wie bereits in Abschnitt 3.1.5 festgestellt wurde, spiegelt sich das Färbungsmerkmal transformierter Zellen sehr gut im Buntton wider. Nach den Ausführungen über die Skalenraumtheorie sowie der Diskussion verschiedener Methoden zur Kantendetektion kristallisiert sich als geeignetes Werkzeug zur Findung der transformierten Zellen eindeutig der Canny-Kantendetektor heraus. Nach [CLMC92] ist er als der beste lineare Kantendetektor anerkannt.

Im Folgenden werden die Ergebnisse der Kantendetektion auf Basis des Bunttons diskutiert.

Ergebnisse

Abbildung 3.12 zeigt eine Pseudofarbdarstellung des Bunttons anhand von zwei Beispielen. Das Betragsquadrat der Bildableitung ist in Abbildung 3.13 und 3.14 für die zwei Skalen $\sigma = 5$ und $\sigma = 12$ dargestellt. Die resultierenden Kanten sind in Abbildung 3.15 und 3.16 zu sehen. Hier ist das Unschärfepinzipp bei der Kantendetektion im Skalenraum zu beobachten: Entweder es treten Fehldetektionen auf (siehe Abbildung 3.15, *linkes Bild*) – die Kanten sind aber genau lokalisiert – oder man erkaufte sich geringe Fehldetektionsraten mit einer starken Delokalisierung der Kanten (siehe Abbildung 3.16).

Um sowohl *zuverlässige* als auch *korrekt lokalisierte* Kanten zu erhalten, werden im nächsten Kapitel aktive Konturmodelle (Schlangen, *Snakes*) behandelt, mit denen sich das Problem auf sehr zuverlässige und effiziente Weise lösen lässt. Zuvor ist jedoch noch ein kleiner Exkurs zur generellen Problematik „Kanten in Farbbildern“ angebracht, um die Wichtigkeit der Wahl eines *geeigneten* Farbraums zu untermauern.

Mehrdimensionale Kantendetektion im RGB-Farbraum (als Vergleich)

Als Standardverfahren könnten bei Mehrkanalbildern die Gradientenvektoren jedes einzelnen Farbkanals miteinander kombiniert werden [Jäh02]. Bei einem RGB-Bild wird dazu die Jacobi-matrix

$$\mathbf{J} = \begin{bmatrix} \frac{\partial I_R}{\partial x} & \frac{\partial I_G}{\partial x} & \frac{\partial I_B}{\partial x} \\ \frac{\partial I_R}{\partial y} & \frac{\partial I_G}{\partial y} & \frac{\partial I_B}{\partial y} \end{bmatrix}^T \quad (3.59)$$

berechnet und die Spur der Matrix

$$\mathbf{S} = \mathbf{J}^T \mathbf{J} \quad (3.60)$$

ausgewertet. Das ergibt die Summe der quadrierten Ableitungskomponenten aller Kanäle.

$$\text{spur}(\mathbf{S}) = \left(\frac{\partial I_R}{\partial x}\right)^2 + \left(\frac{\partial I_G}{\partial x}\right)^2 + \left(\frac{\partial I_B}{\partial x}\right)^2 + \left(\frac{\partial I_R}{\partial y}\right)^2 + \left(\frac{\partial I_G}{\partial y}\right)^2 + \left(\frac{\partial I_B}{\partial y}\right)^2 \quad (3.61)$$

Wie in Abbildung 3.17 zu sehen ist, führt diese Herangehensweise mit dem vorliegenden Bildmaterial nicht ansatzweise zum gewünschten Erfolg. Aufgrund der gemeinsamen Repräsentation (Korrelation) von Helligkeits- und Farbinformation bei der RGB-Darstellung werden bei diesem Ansatz sowohl die Helligkeits- als auch die Farbinformation ausgewertet. Für die Detektion der transformierten Zellen ist jedoch *ausschließlich* die Farbinformation entscheidend. Die Berücksichtigung von Helligkeitsmerkmalen führt hingegen zu unerwünschten Störungen.

Dies untermauert, dass es für die Segmentierung von Farbbildern noch viel weniger als für Grauwertbilder eine allgemein gültige Lösung gibt. Vielmehr müssen für jeden konkreten Problemfall individuelle Lösungen gefunden und angepasst werden. Farbbilder müssen dazu zunächst in ein geeignetes Merkmalsbild überführt werden.

3.3 Verbinden von Kanten mit aktiven Konturmodellen

In den seltensten Fällen ist ein Kantenbild ohne Nachbearbeitung verwendbar. Falsch positive Kanten müssen entfernt und falsch negative Kanten (Unterbrechungen) müssen geschlossen werden. Problematisch ist an dieser Stelle, dass a priori nicht bekannt ist, *welche* Kanten fälschlicherweise auftreten und welche Lücken tatsächlich zu schließen sind.

Handelt es sich bei den zu detektierenden Kanten um *geschlossene* Konturen, so bietet sich zur Lösung des Problems der Einsatz aktiver Konturmodelle, sogenannter Schlangen (*Snakes*), an. Dabei wird zunächst eine geschlossene Kontur mit nur sehr grober Genauigkeit um das zu segmentierende Objekt gelegt und diese Kontur anschließend in einem iterativen Prozess der

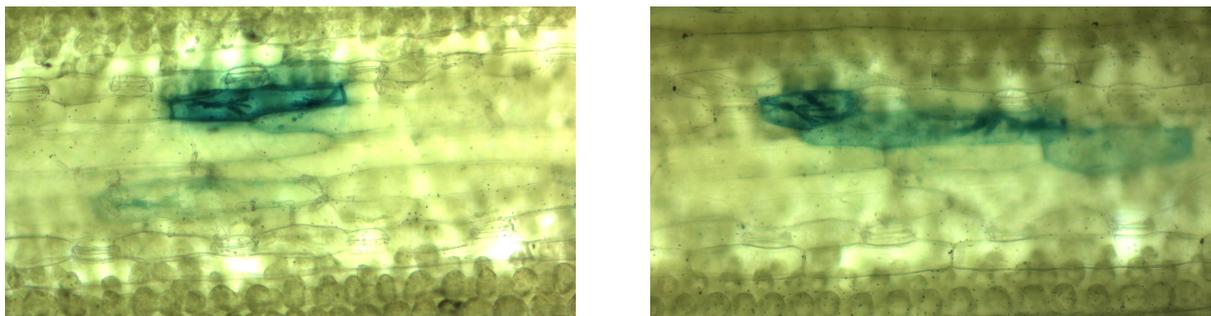


Abbildung 3.11: Ausschnitte aus typischen Mikroskopbildern. *Links*: Eine transformierte Zelle mit zwei Haustorien. *Rechts*: Drei transformierte Zellen; die linke und mittlere Zelle enthalten Haustorien.

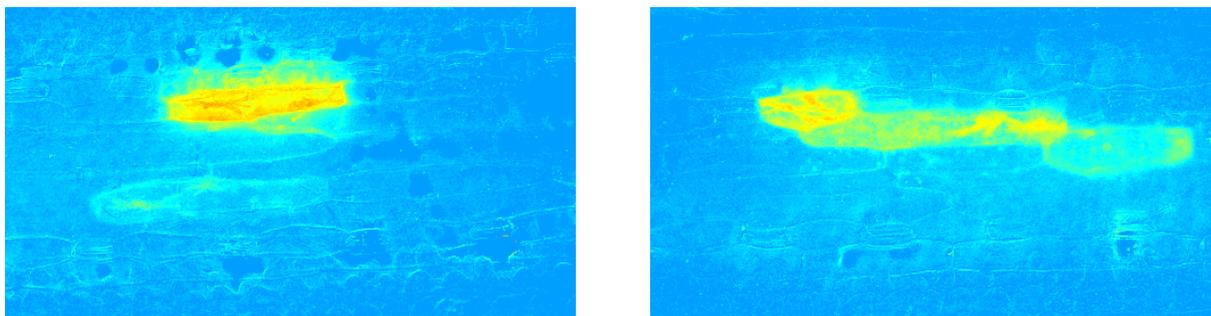


Abbildung 3.12: Pseudofarbdarstellung des Bunttons im Bereich $0 \dots 0.7$. Die transformierten Zellen heben sich in dieser Darstellung deutlich sichtbar vom Hintergrund ab.

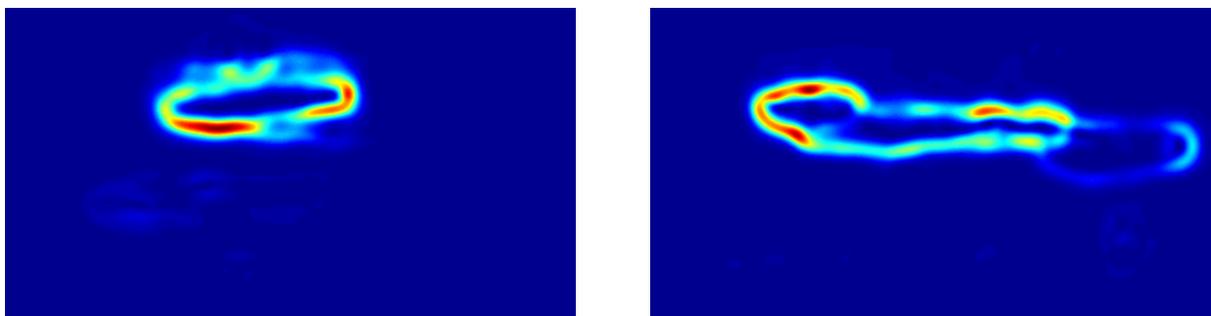


Abbildung 3.13: Betragsgadrat des Bildgradienten von Abbildung 3.12 bei der Skala $\sigma = 5$.

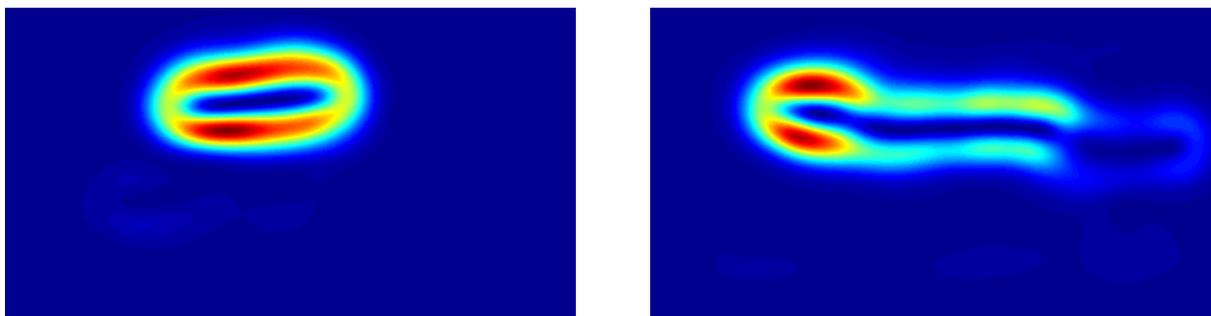


Abbildung 3.14: Betragsgadrat des Bildgradienten von Abbildung 3.12 bei der Skala $\sigma = 12$.

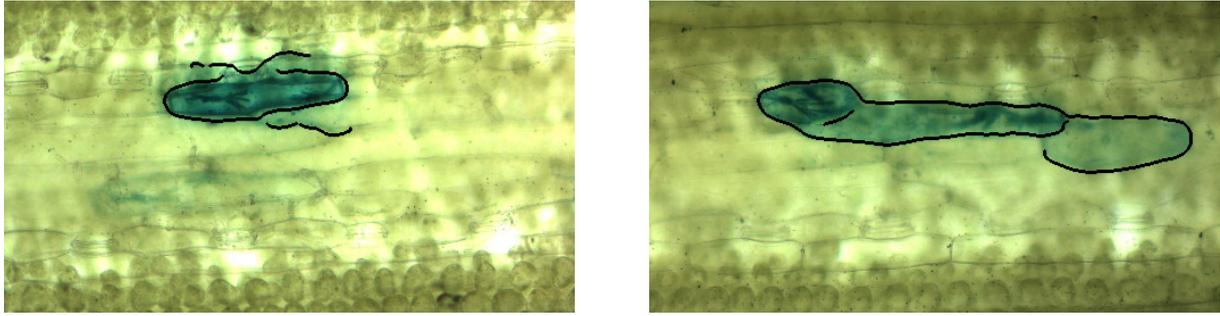


Abbildung 3.15: Detektierte Kanten: Skala $\sigma = 5$, Schwellwerte $[0.16 \quad 0.40]$.

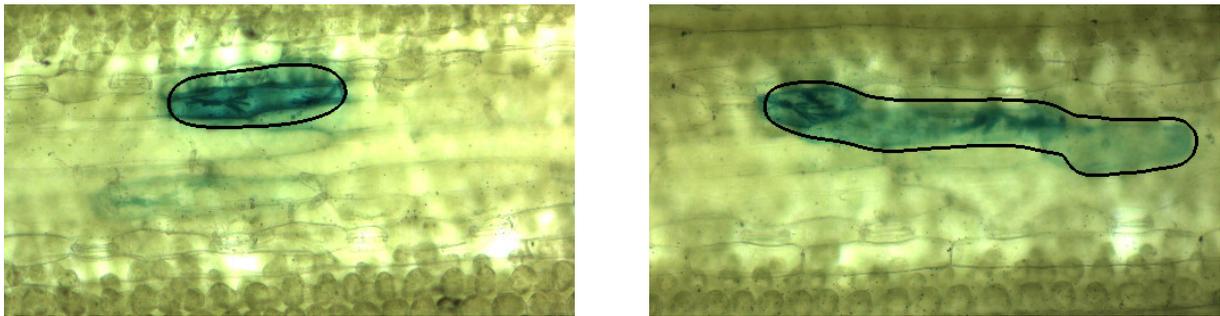


Abbildung 3.16: Detektierte Kanten: Skala $\sigma = 12$, Schwellwerte $[0.16 \quad 0.40]$.

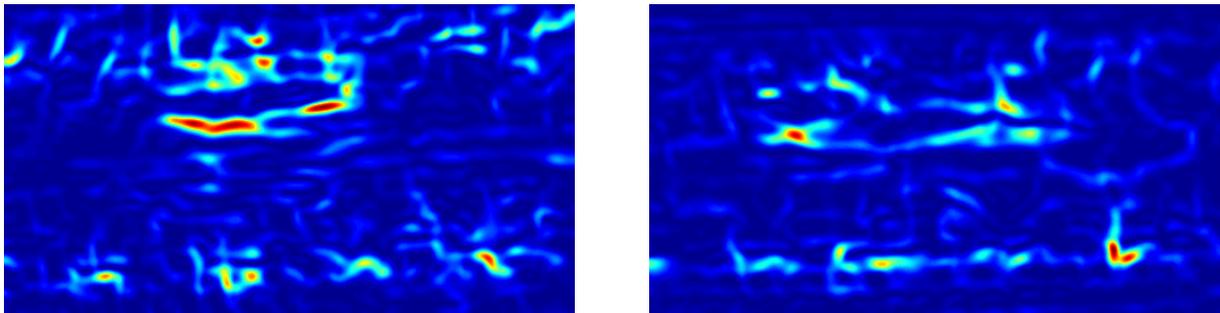


Abbildung 3.17: Betragsquadrat des Gradienten bei Mehrdimensionaler Kantendetektion im RGB-Farbraum: Die transformierten Zellen sind mit diesem Ansatz nicht zu finden.

genauen Objektberandung immer feiner angenähert. Im Fall der transformierten Zellen wird die grobe Objektkontur durch den Canny-Kantendetektor bereitgestellt.

Da aktive Konturmodelle – bedingt durch ihre iterative Arbeitsweise – relativ rechenaufwändig sind und die Bildauswertung damit sehr zeitintensiv wird, soll an dieser Stelle ein neues Lösungsverfahren für aktive Konturmodelle vorgestellt werden, welches spezielle Eigenschaften der zu berechnenden Gleichungen berücksichtigt und zur Lösung die schnelle Fouriertransformation (*Fast Fourier Transform, FFT*) nutzt [IS05]. Auf diese Weise lässt sich die Komplexität der Berechnung von $\mathcal{O}(N^2)$ auf $\mathcal{O}(N \log N)$ reduzieren, was den Einsatz aktiver Konturmodelle nunmehr auch für Hochdurchsatz bzw. Echtzeitanforderungen ermöglicht.

Im Folgenden soll nun erörtert werden, mittels welcher Überlegungen sich der Rechenaufwand für aktive Konturmodelle drastisch verringern lässt und welche Ergebnisse der Einsatz aktiver Konturmodelle bezüglich der Zellsegmentierung liefert.

3.3.1 Grundlagen aktiver Konturmodelle (Snakes)

Eine aktive Kontur ist eine Parameterkurve $v(s) = (x(s), y(s))$, $s \in [0, L]$, die sich – getrieben durch eine äußere „Krafteinwirkung“ seitens des Bildes und unter dem Einfluss ihrer eigenen Dehnungs- und Verbiegungseigenschaften – an bestimmte Bildkonturen anschmiegt. Dieser Ansatz wurde ursprünglich von KASS *et al.* in [KWT88] vorgestellt, erste theoretische und numerische Analysen des zugrunde liegenden Energiefunktionals gehen auf den russischen Mathematiker Tikhonov [Tik63, TA77] zurück.

Die interne Energie einer aktiven Kontur ist gegeben durch das Funktional

$$E_{\text{int}} = \frac{1}{2} \int_0^L \left(\alpha(s) \left| \frac{\partial v(s)}{\partial s} \right|^2 + \beta(s) \left| \frac{\partial^2 v(s)}{\partial s^2} \right|^2 \right) ds, \quad (3.62)$$

wobei die Parameter $\alpha(s)$ und $\beta(s)$ die Elastizität und die Verbiegungssteifigkeit der Parameterkurve beschreiben.

Die externe Energie $E_{\text{ext}}(v(s))$ ist anhand der zu segmentierenden Objekte aus dem Bild abzuleiten. Um die Kontur beispielsweise an Kanten auszurichten, ist für die externe Energie

$$E_{\text{ext}}(v(s)) = -|\nabla I_\sigma(x, y)|^2 \Big|_{(x,y)=v(s)} \quad (3.63)$$

anzusetzen.

Dem Modell der Schlange liegt das physikalische Prinzip einer minimalen Gesamtenergie zugrunde – gesucht ist jene Kontur mit minimaler Gesamtenergie

$$\min_{v(s)} \{E_{\text{int}}(v(s)) + E_{\text{ext}}(v(s))\}. \quad (3.64)$$

Mit den Methoden der Variationsrechnung lässt sich das Minimum dieses Energiefunktionals bestimmen, d.h. jene Kontur finden, welche bestmöglich an das Bildpotenzial angepasst ist. Die Kontur, welche das Energiefunktional nach Gleichung (3.64) minimiert, muss der folgenden Euler-Differentialgleichung genügen:

$$\underbrace{[\alpha(s) v'(s)]'}_{F_{\text{int}}(v(s))} - \underbrace{[\beta(s) v''(s)]''}_{F_{\text{ext}}(v(s))} - \nabla E_{\text{ext}}(v(s)) = 0. \quad (3.65)$$

Diese Gleichung ist erfüllt, wenn ein Gleichgewicht zwischen den internen und externen Kräften F_{int} bzw. F_{ext} herrscht.

$$F_{\text{int}}(v(s)) + F_{\text{ext}}(v(s)) = 0 \quad (3.66)$$

Eine Lösung für diese Differentialgleichung kann mit Hilfe eines Gradientenabstiegs gefunden werden. Dazu wird die Zeitvariable t eingeführt (die Kontur wird zeitabhängig), $v(s) \rightarrow v(s; t)$, und ausgehend von einer Initialkontur $v(s; 0) = v_0(s)$ wird eine zeitliche Entwicklung der Kontur zugelassen, bis sich ein Kräftegleichgewicht eingestellt hat. Dieses Kräftegleichgewicht liegt vor, wenn die Zeitableitung der Kontur verschwindet, also keine Bewegung mehr stattfindet. Die entstehende Bewegungsgleichung ist eine partielle Differentialgleichung vierter Ordnung.

$$\frac{dv(s; t)}{dt} - [\alpha(s) v'(s; t)]' + [\beta(s) v''(s; t)]'' - F_{\text{ext}}(v(s; t)) = 0 \quad (3.67)$$

3.3.2 Klassische Lösung im Ortsbereich

Im klassischen Lösungsansatz nach KASS *et al.* [KWT88] werden sowohl die Ortsableitungen als auch die Zeitableitungen durch finite Differenzen ersetzt. Dazu wird die

- Zeitdiskretisierung mit der Schrittweite τ und die
- Ortsdiskretisierung in N Samples mit der Schrittweite h
 $v_n = (x_n, y_n) = (x(nh), y(nh))$

eingeführt. Dies ergibt die folgende Differenzengleichung:

$$\begin{aligned}
 \frac{v_n^t - v_n^{t-1}}{\tau} - \frac{1}{h^2} [-\alpha_n(v_n^t - v_{n-1}^t) + \alpha_{n+1}(v_{n+1}^t - v_n^t)] \\
 + \frac{1}{h^4} \beta_{n-1}(v_{n-2}^t - 2v_{n-1}^t + v_n^t) \\
 - \frac{2}{h^4} \beta_n(v_{n-1}^t - 2v_n^t + v_{n+1}^t) \\
 + \frac{1}{h^4} \beta_{n+1}(v_n^t - 2v_{n+1}^t + v_{n+2}^t) \\
 - F_{\text{ext}}(v_n^t) = 0.
 \end{aligned} \tag{3.68}$$

Die Tiefstellungen $(\cdot)_n$ und Hochstellungen $(\cdot)^t$ bedeuten dabei die räumlichen bzw. zeitlichen Indizes.

Generell kann diese Differenzengleichung in Matrix-Vektor-Notation geschrieben werden. Die Matrix $\mathbf{K} \in \mathbb{R}^{N \times N}$ wird dabei als die Steifigkeitsmatrix (*Stiffness Matrix*) des Systems bezeichnet.

$$\frac{\mathbf{v}^t - \mathbf{v}^{t-1}}{\tau} + \mathbf{K} \mathbf{v}^t - \mathbf{F}_{\text{ext}}(\mathbf{v}^t) = 0 \tag{3.69}$$

Diese Gleichung kann gelöst werden, indem vereinfachend angenommen wird, dass die externen Kräfte während eines Zeitschritts konstant sind.

$$\mathbf{F}_{\text{ext}}(\mathbf{v}^t) \approx \mathbf{F}_{\text{ext}}(\mathbf{v}^{t-1}) \tag{3.70}$$

Nach elementaren Umformungen mit Hilfe der linearen Algebra ergibt sich als Lösung

$$\mathbf{v}^t = (\mathbf{I} + \tau \mathbf{K})^{-1} (\mathbf{v}^{t-1} + \tau \mathbf{F}_{\text{ext}}(\mathbf{v}^{t-1})). \tag{3.71}$$

Aus Sicht der numerischen Komplexität erfordert dieses Lösungsschema aufgrund der Matrix-Vektor-Multiplikation einen Aufwand von $\mathcal{O}(N^2)$.

Werden die Elastizitäts- und Verbiegungsparameter als konstant angenommen

$$\alpha(s) = \alpha \quad \beta(s) = \beta,$$

so wird die Steifigkeitsmatrix \mathbf{K} *symmetrisch* und *zyklisch* (*zirkulant*).

$$\underbrace{\begin{bmatrix} k_0 & k_1 & k_2 & & & & & & & & k_2 & k_1 \\ k_1 & k_0 & k_1 & k_2 & & & & & & & & & \\ k_2 & k_1 & k_0 & k_1 & k_2 & & & & & & & & \\ & k_2 & k_1 & k_0 & k_1 & k_2 & & & & & & & \\ & & & & & & \ddots & & & & & & \\ k_2 & & & & & & & & & & k_2 & k_1 & k_0 & k_1 \\ k_1 & k_2 & & & & & & & & & k_2 & k_1 & k_0 \end{bmatrix}}_{\mathbf{K}}$$

$$k_0 = \frac{2\alpha}{h^2} + \frac{6\beta}{h^4} \quad k_1 = -\frac{\alpha}{h^2} - \frac{4\beta}{h^4} \quad k_2 = \frac{\beta}{h^4}$$

Die Voraussetzung konstanter Parameter α und β stellt in der Praxis keine besondere Einschränkung dar und wird in der Regel ohnehin angewendet [IP93, XP98, Wei98c].

Im Folgenden wird nun dargestellt [IS05], wie sich der numerische Berechnungsaufwand der Differenzgleichung drastisch reduzieren lässt. Des Weiteren wird gezeigt, wie auf die Approximierung der Ortsableitungen verzichtet werden kann und die Ortsableitungen stattdessen *exakt* ausgedrückt werden können. Darüber hinaus werden die folgenden Betrachtungen einen klareren Blick auf das Regularisierungsprinzip aktiver Konturmodelle eröffnen.

3.3.3 Schnelle Lösung im Spektralbereich mittels FFT

Ausgangspunkt ist die Euler-Differentialgleichung mit konstanten Koeffizienten α und β .

$$\frac{\mathbf{v}^t - \mathbf{v}^{t-1}}{\tau} - \alpha \mathbf{v}^{t''} + \beta \mathbf{v}^{t''''} - \mathbf{F}_{\text{ext}}(\mathbf{v}^{t-1}) = 0 \quad (3.72)$$

Da die Kontur \mathbf{v} in Form einer abgetasteten periodischen Funktion vorliegt, kann sie mit Hilfe der diskreten Fouriertransformation (*Discrete Fourier Transform*) $\text{DFT}\{\cdot\}$ bzw. mit ihrem wesentlich effizienteren Äquivalent, der schnellen Fouriertransformation (*Fast Fourier Transform*) $\text{FFT}\{\cdot\}$, in den Spektralbereich überführt werden. Dasselbe gilt für die externen Kräfte $\mathbf{F}_{\text{ext}}(\mathbf{v}^{t-1})$.

$$\mathbf{v} \quad \circ\bullet \quad \hat{\mathbf{v}} = \text{FFT}\{\mathbf{v}\} \quad (3.73)$$

$$\mathbf{F}_{\text{ext}}(\mathbf{v}^{t-1}) \quad \circ\bullet \quad \hat{\mathbf{F}}_{\text{ext}}(\mathbf{v}^{t-1}) = \text{FFT}\{\mathbf{F}_{\text{ext}}(\mathbf{v}^{t-1})\} \quad (3.74)$$

Die Ortsableitungen transformieren sich dabei anhand des Differentiationstheorems der Fouriertransformation in komplexwertige Multiplikationen. Der dabei entstehende Vektor $\hat{\mathbf{k}}$ wird als Steifigkeitsspektrum (*Stiffness Spectrum*) bezeichnet.

$$\frac{\hat{\mathbf{v}}^t - \hat{\mathbf{v}}^{t-1}}{\tau} + [-\alpha (j2\pi\nu/N)^2 + \beta (j2\pi\nu/N)^4] \odot \hat{\mathbf{v}}^t - \hat{\mathbf{F}}_{\text{ext}}(\mathbf{v}^{t-1}) = 0 \quad (3.75)$$

$$\frac{\hat{\mathbf{v}}^t - \hat{\mathbf{v}}^{t-1}}{\tau} + \underbrace{[\alpha (2\pi\nu/N)^2 + \beta (2\pi\nu/N)^4]}_{\hat{\mathbf{k}}} \odot \hat{\mathbf{v}}^t - \hat{\mathbf{F}}_{\text{ext}}(\mathbf{v}^{t-1}) = 0 \quad (3.76)$$

Die Operation \odot führt die elementweise Multiplikation von Vektoren (das so genannte Hadamard-Produkt) aus, und der Vektor ν ist gegeben durch $\nu = [0 \ 1 \ 2 \ \dots \ N-1]^T$. Nach einigen elementaren Umformungen ergibt sich zunächst

$$\hat{\mathbf{v}}^t = \frac{1}{\mathbf{1} + \tau \hat{\mathbf{k}}} \odot \left(\hat{\mathbf{v}}^{t-1} + \tau \hat{\mathbf{F}}_{\text{ext}}(\mathbf{v}^{t-1}) \right) \quad (3.77)$$

mit dem Vektor $\mathbf{1} = [1 \ 1 \ \dots \ 1]^T$ der Länge N . Bis dato sind sowohl die Kontur als auch die externen Kräfte in Form ihrer *Spektren* gegeben. Um jedoch die externen Kräfte berechnen zu können, müssen die *Ortskoordinaten* der Kontur vorliegen. Daher müssen die Kontursamples in jedem Iterationsschritt zwischen dem Orts- und Spektralbereich hin- und zurücktransformiert werden. Durch Rücktransformation von Gleichung (3.77) aus dem Spektral- in den Ortsbereich erhält man schließlich die Lösung der Bewegungsgleichung.

$$\mathbf{v}^t = \text{IFFT} \left\{ \frac{1}{\mathbf{1} + \tau \hat{\mathbf{k}}} \odot \text{FFT} \left\{ \mathbf{v}^{t-1} + \tau \mathbf{F}_{\text{ext}}(\mathbf{v}^{t-1}) \right\} \right\} \quad (3.78)$$

Dieses Lösungsschema mag auf den ersten Blick etwas ungewöhnlich erscheinen – schließlich erfordert es in jedem Iterationsschritt jeweils eine FFT und eine IFFT. Im Vergleich zur klassischen Lösungsformulierung mittels Matrix-Vektor-Multiplikation nach Gleichung (3.71) ist dieses

Lösungsschema jedoch wesentlich effizienter. Wird die Anzahl der Konturabtastpunkte N derart gewählt, dass sich die Sequenz leicht faktorisieren lässt (im Idealfall wird N zu einer Zweierpotenz gewählt, aber auch größere Primfaktoren sind unproblematisch), so ist die schnelle Fouriertransformation mit einem Aufwand von $\mathcal{O}(N \log N)$ durchführbar. Seit der (Wieder)entdeckung³ der schnellen Fouriertransformation durch Cooley und Tukey [CT65] werden die Algorithmen für eine effiziente Berechnung der diskreten Fouriertransformation beständig verbessert. Eine der erfolgreichsten Software-Bibliotheken ist die FFTW [FJ, FJ05], welche unter der GNU General Public License (GPL) frei verfügbar ist. Später werden Ergebnisse zu den Laufzeiteigenschaften dieses Lösungsschemas unter Nutzung dieser Bibliothek angegeben.

Bei der Berechnung der Fouriertransformation in Gleichung (3.78) sollte weiterhin Folgendes beachtet werden: Die x - und y -Koordinaten der Kontur liegen als *reelle* Zahlen vor. In der Regel sind FFT-Algorithmen jedoch für *komplexwertige* Sequenzen ausgelegt. Daher werden im Falle reeller Eingangsdaten unnötige Rechenoperationen ausgeführt. Gute FFT-Implementierungen (wie z.B. die FFTW) bieten die Möglichkeit, die Algorithmen auf die Reellwertigkeit der Eingangsdaten einzustellen und damit die Rechengeschwindigkeit in etwa zu verdoppeln. Ungeachtet dieser Möglichkeit bietet sich als generelle Lösung die geschickte Darstellung der Koordinaten in komplexwertiger Form an: Dazu beschreibt man sowohl die x - und y -Koordinaten als auch die externen Kräfte durch komplexe Zahlen $z = x + jy$. Auf diese Weise lässt sich in jedem Fall der volle Geschwindigkeitsvorteil einer FFT-basierten Implementierung ausschöpfen.

Auf Basis dieser eben hergeleiteten, sehr effizienten Lösung soll nun untersucht werden, ob sich der Rechenaufwand möglicherweise noch weiter senken lässt.

3.3.4 Lösung durch zyklische Faltung

Die Lösung über den Spektralbereich mittels FFT nach Gleichung (3.78) lässt sich leicht in ein Lösungsschema umformen, welches im Ortsbereich arbeitet und ohne eine Approximation der Ableitungen durch finite Differenzen auskommt. Dazu ist die inverse Fouriertransformation anstatt auf die gesamte Gleichung auf die einzelnen Faktoren anzuwenden. Dabei transformiert sich die Multiplikation in eine zyklische Faltung \circledast .

$$\mathbf{v}^t = \text{IFFT} \left\{ \frac{1}{\mathbf{1} + \tau \hat{\mathbf{k}}} \right\} \circledast (\mathbf{v}^{t-1} + \tau \mathbf{F}_{\text{ext}}(\mathbf{v}^{t-1})) \quad (3.79)$$

Zunächst soll geklärt werden, in welchem Zusammenhang diese Gleichung mit dem klassischen Lösungsschema nach KASS *et al.* [KWT88] in Gleichung (3.71) steht. Wie bereits beschrieben, ist die Steifigkeitsmatrix \mathbf{K} im Falle konstanter Parameter α und β symmetrisch und *zyklisch (zirkulant)*. Zirkulante Matrizen haben folgende besondere Eigenschaft: Ihre Eigenwerte und Eigenvektoren lassen sich mittels der diskreten bzw. schnellen Fouriertransformation bestimmen. Eine zyklische Matrix „wirkt“ wie eine zyklische Faltung. Daher ist ihr Eigenwertspektrum gleich dem Fourierspektrum und ihre Eigenwerte und Eigenvektoren lassen sich mittels der diskreten bzw. schnellen Fouriertransformation berechnen. So ist die Eigenwertzerlegung einer zyklischen Matrix \mathbf{C} gegeben durch

$$\mathbf{C} = \mathbf{W} \mathbf{\Lambda} \mathbf{W}^H \quad (3.80)$$

mit der Fouriermatrix \mathbf{W} und der Diagonalmatrix $\mathbf{\Lambda}$, welche die Fourierkoeffizienten

$$\text{diag}\{\mathbf{\Lambda}\} = \text{FFT}\{\mathbf{C}(:, 1)\} \quad (3.81)$$

enthält. Die Notation $\mathbf{C}(:, 1)$ wählt dabei die erste Spalte der Matrix \mathbf{C} aus⁴. Mit Hilfe dieser Theorie kann eine zyklische Matrix mit einem Aufwand von $\mathcal{O}(N \log N)$ (anstatt $\mathcal{O}(N^3)$) für

³Die schnelle Fouriertransformation wurde bereits von GAUSS beschrieben [Gau66, HJB84].

⁴Aus Gründen einer unmissverständlichen Notation wird an dieser Stelle die Indexschreibweise angewendet, wie sie beispielsweise in MATLAB verwendet wird. Die Operation $\mathbf{C}(:, 1)$ selektiert dabei *alle* Elemente der *ersten* Spalte der Matrix \mathbf{C} .

den allgemeinen Fall) invertiert werden. Die Inverse einer zyklischen Matrix ist wieder zyklisch und ihre erste Spalte ist gegeben durch

$$\mathbf{C}^{-1}(:, 1) = \text{IFFT} \left\{ \frac{1}{\text{FFT}\{\mathbf{C}(:, 1)\}} \right\}. \quad (3.82)$$

Nach Umformung des Faltungskerns aus Gleichung (3.79)

$$\text{IFFT} \left\{ \frac{1}{\mathbf{1} + \tau \hat{\mathbf{k}}} \right\} = \text{IFFT} \left\{ \frac{1}{\mathbf{1} + \tau \text{FFT}\{\mathbf{k}\}} \right\} = \text{IFFT} \left\{ \frac{1}{\text{FFT}\{[1 \ 0 \ \dots \ 0]^T + \tau \mathbf{k}\}} \right\} \quad (3.83)$$

und mit Hilfe der Gleichung (3.82) überzeugt man sich leicht davon, dass der Vektor \mathbf{k} , welcher nach Gleichung (3.76) gegeben ist durch

$$\mathbf{k} = \text{IFFT} \left\{ \underbrace{\alpha (2\pi\nu/N)^2 + \beta (2\pi\nu/N)^4}_{\hat{\mathbf{k}}} \right\}, \quad (3.84)$$

dem *exakten* Steifigkeitskern entspricht, welcher in der klassischen Lösung in Gleichung (3.71) durch finite Differenzen approximiert wurde.

Durch die Überführung der Euler-Differentialgleichung in den Spektralbereich und anschließender Rücktransformation in den Ortsbereich unter Nutzung der zyklischen Faltung lässt sich also ein Lösungsschema formulieren, mit dem sich die Ortsableitungen analytisch ausdrücken lassen. Die Unterschiede zwischen der exakten Lösung und der Approximation durch finite Differenzen lassen sich in der Darstellung der Steifigkeitsspektren in Abbildung 3.18 erkennen: Der Hochpasscharakter der approximierten Lösung (gestrichelt) lässt im Vergleich zum Spektrum der exakten Lösung (durchgezogen) bei hohen Frequenzen deutlich nach.

Im nächsten Schritt soll die numerische Komplexität dieses Lösungsschemas untersucht werden. Der Stabilisierungskern $\text{IFFT}\{1/(1 + \tau \hat{\mathbf{k}})\}$ aus Gleichung (3.79) hat eine Länge von $M = N$ Samples – damit ergibt sich eine Komplexität von $\mathcal{O}(N^2)$. Bei näherer Betrachtung des Stabilisierungskerns (vgl. Abbildung 3.19) fällt auf, dass seine Einhüllende sehr schnell abklingt und damit eine Approximation dieses Faltungskerns mit wenigen Samples

$$M \ll N$$

möglich ist. Durch diese Approximation wächst die Länge des Faltungskerns bei einer Erhöhung der Konturlänge nicht mehr mit und es ergibt sich ein Aufwand von $\mathcal{O}(MN)$ – also eine *lineare* Komplexität.

3.3.5 Diskussion und Vergleich der Lösungsschemata

Obwohl die Idee, die den Schlangen zugrunde liegende Differentialgleichung im Spektralbereich zu lösen, nahe liegend erscheint, wurde sie erst jüngst von Weruaga *et al.* in [WVM04] publiziert. Die Autoren leiten die Lösung der Bewegungsgleichung im Spektralbereich direkt aus dem zu minimierenden Energiefunktional über das Theorem von Parseval ab – ohne den „Umweg“ über das Kräftegleichgewicht der Euler-Differentialgleichung.

Der hier vorgeschlagene „Geradeaus-Lösungsansatz“ [IS05] mittels Fouriertransformation der Euler-Differentialgleichung bildet eine Brücke zwischen dem klassischen Lösungsansatz nach KASS *et al.* [KWT88] und der Formulierung im Spektralbereich nach Weruaga *et al.* [WVM04] und konnte bisher in der Literatur noch nicht aufgefunden werden.

In Anhang D sind Ergebnisse zum tatsächlichen Laufzeitverhalten dargestellt. Im Ergebnis ist das Lösungsschema mittels schnellerer Fouriertransformation die erste Wahl.

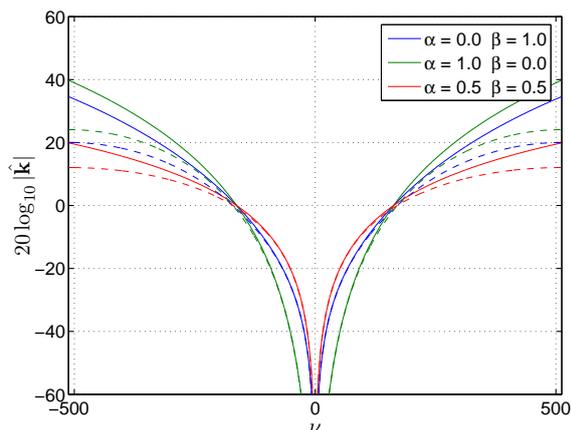


Abbildung 3.18: Steifigkeitsspektrum $\hat{\mathbf{k}}$ in logarithmischer Darstellung für $N = 1024$. Durchgezogene Linie: exakte Lösung; gestrichelte Linie: Approximation mittels finiter Differenzen.

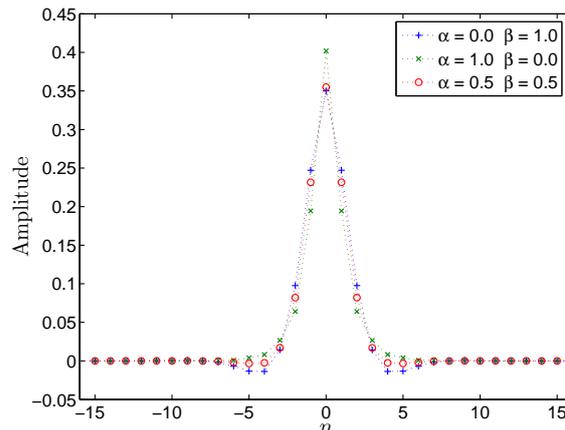


Abbildung 3.19: Ausschnitt des Stabilisierungskerns $\text{IFFT}\{1/(1 + \tau \hat{\mathbf{k}})\}$ für $N = 1024$.

3.3.6 Anwendung zur genauen Bestimmung der Zellberandung

Mit den vorgestellten mathematischen Zusammenhängen sind nun die Grundlagen für eine effiziente Berechnung aktiver Konturmodelle gelegt. Für das Segmentierungsergebnis ist jedoch die Wahl eines dem Problem angepassten Energiefunktionals bzw. Kräftefeldes entscheidend. In der Arbeit von KASS *et al.*, in der das Schlangenmodell ursprünglich vorgestellt wird, finden sich einige Vorschläge zur Orientierung der Kontur an *Linien*, *Kanten*, *Objektenden* sowie zur Problematik *Stereo* und *Bewegung*.

Für die Orientierung der Schlange an transformierten Zellen in den vorliegenden Farbbildern ist zweifellos ein Energiefunktional zu wählen, welches die *Kanteninformation* des Bildes widerspiegelt. Der Standardansatz dazu ist das Funktional nach Gleichung (3.63). Xu [XP98] widmete sich dem Problem des schlechten Konvergenzverhaltens der Schlange an konkaven Objektstellen, die dieses Funktional mit sich bringt, und schlug das GVF-Feld (GVF = Gradient Vector Flow) als Lösung vor. Dieses Feld weist einen sehr weiten „Fangbereich“ auf und zieht selbst eine weit vom Objekt entfernt initialisierte Kontur auch in ausgeprägte konkave Regionen hinein.

Zur genauen Segmentierung der transformierten Zellen hat sich jedoch eine Modifikation des Standardansatzes $E_{\text{ext}}(v(s)) = -|\nabla I_{\sigma}(x, y)|^2$ nach Gleichung (3.63) bzw. des zugehörigen Kräftefeldes

$$F_{\text{ext}} = \nabla E_{\text{ext}} = \nabla(-|\nabla I_{\sigma}(x, y)|^2) \quad (3.85)$$

bewährt. Die Vektoren der an der Schlange angreifenden äußeren Kräfte lassen sich in zwei Komponenten zerlegen – eine Normal- und eine Tangentialkomponente. Für die Bewegung der Schlange sind die Tangentialkomponenten jedoch eher kontraproduktiv; sie dehnen und stauchen die Kontur, anstatt sie zu einer konstruktiven Bewegung zu veranlassen. In der Implementierung von Xu⁵ werden die Positionen der Konturpunkte regelmäßig reorganisiert, damit die Schlange durch die Tangentialkräfte nicht entartet. Dieses Problem lässt sich jedoch von vornherein vermeiden, wenn nur die *Normalkomponente* der Kräftevektoren auf die Schlange wirkt. Zur Berechnung der Normalvektoren der Kontur sei an dieser Stelle auf das später folgende Kapitel 4.3 zur konturbasierten Formanalyse verwiesen.

⁵Die Implementierung kann von <http://iacl.ece.jhu.edu/projects/gvf/> heruntergeladen werden.

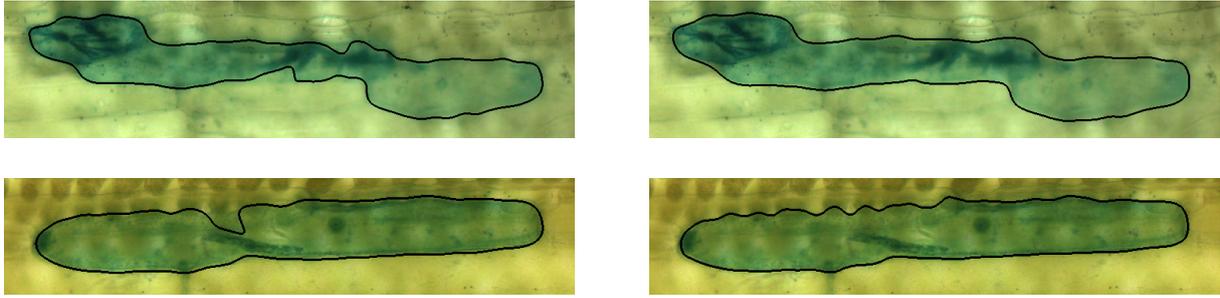


Abbildung 3.20: Einfluss des Farbraumes auf das Segmentierungsergebnis aktiver Konturmodelle. *Links*: Energiefunktional auf Basis des Bunttons (*Hue*). *Rechts*: Energiefunktional auf Basis des adaptiven Farbraums. Der adaptive Farbraum ist deutlich weniger anfällig für Störungen durch Objekte innerhalb der Zelle.

Ergebnisse und Diskussion

Nach der Spezifizierung eines geeigneten äußeren Kräftefeldes bleibt noch zu klären, aus welchen Bilddaten dieses Kräftefeld berechnet werden sollte. Nach den Ausführungen in Abschnitt 3.1.5 stehen sowohl der adaptive Farbraum auf Basis des multivariaten Gaußmodells als auch der Buntton als Kandidaten zur Verfügung.

Es lohnt sich an dieser Stelle, den adaptiven Farbraum einzusetzen. Die Regionen transformierter Zellen liegen nach der Grobsegmentierung mittels Canny's Kantendetektor als rechteckige Bildausschnitte vor, ähnlich dem Ausschnitt nach Abbildung 3.1 (linkes Bild). Innerhalb dieser Ausschnitte können die Parameter des adaptiven Farbraumes sehr zuverlässig und mit vertretbarem Aufwand geschätzt werden. Damit lassen sich potenzielle Probleme umgehen, welche durch Strukturen (z.B. Haustorien) innerhalb der Zelle verursacht werden, welche den Buntton ebenfalls beeinflussen und die Zellfärbung damit möglicherweise „überstrahlen“. Der adaptive Farbraum ist diesbezüglich deutlich weniger anfällig, weil er sich konkret auf das vorliegende Bildmaterial anzupassen vermag. Auf diese Weise lässt sich selbst von sehr schwach gefärbten Zellen die Berandung zuverlässig ermitteln.

Die Initialkontur $v(s;0) = v_0(s)$ bildet die geschlossene Kante, welche vom Canny-Kantendetektor bei großzügiger Glättung (grobe Skala) (vgl. Abbildung 3.16) bereitgestellt wird. Die Schlange wird nun *durch den Skalenraum hindurch* der tatsächlichen Zellgrenze immer feiner angenähert, indem das externe Energiefunktional stufenweise bei immer kleiner werdenden Skalen auf die Kontur wirkt.

3.4 Segmentierung der Haustorien

Nachdem die transformierten Zellen im Bildmaterial gefunden und segmentiert sind, ist das Ziel der nächsten Verarbeitungsstufe, potenzielle Haustorienregionen innerhalb dieser Zellen zu detektieren. Das sichere Herauslösen der Haustorien vom restlichen Zellmaterial ist jedoch aus folgenden Gründen ein recht anspruchsvolles Unterfangen:

- Generell heben sich Haustorien bezüglich ihres Farbkontrasts nur sehr schwach vom restlichen Zellgewebe ab.
- Haustorien zeichnen sich durch eine relativ hohe Variabilität bezüglich ihrer *Lage*, *Größe*, *Form* und insbesondere ihrer *Färbung* aus.
- Dazu kommt, dass die zu untersuchenden Zellen selten „sauber“ sind; häufig sind diverse Verfärbungen innerhalb der Zelle zu beobachten, die *nicht* von Haustorien stammen, sondern anderen Objekten (z.B. dem Zellkern) zuzurechnen sind.

- Häufig erscheinen Bereiche der Zellwand in einer ähnlichen Färbung wie potenzielle Haustorien.

Daher ist davon auszugehen, dass die *alleinige* Verwendung des Merkmals „Farbe“ nur eine begrenzt zuverlässige Erkennung von Haustorien zu liefern vermag.

Auf den nächsten Seiten werden zwei Verfahren vorgestellt, mit denen versucht wird, das sehr spezifische Problem der Haustoriensegmentierung zu lösen. Das erste Verfahren greift dabei die Erkenntnisse aus Abschnitt 3.1 auf und führt eine Clusterung im Farbraum mit Hilfe des Expectation Maximization (EM)-Algorithmus durch. Völlig konträr dazu ist der zweite Ansatz, welcher Methoden der mathematischen Morphologie nutzt, um den Farbkontrast der gesuchten Haustorienregionen zu erhöhen, so dass eine robuste Segmentierung ermöglicht wird.

In jedem Fall muss sichergestellt sein, dass die Haustorien in den Bildern *scharf* abgebildet sind. Daher wird zunächst die Bestimmung der korrekten Fokusebene diskutiert, bevor die Segmentierungsverfahren vorgestellt werden.

3.4.1 Selektion der Fokusebene

Für die Haustoriensegmentierung ist unbedingt ein korrekt fokussiertes Bild notwendig. Da die Blattoberfläche ein leicht hügeliges Profil aufweist, ist eine scharfe Abbildung aller im Bildausschnitt der Kamera liegenden Zellen nicht immer möglich, wie in Abbildung 1.7 auf Seite 8 gezeigt wurde. Bei der Bildaufnahme werden daher mehrere Bilder mit leicht verändertem Fokus in Form eines Z-Stapels bereitgestellt, so dass eine Überdeckung der Schärfeebenen aller im Bildausschnitt liegenden Objekte gewährleistet ist. In der Praxis sind dazu fünf fokale Ebenen ausreichend.

Für die Haustoriensegmentierung ist nun jenes Bild aus dem Z-Stapel auszuwählen, welches die *schärfste* Abbildung der betreffenden Zelle aufweist. Zur Bestimmung der Bildschärfe lassen sich verschiedene Verfahren einsetzen, z.B. die Messung des lokalen Bildgradienten, eine Spektralanalyse mittels Fouriertransformation oder auch die Bestimmung der lokalen Entropie [NKCNAJ01]. Ein sehr häufig verwendetes und praktikables Maß ist der so genannte Tenenbaum-Gradient [Ten70], welcher auch unter der Bezeichnung „Tenengrad“ bekannt ist. Zur Berechnung dieses Schärfemaßes werden die Bildableitungen $\frac{\partial I}{\partial x}$ und $\frac{\partial I}{\partial y}$ (z.B. mittels des Sobel-Operators nach Gleichung (C.8)) bestimmt und deren Quadrate addiert. Über den entsprechenden Bildausschnitt gemittelt, ergibt sich damit eine zuverlässige Abschätzung der Bildschärfe. Bei *Farbbildern* ist an dieser Stelle zunächst mittels einer Farbraumtransformation ein geeignetes Merkmalsbild zu erzeugen, in welchem sich die gesuchten Strukturen möglichst gut vom Hintergrund abheben. Im vorliegenden Fall hat sich dazu als geeignetes Merkmal der Buntton bewährt. Für jede zu untersuchende Zellregion wird für alle fokalen Ebenen der mittlere Tenenbaum-Gradient auf Basis des Bunttons bestimmt und jene Ebene für die Haustoriensegmentierung ausgewählt, für die dieses Maß maximal ist.

Auf Basis dieser korrekt fokussiert abgebildeten Zellregion kann nun nach Haustorien gesucht werden.

3.4.2 Farbbasierte Segmentierung mittels EM-Clustering

Bei diesem Verfahren wird jeder einzelne Pixel des zu untersuchenden rechteckigen Bildausschnitts (*Region of Interest, ROI*) mittels des Expectation Maximization (EM)-Algorithmus den drei Klassen „Hintergrund“, „transformierte Zelle“ und „Haustorienregion“ zugeordnet. Dazu werden die Parameter μ_k und Σ_k der drei Klassen mit Daten von repräsentativen Beispielen initialisiert und schrittweise mittels des EM-Algorithmus (vgl. Abschnitt 3.1.3) an den jeweiligen zu analysierenden Bildausschnitt angepasst.

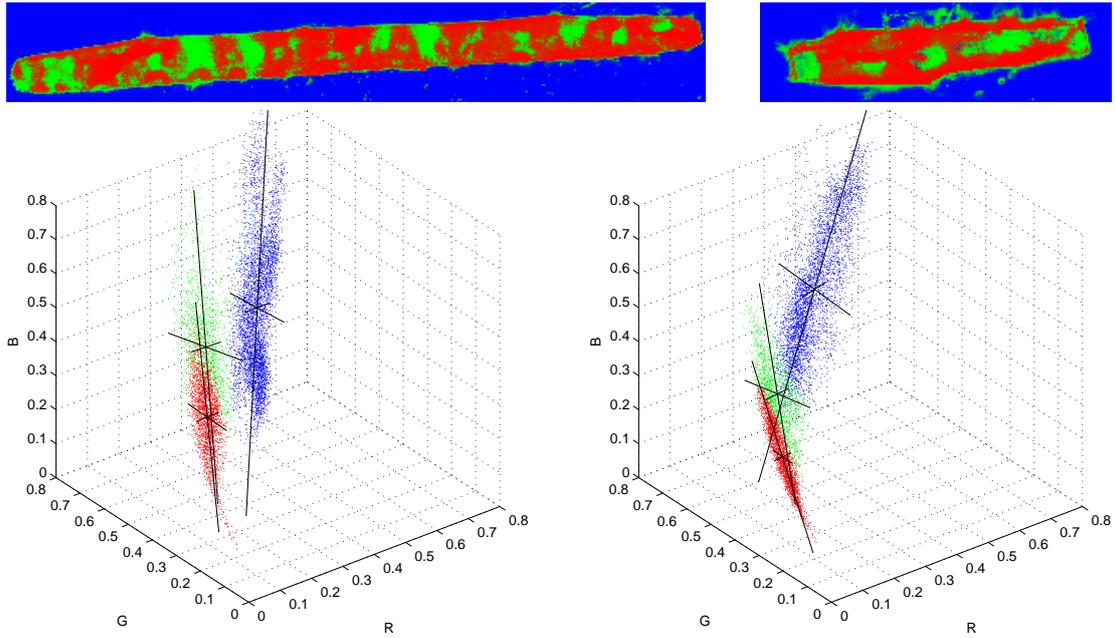


Abbildung 3.21: Segmentierungsergebnisse des *klassischen* EM-Algorithmus in die Klassen „Hintergrund“ (blau), „transformierte Zelle“ (grün) und „Haustorium“ (rot) anhand von zwei Beispielen. Auffällig ist die starke Überbewertung der Klasse „Haustorium“ (rot). Dieses Problem wird durch eine Modifikation des EM-Algorithmus gelöst, deren Ergebnisse in Abbildung 3.22 dargestellt sind.

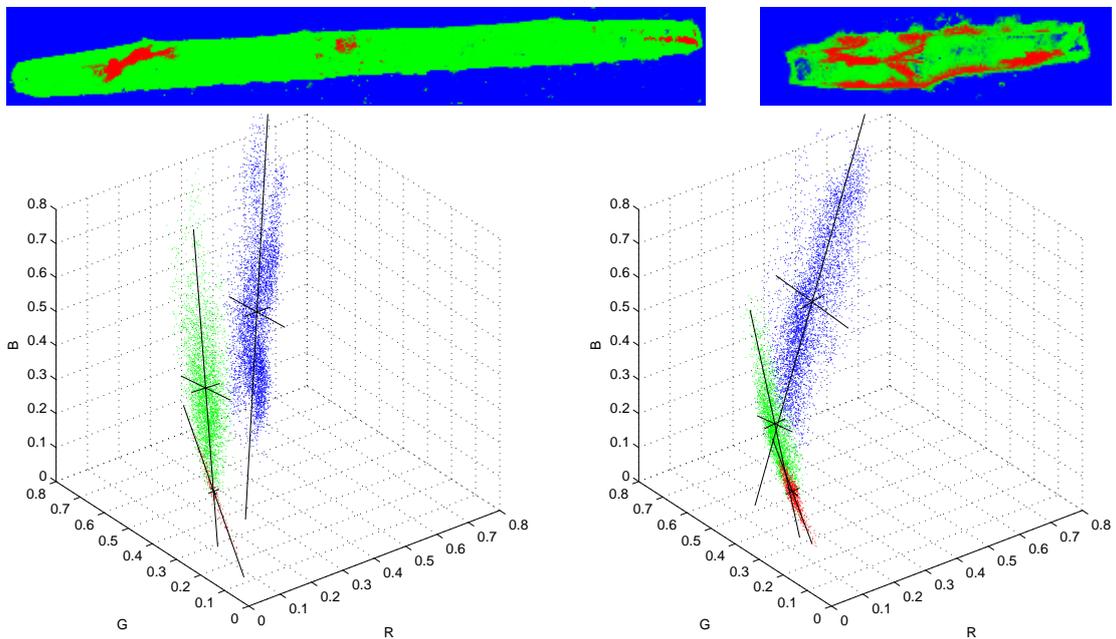


Abbildung 3.22: Segmentierungsergebnisse des *modifizierten* EM-Algorithmus. Im M-Schritt werden nur Samples berücksichtigt, die eine Entscheidungssicherheit von mindestens $R_{min} = 0.65$ aufweisen (vgl. Tabelle 3.1). Gegenüber dem klassischen EM-Algorithmus (siehe Abbildung 3.21) wird ein deutlich verbessertes Segmentierungsergebnis erreicht.

Direkte Anwendung des EM-Algorithmus

Ähnlich wie bei der Generierung des Optimalfarbraumes für die Zellsegmentierung (vgl. Abschnitt 3.1), bei der mit $K = 2$ Klassen gearbeitet wird, gelingt zunächst die Detektion des Hintergrundes sehr gut, wie die Ergebnisse in Abbildung 3.21 zeigen. Bei der Unterscheidung zwischen den Klassen „transformierte Zelle“ und „Haustorium“ wird jedoch die Klasse „Haustorium“ sehr stark überbewertet. Der EM-Algorithmus konvergiert gegen ein Clusterungsergebnis, welches zu keiner brauchbaren Haustoriensegmentierung führt. Der Cluster der Klasse „Haustorium“ hängt mit der Klasse „transformierte Zelle“ sehr eng zusammen, er bildet sozusagen ein kleines „Anhängsel“, und es gelingt mittels des EM-Algorithmus auf diese Weise nicht, diese beiden Klassen korrekt zu trennen. Egal wie präzise die initialen Klassenparameter auch sein mögen, es wandern während des Iterationsprozesses fortlaufend immer mehr Pixel von der Klasse „transformierte Zelle“ in die Klasse „Haustorium“ über, was letztlich zu der beobachteten Überbewertung der Klasse „Haustorium“ führt.

Adaption des EM-Algorithmus – „Constraint-EM-Algorithmus“

Durch einen geschickten Eingriff in den Iterationsprozess gelingt es jedoch, das Segmentierungsergebnis deutlich zu verbessern: Der Lösungsansatz besteht darin, die Klassenparameter während der Iterationen in gewissen Grenzen „festzuklammern“ und damit eine zu starke Abweichung von den Initialwerten zu verhindern.

Während jeder Iteration des EM-Algorithmus wird im E-Schritt (vgl. Gleichung (3.13)) ohnehin zunächst die Wahrscheinlichkeit der Klassenzugehörigkeit $P^t(k | \mathbf{x}_n)$ jedes Samples \mathbf{x}_n bestimmt. Anschließend wird das Sample jener Klasse mit maximaler Klassenzugehörigkeitswahrscheinlichkeit zugeordnet; dabei ist die resultierende Entscheidungssicherheit gegeben durch

$$R = \max_k \{P^t(k | \mathbf{x}_n)\} \in [1/K \dots 1]. \quad (3.86)$$

Unter Nutzung der Entscheidungssicherheit R ist es möglich, Samples, die nur sehr unsicher zugeordnet werden konnten, bei der Parameteraktualisierung im M-Schritt (vgl. Gleichungen (3.14) – (3.16)) zu ignorieren. Dazu wird der Datensatz, der im M-Schritt berücksichtigt wird, auf die Menge der Samples beschränkt, die mit einer Entscheidungssicherheit von mindestens R_{min} zugeordnet wurden.

In Tabelle 3.1 ist eine Testreihe mit den Segmentierungsergebnissen für einen weiten Variationsbereich des Parameters R_{min} dargestellt. In den verwendeten zwei Beispielbildern zeigt sich eine schlagartige Verbesserung der Segmentierung, wenn mit einer Beschränkung auf $R_{min} > 0.60$ gearbeitet wird. Abbildung 3.22 zeigt detaillierte Ergebnisse für $R_{min} = 0.65$. Verglichen mit Abbildung 3.21 zeigt Abbildung 3.22 keine Überbewertung der Klasse „Haustorium“ mehr; vielmehr sind die Segmentierungsergebnisse durchaus vielversprechend. Mit Hilfe dieses „Constraint-EM-Algorithmus“ werden die Haustorien in der links abgebildeten Zelle (abgesehen von einigen wenigen Fehlklassifikationen durch Verfärbungen) korrekt segmentiert. In der rechts abgebildeten Zelle werden die zwei Haustorien ebenfalls korrekt markiert, jedoch treten dort gleichermaßen einige Bereiche hervor, die kein Haustorium darstellen, sondern zur Zellwand gehören. Solche Verfärbungen sind in ihren Farbparametern denen der Haustorien sehr ähnlich und werden daher von farbbasierten Clusterverfahren immer mitdetektiert.

Bisher wurde das Bild in $K = 3$ Klassen unterteilt (segmentiert). Wird die transformierte Zelle jedoch (z.B. mit Hilfe der Methoden aus Abschnitt 3.3) korrekt anhand ihrer Berandung ausgeschnitten, so kann bei der Clusterung auf die Klasse „Hintergrund“ verzichtet und das Bild stattdessen in die zwei Klassen „transformierte Zelle“ und „Haustorium“ zerlegt werden. Jedoch kommt es auch bei $K = 2$ Klassen zu dem oben beschriebenen „Abgleiten“ in Richtung der Klasse „Haustorium“, so dass auch in diesem Fall auf die beschriebene Constraining-Methode zurückgegriffen werden muss.

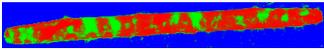
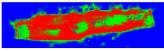
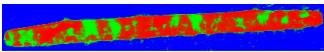
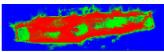
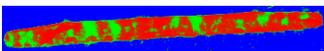
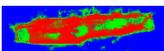
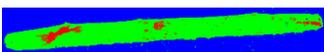
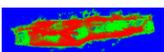
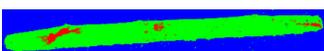
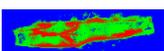
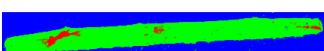
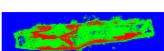
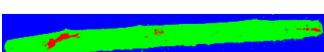
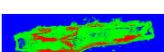
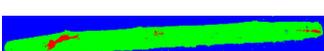
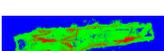
R_{min}	Segmentierungsergebnisse	
0.40		
0.45		
0.50		
0.55		
0.60		
0.65		
0.70		
0.75		

Tabelle 3.1: Segmentierungsergebnisse des modifizierten EM-Algorithmus bei Variation der minimal geforderten Entscheidungssicherheit R_{min} . Dieser Parameter bestimmt, wie zuverlässig ein Sample während der EM-Iteration mindestens zugeordnet sein muss, um für die Parameteraktualisierung im M-Schritt berücksichtigt zu werden.

Validierung der multivariaten Gaußverteilung

Im Gegensatz zum Mischmodell aus Abschnitt 3.1, bei dem das Bild in zwei Klassen unterteilt wurde und die Bereiche „transformierte Zelle“ und „Haustorium“ einer gemeinsamen Klasse zugeteilt wurden, wird das Bild nunmehr in $K = 3$ Farbklassen partitioniert. An dieser Stelle soll wieder die Gültigkeit des zugrunde gelegten multivariaten Gaußmodells überprüft werden. Dazu werden wieder die in Abschnitt 3.1.2 vorgestellten grafischen Methoden angewendet. Abbildung 3.23 und 3.24 zeigen die Histogramme der Hauptkomponenten der Farbcluster der Klassen bzw. die Mahalanobisdistanzen sowohl als Wahrscheinlichkeitsdichtefunktion als auch als Quantil-Quantil-Darstellung (Q-Q-Plot). Beide Testmethoden lassen dabei wieder auf die Zulässigkeit der Modellierung mittels multivariater Gaußverteilung schließen.

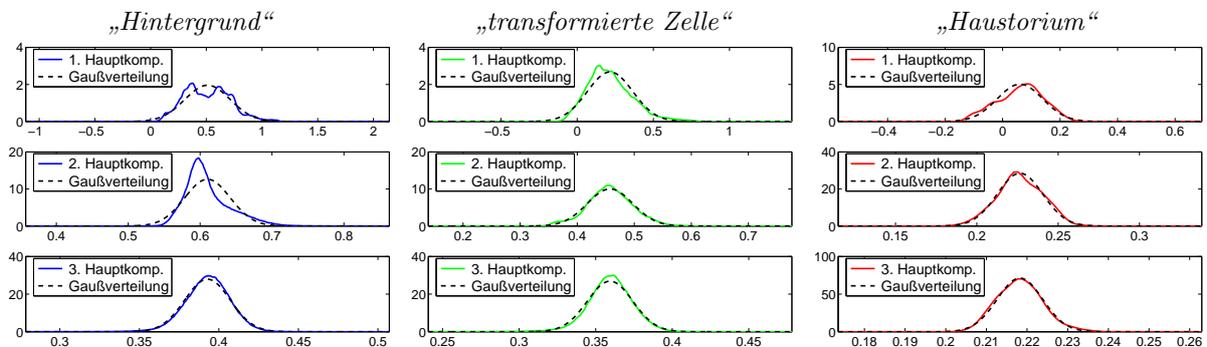


Abbildung 3.23: Verteilungsdichtefunktionen der Hauptkomponenten der Farbcluster von „Hintergrund“, „transformierte Zelle“ und „Haustorium“ für das Bild aus Abbildung 3.22 links.

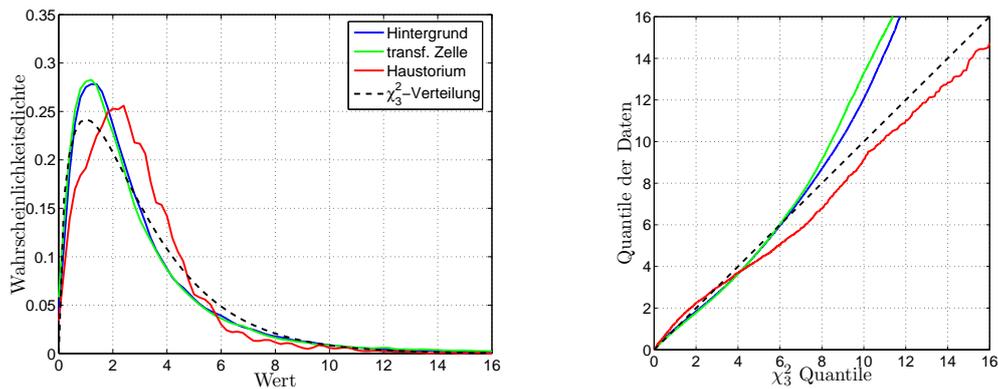


Abbildung 3.24: Wahrscheinlichkeitsdichte der Mahalanobis-Distanz (*links*) und Quantil-Quantil-Darstellung (Q-Q-Plot) (*rechts*).

Ergebnisse und Diskussion der EM-Clustering

Eine Segmentierung mittels Clustering im Farbraum führt bei „gutmütigen“ Daten bereits zu recht vielversprechenden Ergebnissen, kann jedoch die Anforderungen an reales Probenmaterial nicht vollständig befriedigen. Kernpunkt des Problems ist der schlechte Farbkontrast zwischen Haustorien und dem restlichen Zellgewebe, welcher die Unterscheidung von Haustorien und diversen Verfärbungen (z.B. von der Zellwand) schwierig macht. Eine Lösung für ein besseres Segmentierungsverfahren kann nur sein, die räumlichen Beziehungen der Pixel zueinander mit zu berücksichtigen. Im nächsten Abschnitt wird daher ein Segmentierungsverfahren vorgestellt, welches Methoden der mathematischen Morphologie benutzt, um gezielt Regionen hervorzuheben, die ein Haustorium darstellen könnten.

3.4.3 Segmentierung mittels morphologischer Kontrastverstärkung

Grundlagen der mathematischen Morphologie

Im Gegensatz zu linearen Filteroperationen (Faltung), die auf dem Prinzip der gewichteten Addition beruhen, basieren morphologische Operationen [Soi04] auf *nichtlinearen* Filteroperationen aus der Kategorie der Rangordnungsfiler. Solche Filter arbeiten nach dem Prinzip des Vergleichens und Selektierens benachbarter Pixel. Dabei werden die innerhalb einer Maske befindlichen Pixel zunächst gemäß ihrer Werte sortiert, anschließend wird das anhand eines bestimmten Kriteriums ausgewählte Ergebniselement in das Ergebnisbild geschrieben. Ein *Medianfilter* ist beispielsweise dadurch gekennzeichnet, dass das zentrale Element der sortierten Pixelwerte – der Median – als Ergebniselement ausgewählt wird.

Für die Graubildmorphologie werden ein *Maximumfilter* und ein *Minimumfilter* benötigt. Auf diesen Filtern basieren die beiden Fundamentaloperationen *Dilation* und *Erosion*, welche das Fundament der Graubildmorphologie bilden und mit deren Hilfe sich sehr leistungsfähige Werkzeuge realisieren lassen. Dabei beschreibt B das so genannte Strukturelement (*structuring element*), welches als Maske über das Bild geschoben wird.

$$\delta_B(I)_x = \max_{\beta \in B} \{I(x + \beta)\} \quad \text{Dilation} \quad (3.87)$$

$$\epsilon_B(I)_x = \min_{\beta \in B} \{I(x + \beta)\} \quad \text{Erosion} \quad (3.88)$$

Das morphologische *Öffnen* (*opening*) und *Schließen* (*closing*) ergibt sich aus der Kombination von jeweils einer *Dilation* und *Erosion*.

$$\gamma_B = \delta_B[\epsilon_B(I)] \quad \text{Öffnen (opening)} \quad (3.89)$$

$$\phi_B = \epsilon_B[\delta_B(I)] \quad \text{Schließen (closing)} \quad (3.90)$$

Auf Basis des *Öffnens* und *Schließens* lassen sich die so genannten *Zylinderhutoperationen* (*top hat operations*) wie folgt definieren:

$$WTH_B(I) = I - \gamma_B(I) \quad \text{weißer Zylinderhut (white top hat, opening top hat)} \quad (3.91)$$

$$BTH_B(I) = \phi_B(I) - I \quad \text{schwarzer Zylinderhut (black top hat, closing top hat)}. \quad (3.92)$$

Mittels dieser Operationen lässt sich schließlich eine bildinhaltsadaptive Kontrastverstärkung realisieren, wobei der weiße Zylinderhut helle Merkmale und der schwarze Zylinderhut dunkle Merkmale hervorhebt (vgl. Abbildung 3.26).

$$\kappa(I) = I + \underbrace{WTH_B(I)}_{\substack{\text{verstärkte helle} \\ \text{Merkmale}}} - \underbrace{BTH_B(I)}_{\substack{\text{verstärkte dunkle} \\ \text{Merkmale}}} \quad (3.93)$$

Zur Wahl des Strukturelements

Mittels Hintergrundwissen über Form und Größe von gesuchten Bildstrukturen lassen sich morphologische Operationen über eine geeignete Wahl des Strukturelements an das jeweilige Problem anpassen. Haustorien sind jedoch – bedingt durch ihre besonders komplexe und zudem sehr variable Form – recht schwer bzw. unmöglich hinreichend zu parametrieren. An dieser Stelle kommt folgende Eigenschaft der Zylinderhutoperationen zum Tragen: Zylinderhüte arbeiten *konträr*, d.h., sie verstärken Bildregionen, in die das Strukturelement *nicht* hineinpasst. Daher eignet sich zur Hervorhebung der Haustorien ein einfach geformtes Strukturelement, welches geringfügig größer ist als der „Bauch“ eines Haustoriums, beispielweise in Form einer Ellipse, einer Kreisscheibe oder eines Quadrates.

Farbmorphologie: Ansätze und Probleme der Verallgemeinerung

Die bisherigen Definitionen der Graubildmorphologie gelten – wie der Name sagt – nur für Graubilder. Einer „Geradeaus-Erweiterung“ zu einer „Farbbildmorphologie“ steht jedoch eine fundamentale Hürde im Weg: Für Skalare, wie sie innerhalb eines Graubildausschnitts unter der Filtermaske vorliegen, ist die Bildung eines Maximums bzw. Minimums zur Berechnung der Dilation bzw. Erosion offensichtlich. Das Prinzip einer Sortierung lässt sich jedoch nicht direkt von skalarwertigen Daten auf multivariate Daten übertragen, weil keine *eindeutige* Vorschrift zum sortierten Anordnen multivariater Samples existiert.

Es lassen sich jedoch verschiedene Verfahren definieren, mit denen eine Sortierung unter bestimmten Randbedingungen vorgenommen werden kann [Bar76, PV00]:

- *Marginal Ordering (M-ordering)*

Gemäß dieses Ansatzes wird eine *unabhängige* Sortierung der Vektorkomponenten vorgenommen.

$$\begin{aligned} x_{1(1)} &\leq x_{1(2)} \leq \dots x_{1(N)} \\ x_{2(1)} &\leq x_{2(2)} \leq \dots x_{2(N)} \\ &\vdots \\ x_{d(1)} &\leq x_{d(2)} \leq \dots x_{d(N)} \end{aligned} \quad (3.94)$$

Ein Nachteil dieser Sortiervorschrift ist, dass die entstehenden Vektoren – bedingt durch die *komponentenweise* Umsortierung – nicht notwendigerweise in den Originalvektoren vorkommen müssen. Aus Sicht der Bildverarbeitung bedeutet dies, dass das Ergebnisbild keine Untermenge des Ausgangsbildes mehr ist. Es können neue Farbwerte entstehen, die als Artefakte sichtbar werden.

- *Reduced Ordering (R-ordering)*

Beim *Reduced Ordering* werden die multivariaten Samples zunächst anhand einer Transformation $R: \mathbb{R}^d \rightarrow \mathbb{R}$ auf einen Skalar abgebildet (reduziert). Diese Skalare werden dann sortiert und die entstandene Sortierreihenfolge auf die multivariaten Samples übertragen.

$$\begin{aligned} R_{(1)} &\leq R_{(2)} \leq \dots R_{(N)} \\ \mathbf{x}_{[1]} &\leq \mathbf{x}_{[2]} \leq \dots \mathbf{x}_{[N]} \end{aligned} \quad (3.95)$$

Als Transformationsvorschrift kann beispielsweise die Mahalanobis-Distanz benutzt werden. Für weiter führende Überlegungen zur Wahl der Transformationsvorschrift sei an dieser Stelle auf die Literatur [Bar76, PV00] verwiesen.

- *Conditional Ordering (C-ordering)*

Hierbei handelt es sich um einen Spezialfall des *Reduced Ordering*. Die multivariaten Samples werden anhand der Sortierreihenfolge einer Randmenge (beispielsweise anhand der ersten Dimension) sortiert.

Eine allgemein gültige, universelle Lösung zur Übertragung der Graubildmorphologie auf Farbbilder ist mit keiner dieser vorgestellten multivariaten Sortiermethoden möglich. Als unkompliziertestes Verfahren stellt sich jedoch das Vorgehen anhand des *Marginal Ordering* dar, weil sich mit diesem Prinzip die bestehenden Algorithmen der Graubildmorphologie sehr einfach auf Farbbilder übertragen lassen. Dazu sind die einzelnen Kanäle des Bildes *unabhängig voneinander* mit der anzuwendenden morphologischen Operation zu beaufschlagen.

Anwendung der morphologischen Kontrastverstärkung auf Zellbilder

Die per *Marginal Ordering* ausgeführte morphologische Kontrastverstärkung bringt die Segmentierung der Haustorien einen entscheidenden Schritt voran. Generell zeichnen sich Haustorienbereiche durch eine *etwas größere Farbsättigung* und *etwas dunklere Färbung*, verglichen mit dem restlichen Zellgewebe, aus. Diese beiden Merkmale werden mittels der oben beschriebenen morphologischen Methode deutlich verstärkt. Abbildung 3.25 verdeutlicht das Prinzip anhand von vier Beispielbildern: In den ersten beiden Zeilen ist jeweils das Original- und das kontrastverstärkte Farbbild zu sehen. Die Zeilen drei und vier zeigen die zugehörige Farbsättigung (in Pseudofarbdarstellung). Gegenüber dem Originalbild weist das kontrastverstärkte Pendant *für Haustorienbereiche* eine deutlich erhöhte Farbsättigung auf, wobei einige Bereiche den Maximalwert erreichen. An Stellen maximaler Farbsättigung gibt es keine Beträge der Farbe *Rot* mehr; die Farbwerte sind durch die Kontrastverstärkung gegen die „Grün-Blau-Wand“ des RGB-Farbwürfels „gedrückt“ worden. Im Gegensatz zu Haustorienbereichen bleibt die Farbsättigung des *Zellgewebes* durch die Kontrastverstärkung nahezu unbeeinflusst.

Segmentierung mit zwei Schwellwerten

Aufgrund dieses Verstärkungseffekts seitens der Farbsättigung für Haustorienregionen ist nunmehr eine Segmentierung auf Basis eines einkanaligen Bildes (der Farbsättigung) möglich. Prinzipiell ließe sich diese Segmentierung mit einer einfachen Schwellwerthaltung durchführen, jedoch kann die Menge potenzieller Falschdetektionen durch Anwendung von *zwei Schwellwerten*

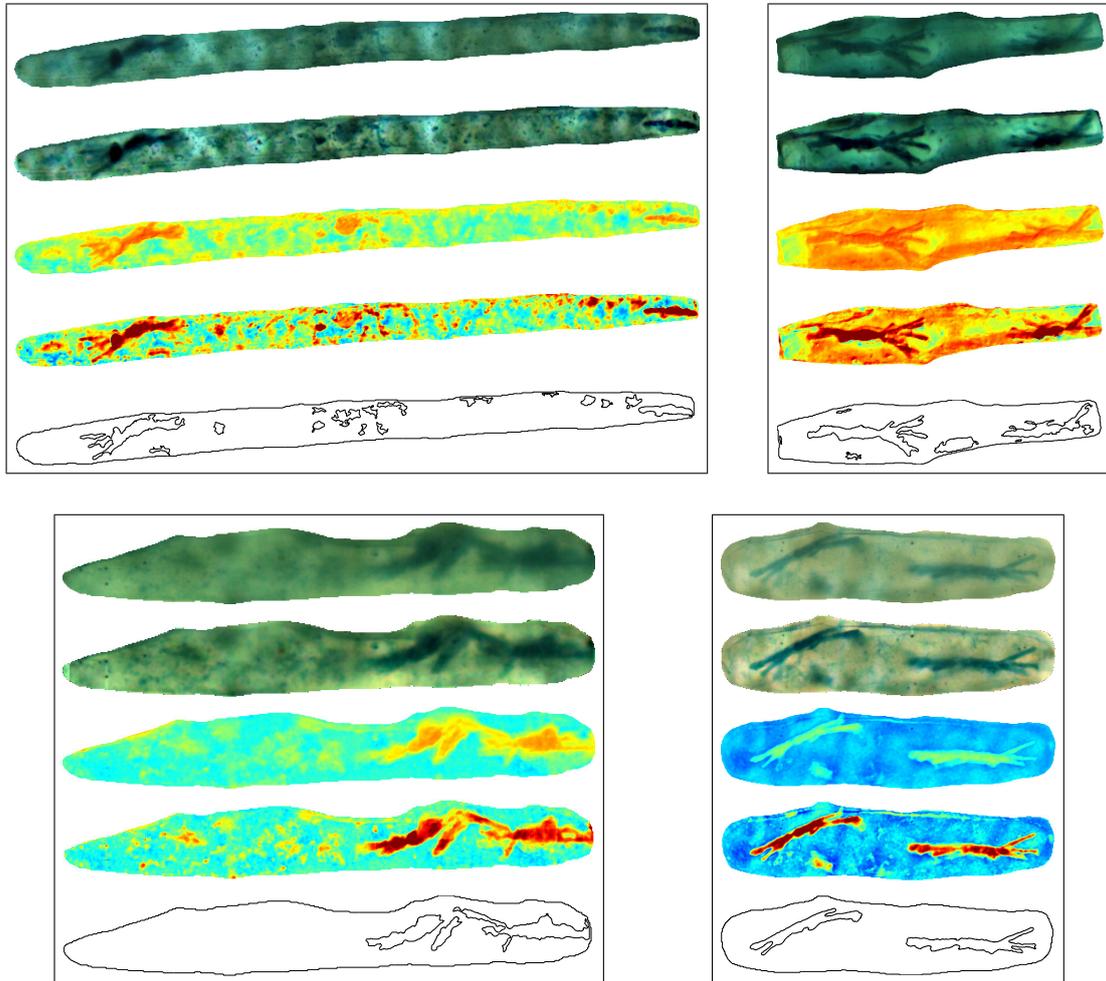


Abbildung 3.25: Vier Beispiele zur Haustoriensegmentierung mittels morphologischer Kontrastverstärkung und anschließender Hysterese-Schwellwerthaltung der Farbsättigung (jeweils von oben nach unten):

- Originalbild nach der Zellsegmentierung
- Ergebnis der morphologischen Kontrastverstärkung (vgl. Abb. 3.26)
- Farbsättigung des Originalbildes
- Farbsättigung des kontrastverstärkten Bildes
- Daraus durch Hysterese-Schwellwerthaltung (vgl. Abb. 3.27) detektierte Objektkonturen.

deutlich reduziert werden: Dazu wird, ausgehend von Bereichen, die eine maximale Sättigung aufweisen, ein Bereichswachstum (*region growing*) durchgeführt, bis ein bestimmter Sättigungsschwellwert unterschritten ist. Dies kann auch als eine Hysterese-Schwellwerthaltung interpretiert werden, bei der alle Regionen zwischen dem oberen und unteren Schwellwert selektiert werden, wenn sie mit einer Region oberhalb des oberen Schwellwertes *verbunden* sind.

Der Gewinn dieses Hystereseverfahrens gegenüber einer einfachen Schwellwerthaltung ist deutlich in Abbildung 3.27 zu erkennen: Viele Bereiche, die zwar den unteren Schwellwert überschreiten, aber den oberen Schwellwert nicht erreichen, werden dadurch ausgeschlossen. Das resultierende Binärbild enthält aus diesem Grund deutlich weniger Falschdetektionen als im Fall einer einfachen Schwellwerthaltung.

Bedingt durch die hohe Variabilität der Daten muss die Bestimmung des unteren Schwellwertes adaptiv anhand des konkret vorliegenden Bildmaterials erfolgen. Dazu wird die Sättigung einer Histogrammanalyse unterzogen. Der obere Schwellwert (Bereiche, von denen aus das Bereichswachstum beginnt) liegt bedingt durch den beobachteten Effekt maximaler Sättigung bei Eins. Eine Schätzung des unteren Schwellwertes ergibt sich als Mittelwert zwischen dem oberen Schwellwert (Eins) und einer Schätzung des Wertes für das Zellmaterial.

Wertung der Ergebnisse

Das hier vorgestellte Verfahren ermöglicht eine zuverlässige Segmentierung der gesuchten Haustorienregionen. Bedingt durch Verfärbungen, welche eine den Haustorien sehr ähnliche Farbcharakteristik aufweisen, werden jedoch gelegentlich auch Bereiche segmentiert, die *kein* Haustorium darstellen. Solche Regionen müssen in einem nachgeschalteten Verarbeitungsschritt anhand weiterer Kriterien, wie z.B. der Form, erkannt und zurückgewiesen werden.

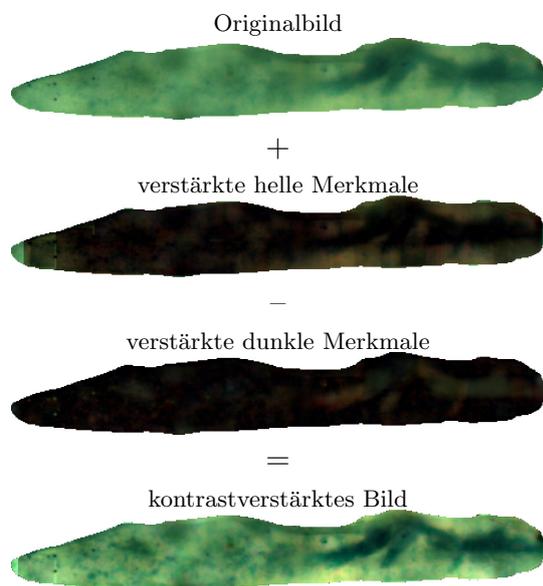


Abbildung 3.26: Morphologische Kontrastverstärkung: Zum Originalbild (Zeile 1) werden die verstärkten hellen Merkmale (Zeile 2) addiert und die verstärkten dunklen Merkmale (Zeile 3) subtrahiert. Dies ergibt das kontrastverstärkte Bild (Zeile 4).

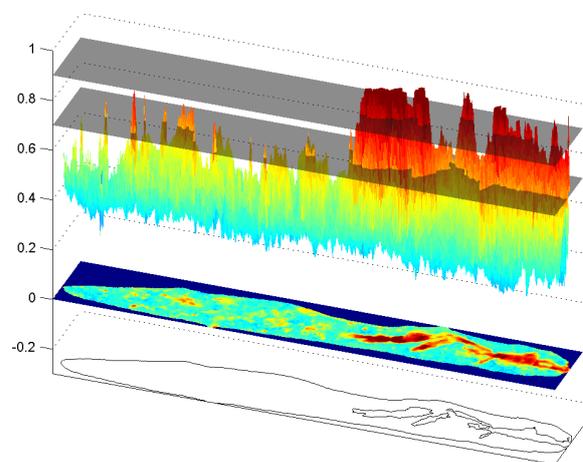


Abbildung 3.27: Hysterese-Schwellwerthaltung: Alle Regionen zwischen dem oberen und unteren Schwellwert werden selektiert, wenn sie mit einer Region oberhalb des oberen Schwellwertes verbunden sind.

3.5 Zusammenfassung der Ergebnisse

Ziel dieses Kapitels war es, die für die Analyse der Mikroskopbilder notwendigen Segmentierungsprobleme zu lösen. Dazu wurde zunächst auf die Segmentierung der transformierten Zellen fokussiert. Es zeigte sich, dass sich sowohl die Färbung des Hintergrundes als auch die der gesuchten Zellen sehr gut mittels einer multivariaten Gaußverteilung modellieren lässt. Mit Hilfe dieses parametrischen Modells war es möglich, das Bild in einen adaptiven Farbraum zu überführen, in dem die beiden Klassen „Hintergrund“ und „transformierte Zelle“ (unter den Randbedingungen des Modells) bestmöglich voneinander separiert werden. Die Schätzung der Modellparameter erfolgte mittels des Expectation Maximization (EM)-Algorithmus. Aus Aufwandsgründen wurde die Zellsegmentierung als zweistufiges Verfahren ausgeführt: Da sich das Merkmal der Zellfärbung sehr gut im Buntton (*Hue*) widerspiegelt, wurde zunächst nur auf ein generelles Auffinden transformierter Zellen (Finden der *Region of Interest, ROI*) auf Basis einer Farbraumtransformation in den HSV-Bereich orientiert. Unter Anwendung des Canny-Kantendetektors war es möglich, die transformierten Zellen mit einer geschlossenen Kontur vom Hintergrund abzugrenzen.

Die entstandenen rechteckigen Bildausschnitte wurden anschließend in einem zweiten Schritt in den adaptiven Farbraum überführt. Mittels eines aktiven Konturmodells erfolgte nun eine iterative Bestimmung der exakten Zellberandung. Bei der Implementierung des aktiven Konturmodells wurde ein neues Verfahren angewendet, welches die schnelle Fouriertransformation (*FFT*) zur Lösung nutzt und damit einen großen Geschwindigkeitsvorteil gegenüber dem klassischen Lösungsverfahren bietet.

Nach der Lösung des Problems der Zellsegmentierung erfolgte die Entwicklung von Segmentierungsverfahren zur Detektion der Haustorien innerhalb transformierter Zellen. Mit Hilfe eines modifizierten Clustering-Verfahrens auf Basis des EM-Algorithmus konnten auf „gutmütigen“ Daten recht zuverlässige Ergebnisse erzielt werden. Als robuster erwies sich ein Verfahren, welches auf einer morphologischen Kontrastverstärkung der Zellregion im Farbbild und anschließender Segmentierung im Sättigungsbereich mittels zwei Schwellwerten (Hysterese-Schwellwerthaltung) basiert. Im Ergebnis liegen damit zuverlässige und effiziente Verfahren sowohl für die Zell- als auch für die Haustoriensegmentierung vor.

Es stellte sich heraus, dass eine sichere Detektion von Haustorien auf alleiniger Basis des Merkmals „Farbe“ nicht gewährleistet werden kann, weil andere Objekte (z.B. der Zellkern oder Verfärbungen) eine sehr ähnliche Farbcharakteristik wie die Haustorien aufweisen. Im Anschluss an die Haustoriensegmentierung muss daher für jede markierte Region überprüft werden, ob es sich tatsächlich um ein Haustorium handelt. Diese Problematik wird nun im Anschluss an dieses Kapitel im zweiten Teil der Arbeit gelöst. Dabei werden die Thematiken Formanalyse, Formbeschreibung, Merkmalsräume und Klassifikation behandelt.

Teil II

Merkmalsextraktion und Klassifikation

Kapitel 4

Formanalyse und Formmerkmale

Das Ergebnis der Bildsegmentierung liegt zunächst wieder in Form eines Bildes vor. Im einfachsten Fall handelt es sich um ein Binärbild, welches zwischen „interessanten“ Regionen und Bildhintergrund unterscheidet. Mittels dieses Binärbildes lässt sich beispielsweise das Originalbild maskieren, um uninteressante Regionen auszublenden.

Im nächsten Schritt soll es nun um die Interpretation der segmentierten Bildinhalte gehen. Dazu soll insbesondere die *Form* der segmentierten Objekte beurteilt werden. Um die Form von Objekten – seien es in diesem konkreten Fall Haustorien – zu beschreiben, sind einerseits möglichst kompakte und aussagekräftige Parameter zu finden, welche die wichtigen Charakteristiken der Objekte widerspiegeln; andererseits sollen diese Parameter hinreichend gut generalisieren, um beispielsweise unabhängig von einfachen geometrischen Transformationen der Objekte zu werden.

Invarianz bezüglich einfacher geometrischer Transformationen

Die wichtigsten geometrischen Transformationen, denen Objekte unterworfen werden können, sind *Translation*, *Rotation* und *Skalierung*. Per Definition ändern diese drei Transformationen nichts an der Form eines Objekts.

Um einen Bildpunkt mit den Koordinaten (x, y) diesen drei genannten Transformationen zu unterziehen, sind die Parameter (t_x, t_y) (Verschiebung), $\phi \in [0 \dots 2\pi)$ (Drehung) und $\alpha \in \mathbb{R}_+$ (Skalierung) entsprechend der Transformationsvorschrift

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \alpha \cdot \begin{bmatrix} \cos \phi & -\sin \phi \\ \sin \phi & \cos \phi \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \end{bmatrix} \quad (4.1)$$

anzuwenden. Translations-, rotations- und skalierungsinvariante Formparameter sind gegenüber dieser Transformationsvorschrift unempfindlich und erlauben, das Objekt sowohl aus verschiedenen Entfernungen als auch aus verschiedenen Positionen zu betrachten.

Dem nachträglichen Hinzufügen der Verschiebungs-, Drehungs- und Skalierungsparameter des betrachteten Objekts zum Parametervektor (falls sie doch von Interesse sein sollten) steht in jedem Fall nichts entgegen und bietet gegenüber transformationsvarianten Parametern den Vorteil, dass die Transformationsparameter nicht in korrelierter (versteckter) Form im Parametervektor vorliegen, sondern direkt greifbar sind.

Klassifizierung von Methoden zur Formanalyse

PAVLIDIS [Pav78] hat die generelle Einteilung von Formanalysetechniken in die Rubriken *regionenbasiert* und *konturbasiert* vorgeschlagen. Während regionenbasierte Verfahren (auch als globale bzw. interne Verfahren bezeichnet) auf die gesamte Fläche des Objekts angewendet werden,

operieren konturbasierte Verfahren (auch als Grenzverfahren bzw. externe Verfahren bezeichnet) nur auf der äußeren Begrenzung des Objekts. Da die Objektkontur die volle Information über die Gestalt eines Objekts enthält, lassen sich mit beiden Ansätzen Verfahren finden, welche *informationserhaltend* arbeiten, d.h., bei denen sich die Objektform mit beliebiger Genauigkeit aus den extrahierten Parametern rekonstruieren lässt.

4.1 Einfache, nicht informationserhaltende Formmerkmale

Die nachfolgend vorgestellten einfachen Formmerkmale sind *nicht* informationserhaltend und verletzen teilweise auch die geforderten Transformationsinvarianzen.

Objektfläche

Für ein in einem Binärbild I mit M Spalten und N Zeilen vorliegendes Objekt lässt sich dessen Fläche (*Area*) leicht abschätzen als

$$A = \sum_{x=1}^M \sum_{y=1}^N I(x, y) \quad \text{mit} \quad I(x, y) \in \{0, 1\}. \quad (4.2)$$

Die so bestimmte Objektfläche ist zwar invariant gegen die geometrischen Operationen Rotation und Translation, jedoch nicht gegen Skalierung. Ihre Einheit bemisst sich anhand der Fläche, welche ein Bildpunkt in natura einnimmt.

Objektumfang

Eine einfache Approximation des Umfangs (*Perimeter*) eines als Binärbild gegebenen Objekts lässt sich aus der Summe seiner Konturpunkte $C(x, y)$ berechnen:

$$P = \sum_{x=1}^M \sum_{y=1}^N C(x, y) \quad \text{mit} \quad C(x, y) \in \{0, 1\}, \quad (4.3)$$

wobei die Kontur C verbunden als 4er-Nachbarschaft (siehe Abbildung 4.1) gegeben sei. Ebenso wie die Objektfläche ist der Umfang zwar invariant bezüglich Rotation und Translation, nicht jedoch gegen Skalierung.

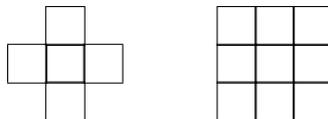


Abbildung 4.1: Nachbarschaft eines Pixels: 4er-Nachbarschaft (*links*); 8er-Nachbarschaft (*rechts*)

Rundheit

Ein sehr einfacher translations- rotations- und skalierungsinvarianter Parameter ist die Rundheit (*Roundness*). Dazu werden die Objektfläche und das Quadrat des Objektumfangs ins Verhältnis gesetzt und normiert, so dass dieses Maß für einen Kreis den Maximalwert Eins annimmt und für jedes komplexer geformte Objekt zunehmend kleiner als Eins wird.

$$R = \frac{4 \cdot \pi \cdot A}{P^2} \quad (4.4)$$

4.2 Regionenbasierte Formanalyse mittels Momentinvarianzen

Mittels der aus der mathematischen Statistik bekannten Momente lassen sich Objekte in Bildern *informationserhaltend* parametrieren. Die Nutzung von Momenten zur Objektklassifikation wurde erstmalig von HU [Hu61, Hu62] vorgeschlagen.

Für ein Grauwertbild $I(x, y)$ sind die Momente m_{pq} der Ordnung $(p+q)$ folgendermaßen definiert:

$$m_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^p y^q I(x, y) \, dx \, dy, \quad \text{mit } p, q = 0, 1, 2, \dots \quad (4.5)$$

Dabei sind die Koordinaten des Bildschwerpunktes gegeben durch

$$\bar{x} = \frac{m_{10}}{m_{00}} \quad \text{und} \quad \bar{y} = \frac{m_{01}}{m_{00}}. \quad (4.6)$$

Bei Grauwertbildern spiegeln diese Momente nicht nur die geometrische Form des Objekts, sondern auch die Verteilung der Grauwerte innerhalb des Objekts wider. Ist dies nicht gewünscht, so kann für die Momentenberechnung statt des Grauwertbildes das entsprechende Binärbild verwendet werden.

4.2.1 Zentralmomente

Durch Berücksichtigung des Bildschwerpunktes (\bar{x}, \bar{y}) ergeben sich die Zentralmomente μ_{pq} .

$$\mu_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \bar{x})^p (y - \bar{y})^q I(x, y) \, dx \, dy, \quad \text{mit } p, q = 0, 1, 2, \dots \quad (4.7)$$

Sie sind invariant bezüglich einer Translation.

4.2.2 Momenten-Tensor

Die Zentralmomente zweiter Ordnung sind aus der mathematischen Statistik in Form der Kovarianzmatrix hinlänglich bekannt. In der Mechanik tritt die Anordnung der Zentralmomente zweiter Ordnung als der so genannte *Trägheitstensor* in Erscheinung, welcher sich analog auch auf die Auswertung von Bildobjekten anwenden lässt [Jäh02].

$$\mathbf{J} = \begin{bmatrix} \mu_{20} & -\mu_{11} \\ -\mu_{11} & \mu_{02} \end{bmatrix} \quad (4.8)$$

Das Objekt wird dabei in erster Näherung als eine Ellipse beschrieben. Die **Hauptachsen** dieser Ellipse lassen sich mit Hilfe einer Eigenwertzerlegung des Trägheitstensors bestimmen und sind gegeben durch

$$2a = 2\sqrt{2} \sqrt{\mu_{20} + \mu_{02} + \sqrt{(\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2}} \quad (4.9)$$

$$2b = 2\sqrt{2} \sqrt{\mu_{20} + \mu_{02} - \sqrt{(\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2}}. \quad (4.10)$$

Ebenso lässt sich aus dem Momententensor die **Richtung der Hauptachse** des Objekts bestimmen. Sie entspricht der Orientierung des zum größten Eigenwert gehörigen Eigenvektors.

$$\phi = \frac{1}{2} \arctan \frac{2\mu_{11}}{\mu_{20} - \mu_{02}} \quad (4.11)$$

Ein weiterer Parameter ist die **Exzentrizität**. Dieses Maß ist begrenzt zwischen den Werten Null (bei einem runden Objekt) und Eins (bei einem linienförmigen Objekt).

$$\epsilon = \frac{(\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2}{(\mu_{20} + \mu_{02})^2} \quad (4.12)$$

4.2.3 Normalisierte Zentralmomente

Werden die Zentralmomente mittels

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^\gamma} \quad \text{mit} \quad \gamma = \frac{p+q}{2} + 1 \quad \text{für} \quad p+q = 2, 3, \dots \quad (4.13)$$

normiert, so entstehen die normalisierten Zentralmomente η_{pq} , die nicht nur translations-, sondern zusätzlich skalierungsinvariant sind.

4.2.4 Hu-Momente

Die zusätzliche Forderung der Rotationsinvarianz führt schließlich auf die folgenden sieben Momente, die nach ihrem Entdecker HU [Hu61, Hu62] benannt sind.

$$\phi_1 = \eta_{20} + \eta_{02} \quad (4.14)$$

$$\phi_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \quad (4.15)$$

$$\phi_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \quad (4.16)$$

$$\phi_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \quad (4.17)$$

$$\phi_5 = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12}) [(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] \\ + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03}) [3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \quad (4.18)$$

$$\phi_6 = (\eta_{20} - \eta_{02}) [(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \\ + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \quad (4.19)$$

$$\phi_7 = (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12}) [(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] \\ + (3\eta_{12} - \eta_{30})(\eta_{21} + \eta_{03}) [3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \quad (4.20)$$

Erweiterung auf affine Momentinvarianzen

REISS zeigt in [Rei91] eine Erweiterung der Momenteninvarianz auf generelle lineare Transformationen des Eingangsbildes. Vor dem Hintergrund einer Parametrierung von Haustorien ist eine solche Invarianz jedoch nicht notwendig und wird daher nicht weiter betrachtet.

4.2.5 Momentinvarianzen nach Flusser

FLUSSER widmet sich in [Flu00] der Frage nach möglichen Abhängigkeiten von Momentinvarianzen untereinander und zeigt, dass unter Verwendung von Momenten bis zur dritten Ordnung eine Basis von lediglich sechs *unabhängigen* Momenteninvarianzen existiert. Als Konsequenz zeigt sich, dass das Hu-Moment ϕ_3 durch

$$\phi_3 = \frac{\phi_5^2 + \phi_7^2}{\phi_4^3} \quad (4.21)$$

ausgedrückt werden kann – also redundant ist.

FLUSSER leitet sein System von Momentinvarianzen auf Basis der komplexen Momente

$$c_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x + jy)^p (x - jy)^q I(x, y) dx dy, \quad \text{mit} \quad p, q = 0, 1, 2, \dots \quad (4.22)$$

her. Für den Übergang zu komplexen *Zentralmomenten* bzw. *normierten Zentralmomenten* gelten die gleichen Überlegungen analog zu Gleichung (4.7) und (4.13).

Unter Verwendung der Momente zweiter und dritter Ordnung schlägt FLUSSER folgenden Satz von Momentinvarianzen vor, wobei die Hu-Momente $\phi_1, \phi_4, \phi_5, \phi_6$ und ϕ_7 mit enthalten sind.

$$\psi_1 = c_{11} = \phi_1 \quad (4.23)$$

$$\psi_2 = c_{21} c_{12} = \phi_4 \quad (4.24)$$

$$\psi_3 = \operatorname{Re}\{c_{20} c_{12}^2\} = \phi_6 \quad (4.25)$$

$$\psi_4 = \operatorname{Im}\{c_{20} c_{12}^2\} \quad (4.26)$$

$$\psi_5 = \operatorname{Re}\{c_{30} c_{12}^3\} = \phi_5 \quad (4.27)$$

$$\psi_6 = \operatorname{Im}\{c_{30} c_{12}^3\} = \phi_7 \quad (4.28)$$

Werden die Momente vierter Ordnung hinzugenommen, erweitert sich der Satz.

$$\psi_7 = c_{22} \quad (4.29)$$

$$\psi_8 = \operatorname{Re}\{c_{31} c_{12}^2\} \quad (4.30)$$

$$\psi_9 = \operatorname{Im}\{c_{31} c_{12}^2\} \quad (4.31)$$

$$\psi_{10} = \operatorname{Re}\{c_{40} c_{12}^4\} \quad (4.32)$$

$$\psi_{11} = \operatorname{Im}\{c_{40} c_{12}^4\} \quad (4.33)$$

Bei den Momentinvarianzen ψ_4, ψ_6, ψ_9 und ψ_{11} handelt es sich nicht um wirkliche Invarianzen (*true invariants*), sondern um schiefe Invarianzen oder Pseudoinvarianzen (*skew invariants, pseudoinvariants*), die bei einer Spiegelung des Objekts das Vorzeichen wechseln. Diese Information kann bei einigen Anwendungen von Nutzen sein, im Fall der Haustorienerkennung ist dies jedoch nicht erwünscht. Aus diesem Grund sollten hier die *Beträge* dieser schiefen Momentinvarianzen als Merkmale verwendet werden.

Die Eigenschaften der Flusser-Momentinvarianzen werden in Kapitel 5 noch eingehend untersucht.

4.3 Konturbasierte Formanalyse mittels Fourierdeskriptoren

Wie bereits eingangs bemerkt, spiegelt die Kontur die gesamte Forminformation eines Objektes wider. Damit genügt es, die Berandung eines Objektes zu analysieren, um Aussagen über seine Form treffen zu können.

Ein sehr leistungsfähiges Werkzeug zur Analyse von Konturfunktionen führt über den Spektralbereich – die so genannten Fourierdeskriptoren. Gegeben sei eine *geschlossene* Konturfunktion in Kartesischen Koordinaten x und y in Form einer Parameterdarstellung bezüglich eines reellwertigen Parameters $t \in [0, T)$. Diese Konturfunktionen sind (aufgrund der Randbedingung *geschlossener* Konturen) periodisch in T . Ohne Beschränkung der Allgemeinheit soll die Kontur entgegen dem Uhrzeigersinn (mathematisch positiv) durchfahren werden.

Aufgrund ihrer Periodizität lassen sich die Konturfunktionen in eine Fourierreihe entwickeln. Das abgetastete Spektrum lässt sich mittels der Fourierkoeffizienten darstellen, welche im Falle von Konturfunktionen als Fourierdeskriptoren bezeichnet werden.

Eine Fourierreihenentwicklung der reellen Konturfunktion lässt sich schreiben als

$$\begin{bmatrix} x(t) \\ y(t) \end{bmatrix} = \begin{bmatrix} \hat{A}_0 \\ \hat{C}_0 \end{bmatrix} + \sum_{\mu=1}^{\infty} \begin{bmatrix} \hat{A}_\mu & \hat{B}_\mu \\ \hat{C}_\mu & \hat{D}_\mu \end{bmatrix} \begin{bmatrix} \cos(2\pi\mu t/T) \\ \sin(2\pi\mu t/T) \end{bmatrix}. \quad (4.34)$$

Die reellen Koeffizienten \hat{A}_0 und \hat{A}_μ und \hat{B}_μ sind dabei der Mittelpunkt und die Amplituden der Harmonischen der Funktion der x -Koordinate. Analoges gilt für die Koeffizienten \hat{C}_0 und \hat{C}_μ und \hat{D}_μ für die Funktion der y -Koordinate.

Eine wesentlich einfachere Darstellung lässt sich mittels der komplexen Fourierreihe erreichen. Dazu werden die x - und y -Koordinaten zur komplexwertigen Funktion

$$z(t) = x(t) + jy(t), \quad t \in [0, T) \quad (4.35)$$

zusammengefasst. Die Darstellung der Kontur als komplexe Fourierreihe ist dann gegeben durch

$$z(t) = \sum_{\mu=-\infty}^{\infty} \hat{Z}_\mu e^{j2\pi\mu t/T} \quad \text{mit} \quad \hat{Z}_\mu = \frac{1}{T} \int_0^T z(t) e^{-j2\pi\mu t/T} dt. \quad (4.36)$$

Die Darstellungen nach Gleichung (4.34) und (4.36) sind gleichwertig, jedoch sollte aus Sicht des Autors der komplexen Darstellung aufgrund der wesentlich einfacheren Handhabung der Vorzug gegeben werden. In Anhang E findet sich eine kurze Herleitung, wie beide Darstellungen ineinander umgerechnet werden können.

Abgetastete Konturen lassen sich ebenfalls in die Spektraldarstellung überführen. Dazu wird die diskrete Fouriertransformation (DFT) verwendet, die in der Praxis durch ihr hochgradig effizientes Pendant, die schnelle Fouriertransformation (*Fast Fourier Transform, FFT*), implementiert wird. Der Abtastvorgang der Konturfunktion sei durch den Abtastoperator $\mathcal{S}_{t_0}\{\cdot\}$ beschrieben, welcher im Intervall t_0 Proben entnimmt. Nach dem Abtastvorgang sei das entstehende diskrete Signal durch N Samples

$$z(n) = \mathcal{S}_{t_0}\{z(t)\}, \quad n \in [0 \dots N-1] \quad (4.37)$$

beschrieben. Die Fourierkoeffizienten treten dann aufgrund des Abtastvorgangs im Ortsbereich periodisch im Intervall $\nu \in [0 \dots N-1]$ bzw. (wegen der Periodizität) $\nu \in [[-N/2] \dots [N/2]-1]$ auf¹. Die komplette Beschreibung der Kontur findet sich also sowohl im Orts- als auch im Spektralbereich in N komplexwertigen Koeffizienten.

$$z(n) = \sum_{\nu=0}^{N-1} \hat{Z}_\nu e^{j2\pi\nu n/N} \quad \hat{Z}_\nu = \frac{1}{N} \sum_{n=0}^{N-1} z(n) e^{-j2\pi\nu n/N} \quad (4.38)$$

Dieser spezielle Anwendungsfall der harmonischen Analyse zur Konturrepräsentation ist unter dem Begriff „Fourierdeskriptoren“ (*Fourier Descriptors*) bekannt und wurde insbesondere in den 70er Jahren vor dem Hintergrund der Erkennung handgeschriebener Buchstaben [Gra72] sowie zur Formanalyse biologischer Objekte [YWB74] untersucht.

Eigenschaften der Spektraldarstellung

Die Spektraldarstellung ist für Konturfunktionen aus den nachfolgend genannten Gründen besonders vorteilhaft bzw. interessant:

- Die Spektraldarstellung bietet eine sehr kompakte Approximation der Konturfunktion in Form von wenigen Koeffizienten, da die wesentliche Forminformation des Objekts in den ersten Koeffizienten konzentriert ist.
- Mittels des Differentiationstheorems der Fouriertransformation lassen sich die Differentiale (Ableitungen) der abgetasteten Konturfunktion analytisch bestimmen.

¹Der Operator $[\cdot]$ rundet das reellwertige Argument auf den nächstgrößeren ganzzahligen Wert auf.

- Aufgrund der inhärenten *periodischen* Natur der Konturfunktion ist eine Behandlung mittels der diskreten Fouriertransformation möglich, ohne die sonst bei aperiodischen Signalen auftretenden Probleme berücksichtigen zu müssen (Fensterung).

Allerdings ist bei Konturen aus Digitalbildern Folgendes zu beachten: Die Konturfunktion wird in der Praxis anhand einer 8er-Nachbarschaft (vgl. Abbildung 4.1) auf einem diskreten Gitter repräsentiert. Da dabei ein Schritt in Diagonalrichtung $\sqrt{2}$ -fach so lang ist wie ein Schritt in x - bzw. y -Richtung, haben die Abtastpunkte nicht den gleichen Abstand – es handelt sich also bei der diskreten Konturfunktion um ein *nicht äquidistant* abgetastetes Signal. Dieses Problem wird unter anderem in [WW80] ausführlich behandelt. Bis heute ist die *korrekte* Bestimmung der Fourierdeskriptoren aus Bildkonturen ein herausforderndes Forschungsthema [Jäh02].

An dieser Stelle sei angemerkt, dass auch eindimensionale Objekte (Linien) mittels Fourierdeskriptoren dargestellt werden können – indem die Linie einmal hin und zurück abgefahren wird.

4.3.1 Signal-Rausch-Verhältnis (SNR) von Konturen

Eine weitere wichtige Eigenschaft von in der praktischen Bildverarbeitung auftretenden Konturen muss unbedingt beachtet werden: Die Konturen sind nicht nur abgetastet, sondern auch *quantisiert* – die x - und y -Koordinaten der Konturpunkte nehmen notwendigerweise immer Werte im Pixelraster des zugrunde liegenden Bildes an. Wie bei jeder Quantisierung von Signalen entsteht dabei ein Fehler; das quantisierte Signal stellt nur eine Näherung des Originalsignals dar. Dieser Quantisierungsfehler soll im Folgenden genauer untersucht werden.

Bekanntermaßen [PM96] lässt sich der Quantisierungsfehler eines linearen Quantisierers in Form eines Quantisierungsrauschens modellieren. Dabei wird das Originalsignal von einem Rauschsignal (weißes Rauschen) mit der Leistung

$$P_q = \Delta^2/12 \quad (4.39)$$

überlagert. Der Parameter Δ ist dabei das Quantisierungsintervall. Aus dem Verhältnis der Wechselleistung \tilde{P}_s des Originalsignals² und der Quantisierungsgeräuschleistung P_q lässt sich das Signal-Rausch-Verhältnis (*Signal-to-noise ratio*, *SNR*) definieren.

$$\varrho = 10 \log_{10} \left(\frac{\tilde{P}_s}{P_q} \right) \quad (4.40)$$

Meist ist die (Wechsel)leistung des Originalsignals nicht bekannt, da in der Regel nur das bereits quantisierte Signal vorliegt. Unter der Annahme, dass das Originalsignal und die Quantisierungsrauschen unkorreliert sind, kann \tilde{P}_s durch die Subtraktion der Quantisierungsgeräuschleistung P_q aus der Leistung des quantisierten Signals geschätzt werden.

Das Signal-Rausch-Verhältnis einer Kontur ist dann gegeben durch

$$\varrho = 10 \log_{10} \left(\frac{1}{2} \left[\frac{\tilde{P}_x - P_q}{P_q} + \frac{\tilde{P}_y - P_q}{P_q} \right] \right) = 10 \log_{10} \left(\frac{\tilde{P}_z - 2P_q}{2P_q} \right), \quad (4.41)$$

wobei $\tilde{P}_x + \tilde{P}_y = \tilde{P}_z$ die Wechselleistung der quantisierten Kontur ist. Mit Hilfe des Theorems von Parseval lässt sich dieser Wert leicht aus der Spektraldarstellung der Kontur berechnen.

$$\tilde{P}_z = \sum_{\nu=1}^{N-1} |\hat{Z}_\nu|^2 \quad (4.42)$$

²Die Gleichleistung darf nicht berücksichtigt werden, da konstante Signale einen konstanten Quantisierungsfehler und kein Quantisierungsrauschen aufweisen.

In Form eines *weißen* Rauschens verteilt sich die Leistung des Quantisierungsfehlersignals gleichmäßig über das gesamte Spektrum. Jede Spektralkomponente (jeder Fourierdeskriptor \hat{Z}_ν , $\nu \neq 0$) der Kontur wird daher mit $\approx 1/N$ der gesamten Rauschleistung beeinflusst. Dieses Hintergrundsignal findet sich im Spektrum bei einem Pegel von etwa $\varrho + 10 \log_{10}(N)$ dB unterhalb des Signalpegels $10 \log_{10}(\hat{P}_s)$ wieder.

4.3.2 Invarianzeigenschaften der Spektraldarstellung

Operation / Eigenschaft	Kontur (Ortsbereich)	Fourierdeskriptoren
Konturdarstellung	z_n	\hat{Z}_ν
Schwerpunkt der Kontur	$\frac{1}{N} \sum_{n=0}^{N-1} z_n$	\hat{Z}_0
Translation um Δz , $\Delta z \in \mathbb{C}$	$z_n + \Delta z$	$\hat{Z}_0 + \Delta z$
Skalierung um α , $\alpha \in \mathbb{R}_+$	$z_n \cdot \alpha$	$\hat{Z}_\nu \cdot \alpha$
Rotation in der komplexen Ebene um den Winkel ϕ	$z_n \cdot e^{j\phi}$	$\hat{Z}_\nu \cdot e^{j\phi}$
Rotation um den Schwerpunkt \hat{Z}_0 um den Winkel ϕ	$(z_n - \hat{Z}_0) \cdot e^{j\phi} + \hat{Z}_0$	$\hat{Z}_\nu \cdot e^{j\phi} \quad \forall \nu \neq 0$
Verschiebung der Parametrierung um k	z_{n-k}	$\hat{Z}_\nu \cdot e^{-j2\pi\nu k/N}$
Eigenschaften bei m -facher Rotationssymmetrie	$z_n = (z_n - \hat{Z}_0) \cdot e^{j\frac{2\pi}{m}} + \hat{Z}_0$	$\hat{Z}_{1 \pm \nu m} \neq 0$

Tabelle 4.1: Invarianzeigenschaften von Fourierdeskriptoren.

Die Fourierdeskriptoren beinhalten (ähnlich wie eine momentenbasierte Formbeschreibung, vgl. Abschnitt 4.2) die vollständige Information zur Rekonstruktion des Objekts. Unter Nutzung der bekannten Eigenschaften und Grundgesetze der Fouriertransformation [KI89] lassen sich leicht die nötigen Schritte ersehen, um zu einer translations-, rotations- und skalierungsinvarianten Objektbeschreibung im Spektralbereich zu gelangen [Jäh02]:

1. Translationsinvarianz

Alle Spektralkoeffizienten, außer \hat{Z}_0 , welcher den Objektmittelpunkt beschreibt, sind translationsinvariant.

2. Skalierungsinvarianz

Der Spektralkoeffizient \hat{Z}_1 ist für den „Hauptkreis“

$$\hat{Z}_1 \cdot e^{j2\pi t/T} \quad t \in [0, T) \quad \text{bzw.} \quad \hat{Z}_1 \cdot e^{j2\pi n/N} \quad n \in [0, 1, \dots, N-1]$$

verantwortlich. Bei Konturen, die mathematisch positiv (entgegen dem Uhrzeigersinn) abgefahren werden und die eine Fläche ungleich Null aufweisen, ist dieser Koeffizient immer ungleich Null. So führt die Normierungsvorschrift

$$\frac{\hat{Z}_\nu}{|\hat{Z}_1|} \tag{4.43}$$

auf eine skalierungsinvariante Objektbeschreibung.

3. Rotationsinvarianz bezüglich Objektdrehung

Die Ausrichtung des Objekts wird durch die Lage der Hauptellipse bestimmt, welche durch die Koeffizienten \hat{Z}_{-1} und \hat{Z}_1 gegeben ist. So dreht die Operation

$$\hat{Z}_\nu \cdot e^{-j\frac{\hat{Z}_{-1} + \hat{Z}_1}{2}} \tag{4.44}$$

das Objekt in „Normallage“. Die lange Hauptachse ist dann parallel zur x-Achse ausgerichtet.

4. Rotationsinvarianz bezüglich einer Verschiebung des Startpunktes

Die Verschiebung des Startpunktes der Kontur führt im Spektrum zu einer Phasenverschiebung (Verschiebungssatz der Fouriertransformation) gemäß

$$z(t - \tau) \quad \circ \bullet \quad \hat{Z}_\mu \cdot e^{-j2\pi\mu\tau/T} \quad \text{bzw.} \quad z_{n-k} \quad \circ \bullet \quad \hat{Z}_\nu \cdot e^{-j2\pi\nu k/N}. \quad (4.45)$$

Durch eine Phasennormierung gemäß

$$\hat{Z}_\nu \cdot e^{-j\nu\varphi_1} \quad \text{mit} \quad \varphi_1 = \arg \left\{ \hat{Z}_1 \right\} \quad (4.46)$$

werden nach vorheriger Anwendung von Gleichung (4.44) die Koeffizienten \hat{Z}_1 und \hat{Z}_{-1} reellwertig. Unter Berücksichtigung von Gleichung (4.43) gilt letztlich $\hat{Z}_1 = 1$ und $\hat{Z}_{-1} \leq 1$.

Bei Anwendung der Normierungen nach Gleichung (4.43), (4.44) und (4.46) bleibt die Forminformation in den Fourierdeskriptoren vollständig erhalten.

Eine direkte Nutzung dieser (normierten) Spektralkoeffizienten zur Objektklassifikation erweist sich jedoch als recht schwierig: Prinzipiell könnten die Spektralkoeffizienten zwar direkt in Form von Real- und Imaginärteil als Merkmalsvektor benutzt werden, jedoch würden anschließende Merkmalstransformationen die *Phaseninformation* (durch unterschiedliche Skalierungen der Komponenten) verfälschen und die Objektbeschreibung damit signifikant stören. Eine andere Möglichkeit ist die Darstellung in Form von Betrag und Phase. Dabei führt jedoch die periodische Natur der Phaseninformation zu Problemen.

Wünschenswert sind *reellwertige* Merkmale, welche signifikante Formeigenschaften des Objekts widerspiegeln. Der einfachste Weg zu solchen Merkmalen führt über die Betragsquadratbildung des Spektrums. Die entstehenden Merkmale beschreiben dann die lineare Selbstähnlichkeit des Objekts in Form des Spektrums seiner Autokorrelationsfunktion.

4.3.3 Bedeutung der Betragsinformation des Spektrums

Die Autokorrelationsfunktion (AKF) [KI89] eines komplexen periodischen Signals $z(t)$ (mit der Primitivperiode T) bezüglich eines Verschiebungsparameters τ ist definiert als

$$\phi(\tau) = \frac{1}{T} \int_0^T z(t) z^*(t + \tau) dt. \quad (4.47)$$

Die zugehörige Spektralfunktion ist das Leistungsdichtespektrum $\Phi(f) \bullet \circ \phi(\tau)$. Das Leistungsdichtespektrum periodischer Signale berechnet sich durch Betragsquadratbildung der Koeffizienten des diskreten Amplitudenspektrums.

$$\Phi(f) = \sum_{\mu=-\infty}^{\infty} |\hat{Z}_\mu|^2 \delta(f - \mu f_0) \quad \text{mit} \quad f_0 = \frac{1}{T} \quad (4.48)$$

$$\Phi_\mu = |\hat{Z}_\mu|^2 \quad (4.49)$$

Die AKF hat die gleiche Periode T wie das zugehörige periodische Signal, außerdem gilt

$$\text{Re}\{\phi(\tau)\} = \text{Re}\{\phi(-\tau)\} \quad (4.50)$$

$$-\text{Im}\{\phi(\tau)\} = \text{Im}\{\phi(-\tau)\}. \quad (4.51)$$

Durch eine Betragsbildung der Fourierdeskriptoren geht also die tatsächliche Forminformation des Objekts verloren – lediglich die Information über die lineare Selbstähnlichkeit (Autokorrelation) bleibt erhalten. Der zugehörige Merkmalsvektor ist

$$\left[\frac{|\hat{Z}_{-n}|^2}{|\hat{Z}_1|^2} \quad \frac{|\hat{Z}_{-(n-1)}|^2}{|\hat{Z}_1|^2} \quad \dots \quad \frac{|\hat{Z}_{-1}|^2}{|\hat{Z}_1|^2} \quad \frac{|\hat{Z}_2|^2}{|\hat{Z}_1|^2} \quad \dots \quad \frac{|\hat{Z}_n|^2}{|\hat{Z}_1|^2} \right] \quad \text{mit } n \leq \lceil N/2 \rceil - 1. \quad (4.52)$$

Diese Merkmale werden unter anderem in [TSV⁺98] zur Objektklassifikation verwendet. Ein anderer, fortschrittlicherer Ansatz zur Gewinnung von Merkmalen, die die Selbstähnlichkeit von Objekten beschreiben, wurde von GRANLUND [Gra72] vorgeschlagen.

4.3.4 Spektralinvarianzen nach Granlund

Betrachten wir die Fourierdeskriptoren \hat{Z}_μ eines beliebigen Objekts, welche durch Skalierung um den Faktor α , Rotation um den Winkel ϕ und Verschiebung des Startpunktes um τ aus den Fourierdeskriptoren $\hat{Z}_\mu^{(o)}$ hervorgehen [Gra72].

$$\hat{Z}_\mu = \hat{Z}_\mu^{(o)} \cdot \alpha e^{j\phi} e^{j2\pi\mu\tau/T} \quad (4.53)$$

Offensichtlich ist der Ausdruck

$$b_n = \frac{\hat{Z}_{1+n} \hat{Z}_{1-n}}{\left(\hat{Z}_1\right)^2} \quad n = 2, 3, \dots \quad (4.54)$$

$$= \frac{\hat{Z}_{1+n}^{(o)} \alpha e^{j\phi} e^{j2\pi(1+n)\tau/T} \hat{Z}_{1-n}^{(o)} \alpha e^{j\phi} e^{j2\pi(1-n)\tau/T}}{\left(\hat{Z}_1^{(o)} \alpha e^{j\phi} e^{j2\pi\tau/T}\right)^2} \quad (4.55)$$

$$= \frac{\hat{Z}_{1+n}^{(o)} \hat{Z}_{1-n}^{(o)}}{\left(\hat{Z}_1^{(o)}\right)^2} \quad (4.56)$$

unabhängig von α , ϕ und τ . Des Weiteren wird der Deskriptor \hat{Z}_0 nicht verwendet – die entstehenden Merkmale sind also nicht nur skalierungs- und rotationsinvariant, sondern auch translationsinvariant. Ähnlich wie die im vorigen Unterkapitel betrachteten Momentinvarianzen nach HU bzw. FLUSSER (vgl. Kap. 4.2.4 und 4.2.5) beinhalten diese Spektralinvarianzen nicht mehr die *vollständige* Information über die Form, so dass die Kontur daraus nicht wieder rekonstruiert werden kann. So handelt es sich bei den Invarianzen nach GRANLUND um Merkmale, welche von der *Rotationssymmetrie* des Objekts abhängen. In Tabelle 4.2 sind die Zusammenhänge zwischen m -facher Rotationssymmetrie und den Beträgen der Invarianzen b_n qualitativ zusammengefasst.

4.3.5 Oberwellenhaltigkeit, „Klirrfaktor“

Der aus der harmonischen Analyse bekannte Klirrfaktor (*Total Harmonic Distortion, THD*) kann bei Anwendung auf die Fourierdeskriptoren ebenfalls als Formmerkmal herangezogen werden. Im Kontext der Konturanalyse beschreibt er die Abweichung des Objekts von einer Ellipse und nimmt Werte zwischen Null und Eins an.

$$k = \sqrt{\frac{\sum_{\nu=2}^{N-1} |\hat{Z}_\nu|^2}{\sum_{\nu=1}^{N-1} |\hat{Z}_\nu|^2}} \quad (4.57)$$

	Rotationssymmetrie m des Objekts					
	$m = 1$	$m = 2$	$m = 3$	$m = 4$	$m = 5$	$m = 6$
	(360°)	(180°)	(120°)	(90°)	(72°)	(60°)
b_2	+	+	○	○	○	○
b_3	+	○	+	○	○	○
b_4	+	+	○	+	○	○
b_5	+	○	○	○	+	○
b_6	+	+	+	○	○	+
b_7	+	○	○	○	○	○
b_8	+	+	○	+	○	○
b_9	+	○	+	○	○	○
b_{10}	+	+	○	○	+	○
b_{11}	+	○	○	○	○	○
b_{12}	+	+	+	+	○	+

Tabelle 4.2: Signifikante (+) und insignifikante (○) Beträge der Granlund-Invarianzen b_n nach Gleichung (4.54) bei m -facher Rotationssymmetrie des Objekts.

4.3.6 Merkmale aus den Differentialen der Kontur

Die Funktion $z(t)$ kann in Analogie zur Physik als die Bewegung eines Teilchens auf einer geschlossenen Bahn interpretiert werden. Seine Geschwindigkeit $v(t)$ und Beschleunigung $a(t)$ berechnet sich durch

$$v(t) = \dot{z}(t) = \frac{dz(t)}{dt} \quad \text{und} \quad a(t) = \ddot{z}(t) = \frac{d^2z(t)}{dt^2}. \quad (4.58)$$

Aus diesen Funktionen lassen sich wichtige Eigenschaften der Kontur, z.B. der Tangentialwinkel und die Krümmung (*Curvature*), berechnen.

Mittels des Differentiationstheorems der Fouriertransformation ist die Bestimmung der Ableitungen von $z(t)$ auf sehr elegante Art möglich [FCCJ01].

$$z(t) \circ \bullet \hat{Z}_\mu \quad \mathcal{S}\{z(t)\} = z(n) \circ \bullet \hat{Z}_\nu \quad (4.59)$$

$$\dot{z}(t) \circ \bullet j 2\pi\mu/T \hat{Z}_\mu \quad \mathcal{S}\{\dot{z}(t)\} = \dot{z}(n) \circ \bullet j 2\pi\nu/N \hat{Z}_\nu \quad (4.60)$$

$$\ddot{z}(t) \circ \bullet -(2\pi\mu/T)^2 \hat{Z}_\mu \quad \mathcal{S}\{\ddot{z}(t)\} = \ddot{z}(n) \circ \bullet -(2\pi\nu/N)^2 \hat{Z}_\nu \quad (4.61)$$

Tangentialwinkel

Der Tangentialwinkel der Kontur lässt sich mit Hilfe des physikalischen Äquivalents als die aktuelle Bewegungsrichtung des Partikels auf seiner geschlossenen Bahn interpretieren.

$$\varphi(t) = \arctan\left(\frac{\dot{y}(t)}{\dot{x}(t)}\right) = \arg\{\dot{z}(t)\} \quad (4.62)$$

Krümmung

Die Krümmung (*Curvature*) κ ist definiert als die Änderung des Tangentialwinkels φ bezogen auf die Bogenlänge s .

$$\kappa(t) = \frac{d\varphi}{ds} = \frac{\frac{d\varphi(t)}{dt}}{\frac{ds(t)}{dt}} = \frac{\dot{\varphi}(t)}{\dot{s}(t)} \quad (4.63)$$

Mittels der Definitionsgleichungen für den Tangentialwinkel und die Bogenlänge lässt sich die Gleichung leicht herleiten.

$$\varphi(t) = \arctan\left(\frac{\dot{y}(t)}{\dot{x}(t)}\right), \quad \dot{\varphi}(t) = \frac{\dot{x}(t)\ddot{y}(t) - \dot{y}(t)\ddot{x}(t)}{\dot{x}^2(t) + \dot{y}^2(t)} \quad (4.64)$$

$$s(t) = \int \sqrt{\dot{x}^2(t) + \dot{y}^2(t)} dt, \quad \dot{s}(t) = \sqrt{\dot{x}^2(t) + \dot{y}^2(t)} \quad (4.65)$$

$$\kappa(t) = \frac{\dot{x}(t)\ddot{y}(t) - \dot{y}(t)\ddot{x}(t)}{[\dot{x}(t) + \dot{y}(t)]^{3/2}} \quad (4.66)$$

Die Krümmungsfunktion spiegelt wichtige Merkmale einer Kontur wider:

- Stark nach außen gebogene (konvexe) Bereiche verursachen eine stark positive Krümmung.
- Im Gegenzug verursachen nach innen gebogene (konkave) Regionen eine *negative* Krümmung.
- Bereiche geradlinigen Verlaufs weisen eine Krümmung nahe Null auf.

Anschaulich ist die Krümmung $\kappa(t)$ der Kehrwert des Radius' des Kreises, der die Kurve an der betrachteten Stelle am engsten berührt (*Osculating Circle*).

Verbiegungsenergie

Von YOUNG *et al.* [YWB74] wurde das Maß der Verbiegungsenergie (*Bending Energy*) zur Formanalyse vorgeschlagen.

$$E = \frac{1}{T} \int_0^T |\kappa(t)|^2 dt \quad (4.67)$$

Es handelt sich dabei um ein *globales, nicht informationserhaltendes* Maß, welches für relativ einfach geformte, eher runde Objekte einen kleinen Wert annimmt, der mit zunehmender Komplexität (in Form von Ein- und Ausbuchtungen) steigt. Die Autoren beweisen in [YWB74], dass ein Kreis das Objekt mit minimaler Verbiegungsenergie ist.

Vor dem Hintergrund der Haustoriendetektion erscheint ein solches Maß als besonders geeignet, weil sich darin der Verzweigungsgrad (welcher bei Haustorien durch die Fingerstrukturen verursacht wird) widerspiegelt.

4.3.7 Kontur-Skalenraum

Ebenso wie das Bild selbst können auch Konturen bei verschiedenen Auflösungsstufen betrachtet werden. Mit Hilfe der Skalenraumtheorie (vgl. Abschnitt 3.2.1) lässt sich die Konturfunktion bei verschiedenen Detailgraden angeben, wobei mit zunehmender Glättung immer mehr Details verschwinden und das Objekt im Grenzfall zu einer Ellipse „entartet“. Die Überführung der Kontur in den Kontur-Skalenraum erfolgt durch die Glättung mit einer Gaußfunktion entsprechender Breite σ .

$$z_\sigma(t) = z(t) \circledast g_\sigma(t) = z(t) \circledast \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{t^2}{2\sigma^2}} \quad \circ\bullet \quad \hat{Z}_\mu \cdot e^{-2\frac{(\pi\sigma\mu)^2}{T^2}} \quad (4.68)$$

Insbesondere bei Operationen mit Kontur*differentialen* (z.B. bei der Berechnung der Krümmung) ist eine Glättung der Kontur gemäß dieser Überlegungen unabdingbar. Durch die Quantisierung im Pixelraster ist die Konturfunktion von einem breitbandigen Rauschsignal überlagert (vgl. Abschnitt 4.3.1), welches mit zunehmender Frequenz (bei feinen Strukturen) dominiert. Die Tiefpassfilterung dämpft diesen Effekt des Quantisierungsrauschens erheblich, so dass ein zuverlässig auswertbares Signal entsteht.

Normalized Multiscale Bending Energy (NMBE)

Unter Berücksichtigung der Skalenraumtheorie wurde von Cesar und Costa [CJFC97] das Konzept der normierten Multiskalen-Verbiegungsenergie (*Normalized Multiscale Bending Energy, NMBE*) Ψ vorgeschlagen. Da die Krümmung nicht skalierungsinvariant ist (sie hat die Dimension 1/Radius des *Osculating Circle*), wird sie derart normiert, dass ein Kreis (unabhängig von seiner Größe) immer eine NMBE von Eins aufweist. Mit Hilfe der Konturlänge

$$L = \int_0^T |\dot{z}(t)| dt \quad \text{bzw. im diskreten Fall} \quad L = \sum_{n=0}^{N-1} |\dot{z}(n)| \quad (4.69)$$

lässt sich die NMBE angeben als

$$\Psi(\sigma) = \left(\frac{L}{2\pi}\right)^2 \frac{1}{T} \int_0^T |\kappa_\sigma(t)|^2 dt \quad \text{für den kontinuierlichen Fall und} \quad (4.70)$$

$$\Psi(\sigma) = \left(\frac{L}{2\pi}\right)^2 \frac{1}{N} \sum_{n=0}^{N-1} |\kappa_\sigma(n)|^2 \quad \text{für den diskreten Fall.} \quad (4.71)$$

4.4 Weiteres Vorgehen

Die auf den letzten Seiten vorgestellten Formmerkmale werden im nächsten Schritt zu einem Merkmalsvektor zusammengefasst. Damit lässt sich jedes Objekt anhand einer Menge von reellen Zahlen – in Form seines Merkmalsvektors – beschreiben. Ziel ist die möglichst gute Charakterisierung der bei der Haustoriensegmentierung markierten Objekte auf Basis der vorgestellten Formmerkmale, um daraus eine zuverlässige Aussage treffen zu können, ob es sich bei dem betreffenden Objekt tatsächlich um ein Haustorium handelt oder nicht. Die dazu notwendigen Methoden der *Klassifikation* werden im nächsten Kapitel betrachtet.

Kapitel 5

Merkmalsauswahl und Klassifikation

Im Ergebnis der Formanalyse liegt für jedes in den transformierten Zellen detektierte Objekt eine Menge von Merkmalen – ein sogenannter Merkmalsvektor – vor. Die Aufgabe besteht nun darin, anhand der extrahierten Merkmale Schlussfolgerungen über einen Haustorienbefall der betrachteten Zelle zu ziehen, also die detektierten Objekte den zwei Klassen „*Haustorium*“ bzw. „*kein Haustorium*“ zuzuordnen.

Diese Zuordnung wird von einem *Klassifikator* übernommen. Es existiert eine große Menge von Klassifikationsverfahren, einige Vertreter seien an dieser Stelle genannt:

- Statistische Klassifikatoren [DHS01, Fis36]
- k-Nächste-Nachbarn (k-NN) Klassifikator [DHS01, CH67]
- Neuronale Netze [Bis95, Hay99, DHS01]
- Support-Vektor-Maschinen (*Support Vector Machines, SVM*) [CST00, SBV95]

Da jedes Klassifikationsverfahren bestimmte Vorzüge aufweist, ist es wichtig, anhand der konkreten Problemstellung einen *geeigneten* Klassifikator auszuwählen. Nicht immer sind dabei universelle Verfahren mit vielen Parametrierungsmöglichkeiten (dazu gehören z.B. Neuronale Netze und Support-Vektor-Maschinen) die erste Wahl, weil das Finden einer möglichst optimalen Parametrierung für den konkreten Datensatz sehr aufwändig und zeitintensiv sein kann. Von TAUTENHAHN [Tau04, TIS05] wurden diesbezüglich einige Untersuchungen zur Haustorienklassifikation unter Verwendung verschiedener Klassifikationsverfahren durchgeführt.

Bei den nachfolgenden Betrachtungen wird bewusst versucht, mit möglichst *einfachen* Verfahren auszukommen, welche auf eine universell einstellbare Parametrierung verzichten. Um trotzdem gute Ergebnisse zu erhalten bzw. um solche einfachen Verfahren überhaupt anwenden zu können, werden höhere Anforderungen an den zugrunde liegenden Merkmalsraum gestellt.

5.1 Anforderungen an Merkmale und Trainingsdaten

Generell hängt die Zuverlässigkeit einer Klassifikation nicht nur vom Klassifikator, sondern in sehr hohem Maße von der Qualität der dem Klassifikator präsentierten Merkmale ab. Ein guter Merkmalsvektor sollte folgende Eigenschaften aufweisen [Cas96]:

- **Diskriminanz**
Die Merkmalswerte für Objekte verschiedener Klassen sollen sich möglichst signifikant unterscheiden. Kein Klassifikator – wie ausgeklügelt er auch immer sei – vermag zwischen Klassen mit identischen Merkmalswerten zu unterscheiden.

- **Zuverlässigkeit**

Die Merkmale innerhalb einer Klasse sollen möglichst ähnlich sein. In Verbindung mit der vorher genannten Eigenschaft bedeutet dies, dass sich im Merkmalsraum kompakte, deutlich unterscheidbare Cluster bilden.

- **Unabhängigkeit**

Die Merkmale sollen möglichst unabhängig voneinander sein und keine Korrelation aufweisen. Diese Forderung wird durch die vierte gewünschte Eigenschaft erklärt:

- **Kompaktheit**

Die Komplexität eines Erkennungssystems wächst rapide mit der Anzahl der berücksichtigten Merkmale. Die Menge der benötigten Beispiele, um einen Klassifikator zuverlässig zu trainieren, wächst exponentiell mit der Anzahl der berücksichtigten Merkmale (*Curse of Dimensionality*) [Bis95]. Unter Umständen führt das „Aufblähen“ des Merkmalsvektors mit unzuverlässigen, verrauschten Merkmalen oder auch solchen, die mit bereits berücksichtigten Merkmalen eine hohe Korrelation aufweisen, zu einer Verschlechterung der Erkennungsrate.

Für die Lernphase muss ein geeigneter Trainingsdatensatz bereitgestellt werden, der folgenden Anforderungen genügen sollte [Cas96]:

- *Repräsentative* Trainingsdaten beinhalten Beispiele aller möglichen Objekttypen – inklusive selten auftretender Erscheinungen. Ein mit repräsentativen Daten trainierter Klassifikator weist bei der Klassifikation neuer, unbekannter Daten eine ähnliche Leistungsfähigkeit auf wie bei der Klassifikation der Trainingsdaten. Daher ist ein guter Trainingsdatensatz die Basis für die Beurteilung der Leistungsfähigkeit eines Erkennungssystems.
- Der Trainingsdatensatz darf *keine Klassifikationsfehler* aufweisen, da ansonsten die Entscheidungsparameter des Klassifikators fehlerhaft (*biased*) sind.

5.1.1 Datensatz zur Beschreibung der Haustorien

Für die Erstellung eines repräsentativen Datensatzes zur Beschreibung der Haustorien wurden einige tausend Mikroskopbilder mit den in Kapitel 3 vorgestellten Segmentierungsalgorithmen prozessiert, d.h., es wurden zunächst die gefärbten transformierten Zellen detektiert und anschließend die darin enthaltenen potenziellen Haustorienregionen segmentiert. Bei einer anschließenden manuellen Betrachtung der Segmentierungsergebnisse ließen sich einige hundert Zellen mit Haustorien identifizieren. Die Haustorienregionen wurden dabei von den Algorithmen zuverlässig markiert. Darüber hinaus traten erwartungsgemäß Markierungen auf, welche nicht zu Haustorien gehören, sondern von Verfärbungen innerhalb der Zelle herrühren.

In einer Kopie der zugehörigen Bilder wurden die Haustorien mit einer eindeutigen, automatisch auswertbaren Farbmarkierung versehen (manuelle Annotation), so dass jederzeit eine Abfrage der tatsächlichen Klassenzugehörigkeit der segmentierten Objekte auf Basis der *Rohdaten* möglich ist. Im Ergebnis umfasst dieser manuell annotierte Datensatz 500 Objekte der Klasse „*Haustorium*“ und 500 Objekte der Klasse „*kein Haustorium*“. Diese Objekte gehören zu 380 Zellen mit mindestens einem Haustorium sowie 361 Zellen ohne Haustorium. Alle nachfolgenden Analysen und Ergebnisse basieren auf diesem Datensatz.

5.1.2 Aufteilung des Datensatzes in Trainings- und Testmenge

Um die Leistungsfähigkeit eines Klassifikators auf *unbekannten* Daten beurteilen zu können, ist eine Aufteilung des Datensatzes in eine *Trainingsmenge* und eine zur Trainingsmenge disjunkten *Testmenge* notwendig. Ein Standardverfahren hierfür ist die *k*-fache Kreuzvalidierung [Bis95].

Die Stichprobenmenge wird dazu in k zufällig ausgewählte, disjunkte, etwa gleich mächtige Teilmengen zerlegt. In den nun folgenden k Durchläufen wird jeweils die k -te Menge als Testmenge verwendet, die verbleibenden Daten bilden die Trainingsmenge.

Optional kann die k -fache Kreuzvalidierung mehrfach (r Wiederholungen) mit einer jeweils anderen, zufälligen Aufteilungen der Daten durchgeführt werden, was zu einer Verbesserung der Stabilität der Messwerte führt. Dies wird als k -fache Kreuzvalidierung mit r Wiederholungen bezeichnet.

5.2 Merkmalstransformation

Zusätzlich zu den eingangs formulierten Anforderungen an Merkmale ist eine möglichst *homogene* Ausfüllung des Merkmalsraums erwünscht. Dazu ist prinzipiell die Anwendung jeder invertierbaren, nichtlinearen Transformation (z.B. Wurzeln, Logarithmen, etc.) der Eingangsdaten erlaubt. Dies lässt sich am besten an einem Beispiel einsehen: In Abbildung 5.1 sind die Streudiagramme (*Scatter Plots*) der Flusser-Momentinvarianzen ψ_1 , ψ_2 und ψ_7 (ohne Anwendung einer Normierung) dargestellt. Im *linken* Diagramm sind die Punktwolken augenscheinlich bei kleinen Wertebereichen dicht gedrängt und weisen bei größeren Wertebereichen vermeintliche Ausreißer auf. Von einer homogenen Verteilung im Merkmalsraum kann keine Rede sein.

Man beachte an dieser Stelle, dass Momente verschiedener Ordnung auch verschiedene physikalische Einheiten aufweisen. Konkret treten bei dem betrachteten Beispiel der Momentinvarianzen verschiedene Potenzen von *Längeneinheiten* auf. Um diese Merkmale auf die Grundeinheit „Länge“ zurückzutransformieren, bietet sich daher das Ziehen der Wurzel entsprechender Ordnung an. So wurden auf die exemplarisch ausgewählten Merkmale ψ_1 , ψ_2 und ψ_7 die Transformationen

$$\sqrt{\psi_1} \quad \sqrt[6]{\psi_2} \quad \sqrt[4]{\psi_7} \quad (5.1)$$

angewendet. Danach zeigt sich in Abbildung 5.1 (*rechts*) eine wesentlich homogenere Verteilung der Werte im Merkmalsraum. Des Weiteren ist nun deutlich die *lineare* Korrelationen zwischen $\sqrt{\psi_1}$ und $\sqrt[4]{\psi_7}$ erkennbar, welche sich im linken Diagramm (ohne Transformation) als *nichtlineare* Korrelationen andeutet. So ist eine geeignete nichtlineare Merkmalstransformation nicht nur vor dem Hintergrund einer möglichst homogenen Ausfüllung des Merkmalsraumes motiviert, sondern sie ist auch ein notwendiger Vorverarbeitungsschritt für eine aufwandsgünstige Dimensionsreduktion mittels linearer Korrelationsanalyse (siehe Abschnitt 5.4).

Diese exemplarischen Betrachtungen zeigen die Bedeutung dieses Vorverarbeitungsschritts. Generell ist zu beachten, dass die auf den Merkmalen durchgeführten nichtlinearen Transformationen nicht an die konkrete Aufgabenstellung der Haustorienerkennung gebunden sind. Es handelt sich vielmehr um allgemein gültige Ergebnisse, welche auch auf ähnliche Klassifikationsaufgaben angewendet werden können (und sollten).

Die vollständige Liste der nichtlinearen Merkmalstransformationen ist in Tabelle 5.1 angegeben. Zuvor erfolgen jedoch noch einige weitere Betrachtungen zur Vorverarbeitung des Merkmalsvektors, bevor das Thema „Merkmalstransformation“ anhand weiterer Merkmale weiter konkretisiert wird.

5.3 Merkmalsnormalisierung

Prinzipbedingt liegen die Werte der berechneten Merkmale in sehr unterschiedlichen Bereichen. So weisen einige Merkmale (z.B. Rundheit oder Exzentrizität) einen Wertebereich zwischen Null und Eins auf, während beispielsweise die Objektfläche einige Größenordnungen darüber liegen kann. Durch eine geeignete Normierung der Merkmalswerte bekommt jede Dimension des Merkmalsraums einen ähnlich großen Wertebereich.

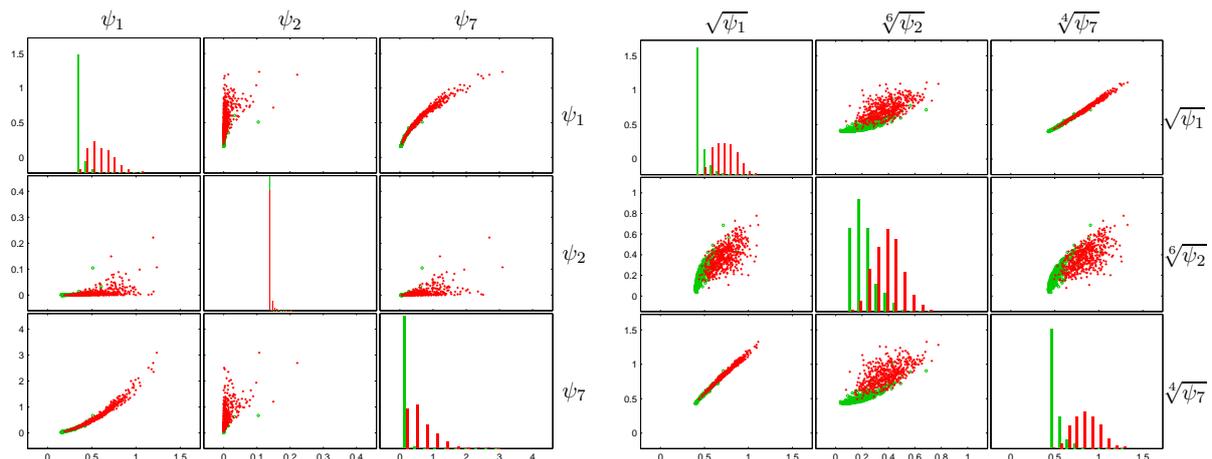


Abbildung 5.1: Einfluss einer nichtlinearen Transformation auf die Verteilung im Merkmalsraum. Dargestellt sind die Streudiagramme (*scatter plots*) einiger Flusser-Momentinvarianzen.

Links: Ohne Transformation.

Rechts: Nach der nichtlinearen Transformation gemäß Gleichung (5.1).

Weit verbreitet ist die Min-Max-Normalisierung, bei der sich die Merkmalswerte in einen Hyperquader mit der unteren (*lower*) und oberen (*upper*) Begrenzung \mathbf{l} und \mathbf{u} aufhalten.

$$\tilde{\mathbf{x}} = \mathbf{l} + (\mathbf{u} - \mathbf{l}) \cdot \frac{\mathbf{x} - \min_{\mathbf{x}}}{\max_{\mathbf{x}} - \min_{\mathbf{x}}} \quad (5.2)$$

Jedoch werden die Minima $\min_{\mathbf{x}}$ und Maxima $\max_{\mathbf{x}}$ stark von Ausreißern (*Outlier*) beeinflusst, so dass diese Methode nur eingeschränkt zu empfehlen ist.

Weitaus robuster ist eine Skalierung derart, dass die Merkmale einen Mittelwert von Null und eine Standardabweichung von Eins aufweisen (*Z-Score*).

$$\tilde{\mathbf{x}} = \frac{\mathbf{x} - \boldsymbol{\mu}}{\boldsymbol{\sigma}} \quad (5.3)$$

Ebenso wie in Gleichung (5.2) ist die Division elementweise durchzuführen, der Vektor $\boldsymbol{\sigma}$ enthält dabei die Standardabweichungen der einzelnen Komponenten des Merkmalsvektors.

5.4 Transformation der Merkmale in den Eingaberaum

Nach den recht umfassenden Ausführungen zum Thema „Formanalyse und Formmerkmale“ in Kapitel 4 steht nun eine *sehr große Anzahl* von Merkmalen zur Verfügung, welche die Formeigenschaften der segmentierten Objekte auf unterschiedliche Art widerspiegeln. Vor dem Hintergrund der Forderung, eine möglichst *kompakte* Beschreibung für die zu klassifizierenden Objekte zu finden, stellt sich als nächstes die Frage, welche Untermenge der Merkmale sich besonders gut zur Lösung des Klassifikationsproblems eignet. Bekanntermaßen vergrößert die Hinzunahme von Eingangsdimensionen die Komplexität des Klassifikationssystems auf exponentielle Weise (*Curse of Dimensionality*) [Bis95]. Die Optimierungsaufgabe besteht nun darin, einen Eingaberaum¹ zu finden, welcher möglichst wenig Dimensionen aufweist und mit dem trotzdem eine hinreichend gute Klassifikation erreicht werden kann. Entscheidend für die Qualität der Klassifikation ist letztlich die Separierbarkeit der Klassen im Merkmalsraum – die Diskriminanz.

Für die Bereitstellung eines guten Eingaberaums bieten sich zwei verschiedene Strategien an, welche sich beide mittels einer Matrix-Vektor-Multiplikation ausdrücken lassen:

¹Unter dem Eingaberaum ist jener transformierte „Merkmalsraum“ zu verstehen, mit dem der Klassifikator letztendlich beaufschlagt wird.

1. Selektion von Merkmalen

Aus dem Satz der vorhandenen Merkmale werden diejenigen ausgewählt, die für die Unterscheidung der Klassen besonders relevant sind. Allerdings ist das vollständige „Durchtesten“ aller Kombinationen nur für kleine Merkmalsätze praktikabel, denn der Aufwand, aus d Merkmalen die Kombination der besten d' Merkmale herauszufinden, bedeutet einen Suchaufwand von $\binom{d}{d'}$. Daher können für diese Vorgehensweise in der Praxis nur suboptimale Techniken bzw. Heuristiken herangezogen werden.

2. Linearkombination von Merkmalen

Ein allgemeineres Vorgehen ist die Linearkombination von Merkmalen. Jede Dimension des Eingaberaums repräsentiert dann eine gewichtete Summe aller Merkmale. Die Beträge der Komponenten des Wichtungsvektors können dabei als die Relevanz des zugehörigen Merkmals interpretiert werden. Bekannte Methoden sind die Hauptachsentransformation (*Principal Component Analysis, PCA*) und die lineare Diskriminanzanalyse (*Linear Discriminant Analysis, LDA*).

Für eine Abschätzung der Klassifikationsgüte lassen sich unter bestimmten Randbedingungen geschlossene Gleichungen angeben. So ist zunächst eine grobe Analyse von Merkmalskombinationen möglich, ohne tatsächlich einen Klassifikator trainieren und testen zu müssen. Dies führt auf eine wichtige Kenngröße von Klassen im Merkmalsraum: die Bhattacharyya-Distanz.

5.4.1 Bhattacharyya-Distanz

Der Bayes-Klassifikationsfehler auf Grundlage zweier Verteilungen ist gegeben durch [Fuk94, DHS01]

$$\epsilon = \int \min \{P_1 p(\mathbf{x} | \omega_1), P_2 p(\mathbf{x} | \omega_2)\} d\mathbf{x}. \tag{5.4}$$

Mittels der Ungleichung

$$\min \{a, b\} \leq a^s b^{1-s} \quad \text{mit} \quad 0 \leq s \leq 1 \quad \text{und} \quad a, b \geq 0 \tag{5.5}$$

lässt sich eine obere Schranke (*upper bound*) ϵ_u für den Klassifikationsfehler angeben. Diese wird auch als Chernoff-Schranke (*Chernoff bound*) [Che52] bezeichnet.

$$\epsilon_u = P_1^s P_2^{1-s} \int p^s(\mathbf{x} | \omega_1) p^{1-s}(\mathbf{x} | \omega_2) d\mathbf{x} \quad \text{mit} \quad 0 \leq s \leq 1 \tag{5.6}$$

Der Parameter s muss dabei durch Minimierung von ϵ_u gefunden werden.

Im Fall normalverteilter Daten lässt sich folgende geschlossene Lösung angeben:

$$\int p^s(\mathbf{x} | \omega_1) p^{1-s}(\mathbf{x} | \omega_2) d\mathbf{x} = e^{-\mu(s)} \quad \text{mit} \tag{5.7}$$

$$\mu(s) = \frac{1}{2} s(1-s)(\boldsymbol{\mu}_2 - \boldsymbol{\mu}_1)^T [s \boldsymbol{\Sigma}_1 + (1-s) \boldsymbol{\Sigma}_2]^{-1} (\boldsymbol{\mu}_2 - \boldsymbol{\mu}_1) + \frac{1}{2} \ln \left(\frac{|s \boldsymbol{\Sigma}_1 + (1-s) \boldsymbol{\Sigma}_2|}{|\boldsymbol{\Sigma}_1|^s |\boldsymbol{\Sigma}_2|^{1-s}} \right). \tag{5.8}$$

Der Ausdruck $\mu(s)$ ist dabei die sogenannte Chernoff-Distanz (*Chernoff distance*).

Wird $s = 0.5$ gewählt, so ergibt sich

$$\epsilon_u = \sqrt{P_1 P_2} \int p(\mathbf{x} | \omega_1) p(\mathbf{x} | \omega_2) d\mathbf{x} = \sqrt{P_1 P_2} e^{-\mu(0.5)} \tag{5.9}$$

mit der so genannten Bhattacharyya-Distanz [Bha43]

$$\mu(0.5) = \underbrace{\frac{1}{8}(\boldsymbol{\mu}_2 - \boldsymbol{\mu}_1)^T \left[\frac{\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2}{2} \right]^{-1} (\boldsymbol{\mu}_2 - \boldsymbol{\mu}_1)}_{\mu_1} + \underbrace{\frac{1}{2} \ln \left(\frac{|\frac{\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2}{2}|}{\sqrt{|\boldsymbol{\Sigma}_1| |\boldsymbol{\Sigma}_2|}} \right)}_{\mu_2}. \quad (5.10)$$

Die Bhattacharyya-Distanz $\mu(0.5)$ ist ein wichtiges Maß für den Abstand zweier Verteilungen im Merkmalsraum. Dabei repräsentieren die Terme μ_1 und μ_2 die Klassenseparabilität aufgrund unterschiedlicher Klassenmittelwerte (Linearterm μ_1) bzw. unterschiedlicher Kovarianz (Kovarianzterm μ_2). Bei gleicher Kovarianz ($\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2$) sind die Klassen *linear separierbar* und der Term μ_2 verschwindet. Außerdem ergibt sich in diesem Fall der optimale Wert des Parameters s der Chernoff-Distanz zu $s = 0.5$ [Fuk94]. Damit ist für $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2$ die Bhattacharyya-Distanz gleich der optimalen Chernoff-Distanz.

Anwendbarkeit dieser Werkzeuge

Im Fall normalverteilter Daten lässt sich mittels der Bhattacharyya-Distanz auf sehr einfache Weise der zu erwartende Bayes-Klassifikationsfehler abschätzen. Etwas genauere Ergebnisse sind durch die Verwendung der Chernoff-Distanz erreichbar – allerdings muss der Parameter s dann für jeden konkreten Fall numerisch bestimmt werden. Weichen die Verteilungen jedoch merklich von der Gaußverteilung ab, so sind diese auf Basis von Mittelwert und Kovarianz ermittelten Ergebnisse nicht mehr sinnvoll interpretierbar [DHS01]. Unter Umständen können Abweichungen von der Gaußverteilung durch nichtlineare Transformationen kompensiert werden, so dass eine Anwendung dieser Distanzmaße dennoch ermöglicht wird. Dies wird in Abschnitt 5.5.3 konkret am Beispiel der Flusser-Momentinvarianzen demonstriert.

5.4.2 Hauptachsentransformation (PCA)

Um Korrelationen zwischen Dimensionen zu entfernen, kann eine Hauptkomponentenanalyse [Bis95, Jol02] (*Principal Component Analysis, PCA*)² durchgeführt werden. Die Eigenwertzerlegung der (symmetrischen, reellwertigen) Kovarianzmatrix $\boldsymbol{\Sigma} = \mathbf{V}\boldsymbol{\Lambda}\mathbf{V}^{-1} = \mathbf{V}\boldsymbol{\Lambda}\mathbf{V}^T$ liefert ein neues Koordinatensystem (Orthonormalsystem), welches in Richtung der *maximalen Varianz* im Datenraum orientiert ist. Das Eigenwertspektrum gibt dabei die Varianz entlang der einzelnen Hauptachsen an. Durch Projektion der Daten in die zu den d' stärksten Eigenwerten gehörenden Hauptachsen wird eine Dimensionsreduktion erreicht: Aus dem d -dimensionalen Merkmalsraum wird ein d' -dimensionaler Eingaberaum.

Da jedoch bei der Berechnung der Hauptachsen keine Information über die Klassenzugehörigkeit der Daten ausgewertet wird, kann bei dieser Art der Dimensionsreduktion wichtige Information (Diskriminanz) verloren gehen, denn große *Varianz* im Merkmalsraum muss nicht gleichbedeutend mit großer *Diskriminanz* sein. Dieses Problem wird durch ein anderes lineares Verfahren gelöst – die lineare Diskriminanzanalyse (LDA).

5.4.3 Lineare Diskriminanzanalyse (LDA)

Ebenso wie die Hauptachsentransformation basiert die lineare Diskriminanzanalyse (*Linear Discriminant Analysis, LDA*) auf den Momenten zweiter Ordnung (Kovarianzen). Während jedoch die Hauptachsentransformation ein Orthonormalsystem in Richtung der größten *Varianz* liefert, ist das Ziel der LDA die Maximierung der *Diskriminanz*. Dazu werden verschiedene Kovarianzmatrizen benötigt: Die Intraklassen-Kovarianzmatrix (*Within-class covariance matrix*) gibt die

²Die Hauptkomponentenanalyse wurde ursprünglich von Pearson [Pea01] beschrieben. Andere Bezeichnungen für die PCA sind Karhunen-Loève-Transformation (KLT) [Loè77, Loè78] und Hotelling-Transformation [Hot33].

Streuung der Samples einer Klasse um den Klassenmittelpunkt $\boldsymbol{\mu}_k$ an.

$$\mathbf{S}_w = \sum_{k=1}^K P_k \mathbb{E}[(\mathbf{x} - \boldsymbol{\mu}_k)(\mathbf{x} - \boldsymbol{\mu}_k)^T | \omega_k] = \sum_{k=1}^K P_k \boldsymbol{\Sigma}_k \quad (5.11)$$

Die Interklassen-Kovarianzmatrix (*Between-class covariance matrix*) misst die Streuung der Klassenmittelpunkte $\boldsymbol{\mu}_k$ um den gemeinsamen Mittelwert $\boldsymbol{\mu}_0$ aller Samples.

$$\mathbf{S}_b = \sum_{k=1}^K P_k (\boldsymbol{\mu}_k - \boldsymbol{\mu}_0)(\boldsymbol{\mu}_k - \boldsymbol{\mu}_0)^T \quad (5.12)$$

Die Misch-Kovarianzmatrix (*Mixture covariance matrix*) gibt letztendlich die Streuung aller Samples (unabhängig von ihrer Klasse) an.

$$\mathbf{S}_m = \mathbb{E}[(\mathbf{x} - \boldsymbol{\mu}_0)(\mathbf{x} - \boldsymbol{\mu}_0)^T] = \mathbf{S}_w + \mathbf{S}_b \quad (5.13)$$

Der zu findende Unterraum soll nun das Verhältnis der Interklassen- zu Intraklassen-Kovarianzen maximieren.

$$C = \max_{\mathbf{d}} \frac{\mathbf{d}^T \mathbf{S}_b \mathbf{d}}{\mathbf{d}^T \mathbf{S}_w \mathbf{d}} \quad (5.14)$$

Die Lösung lässt sich über das verallgemeinerte Eigenwertproblem

$$\mathbf{S}_b \mathbf{d}_1 = \mathbf{S}_w \lambda_1 \mathbf{d}_1 \quad (5.15)$$

bzw. (da \mathbf{S}_w vollen Rang hat und damit invertierbar ist) über das Eigenwertproblem

$$\mathbf{S}_w^{-1} \mathbf{S}_b \mathbf{d}_1 = \lambda_1 \mathbf{d}_1 \quad (5.16)$$

finden, wobei die Eigenwerte anhand ihrer Größe sortiert sind $\lambda_1 \geq \lambda_2 \geq \dots \lambda_d$.

Der maximale Rang von \mathbf{S}_b ist jedoch $\text{rg}(\mathbf{S}_b) = K - 1$, daher lässt sich für ein Zweiklassenproblem auf diese Weise nur *ein einziger* Diskriminanzvektor finden – die Daten können also vom d -dimensionalen Raum in den eindimensionalen Raum projiziert werden.

Wünschenswert ist jedoch, ein Orthonormalsystem von Diskriminanzvektoren zu bestimmen, welches eine Abbildung auf eine *beliebige* Anzahl von $1 \dots d'$ Dimensionen zulässt. Ein entsprechendes Verfahren wird von DUCHENE und LECLERCQ in [DL88] beschrieben.

5.5 Untersuchung des Merkmalsraums

Mittels der in den Abschnitten 5.2, 5.3 und 5.4 vorgestellten Werkzeuge sollen im Folgenden konkrete Untersuchungen zu den in Kapitel 4 vorgestellten Merkmalen durchgeführt werden. Wichtigstes Maß ist dabei die erreichte Trennschärfe (Diskriminanz). Diese lässt sich in Form der Bhattacharyya-Distanz sowohl für *einzelne* Merkmale als auch für *Merkmalskombinationen* angeben. Zur Untermauerung der auf die Merkmale angewendeten nichtlinearen Transformationen (siehe Tabelle 5.1) werden außerdem Streudiagramme (*Scatter Plots*) angegeben, welche die dadurch erreichte Entzerrung der Verteilung im Merkmalsraum verdeutlichen.

5.5.1 Trennschärfe einzelner Merkmale

In Abbildung 5.2 ist die Trennschärfe jedes einzelnen Merkmals anhand der Bhattacharyya-Distanz $\mu(0.5)$ sowie deren Linearterm μ_1 nach Gleichung (5.10) dargestellt: Auffällig ist, dass kein Merkmal eine besonders herausragende Trennschärfe aufweist. In der Klasse der einfachen

Dimension	Merkmal	Bedeutung	Transformation	Referenz
Einfache Merkmale				
1	S	mittlere Farbsättigung	S	(A.17)
2	A	Fläche	$\tilde{A} = \sqrt{A}$	(4.2)
3	P	Umfang	P	(4.3)
4	$2a$	Lange Hauptachse	$2a$	(4.9)
5	$2b$	Kurze Hauptachse	$2b$	(4.10)
6	k	Klirrfaktor	k	(4.57)
7	Ψ	Verbiegungsenergie (NMBE)	$\tilde{\Psi} = \sqrt{1/\Psi}$	(4.71)
Momentinvarianzen aus dem Binärbild				
8	ψ_1	Flusser-Moment 1	$\tilde{\psi}_1 = \sqrt{\psi_1}$	(4.23)
9	ψ_2	Flusser-Moment 2	$\tilde{\psi}_2 = \sqrt[6]{\psi_2}$	(4.24)
10	ψ_3	Flusser-Moment 3	$\tilde{\psi}_3 = \sqrt[8]{ \psi_3 }$	(4.25)
11	ψ_4	Flusser-Moment 4	$\tilde{\psi}_4 = \sqrt[8]{ \psi_4 }$	(4.26)
12	ψ_5	Flusser-Moment 5	$\tilde{\psi}_5 = \sqrt[12]{ \psi_5 }$	(4.27)
13	ψ_6	Flusser-Moment 6	$\tilde{\psi}_6 = \sqrt[12]{ \psi_6 }$	(4.28)
14	ψ_7	Flusser-Moment 7	$\tilde{\psi}_7 = \sqrt[4]{\psi_7}$	(4.29)
15	ψ_8	Flusser-Moment 8	$\tilde{\psi}_8 = \sqrt[10]{ \psi_8 }$	(4.30)
16	ψ_9	Flusser-Moment 9	$\tilde{\psi}_9 = \sqrt[10]{ \psi_9 }$	(4.31)
17	ψ_{10}	Flusser-Moment 10	$\tilde{\psi}_{10} = \sqrt[16]{ \psi_{10} }$	(4.32)
18	ψ_{11}	Flusser-Moment 11	$\tilde{\psi}_{11} = \sqrt[16]{ \psi_{11} }$	(4.33)
Momentinvarianzen aus dem Sättigungsbild				
19...29	$\psi_1 \dots \psi_{11}$	Flusser-Momente 1...11	wie beim Binärbild	(4.23)...(4.33)
Merkmale aus dem Spektrum der Objektkontur				
30...56	$\hat{Z}_{-14} \dots \hat{Z}_{14}$	Spektralkoeffizienten	$\tilde{Z}_\nu = \left \frac{\hat{Z}_\nu}{\hat{Z}_1} \right ^{\frac{1}{ \nu }}$	(5.17)
57...83	$b_2 \dots b_{28}$	Granlund-Invarianzen	$\tilde{b}_n = \sqrt{\tilde{Z}_{1+n} \tilde{Z}_{1-n}}$	(5.18)

Tabelle 5.1: Merkmale zur Haustorienerkennung und angewendete nichtlineare Transformationen zur Entzerrung der Verteilung im Merkmalsraum.

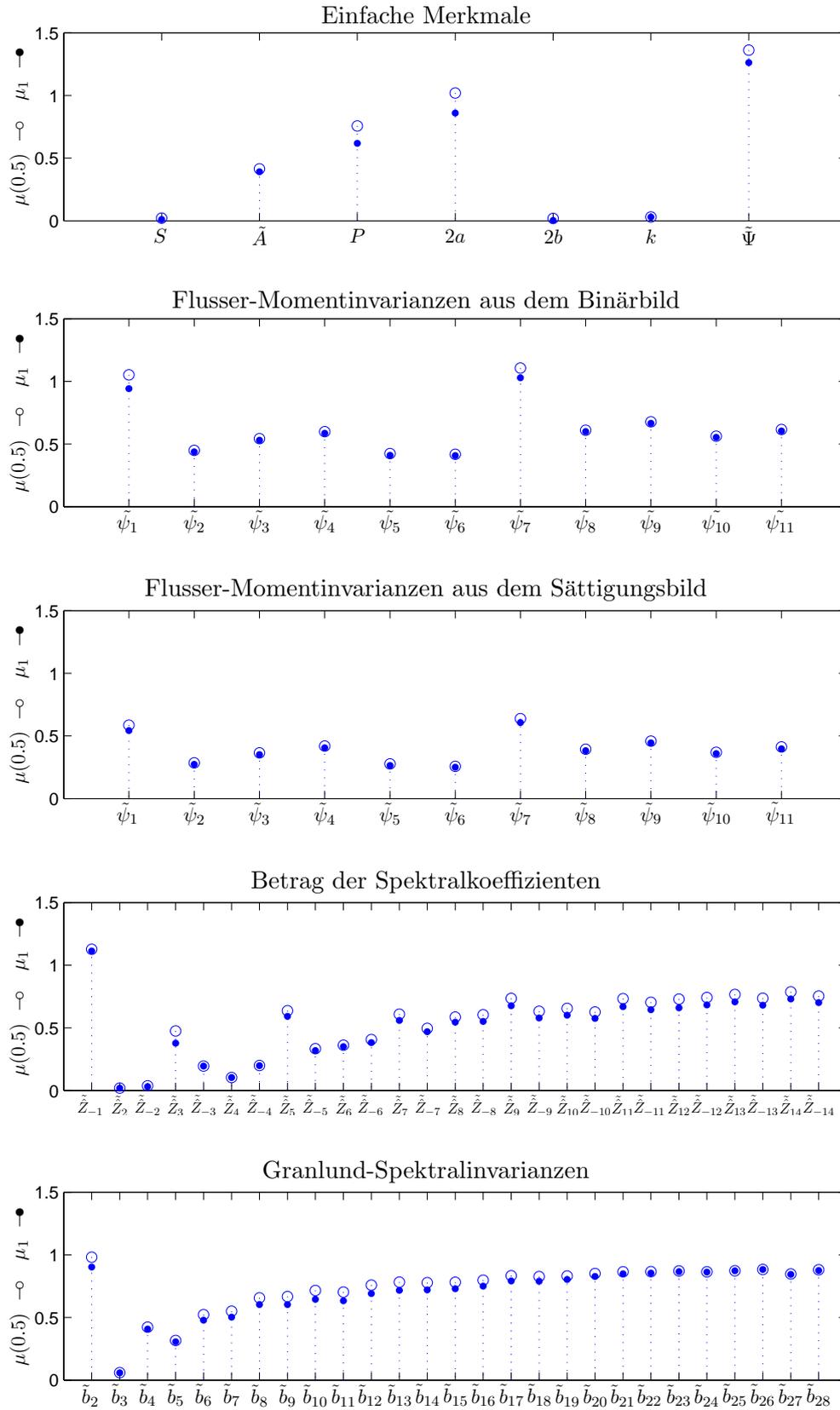


Abbildung 5.2: Bhattacharyya-Distanz $\mu(0.5)$ (○) und deren Linearterm μ_1 (●) jedes einzelnen Merkmals aus Tabelle 5.1.

Merkmale ist die *Normalized Multiscale Bending Energy (NMBE)* Ψ das treffendste Merkmal. Bei den Flusser-Momentinvarianzen heben sich ψ_1 und ψ_7 hervor. Im Spektralbereich ist offenbar der Koeffizient \hat{Z}_{-1} (welcher sich auch in der Granlund-Invarianz b_2 widerspiegelt) besonders diskriminanz. Das Merkmal „Farbsättigung“ S , auf dessen Basis die Haustorienobjekte detektiert werden (vgl. Abschnitt 3.4), ist offenbar bereits durch die Segmentierungsalgorithmen erschöpfend ausgewertet – es weist fast keine Diskriminanz auf.

5.5.2 Trennschärfe von Merkmalskombinationen

Die Untersuchung von Merkmalskombinationen kann aufgrund der Vielfalt von Kombinationsmöglichkeiten sehr schnell unübersichtlich werden. Deshalb wurden die folgenden Merkmale zu Klassen zusammengefasst und einige Kombinationen dieser Klassen untersucht.

- Momentinvarianzen (sowohl auf Basis des Binärbildes als auch des Sättigungsbildes)
- Spektralkoeffizienten (Fourierdeskriptoren)
- Übrige (einfache) Merkmale

In Tabelle 5.2 ist die dabei erreichte Trennschärfe in Form der Bhattacharyya-Distanz angegeben. Vor der Berechnung der Bhattacharyya-Distanz wurde der Merkmalsraum jeweils mittels einer linearen Diskriminanzanalyse (LDA) in einen eindimensionalen Eingaberaum transformiert – die Trennschärfe wurde also auf der Achse maximaler Diskriminanz ermittelt. Dies ist einerseits notwendig, andererseits jedoch auch hinreichend: In Abbildung 5.3 ist die Projektion des vollständigen Merkmalsraums (alle Merkmale aus Tabelle 5.1) mittels LDA in den 2D-Raum dargestellt. Beide Dimensionen weisen eine sehr starke Korrelation auf, so dass die Hinzunahme weiterer Diskriminanzachsen keinen signifikanten Informationsgewinn mit sich bringt. *Notwendig* ist eine Dimensionsreduktion deshalb, weil der Wert der Bhattacharyya-Distanz für höherdimensionale Eingaberäume immer unzuverlässiger wird. Konkret ist es der Kovarianzterm μ_2 in Gleichung (5.10), welcher signifikant zu diesem Fehler beiträgt, da durch die mit der Anzahl der Dimensionen exponentiell zunehmende „Ausdünnung“ der Punktwolke im Merkmalsraum (*Curse of Dimensionality*) die Schätzung der Kovarianzen immer fehlerbehafteter wird. Die berechneten Werte in Abbildung 5.4 verdeutlichen dies.

5.5.3 Streudiagramme der Merkmale

Flusser-Momentinvarianzen

In Abschnitt 5.2 wurde gezeigt, dass die Verteilung der Flusser-Momentinvarianzen im Merkmalsraum mittels nichtlinearer Transformationen (Wurzeloperationen) entzerrt bzw. homogenisiert werden kann. Abbildung F.2 *oben* (auf Seite 124) zeigt die Streudiagramme der nichtlinear transformierten Invarianzen $\psi_1 \dots \psi_{11}$. Mit Ausnahme von ψ_1 , ψ_2 und ψ_7 können sowohl positive als auch negative Werte auftreten. Deshalb wurde die Wurzel aus den Beträgen gezogen und das Vorzeichen separat betrachtet. Im Ergebnis ergeben sich multimodale Verteilungen im Merkmalsraum. Bei näherer Betrachtung der Streudiagramme fällt auf, dass die Berücksichtigung des Vorzeichens *keinen* Gewinn an Diskriminanz zwischen den Klassen bringt. Aus diesem Grund wurden in Abbildung F.2 *unten* nur die Beträge betrachtet. Durch die nunmehr *unimodale* Verteilung im Merkmalsraum ist eine Approximation mittels einer Gaußverteilung möglich.

Merkmale aus dem Spektralbereich (Fourierdeskriptoren)

In den Abschnitten 4.3.3 und 4.3.4 wurden zwei Möglichkeiten vorgestellt, um zu Formbeschreibungen mittels Fourierdeskriptoren zu gelangen: Einerseits können direkt die auf \hat{Z}_1 normierten

Symbol	Merkmale	$\mu(0.5)$	μ_1
○	Einfache Merkmale $[S \ \tilde{A} \ P \ 2a \ 2b \ k \ \tilde{\Psi}]$	1.74	1.73
+	Flusser-Momentinvarianzen $\tilde{\psi}_1 \dots \tilde{\psi}_{11}$ aus dem Binärbild	1.27	1.23
×	Flusser-Momentinvarianzen $\tilde{\psi}_1 \dots \tilde{\psi}_{11}$ aus dem Sättigungsbild	0.95	0.93
◁	Granlund-Spektralinvanzen $\tilde{b}_2 \dots \tilde{b}_{12}$	1.18	1.12
▷	Granlund-Spektralinvanzen $\tilde{b}_2 \dots \tilde{b}_{28}$	1.38	1.37
▽	Spektralkoeffizienten $[\tilde{Z}_{-6} \ \tilde{Z}_{-5} \ \dots \ \tilde{Z}_{-1} \ \tilde{Z}_2 \ \dots \ \tilde{Z}_5]$	1.46	1.45
△	Spektralkoeffizienten $[\tilde{Z}_{-14} \ \tilde{Z}_{-13} \ \dots \ \tilde{Z}_{-1} \ \tilde{Z}_2 \ \dots \ \tilde{Z}_{14}]$	1.59	1.59
*	Kombination + ×	1.58	1.54
⊗	Kombination ○ + ×	2.15	2.14
□	Kombination ○ ▷	1.97	1.97
◇	Kombination ○ △	1.94	1.94
★	Kombination ○ ▷ △	2.08	2.08
●	Kombination ○ + × △	2.33	2.33

Tabelle 5.2: Bhattacharyya-Distanz $\mu(0.5)$ und deren Linearterm μ_1 verschiedener Merkmalskombinationen auf der Achse maximaler Diskriminanz (1. Diskriminanzachse). Die Symbole werden in den Diagrammen auf Seite 89 und 90 verwendet.

Beträge der Fourierdeskriptoren gemäß Gleichung (4.52) verwendet werden, ein anderer Ansatz führt über die Granlund-Spektralinvanzen nach Gleichung (4.54). Da die Einhüllende des Spektrums der betrachteten Objekte rasch abklingt, werden die Spektralkoeffizienten bei beiden Ansätzen vor der Verwendung folgendermaßen entzerrt:

$$\tilde{Z}_\nu = \left| \frac{\hat{Z}_\nu}{\hat{Z}_1} \right|^{\frac{1}{|\nu|}} . \tag{5.17}$$

Die entzerrten Granlund-Invarianzen ergeben sich dann zu

$$\tilde{b}_n = \sqrt{\tilde{Z}_{1+n} \tilde{Z}_{1-n}} . \tag{5.18}$$

Abbildung F.3 (auf Seite 125) zeigt die entsprechenden Streudiagramme.

Übrige Merkmale

Quantitativ stellen die Moment- bzw. Spektralinvanzen den Großteil der verfügbaren Merkmale dar. Weitere wichtige Information ist aber auch in den „einfachen“ Merkmalen enthalten, die am Anfang der Tabelle 5.1 aufgelistet sind. Sie enthalten unter anderem absolute Größenangaben (sind also nicht skalierungsinvariant) sowie auch Farbinformation. Abbildung F.1 (auf Seite 123) zeigt die zugehörigen Streudiagramme.

5.5.4 Schlussfolgerungen

Die Klassifikation der in den Zellen gefundenen Objekte in die Klassen „Haustorium“ bzw. „kein Haustorium“ ist nicht mittels eines alleinigen, „bahnbrechenden“ Merkmals möglich. Das

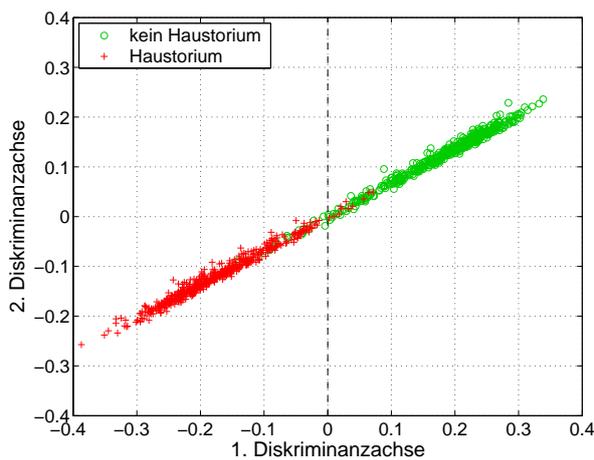


Abbildung 5.3: Projektion des gesamten Merkmalsraums (alle in Tabelle 5.1 aufgeführten Merkmale) in den 2D-Raum mittels linearer Diskriminanzanalyse (LDA). Beide Dimensionen sind sehr stark korreliert.

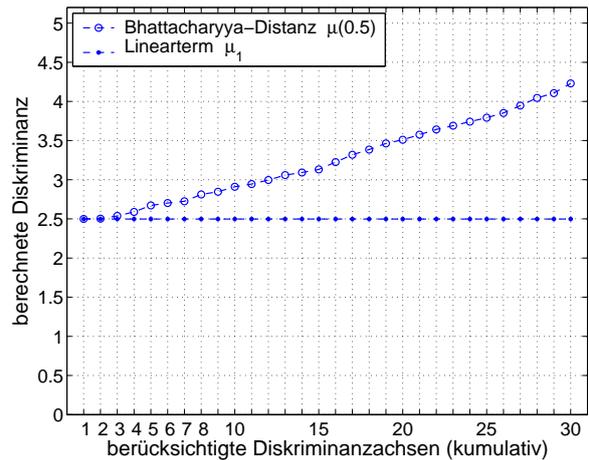


Abbildung 5.4: Berechnete Bhattacharyya-Distanz bei wachsendem Eingaberaum auf Basis der LDA. Während der Linearterm μ_1 auch für hochdimensionale Räume stabile Ergebnisse liefert, führt der Kovarianzterm μ_2 zu immer unzuverlässigeren Ergebnissen.

herausragende Merkmal für solche Objekte existiert offenbar nicht. Vielmehr führt die *Kombination* vieler Merkmale (alle verwendeten Merkmale sind in Tabelle 5.1 aufgeführt) zu einem guten Gesamtergebnis.

Der vergleichsweise große Aufwand, welcher auf den letzten Seiten zur Bereitstellung guter Merkmale betrieben wurde (Anwendung nichtlinearer Transformationen zur Entzerrung und Annäherung der Verteilung an eine Normalverteilung), zahlt sich nun an dieser Stelle aus: Aufgrund der unimodalen und recht kompakten Verteilung der Klassen im Merkmalsraum lässt sich das Problem mit Hilfe der linearen Diskriminanzanalyse (LDA) nahezu verlustfrei in den *eindimensionalen* Raum transformieren. Dies ermöglicht eine sehr einfache, aufwandsgünstige und robuste Klassifikation.

5.6 Klassifikationsergebnisse

Durch die nahezu perfekte Abbildung des hochdimensionalen Merkmalsraums in einen eindimensionalen Raum (unter der Randbedingung der Erhaltung maximaler Diskriminanz) kondensiert das ehemals sehr komplexe Klassifikationsproblem auf die einfache Anwendung einer skalaren Entscheidungsschwelle. Das Klassifikationsergebnis ist damit nur noch von diesem einen Parameter abhängig.

5.6.1 Confusion Matrix

Zur Beurteilung der Güte einer Klassifikation lässt sich die so genannte *Confusion Matrix* aufstellen. Sie enthält die Informationen über die tatsächlichen und die vom Klassifizierer ermittelten Klassen. Für den Fall der Haustorienklassifikation sieht diese Matrix folgendermaßen aus:

	Objekt ist ein Haustorium	Objekt ist <i>kein</i> Haustorium
klassifiziert als „ <i>Haustorium</i> “	richtig positiv (<i>a</i>)	falsch positiv (<i>b</i>)
klassifiziert als „ <i>kein Haustorium</i> “	falsch negativ (<i>c</i>)	richtig negativ (<i>d</i>)

Aus der Anzahl der richtig positiv, falsch negativ, falsch positiv und richtig negativ klassifizierten Objekte lassen sich folgende Aussagen ableiten:

- Die **Sensitivität**, auch als *recall* oder *true positive rate* TP bezeichnet, gibt die Quote der korrekt als „*Haustorium*“ erkannten Objekte an.

$$TP = \frac{a}{a + c} \quad (5.19)$$

- Die **Spezifität**, auch als *true negative rate* TN bezeichnet, gibt die Quote der korrekt als „*kein Haustorium*“ erkannten Objekte an.

$$TN = \frac{d}{b + d} \quad (5.20)$$

Entsprechend ist $1 - TN$ die Quote der falsch positiven Klassifikationen.

- Die **Korrektklassifikationsrate**, auch als *accuracy* AC bezeichnet, gibt die Quote der richtig klassifizierten Objekte (richtig positiv bzw. richtig negativ) an.

$$AC = \frac{a + d}{a + b + c + d} \quad (5.21)$$

Zellklassifikation

Für das Endergebnis aus Sicht der Anwendung ist jedoch nicht die Qualität der Objektklassifikation entscheidend, sondern vielmehr die der *Zell*klassifikation: Eine Zelle kann mehrere Objekte enthalten. Liegt mindestens ein *Haustorium* vor, so gilt die Zelle als infiziert. Aus biologischer Sicht ist es unerheblich, ob eine Zelle mit *einem* oder mit *mehreren* *Haustorien* infiziert ist. Die entsprechende *Confusion Matrix* ist:

	Zelle ist infiziert	Zelle ist <i>nicht</i> infiziert
mind. ein Objekt als „ <i>Haustorium</i> “ klassifiziert	richtig positiv (a)	falsch positiv (b)
kein Objekt als „ <i>Haustorium</i> “ klassifiziert	falsch negativ (c)	richtig negativ (d)

Dabei gelten die gleichen Zusammenhänge zur Sensitivität, Spezifität und Korrektklassifikationsrate wie bei der Objektklassifikation.

5.6.2 Ergebnisse

Auf den folgenden Seiten sind die erreichten Klassifikationsergebnisse unter Verwendung der in Tabelle 5.2 aufgeführten Merkmalskombinationen dargestellt. Abbildung 5.5 zeigt zunächst die optimalen Arbeitspunkte des Klassifikators im ROC-Diagramm (*Receiver Operating Characteristic, ROC*) bei der Klassifikation des Trainingsdatensatzes. In diesen Diagrammen sind die einzelnen Ergebnisse aufgrund des Maßstabs nicht zu unterscheiden – sie sollen lediglich zur Orientierung dienen und verdeutlichen, dass sich die Klassifikationsergebnisse nahe des theoretischen Optimums in der linken oberen Ecke der Diagramme befinden. Im Folgenden werden ausschließlich *Ausschnitte* dieser Diagramme betrachtet. So zeigt Abbildung 5.6 einen Ausschnitt der ROC-Kurven bei Klassifikation auf den Trainingsdaten. Abbildung 5.7 gibt die zugehörigen Korrektklassifikationsraten im optimalen Arbeitspunkt an.

Um eine Aussage zur Klassifikationsgüte auf *unbekannten* Daten ableiten zu können, wurde eine 2-fache Kreuzvalidierung mit 500 Wiederholungen durchgeführt (vgl. Abschnitt 5.1.2). Die Ergebnisse sind in den Abbildungen 5.9 und 5.10 dargestellt. Zusätzlich stellt Abbildung 5.8 exemplarisch die Streuung der Datenpunkte im ROC-Diagramm anhand zweier Merkmalskombinationen dar.

Aus den Diagrammen lassen sich folgende Schlüsse ziehen:

- Generell lassen sich mit den untersuchten Merkmalskombinationen Korrekturklassifikationsraten von mehr als 90 % erreichen.
- Erwartungsgemäß sind die Ergebnisse bei einer Klassifikation auf dem Trainingsdatensatz (siehe Abbildung 5.6 und 5.7) etwas besser als im Fall disjunkter Trainings- und Testdaten (k -fache Kreuzvalidierung, siehe Abbildung 5.9 und 5.10). Praktisch relevant sind jedoch nur die Ergebnisse disjunkter Datensätze.
- Mittels einfacher Formmerkmale (Symbol \circ), wobei auch absolute Größenangaben der Objekte (wie Umfang, Fläche und Länge der Hauptachsen) enthalten sind, lassen sich die zu klassifizierenden Haustorien bereits recht gut von Störungen unterscheiden. Dieses Ergebnis deckt sich mit der beobachteten recht großen Trennschärfe dieser Merkmale (vgl. Abbildung 5.2 *oben*).
- Bezüglich der Flusser-Momentinvarianzen fällt auf, dass die Beschreibung auf Basis des Binärbildes (Symbol $+$) wesentlich bessere Ergebnisse liefert als das Pendant auf dem Sättigungsbild (Symbol \times). Die *Kombination* beider Merkmalsvektoren liefert jedoch sehr gute Ergebnisse (Symbol $*$).
- Die Frage, ob eine invariante Formbeschreibung der Objektkontur besser mittels Grund-Spektralinvarianzen oder direkt durch die Beträge der Fourierdeskriptoren erfolgen sollte, entscheidet sich anhand der ROC-Kurven recht eindeutig zugunsten einer direkten Nutzung der Spektralkoeffizienten.
- Die besten Klassifikationsergebnisse ergeben sich durch die Merkmalskombination
 - einfache Merkmale
 - + Flusser-Momentinvarianzen $\tilde{\psi}_1 \dots \tilde{\psi}_{11}$ aus dem Binärbild
 - \times Flusser-Momentinvarianzen $\tilde{\psi}_1 \dots \tilde{\psi}_{11}$ aus dem Sättigungsbild
 - \triangle Spektralkoeffizienten $[\tilde{Z}_{-14} \quad \tilde{Z}_{-13} \quad \dots \quad \tilde{Z}_{-1} \quad \tilde{Z}_2 \quad \dots \quad \tilde{Z}_{14}]$.

Diese Kombination ist in den Diagrammen mit dem Symbol \bullet gekennzeichnet.

5.7 Fazit

In den letzten beiden Kapiteln wurden Methoden zur objektbasierten Klassifikation auf Basis von Formmerkmalen zur Erkennung von Haustorien vorgestellt. Dies war notwendig, weil unter alleiniger Verwendung einer farbbasierten Objektsegmentierung keine sichere Erkennung der gesuchten Haustorien gewährleistet werden kann. Verfärbungen innerhalb der transformierten Zellen, welche eine ähnliche Farbcharakteristik wie Haustorien aufweisen, werden bei der Segmentierung ebenfalls markiert und müssen mittels einer nachgeschalteten Klassifikation von den tatsächlich gesuchten Objekten unterschieden werden. Dies erfolgt auf Basis eines Merkmalsvektors (bestehend aus den in Tabelle 5.1 angegebenen Merkmalen), welcher jedes segmentierte Objekt hinsichtlich seiner Formeigenschaften parametrisiert. Bei dieser Beschreibung wurde großer Wert auf eine möglichst leicht zu handhabende Verteilung der Merkmale im Merkmalsraum gelegt (im Idealfall sollte eine Gaußverteilung entstehen), um mit *einfachen* Klassifikationsmethoden auszukommen. Die dazu verwendeten nichtlinearen Transformationen zur Entzerrung der Verteilung der Merkmale im Merkmalsraum sind allgemein gültig und lassen sich auch auf andere Klassifikationsprobleme übertragen. Dieser Aspekt der Merkmalsentzerrung verdient nach Ansicht des Autors in der Literatur eine größere Beachtung, weil sich damit die Anforderungen an den nachgeschalteten Klassifikator reduzieren lassen.

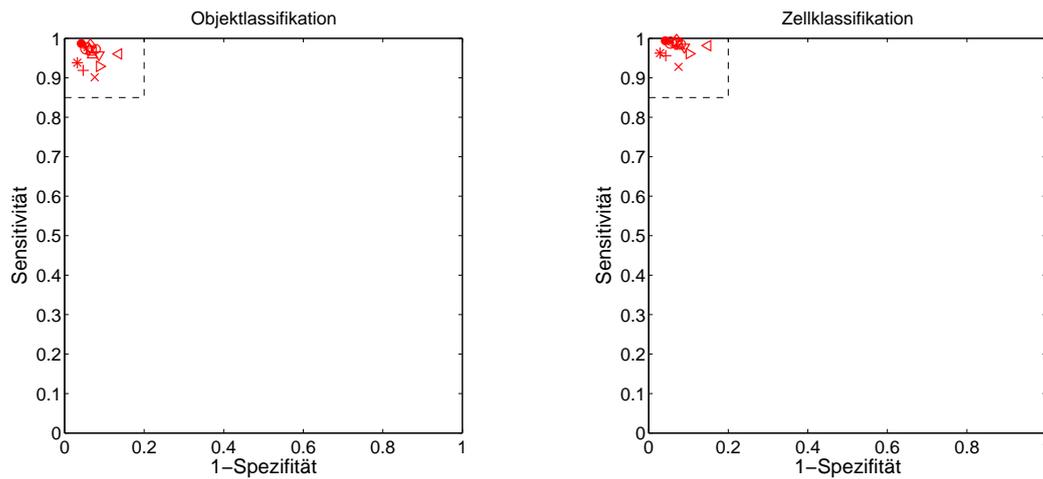


Abbildung 5.5: Gesamtansicht der ROC-Diagramme. Die Klassifikationsergebnisse befinden sich nahe des Optimums in der linken oberen Ecke. In den folgenden Diagrammen wird immer der gekennzeichnete Bildausschnitt dargestellt.

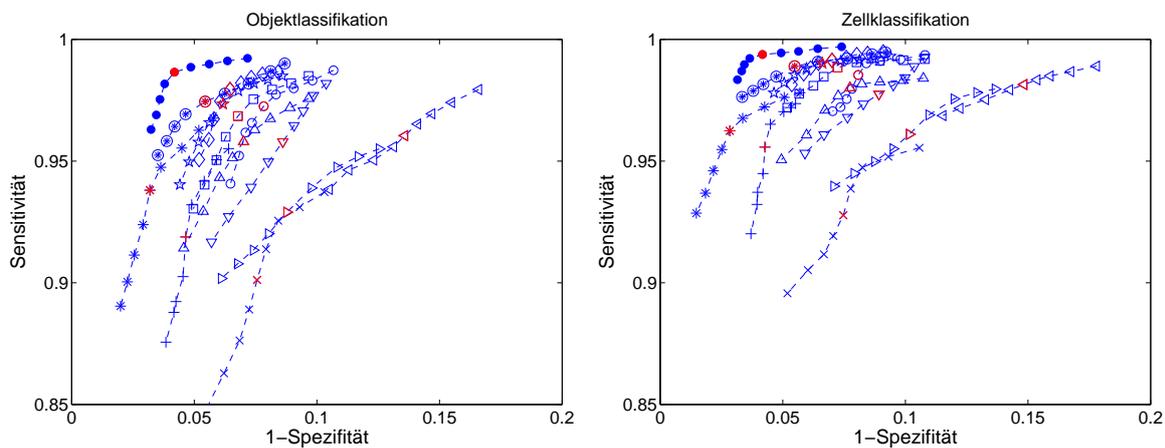


Abbildung 5.6: ROC-Diagramme der Objekt- (*links*) und Zellklassifikation (*rechts*) auf dem Trainingsdatensatz. Der optimale Arbeitspunkt des Klassifikators ist **rot** hervorgehoben.

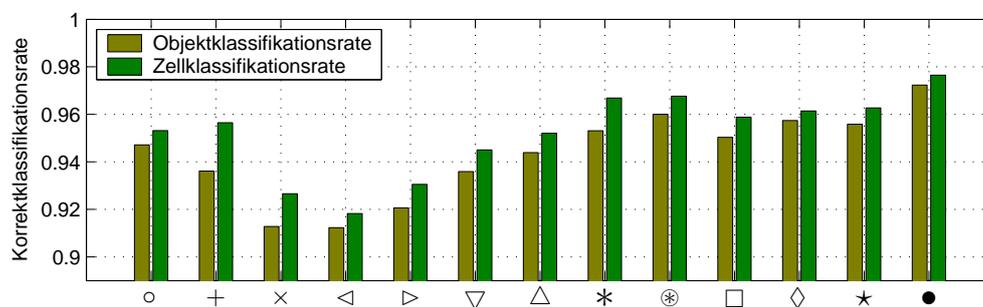


Abbildung 5.7: Korrektklassifikationsrate auf dem Trainingsdatensatz im optimalen Arbeitspunkt des Klassifikators.

Die Zuordnung der Symbole zu den verwendeten Merkmalsvektoren ist in Tabelle 5.2 auf Seite 85 angegeben.

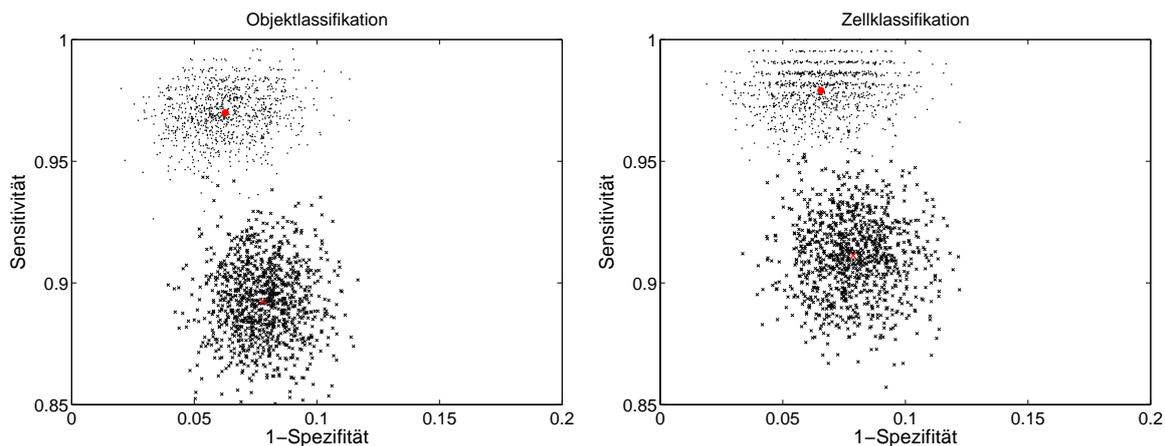


Abbildung 5.8: Verdeutlichung des Zufallscharakters der Klassifikationsergebnisse am Beispiel der Merkmalskombinationen \times und \bullet . In Abbildung 5.9 sind die *Mittelwerte* dieser Zufallsprozesse angegeben.

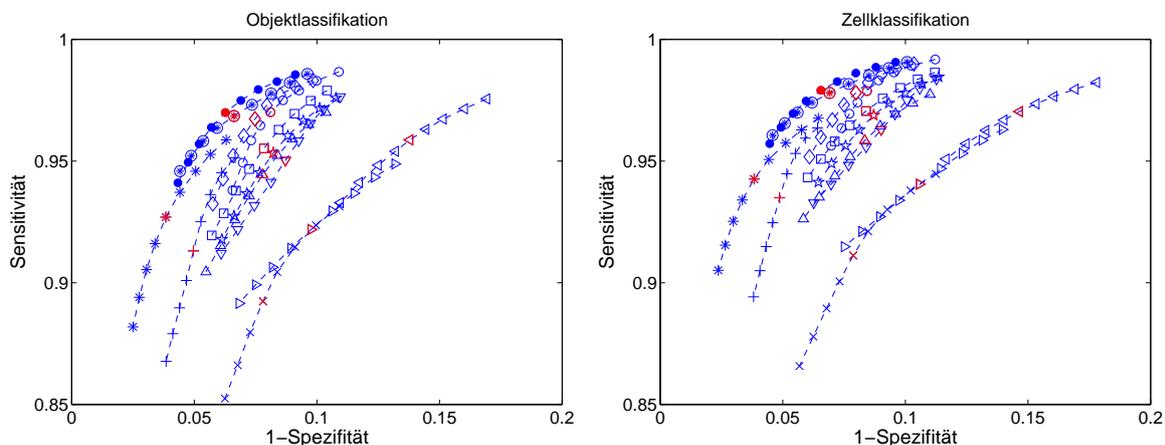


Abbildung 5.9: ROC-Diagramme der Objekt- (*links*) und Zellklassifikation (*rechts*) bei 2-facher Kreuzvalidierung mit 500 Wiederholungen. Der optimale Arbeitspunkt des Klassifikators ist **rot** hervorgehoben.

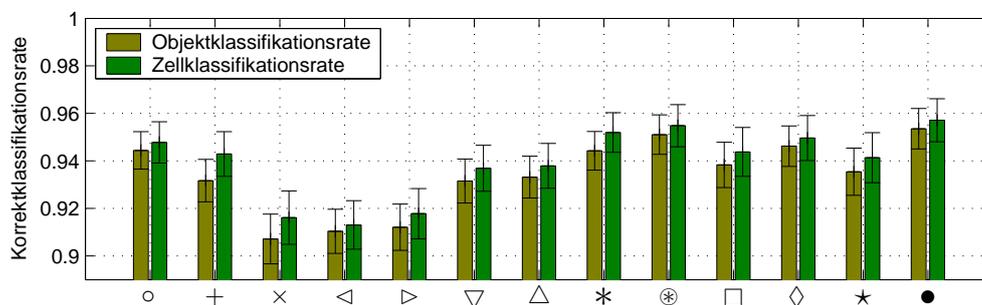


Abbildung 5.10: Korrektklassifikationsrate bei 2-facher Kreuzvalidierung mit 500 Wiederholungen im optimalen Arbeitspunkt des Klassifikators.

Die Zuordnung der Symbole zu den verwendeten Merkmalsvektoren ist in Tabelle 5.2 auf Seite 85 angegeben.

Im vorliegenden Fall konnte das Klassifikationsproblem mittels eines linearen statistischen Klassifikators gelöst werden. Die adäquate Methode ist eine Transformation des hochdimensionalen Merkmalsraums mittels linearer Diskriminanzanalyse (LDA) in einen eindimensionalen Eingaberaum ($d' = 1$). Ein höherdimensionaler Eingaberaum bringt keinen zusätzlichen Gewinn. Die auf die erste Diskriminanzachse projizierten Datenpunkte lassen sich direkt durch die Anwendung einer Schwelle mit dem Wert Null klassifizieren. So kondensiert das Klassifikationsproblem auf die Anwendung einer Entscheidungsschwelle im eindimensionalen Raum.

Mit der Klassifikation der in den transformierten Zellen detektierten Objekte ist die Verarbeitung der Mikroskopbilder abgeschlossen. Die Teilschritte der gesamten Bildverarbeitungskette sind in Abbildung 5.11 noch einmal zusammengefasst.

Auf den folgenden Seiten werden nun im dritten und letzten Teil der Arbeit die eingangs in Abschnitt 1.3 vorgestellten Vorüberlegungen zum Ablauf des Screenings aufgegriffen und das Zusammenspiel der beteiligten Komponenten erläutert.

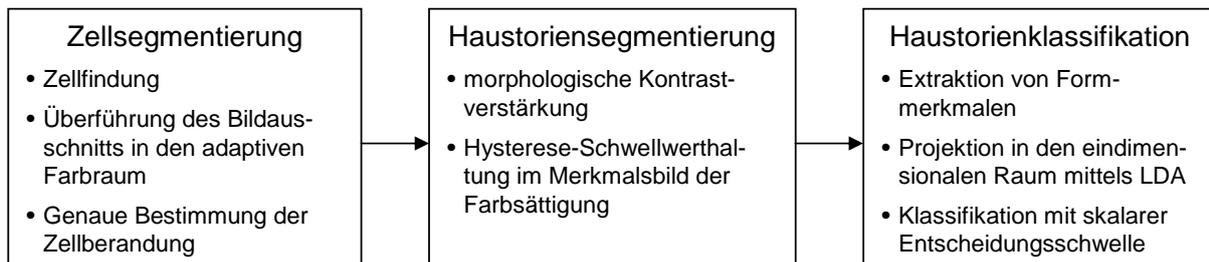


Abbildung 5.11: Bildverarbeitungskette zur Analyse der Mikroskopbilder auf Haustorien des Mehлтаupilzes in transformierten Zellen.

Teil III

Praktische Realisierung des Screeningsystems

Kapitel 6

Bildaufnahmesteuerung und Datenverwaltung

Eingangs wurden in Abschnitt 1.3 die Hardwarekomponenten des Screeningsystems vorgestellt und einige praktische Aspekte zur Bildaufnahme dargelegt. Eine wesentliche Vorüberlegung war, den Bildaufnahmeprozess aus Zeit- und Aufwandsgründen in zwei Stufen auszuführen:

1. Grobe Aufnahme des gesamten Objektträgers bei kleiner Vergrößerung (5×-Objektiv) in Form einer lückenlosen Kachelung mit geringer Bildauflösung (432×342 Pixel)
2. Detailaufnahmen von transformierten Zellen mit dem 10×-Objektiv unter Verwendung einer Bildauflösung von 1300×1030 Pixel.

Gegenstand dieses Kapitels sind sowohl die zwischen diesen beiden Teilprozessen liegenden Verarbeitungsschritte als auch die Darlegung des generellen Zusammenspiels der beteiligten Komponenten. Im Endergebnis steht ein System, bei dem der Nutzer lediglich die Objektträger in den X-Y-Tisch einlegen und das System starten muss.

Der grobe zeitlich-örtliche Ablauf eines Screeningdurchlaufs ist in Abbildung 6.1 dargestellt. Die beteiligten Komponenten sind dabei:

- das **Mikroskop**

Dies umfasst sowohl sämtliche Hardwarekomponenten des Mikroskops als auch einen PC, auf dem die Software „AxioVision“ zur Steuerung der motorischen bzw. elektronischen Komponenten des Mikroskops (X-Y-Tisch, Fokus, Beleuchtung, Objektive, Kamera, etc.) läuft. Des Weiteren findet an dieser Stelle die Mensch-Maschine-Interaktion statt: Der Nutzer legt zunächst die Proben in den X-Y-Tisch ein; anschließend parametriert und startet er das System mittels einer grafischen Nutzeroberfläche.

- der **Server**

Über eine Netzwerkverbindung findet die Kommunikation zwischen dem Server und dem PC des Mikroskops statt. Der Server stellt einerseits ein Netzlaufwerk mit ca. 1.3 TB Speicherkapazität (*File Server*) bereit, auf dem sämtliche Bilder abgelegt werden. Seine wichtigste Aufgabe ist jedoch die Bereitstellung der Bildauswertung in Form von automatischen Diensten (*Compute Server*). Diese vereinen die in dieser Arbeit vorgestellten Algorithmen und Methoden in Form von zwei Software-Paketen, welche Folgendes leisten:

- Auswertung der Vorschaubilder und Bereitstellung der Steuerinformation für die anschließenden Detailaufnahmen
- Auswertung der Detailbilder und Bereitstellung der Ergebnisse zum Haustorienbefall der untersuchten transformierten Zellen.

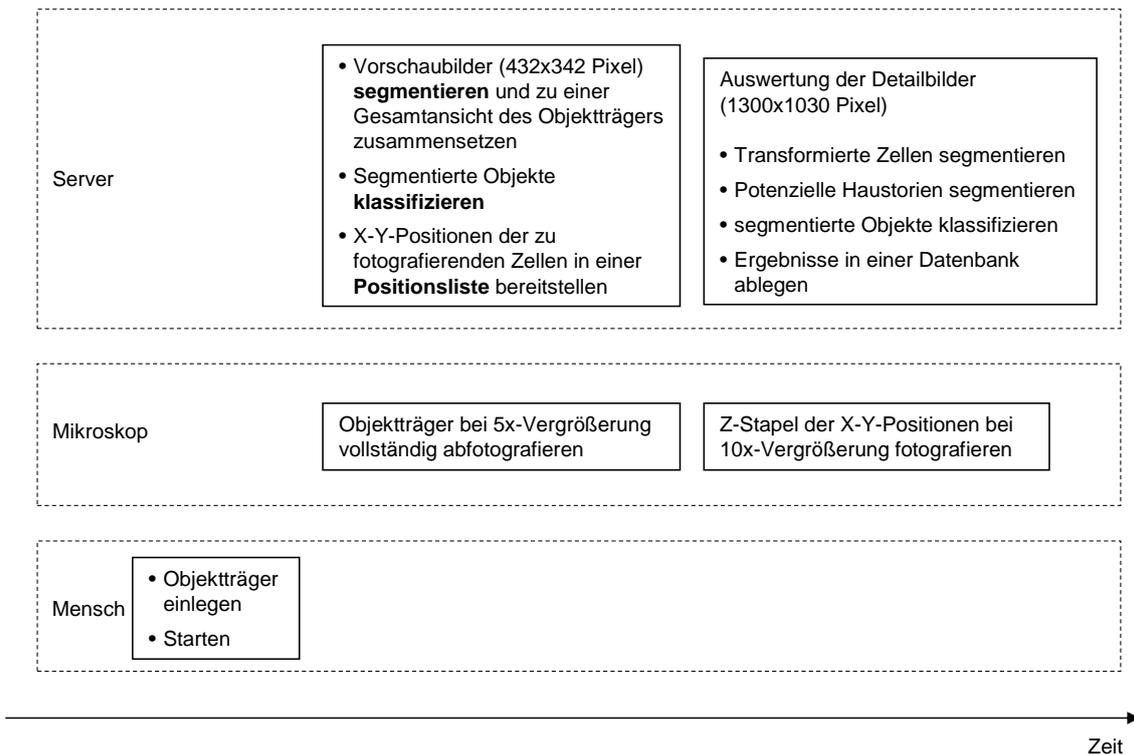


Abbildung 6.1: Zeitlich-örtlicher Ablauf des Bildaufnahme- und Analyseprozesses. Die Einteilung der Zeitachse ist nicht maßstabsgerecht, der tatsächliche Zeitbedarf ist in Tabelle 1.2 dargelegt.

Auf den nächsten Seiten wird zunächst die Verarbeitung der Vorschaubilder vorgestellt, wobei auf bereits erprobte Methoden aus Kapitel 3 und 4 zurückgegriffen werden kann. Anschließend wird auf Basis der in den Vorschaubildern detektierten „interessanten“ Objekte eine Liste der im Detail zu fotografierenden Bereiche erstellt. Diese Liste ist hinsichtlich zweier Kriterien zu optimieren: Zum einen soll versucht werden, mit möglichst wenig im Detail zu fotografierenden Regionen auszukommen, zum anderen ist die vom Tisch zurückgelegte Wegstrecke zu minimieren. Beide Probleme lassen sich entkoppelt voneinander behandeln. Im Abschluss des Kapitels wird der Informationsfluss zwischen den Systemkomponenten erklärt sowie die Speicherung und Präsentation der Analyseergebnisse diskutiert.

6.1 Auswertung der Vorschaubilder

Aus Gründen einer zeiteffizienten Bildaufnahme (es soll möglichst kein Zeitverzug zwischen den Vorschauaufnahmen und den Detailaufnahmen entstehen) wird die Auswertung der Vorschaubilder – soweit möglich – parallel zur Bildaufnahme ausgeführt (*pipelining*). Da dabei die zu findenden transformierten Zellen über die Einzelbildgrenzen hinausragen können, sind diesem Vorgehen jedoch Grenzen gesetzt: Spätestens die Klassifikation der detektierten Objekte kann nicht mehr auf Basis der einzelnen Bildkacheln erfolgen, sondern bedarf einer Gesamtansicht des Objektträgers. Diese Gesamtansicht kann jedoch erst fertig gestellt werden, wenn das letzte Einzelbild vorliegt.

Segmentierung

Aus Kapitel 3 ist bekannt, dass sich die Färbung der transformierten Zellen sehr gut im Buntton (*Hue*) widerspiegelt und damit eine aufwandsgünstige Segmentierung durchgeführt werden kann.

Um die Einzelbilder parallel zur laufenden Bildaufnahme zu segmentieren, kommen nur einfache, pixelbasierte Verfahren wie eine globale Schwellwerthaltung in Frage, weil die gesuchten Objekte über den Bildausschnitt hinausragen können. Eine Kantendetektion ist damit beispielsweise ausgeschlossen. Die Anforderungen an die Qualität der Segmentierung sind jedoch – verglichen mit jenen bei der Auswertung der Detailbilder – wesentlich geringer, so dass daraus kein Nachteil entsteht. Beispielsweise ist *keine* exakte Bestimmung der Zellberandung notwendig, vielmehr genügt eine Segmentierung, welche die tatsächliche Objektform näherungsweise wiedergibt. In Abbildung 6.2 ist die Segmentierung dargestellt: Zunächst wird das RGB-Bild mit einem Gaußfilter der Breite $\sigma \approx 3$ geglättet, um Störungen zu unterdrücken. Anschließend werden zwei Schwellwerte auf den Buntton angewendet. Ausgehend von Regionen, die über den *oberen* Schwellwert hinausragen, wird ein Bereichswachstum durchgeführt, bis der *untere* Schwellwert unterschritten wird (Hysterese-Schwellwerthaltung). Gegenüber einer einfachen Schwellwerthaltung erreicht dieses Verfahren deutlich robustere Ergebnisse.

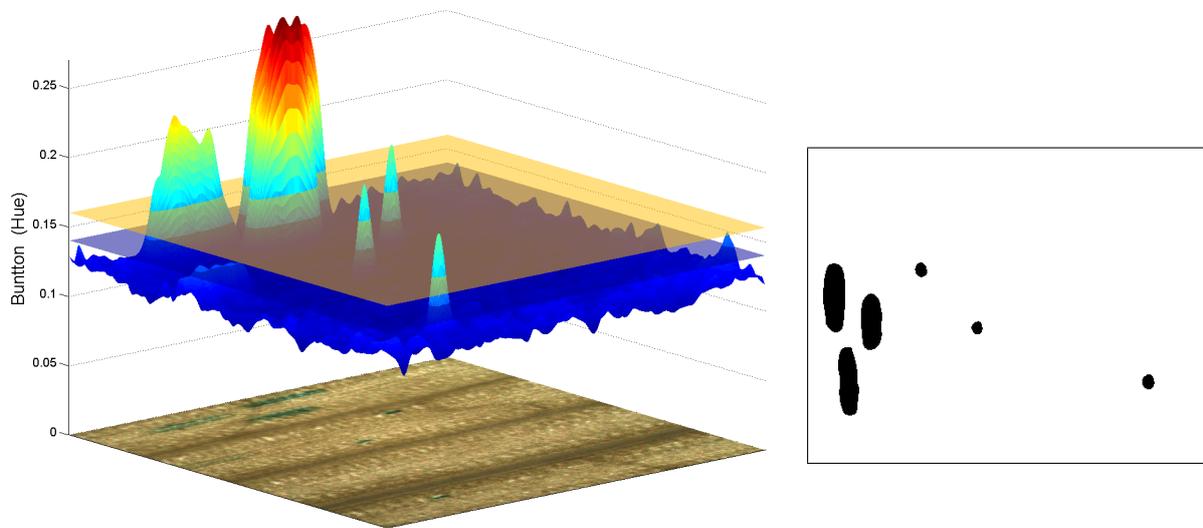


Abbildung 6.2: Segmentierung der Vorschaubilder durch Anwendung von zwei Schwellwerten auf den Buntton des geglätteten Bildes (Glättung in RGB mit einem Gaußtiepass, $\sigma \approx 3$). Es werden alle Regionen zwischen den Schwellwerten segmentiert, wenn sie mit einer Region oberhalb des oberen Schwellwertes verbunden sind (Hysterese-Schwellwerthaltung). Im entstehenden Binärbild (*rechts*) sind deutlich die gefärbten, transformierten Zellen markiert. Diese Segmentierung wird nacheinander für alle Vorschaubilder durchgeführt, anschließend werden die Bilder zu einer großen Gesamtansicht des Objektträgers zusammengesetzt (vgl. Abbildung 6.3).

Erzeugung einer Gesamtansicht des Objektträgers

Diese Segmentierung wird nacheinander auf allen Vorschaubildern ausgeführt, sobald das Eingangsbild auf dem Server zur Verarbeitung bereitsteht. Die Segmentierungsergebnisse werden dabei schrittweise (Bild für Bild) zu einer Gesamtansicht des Objektträgers zusammengesetzt. Das resultierende Binärbild besteht aus 15×38 Einzelbildern (vgl. Abbildung 1.5) mit einer Auflösung von jeweils 432×342 Pixel und hat damit eine beachtliche Größe von 6480×12996 Pixel.

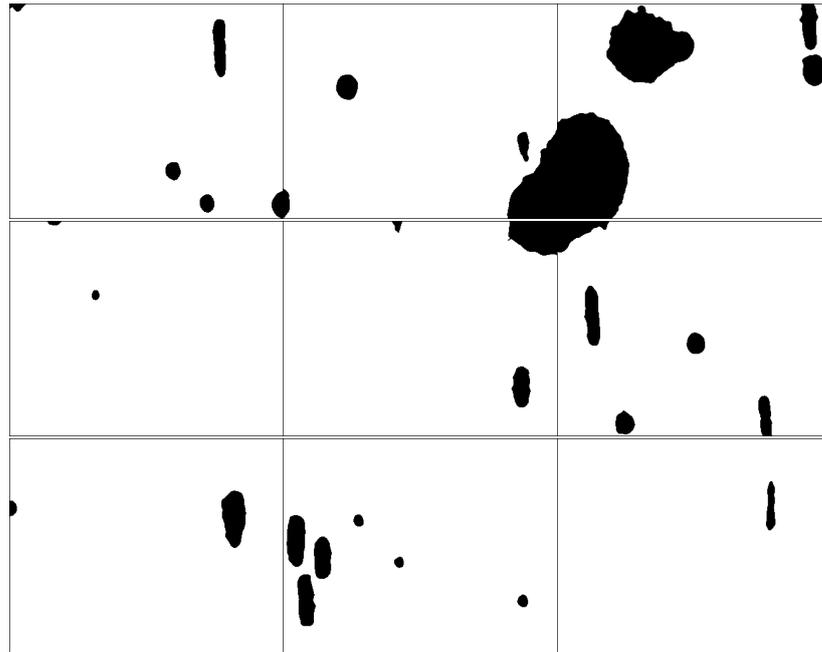


Abbildung 6.3: Segmentierte Vorschaubilder aus Abbildung 1.6. Gefärbte Bereiche werden durch die Hysterese-Schwellwerthaltung im Buntton sicher segmentiert. Die Genauigkeit der Segmentierung ist für eine anschließende Klassifikation der Objekte ausreichend.

Objektklassifikation

Im nächsten Schritt werden auf Basis dieser Gesamtansicht des Objektträgers die darauf segmentierten Objekte *klassifiziert*. Dieser Schritt ist notwendig, weil nicht alle gefärbten Bereiche *auswertbare* Zellen markieren. Bei einem Blick zurück auf Abbildung 1.6 (in Abschnitt 1.3 auf Seite 8) fällt auf, dass die großen „Flecken“ in Abbildung 6.3 zu Regionen gehören, in denen der Indigofarbstoff über weite Bereiche der Blattoberfläche diffundiert ist (ausgelaufener Farbstoff). Solche Regionen werden keiner Analyse unterzogen. Ebenso müssen die transformierten Zellen bestimmte Größenkriterien erfüllen, um auswertbar zu sein. So sind einige gefärbte Zellen erkennbar, die auffällig klein sind. Diese sind Schließzellen, welche sich an den Spaltöffnungen befinden. Sie werden ebenfalls nicht ausgewertet. Ein weiteres Kriterium für den Ausschluss einer Zelle ist eine zu dunkle Färbung. In Tabelle 6.1 sind die für die Klassifikation verwendeten Parameter zusammengefasst.

Merkmals	Minimumwert	Maximumwert
Objektlänge	30 Pixel	170 Pixel
Objektbreite	10 Pixel	40 Pixel
Mittlere Helligkeit (Skala 0...1)	0.2	0.6

Tabelle 6.1: Parameter zur Klassifikation der segmentierten Zellen. Die Größenangaben beziehen sich auf die Auflösung der Vorschaubilder (432×342 Pixel).

6.2 Festlegung der Aufnahmepositionen für die Detailbilder

Nach der Segmentierung der Vorschaubilder und der Klassifikation der detektierten Objekte liegt eine Abbildung des gesamten Objektträgers in Form eines großes Binärbildes (6480×12996

Pixel) vor, in dem die zu untersuchenden transformierten Zellen markiert sind.

Diese markierten Objekte sind im nächsten Schritt mit dem 10×-Objektiv hochauflösend zu fotografieren. Das Mikroskop benötigt dazu die entsprechende Steuerinformation in Form der vom X-Y-Tisch anzufahrenden Koordinaten, an welchen eine Bildaufnahme erfolgen soll. Dabei ergeben sich zwei wesentliche Aspekte zu Optimierung dieses Prozesses:

1. Aus Gründen einer zeit- und speicherplatzeffizienten Aufnahme der Detailbilder sollte Wert darauf gelegt werden, benachbarte Zellen, die gemeinsam auf ein Bild passen, auch entsprechend gemeinsam zu fotografieren. Zur Untermauerung des Zeit- und Speicherplatzaspekts sei an dieser Stelle nochmals auf die Tabellen 1.2 und 1.3 verwiesen. Demnach verursacht jede unnötige Aufnahmeposition durch die nötigen Schritte
 - Tischpositionierung
 - Autofokus
 - Aufnahme des Bildstapels

einen Zeitverlust von bis zu 15 Sekunden und belegt ca. 8.5 MB zusätzlichen Speicherplatz.

2. Die insgesamt vom X-Y-Tisch zurückzulegende Wegstrecke wird durch die *Reihenfolge* der Aufnahmepositionen bestimmt. Eine Minimierung dieser Wegstrecke kommt nicht nur einer Zeitersparnis, sondern auch der Schonung der mechanischen Teile zugute. Außerdem verringert sich dadurch die Geräuschentwicklung der Tischmotoren, was zu einer geringeren Belästigung von im Labor anwesenden Personen führt.

In den nächsten Abschnitten werden entsprechende Verfahren zur Optimierung dieser beiden Aspekte der Bildaufnahme diskutiert.

6.2.1 Minimierung der Anzahl der nötigen Aufnahmepositionen

Anhand der geometrischen Ausdehnung der einzelnen Objekte lässt sich leicht bestimmen, welche Zellen gemeinsam auf ein Bild passen. Um den Objektmittelpunkt herum ergibt sich eine Maske, um die das Objekt bzw. das Sichtfenster verschoben werden kann, ohne dass Teile den Sichtbereich verlassen. Liegt der Mittelpunkt des Sichtfensters im Überlappungsbereich dieser Verschiebungsmasken, so werden die zugehörigen Objekte vollständig abgebildet. Dies wird in Abbildung 6.4 schematisch dargestellt.

Auf Basis dieser einfachen geometrischen Betrachtungen lässt sich das Problem graphentheoretisch [Jun05] formulieren: Gegeben sei ein ungerichteter Graph ohne Mehrfachkanten¹ $G = (V, E)$, bestehend aus den Knoten (*Vertices*) V und Kanten (*Edges*) E . Dabei sind die Knoten V des Graphen die zu fotografierenden Objekte (die transformierten Zellen), und zwischen den Knoten existiert dann eine Kante, wenn die beiden Objekte gemeinsam auf ein Bild passen.

Die Lösung der Aufgabe

- „Fotografiere *alle* Objekte mit einer *minimalen* Anzahl von Bildern“

erfordert eine geschickte Clusterung des Graphen. Alle *vollständigen*² Graphen stellen dabei mögliche Konstellationen für eine Bildaufnahme dar.

Von FOWLER *et al.* [FPT81] wurde gezeigt, dass diese Optimierungsaufgabe zur Komplexitätsklasse der NP-vollständigen Probleme gehört. Damit wächst der für das Finden der Optimallösung erforderliche Zeitaufwand in Abhängigkeit von der Problemgröße mehr als polynomial.

¹Ungerichtete Graphen ohne Mehrfachkanten nennt man auch häufig *schlicht* oder *einfach*.

²Ein vollständiger Graph ist ein ungerichteter Graph ohne Mehrfachkanten, in dem jeder Knoten mit jedem anderen Knoten durch eine Kante verbunden ist.

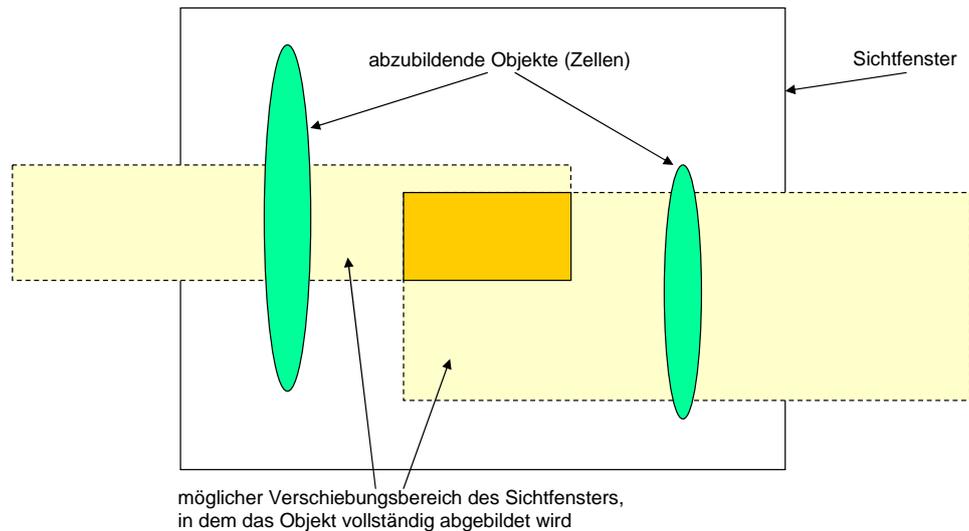


Abbildung 6.4: Voraussetzungen zum gemeinsamen Fotografieren mehrerer Objekte. Die zwei Zellen passen im Überlappungsbereich der zwei Verschiebungsmasken vollständig gemeinsam auf das Bild.

In der Praxis ist der entstehende Graph nur sehr schwach verbunden, da der Überdeckungsbereich aufgrund des stark eingeschränkten Sichtfeldes der Kamera sehr klein ist. Zur Lösung wird ein unüberwachtes Clustering auf Basis eines Greedy-Verfahrens (*greedy algorithm*) [Sch06] eingesetzt.

6.2.2 Minimierung der vom X-Y-Tisch zurückzulegenden Wegstrecke

Das Problem der Minimierung einer Wegstrecke ist als recht populäres und alltägliches Problem unter dem Begriff „Problem des Handlungsreisenden“ (*Traveling Salesman Problem, TSP*) [JM97] bekannt. Es gehört zur Klasse der NP-vollständigen Probleme und ist damit in der Praxis für eine große Anzahl von Positionen nach heutigem Kenntnisstand nicht exakt lösbar.

Das Finden der Optimallösung ist jedoch vor dem Hintergrund der Aufgabenstellung nicht notwendig. Vielmehr genügt an dieser Stelle der Einsatz von Heuristiken, mit denen sich das Problem in wenigen Sekunden näherungsweise lösen lässt. Der Einsatz *schneller* Algorithmen ist an dieser Stelle wichtiger als das Finden einer besseren Lösung, um die durch die Wegstreckenverringerung eingesparte Zeit nicht wieder durch das „Warten auf die optimierte Positionsliste“ zu verlieren. Im vorliegenden Fall wurde das Problem mit Hilfe eines Genetischen Algorithmus (GA) [Gol04] gelöst, welcher eine hinreichend optimierte Permutation der Positionsliste in weniger als 5 Sekunden liefert [Sch06] (unter der realistischen Annahme von etwa 150 Aufnahmepositionen auf dem Objektträger).

6.3 Austausch der Steuer- und Nutzinformation

Die *Bildaufnahme* und *-auswertung* (sowohl die Auswertung der Vorschaubilder als auch die der Detailbilder) finden an räumlich getrennten Plätzen auf verschiedenen Hardwarekomponenten statt (siehe Abbildung 6.1). Der Datenaustausch zwischen dem Steuer-PC des Mikroskops und dem File- und Compute-Server erfolgt auf Dateiebene über ein im Netzwerk freigegebenes Verzeichnis des Servers (*shared directory*). In diesem Verzeichnis werden alle experimentbezogenen Daten strukturiert abgelegt. Bei diesen Daten ist zwischen vorübergehenden (temporären) Daten und dauerhaften (persistenten) Informationen zu unterscheiden. Zu den persistenten Daten gehören:

- biologische Parameter des Experiments
- Vorschaubilder
- Detailbilder
- generierte Positionsliste für die Detailaufnahmen
- Protokoll sämtlicher Steuerbefehle des Mikroskops (*log file*).

Über temporäre Dateien werden einerseits Steuerbefehle zwischen den Komponenten ausgetauscht, andererseits wird auf diese Weise der aktuelle Status der Auswertelgorithmen an die grafische Nutzerschnittstelle (*Graphical User Interface, GUI*) übermittelt.

Die Nutzerschnittstelle befindet sich auf dem Steuer-PC des Mikroskops. Hier gibt der Experimentator die Parameter des Experiments an und startet nach dem Einlegen der Objektträger in den X-Y-Tisch das Screening. Ein Bildschirmfoto der Nutzerschnittstelle ist in Abbildung 6.5 zu sehen.

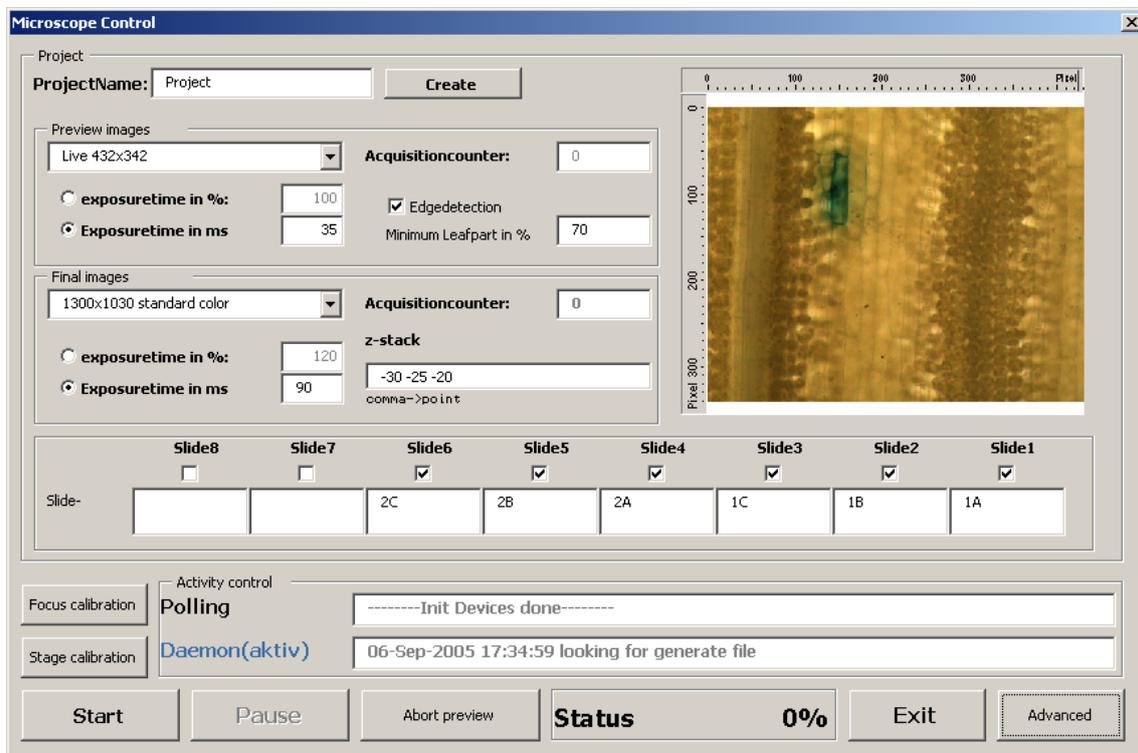


Abbildung 6.5: Grafische Nutzerschnittstelle (*Graphical User Interface, GUI*) zur Parametrierung und Überwachung des Screenings. Dieses Modul läuft auf dem Steuer-PC des Mikroskops und wurde mit Hilfe der Scriptingmöglichkeiten der Mikroskop-Steuersoftware „AxioVision“ mittels *Visual Basic for Applications (VBA)* implementiert.

6.4 Informationsspeicherung und -zugriff

Die Speicherung der Analyseergebnisse in Form von Dateien (*flat files*) ist vor dem Hintergrund einer komfortablen und flexiblen Datenauswertung nur eingeschränkt empfehlenswert. Wesentlich geschickter ist die Archivierung aller Informationen in strukturierter Form in einer Datenbank. Damit lassen sich die Ergebnisse unter verschiedensten Gesichtspunkten mit Hilfe

einfach zu formulierender Datenbankabfragen auswerten. Diese Datenbanklösung befindet sich momentan noch in der Entwicklung. Das Entity-Relationship-Diagramm ist im Anhang G auf Seite 127 zu finden.

Kapitel 7

Zusammenfassung

Diese Arbeit wurde aus einer biologischen Fragestellung motiviert, bei der nach genetisch bedingten Ursachen natürlicher Abwehrmechanismen von Kulturpflanzen gegen Schädlingsbefall gefragt wird. Konkret handelt es sich dabei um Untersuchungen zum Resistenzverhalten von Kulturgerste (*Hordeum vulgare*) gegen den echten Mehltau (*Blumeria graminis f. sp. hordei*). Vor dem Hintergrund dieser Fragestellung sind dazu im Rahmen einer großen Serie von Experimenten sehr aufwändige Screeningaufgaben zu bewältigen, welche das Potenzial einer manuellen Begutachtung bei weitem übersteigen. Deshalb wurde die Entwicklung eines auf diese Fragestellung zugeschnittenen, vollautomatischen Hochdurchsatz-Screeningsystems im Rahmen eines vom BMBF geförderten Bioinformatik-Projekts verfolgt.

Die vorliegende Dissertation beschreibt die ingenieurtechnische Realisierung dieses Systems. Ziel war die vollautomatische Auswertung von präparierten, auf Objektträgern vorliegenden Gerstenblättern. Bei den Prüfobjekten handelt es sich um einzelne gefärbte, genetisch transformierte Zellen der oberen Blattepidermis, welche auf das Vorhandensein von **Haustorien** des Mehltaupilzes zu untersuchen sind. Über die statistische Häufigkeit des haustorialen Befalls, angegeben durch den Anfälligkeitsindex (*susceptibility index*), lassen sich leicht Rückschlüsse auf die Funktion des Abwehrverhaltens der Pflanze gegen Eindringversuche des Pilzes ziehen.

Die für die Realisierung zur Verfügung stehende Hardwarebasis umfasste ein per Computer steuerbares Lichtmikroskop, ausgestattet mit CCD-Kamera und motorischem X-Y-Tisch mit Platz für bis zu acht Objektträger. Seine Funktionen werden gemeinsam mit den Eigenschaften des Probenmaterials im Einleitungsteil der Arbeit in Kapitel 1 vorgestellt. Da dem Merkmal „Farbe“ zur Auswertung der Prüfobjekte eine sehr große Bedeutung zukommt und ein bis dato manuelles Auswertungsverfahren automatisiert werden soll, beginnt die Arbeit in Kapitel 2 mit einigen grundlegenden Betrachtungen zur menschlichen Farbwahrnehmung und der Repräsentation von Farbinformation in technischen Systemen. Unterstützend sind dazu einige Farbraumtransformationen im Anhang A angegeben.

Kapitel 3 löst die zentralen Probleme zur Segmentierung des zu untersuchenden Bildmaterials. Dazu wurde zunächst ein adaptiver Farbraum auf Basis eines parametrischen statistischen Modells hergeleitet, der auf die Färbung transformierter Zellen und des Blatthintergrundes abgestimmt ist. Dieser adaptive Farbraum stellt ein einkanaliges Bild mit maximaler Trennschärfe zwischen den beiden zu unterscheidenden Klassen bereit, auf das alle bekannten Methoden der klassischen Graubildverarbeitung angewendet werden können. Da die Schätzung der Modellparameter mittels des Expectation Maximization (EM)-Algorithmus sehr rechenaufwändig ist, wurde zur Zellsegmentierung ein zweistufiges Verfahren konzipiert: Zunächst wird mit dem Canny-Kantendetektor eine grobe Segmentierung auf Basis des Bunntons (*Hue*) durchgeführt (Finden der *Region of Interest, ROI*). Die gesuchten Zellen werden dabei mit einer geschlossenen Kontur markiert. Zur Verfeinerung der Segmentierung werden die markierten rechteckigen Bildausschnitte im zweiten Schritt in den adaptiven Farbraum überführt. Unter Nutzung eines aktiven Kontur-

modells lässt sich nun die Berandung der Zelle genau bestimmen. An dieser Stelle stellte sich heraus, dass die den aktiven Konturmodellen zugrunde liegende Differentialgleichung aufgrund ihrer zyklischen Struktur effizient unter Nutzung der schnellen Fouriertransformation (*Fast Fourier Transform, FFT*) gelöst werden kann. Diese Erkenntnis ist von universeller Bedeutung und ermöglicht den Einsatz aktiver Konturmodelle auch für Hochdurchsatz- und Echtzeitanwendungen, da sich auf diese Weise die numerische Komplexität von $\mathcal{O}(N^2)$ auf $\mathcal{O}(N \log N)$ reduziert. Anhang D zeigt unterstützend dazu praktische Messergebnisse zum Laufzeitverhalten verschiedener Lösungsstrategien.

Die im Anschluss an die Zellsegmentierung durchzuführende Detektion der Haustorien gestaltete sich als besonders herausfordernd. Zur Segmentierung wird zunächst ihr sehr geringer Farbkontrast zum Zellhintergrund mittels einer inhaltsadaptiven Kontrastverstärkung auf Basis morphologischer Zylinderhutoperationen (*morphological top-hats*) erhöht. Anschließend ist eine robuste Binarisierung mittels zweier Schwellwerte (Hysterese-Schwellwerthaltung) im Farbsättigungskanal durchzuführen. Aufgrund des sehr geringen Farbkontrasts zum Zellhintergrund sowie des Auftretens diverser Störungen innerhalb von Zellen (Verfärbungen) ist eine *sichere* Haustoriendetektion auf alleiniger Basis von Farbinformation jedoch nicht zu gewährleisten.

Deswegen wurde zur Überprüfung der segmentierten potenziellen Haustorienregionen eine objektbasierte Klassifikation nachgeschaltet. Kapitel 4 stellt die dazu notwendigen Formmerkmale bereit und deckt dabei ein breites Spektrum (von einfachen Merkmalen über Momentinvarianzen bis hin zu Fourierdeskriptoren und daraus abgeleiteten Charakteristiken) ab. Besonderer Wert wurde bei der anschließenden Klassifikation in Kapitel 5 auf die Anwendbarkeit eines möglichst *einfachen* Klassifikationsverfahrens gelegt. Dazu wurde die Verteilung jedes einzelnen Merkmals im Merkmalsraum gründlich untersucht und diese mittels angepasster nichtlinearer Transformationen derart entzerrt, dass sie dem Ideal einer Normalverteilung möglichst nahe kommt. Das gesteckte Ziel wurde erreicht – im Ergebnis lässt sich der hochdimensionale Merkmalsraum nahezu perfekt mittels einer linearen Diskriminanzanalyse (LDA) in den eindimensionalen Raum abbilden. Die Klassifikation kondensiert damit zu einer einfachen Anwendung einer skalaren Entscheidungsschwelle.

Mit diesen Methoden ist die automatische Analyse der zu untersuchenden Mikroskopbilder abgedeckt. Kapitel 6 spannt schließlich den Bogen zurück zu der in der Einleitung vorgestellten Mikroskophardware und widmet sich dem Problem der automatischen Bildaufnahmesteuerung sowie der Verwaltung, Dokumentation und Präsentation der Roh- und Ergebnisdaten. Die mikroskopseitige Steuerung konnte mit vertretbarem Aufwand über die Scriptingfähigkeiten der Mikroskop-Steuersoftware „AxioVision“ (mittels der optional erhältlichen Programmierschnittstelle *Visual Basic for Applications, VBA*) realisiert werden. Das Steuerprogramm ist gleichzeitig die grafische Nutzerschnittstelle (*Graphical User Interface, GUI*), mit der der Experimentator das Screening parametrisiert und steuert. Die Steuersoftware kommuniziert über eine Dateischnittstelle mit den entwickelten Bildauswertungsalgorithmen, welche auf einem dedizierten *File-* und *Compute Server* bereitgestellt werden und das Bildmaterial auf Anforderung vollautomatisch auswerten.

Im Ergebnis läuft die Durchführung eines Screenings folgendermaßen ab: Der Experimentator legt bis zu acht Objektträger in die vorgesehenen Halterungen des X-Y-Tisches ein, gibt die das Experiment beschreibenden Parameter in der grafischen Nutzerschnittstelle ein und startet mittels Knopfdruck das Screening. Damit beginnt ein zweistufiger Bildaufnahmeprozess: In der ersten Stufe werden alle Objektträger lückenlos mäanderförmig abgefahren und bei kleiner Vergrößerung fotografiert. In den entstehenden Vorschaubildern wird parallel dazu eine pixelbasierte Segmentierung durchgeführt, mit der gefärbte Bereiche (transformierte Zellen) detektiert werden. So entsteht Bild für Bild eine Binäransicht des Objektträgers, in der gefärbte Regionen markiert sind. Mittels einer Objektklassifikation werden anschließend interessante Objekte (transformierte Zellen bestimmter Größe) herausgesucht, für die eine detaillierte Betrachtung

infrage kommt. Aus diesen Informationen wird eine Liste von x - und y -Positionen (im Koordinatensystem des X-Y-Tisches) generiert, so dass in der zweiten Stufe alle gewünschten Objekte mit einer minimalen Anzahl von Bildaufnahmen hochauflösend fotografiert werden. Die entstehenden Detailbilder werden anschließend mit den in dieser Arbeit entwickelten Analysealgorithmen prozessiert und die Ergebnisse zum „haustorialen Befund“ jeder untersuchten Zelle in einer Datenbank abgelegt. Zusammen mit den in dieser Datenbank ebenfalls hinterlegten Parametern zur durchgeführten genetischen Transformation der Zellen (Gen-Konstrukte) ist damit eine sehr komfortable und universelle Auswertung der Experimente möglich. Diese Datenbanklösung ist momentan noch in der Entwicklung. Dem Experimentator steht somit ein Werkzeug zur Verfügung, welches bisher manuell durchgeführte, sehr zeitaufwändige Auswertungen deutlich vereinfacht.

Mit der erreichten Korrekturklassifikationsrate von über 90 % handelt es sich um ein Laborassistenzsystem, welches sich prinzipiell völlig autark betreiben lässt. Das Gros der Zellen wird automatisch mit hinreichender Genauigkeit¹ in „*Haustorien beinhaltend*“ und „*keine Haustorien beinhaltend*“ klassifiziert. Während der Einführungsphase in den täglichen Laborbetrieb sollte jedoch eine (zumindest stichprobenartige) manuelle Nachkontrolle erfolgen, um die Erkennungssicherheit im praktischen Einsatz zu verifizieren. Dies gilt insbesondere bei Klassifikationen nahe der Entscheidungsschwelle. Unsichere Fälle können vom Experimentator durch Betrachtung der betreffenden Bilddaten am Monitor binnen weniger Sekunden überprüft werden. Diese Möglichkeit der Kontrolle ist mit dem entwickelten System jederzeit gegeben, so dass auch bei einer möglicherweise in Einzelfällen geforderten „absoluten Genauigkeit“ eine zusätzliche manuelle Klassifikation am Bildschirm erfolgen kann. In jedem Fall entfällt für den Experimentator das anstrengende und ermüdende stundenlange Sitzen am Mikroskop. Für den Laborbetrieb bedeutet dies nicht nur eine drastische Zeitersparnis, sondern auch eine deutliche Verbesserung der Arbeitsbedingungen. Letztendlich werden durch den Einsatz dieses Systems Arbeiten an neuen biologischen Fragestellungen überhaupt erst realistisch durchführbar.

Verbesserung des Screeningsystems

Vor dem Hintergrund einer Hochdurchsatz-Anwendung des Systems spielt der zur Abarbeitung eines Screenings benötigte Zeitaufwand eine wichtige Rolle. Limitierender Faktor ist dabei die Aufnahme der hochaufgelösten Detailbilder. Im Rückblick auf Tabelle 1.2 fällt auf, dass eine verbesserte Bildaufnahmestrategie bei der Aufnahme des Z-Stapels eine erhebliche Zeiteinsparung mit sich bringen würde. Aufgrund des gewissen zufälligen Fokussierungsfehlers der in der Steuersoftware „AxioVision“ enthaltenen Autofokus-Algorithmen sowie der Notwendigkeit der Erweiterung der Schärfentiefe, um die hügelige Blattoberfläche vollständig abzubilden, müssen bislang jeweils fünf Bilder je Position eingezogen werden. Eine Implementierung eigener, inhaltsadaptiver Fokussierungsalgorithmen könnte den Zeitaufwand für die Bildaufnahme drastisch reduzieren. Dazu wären die zu fotografierenden transformierten Zellen in der Steuersoftware des Mikroskops zu detektieren und eine Fokussteuerung in Abhängigkeit der Lage der detektierten Objekte vorzunehmen. Im Idealfall ließe sich damit die Bildaufnahme mit einem *einzigem* Bild erledigen. Dieser Ansatz wird auf jeden Fall im weiteren Verlauf des Projekts verfolgt werden.

Vor dem Hintergrund einer platzsparenden Archivierung der Bilder lässt sich eine weitere Optimierung des Systems durch eine geeignete Kompression des Bildmaterials realisieren. Bisher wird dazu das verlustfrei komprimierende PNG-Format verwendet (vgl. Tabelle 1.3). SEIFFERT hat ein Verfahren vorgeschlagen [Sei04], welches unterschiedliche Bildbereiche vor der Kompression klassifiziert und anschließend mit jeweils speziell angepassten Kompressionsverfahren behandelt. Das hier vorliegende Bildmaterial, welches sich ständig wiederholende Bildinhalte aufweist, ist

¹Vor dem Hintergrund vieler Unwägbarkeiten im Laborablauf bei der Vorbereitung der Proben ist eine Korrekturklassifikationsrate von mehr als 90 % ein sehr zufriedenstellendes Ergebnis.

für eine solche inhaltsadaptive Kompression sehr gut geeignet. Daher ist eine Verwendung dieser Kompressionsalgorithmen für das System geplant.

Weiter- und Wiederverwendung von Systemkomponenten

Die Bestimmung des Anfälligkeitsindex (*susceptibility index*) anhand von Haustorien ist *eine* Möglichkeit, die Interaktion des Mehltaupilzes mit seiner Wirtspflanze zu beurteilen. Eine andere Möglichkeit liegt in der Analyse des sekundären Hyphenwachstums: Über die Größe der sich auf der Blattoberfläche ausbreitenden Pilzkolonien lassen sich ebenfalls Aussagen zum Wachstum des Pilzes und damit zum Resistenzverhalten der Pflanze treffen. Zur computergestützten Vermessung des Hyphenwachstums wurden bereits einige Untersuchungen durchgeführt [SS05]. Die Bildaufnahme muss bei dieser Analyse jedoch noch manuell erfolgen. Um das System im Sinne einer automatischen Arbeitsweise zu komplettieren, erscheint es adäquat, Teile der im Rahmen dieser Arbeit entwickelten Ansteueralgorithmen für das Mikroskop wieder zu verwenden und mit dem bereits vorliegenden Know-how eine automatische Bildaufnahme für die Vermessung des Hyphenwachstums auf der Blattoberfläche zu implementieren.

Anhang A

Farbräume – Eine Zusammenstellung

A.1 Lineare Transformationen von $R'G'B'$

A.1.1 $Y'UV$ -Farbraum (PAL-Fernsehnorm)

Das $Y'UV$ -System wird hauptsächlich im außeramerikanischen Raum in Form der PAL¹-Farbfernsehnorm benutzt.

$$Y' = 0.299 R' + 0.587 G' + 0.114 B' \quad (\text{A.1})$$

$$U = 0.492 (B' - Y') \quad (\text{A.2})$$

$$V = 0.877 (R' - Y') \quad (\text{A.3})$$

A.1.2 $Y'IQ$ -Farbraum (NTSC-Fernsehnorm)

Die Farbdifferenzsignale des für die amerikanische Fernsehnorm NTSC² verwendeten $Y'IQ$ -Farbraums sind gegenüber dem $Y'UV$ -System um 33° verdreht.

$$\begin{bmatrix} I \\ Q \end{bmatrix} = \begin{bmatrix} \cos 33^\circ & -\sin 33^\circ \\ \sin 33^\circ & \cos 33^\circ \end{bmatrix} \begin{bmatrix} V \\ U \end{bmatrix} \quad (\text{A.4})$$

A.1.3 $Y'D_bD_r$ -Farbraum (SECAM-Fernsehnorm)

Für den hauptsächlich in Frankreich etablierten Fernsehstandard SECAM³ kommt das $Y'D_bD_r$ -System zum Einsatz, dessen Farbdifferenzkomponenten skalierte Versionen der Farbdifferenzkomponenten des $Y'UV$ -Systems sind.

$$D_b = 3.059 U = 1.505 (B' - Y') \quad (\text{A.5})$$

$$D_r = -2.169 V = -1.902 (R' - Y') \quad (\text{A.6})$$

¹PAL – *Phase Alternating Line*

²NTSC – *National Television System(s) Committee*. Aufgrund der Schwächen dieses Standards bezüglich der Farbkonstanz wird er auch gern „hochachtungsvoll“ mit *“Never The Same Color”* oder *“Never Tested Since Christ”* übersetzt.

³SECAM – *Séquentiel Couleur avec Mémoire*. Von amerikanischen Ingenieuren wird SECAM gern scherzhaft mit *“System Essentially Contrary to the American Method”* übersetzt.

A.1.4 $Y'P_bP_r$ -Farbraum (Komponentensignal bei analogem Video)

Der $Y'P_bP_r$ -Farbraum wird für das Komponentensignal bei der analogen Videoübertragung verwendet.

$$Y' = Y_r R' + Y_g G' + Y_b B' \quad (\text{A.7})$$

$$P_b = (B' - Y') / (2 - 2Y_b) \quad (\text{A.8})$$

$$P_r = (R' - Y') / (2 - 2Y_r) \quad (\text{A.9})$$

$$Y_r = 0.299 \quad Y_g = 0.587 \quad Y_b = 0.114$$

A.1.5 $Y'C_bC_r$ -Farbraum (Komponentensignal bei digitalem Video)

Die Quantisierung des $Y'P_bP_r$ -Farbraumes auf 8 bit (Wertebereich 0 . . . 255) führt zum $Y'C_bC_r$ -Farbraum, welcher für digitale Videogeräte spezifiziert wurde und auch in den Standards JPEG und MPEG verwendet wird.

$$Y' = 219 Y'_{Y'P_bP_r} + 16 \quad (\text{A.10})$$

$$C_b = 224 P_b + 128 \quad (\text{A.11})$$

$$C_r = 224 P_r + 128 \quad (\text{A.12})$$

A.1.6 $I_1I_2I_3$ - Ohtas Farbraum

Der von Ohta [OKS80] spezifizierte $I_1I_2I_3$ -Farbraum wurde empirisch auf Basis einer Menge von Farbbildern bestimmt und führt bei vielen Bildern zu einer guten Dekorrelation der Kanäle.

$$I_1 = (R' + G' + B') / 3 \quad (\text{A.13})$$

$$I_2 = (R' - B') / 2 \quad (\text{A.14})$$

$$I_3 = (2G' - R' - B') / 4 \quad (\text{A.15})$$

A.2 Wahrnehmungsorientierte Farbräume – HSV

Bei den wahrnehmungsorientierten Farbräumen HSI , HSV und HLS existieren in der Literatur zum Teil widersprüchliche Definitionen bzw. Bezeichnungen.

Die in dieser Arbeit verwendeten Bezeichnungen Buntton (*Hue*) und Sättigung (*Saturation*) beziehen sich auf die folgende Definition des HSV -Farbraums.

$$V = \max(R', G', B') \quad (\text{A.16})$$

$$S = \begin{cases} \frac{\max(R', G', B') - \min(R', G', B')}{\max(R', G', B')} & \text{für } \max(R', G', B') \neq 0 \\ 0 & \text{sonst} \end{cases} \quad (\text{A.17})$$

$$H = \begin{cases} \frac{1}{6} \frac{G' - B'}{S} & V = R' \\ \frac{1}{3} + \frac{1}{6} \frac{B' - R'}{S} & V = G' \\ \frac{2}{3} + \frac{1}{6} \frac{R' - G'}{S} & V = B' \end{cases} \quad (\text{A.18})$$

A.3 Wahrnehmungsgetreue Farbräume

Bei wahrnehmungsgetreuen Farbräumen stimmt der euklidische Abstand zweier Farben in diesem Raum annähernd mit dem empfindungsgemäßen „Farbabstand“ überein. Dieses Ziel wird von den im Jahr 1976 von der CIE definierten Farbräumen $L^*a^*b^*$ und $L^*u^*v^*$ näherungsweise verwirklicht. Basis ist der XYZ -Farbraum, welcher sich aus dem intensitätslinearen RGB -Farbraum mittels einer Matrix-Vektor-Multiplikation nach Gleichung (2.5) berechnen lässt.

A.3.1 CIE $L^*a^*b^*$

Die empfundene Helligkeit (*Lightness*) berechnet sich durch

$$L^* = \begin{cases} 116 \left(\frac{Y}{Y_0}\right)^{\frac{1}{3}} - 16 & \text{für } \frac{Y}{Y_0} > 0.008856 \\ 903.3 \left(\frac{Y}{Y_0}\right) & \text{für } \frac{Y}{Y_0} \leq 0.008856 \end{cases} \quad (\text{A.19})$$

und nimmt Werte zwischen $L^* = 0$ (schwarz) und $L^* = 100$ (weiß) an. Die Chrominanzkomponenten berechnen sich durch

$$a^* = 500 \left[f\left(\frac{X}{X_0}\right) - f\left(\frac{Y}{Y_0}\right) \right] \quad (\text{A.20})$$

$$b^* = 200 \left[f\left(\frac{Y}{Y_0}\right) - f\left(\frac{Z}{Z_0}\right) \right] \quad (\text{A.21})$$

$$\text{mit } f = \begin{cases} t^{\frac{1}{3}} & \text{für } \frac{Y}{Y_0} > 0.008856 \\ 7.787t + \frac{16}{116} & \text{für } \frac{Y}{Y_0} \leq 0.008856 \end{cases} \quad (\text{A.22})$$

Dabei geben X_0 , Y_0 und Z_0 den Referenzweißpunkt an.

A.3.2 CIE $L^*u^*v^*$

Die Helligkeitskomponenten des $L^*u^*v^*$ - und $L^*a^*b^*$ -Farbraums sind gleich.

$$L^* = \begin{cases} 116 \left(\frac{Y}{Y_0}\right)^{\frac{1}{3}} - 16 & \text{für } \frac{Y}{Y_0} > 0.008856 \\ 903.3 \left(\frac{Y}{Y_0}\right) & \text{für } \frac{Y}{Y_0} \leq 0.008856 \end{cases} \quad (\text{A.23})$$

Die Chrominanzkomponenten u^* und v^* berechnen sich durch

$$u^* = 13 L^* (u' - u_0), \quad \text{mit } u' = \frac{4X}{X + 15Y + 3Z}, \quad u_0 = \frac{4X_0}{X_0 + 15Y_0 + 3Z_0} \quad (\text{A.24})$$

$$v^* = 13 L^* (v' - v_0), \quad \text{mit } v' = \frac{9Y}{X + 15Y + 3Z}, \quad v_0 = \frac{9Y_0}{X_0 + 15Y_0 + 3Z_0}. \quad (\text{A.25})$$

Anhang B

Zellfindung mit anderen Werkzeugen

B.1 VOCUS – Selbständiges Finden interessanter Regionen

In den letzten Jahren wurden vermehrt Anstrengungen unternommen, die menschliche Fähigkeit, komplexe Szenen in sehr kurzer Zeit erfassen und interpretieren zu können, zu modellieren und künstlich nachzubilden [IKN98]. Ein ganz wichtiger Verarbeitungsschritt ist dabei die Lenkung des Aufmerksamkeitsfokus (*Focus of Attention, FOA*) auf wichtige Szenenbestandteile. Ähnlich verhält es sich in vielen Problemstellungen der Bildverarbeitung: Im Sinne einer aufwandsgünstigen Verarbeitung liegt es nahe, in einer komplexen Szene zunächst alle uninteressanten Bereiche zu eliminieren bzw. nur solche „interessanten Regionen“ (*Regions of Interest, ROI*) aufzuspüren, die zur Lösung der konkreten Problemstellung infrage kommen.

Ein Vertreter eines solchen Systems, welches selbständig interessante Regionen in Bildern detektiert, ist VOCUS (*Visual Object Detection with a Computational Attention System*) [FBR05]. VOCUS verfügt sowohl über eine Von-unten-nach-oben- (*Bottom-Up*) als auch über eine Von-oben-nach-unten-Methode (*Top-Down*). Im Bottom-Up-Modus gibt es keine Vorgaben, was im Bild gesucht werden soll. Interessante Regionen werden vom System selbständig anhand der Merkmale *Intensität*, *Orientierung* und *Farbe* unter Nutzung des $L^*a^*b^*$ -Farbraums identifiziert. Eine Verbesserung der Erkennungsrate lässt sich mittels des Top-Down-Ansatzes erreichen. Hier werden dem System in einem Trainingsprozess zunächst Beispiele für zu findende Regionen vorgelegt. Das System entscheidet anhand der daraus berechneten Merkmale selbständig, wodurch die Zielregionen charakterisiert sind.

Für die Problemstellung des Findens der transformierten Zellen in den Detailbildern wurden VOCUS in beiden Modi (sowohl Bottom-Up als auch Top-Down) einige Bilder zur Regionenfindung vorgelegt. Die Ergebnisse sind in Abbildung B.1 dargestellt. Es zeigt sich, dass VOCUS im Bottom-Up-Modus (ohne Training, dargestellt in der zweiten Bildzeile) oft die transformierten Zellen als „wichtig“ erachtet, es werden jedoch teilweise auch andere Regionen markiert. Für den Top-Down-Ansatz wurde VOCUS auf die grünlich gefärbten Zellen trainiert. Die dabei entstandene Merkmalswichtung ist in Tabelle B.1 angegeben. Gemäß dieser Wichtung charakterisiert VOCUS die Zielregionen als

- dunkel auf hellem Hintergrund
- ohne ausgeprägte Orientierung
- mit einem sehr hohen Grünanteil und einem hohen Blauanteil.

Im Ergebnis der Regionenfindung im Top-Down-Modus (dritte Bildzeile von Abbildung B.1) zeigt sich eine deutliche Verbesserung gegenüber dem Bottom-Up-Ansatz: Es werden keine zum Hintergrund gehörigen Regionen mehr detektiert; statt dessen werden *alle* transformierten Zellen mit einen bzw. mehreren FOA markiert.

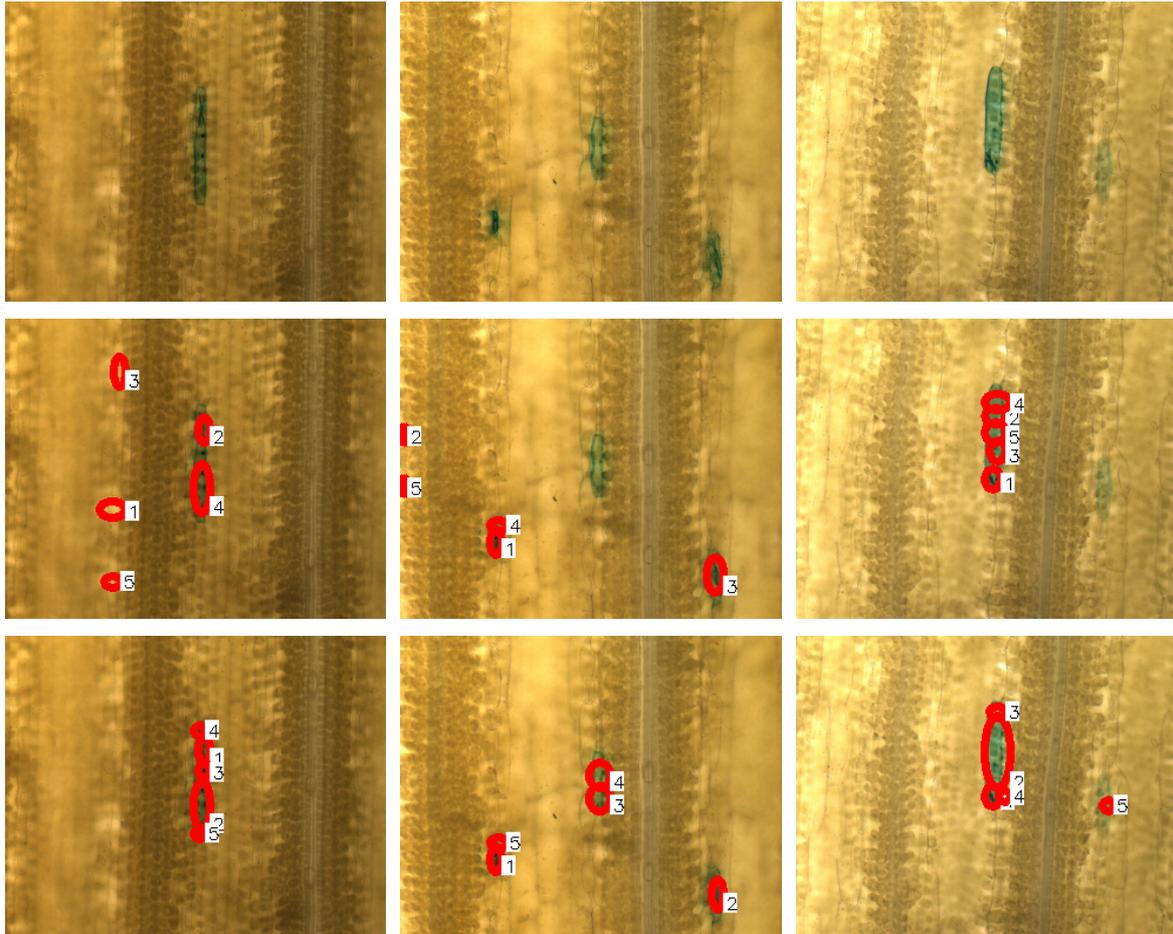


Abbildung B.1: Von VOCUS gefundene interessante Regionen (*Foci of Attention, FOA*).
Zeile 1: Drei typische Detailaufnahmen mit transformierten Zellen.
Zeile 2: Detektierte FOA anhand des Bottom-Up-Ansatzes (ohne Training).
Zeile 3: FOA mittels Top-Down-Ansatz (zuvoriges Training auf die gefärbten Zellen).

Merkmale	Wichtung
Intensität hell auf dunkel	0.0183
Intensität dunkel auf hell	4.0367
Orientierung \leftrightarrow	2.2976
Orientierung \nearrow	2.6532
Orientierung \downarrow	2.9223
Orientierung \nwarrow	2.1333
Farbe Grün	18.0006
Farbe Blau	10.8865
Farbe Rot	0.0114
Farbe Gelb	0.0121
Auffälligkeit Intensität	2.1234
Auffälligkeit Orientierung	2.6529
Auffälligkeit Farbe	9.4932

Tabelle B.1: VOCUS' Merkmalswichtung nach dem Training auf die gefärbten Zellen.

B.2 SIOX – Selbständiges Freistellen von Objekten

Das Segmentieren (auch Herauslösen bzw. Freistellen genannt) ist ein fundamentales Problem der Bildverarbeitung, für das keine allgemein gültige Universallösung existiert. Dennoch stehen Werkzeuge zur Verfügung, mit denen auch komplexe Strukturen mit relativ wenig manueller Hilfe vom (ebenfalls komplex strukturierten) Bildhintergrund herausgelöst werden können.

Eines der bekanntesten Werkzeuge dieser Art ist sicherlich die *Magic Wand* aus Adobes Photoshop [Inc02], welche nach dem Prinzip des Bereichswachstums (*Region Growing*) unter Berücksichtigung von Farbtoleranzen arbeitet. Dieses Verfahren funktioniert sehr gut für die Herauslösung homogen gefärbter Objekte, z.B. bei Comics, stößt jedoch bei natürlichen Szenen schnell an seine Grenzen.

Jüngst wurde von Friedland *et al.* [FJKR05, FJR05] ein sehr ausgeklügeltes Verfahren zur Herauslösung komplexer Objekte aus Bildern vorgeschlagen und als Zusatzmodul (*Plug-in*) für das freie Bildverarbeitungsprogramm GIMP (*GNU Image Manipulation Program*) [GIM] bereitgestellt. Dieses Verfahren namens SIOX (*Simple Interactive Object Extraction*) funktioniert nach dem Prinzip einer Farbclusterung im $L^*a^*b^*$ -Farbraum mit anschließender regionenbasierter Nachbearbeitung. Der Nutzer muss lediglich einige repräsentative Bereiche des herauszulösenden Objekts markieren und kann das Segmentierungsergebnis dann nötigenfalls iterativ weiter verfeinern. In Abbildung B.2 sind die mittels SIOX erzielten Segmentierungsergebnisse dargestellt. Zum Training des Algorithmus wurden Bildbereiche mit schwach gefärbten transformierten Zellen markiert (schwarze Bereiche in der oberen Bildzeile). Im Ergebnis werden die gefärbten Zellen vom Hintergrund (untere Bildzeile, hier violett eingefärbt dargestellt) herausgelöst. SIOX segmentiert die in den Bildern enthaltenen transformierte Zellen, jedoch ist die Zellberandung recht ausgefranst.

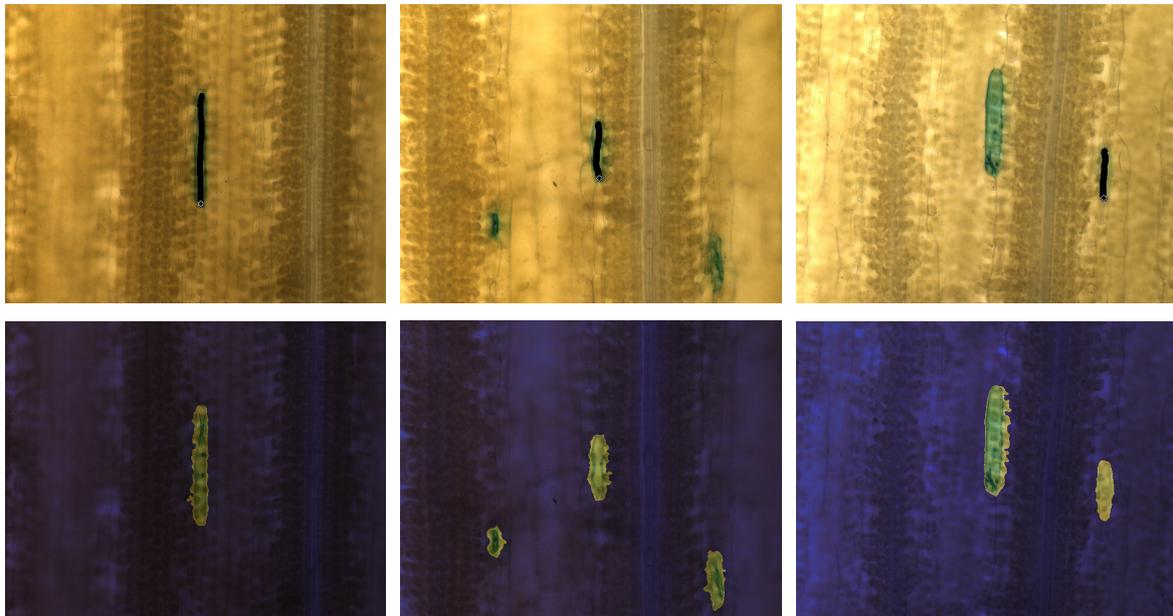


Abbildung B.2: Interaktive Segmentierung transformierter Zellen mittels SIOX. Es wurden die gleichen Bilder wie in Abbildung B.1 verwendet.

Zeile 1: Markierung von Bereichen mit typischer Färbung (in Schwarz dargestellt).

Zeile 2: Freigestellte Objekte.

Anhang C

Einige Kantendetektoren erster Ordnung

Diskreter Gradient Als einfachster Ansatz zur Berechnung eines diskreten Gradientenvektors lässt sich die Rückwärts-, Vorwärts- sowie die symmetrische Differenz definieren [Jäh02].

$$I_x = \frac{\partial I(x, y)}{\partial x} \approx \frac{I(x, y) - I(x - \Delta x, y)}{\Delta x} \quad (\text{Rückwärtsdifferenz}) \quad (\text{C.1})$$

$$\approx \frac{I(x + \Delta x, y) - I(x, y)}{\Delta x} \quad (\text{Vorwärtsdifferenz}) \quad (\text{C.2})$$

$$\approx \frac{I(x + \Delta x, y) - I(x - \Delta x, y)}{2\Delta x} \quad (\text{symmetrische Differenz}) \quad (\text{C.3})$$

Die entsprechenden Filterkerne sind

$$G_x^{(r)} = \begin{bmatrix} 0 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad G_x^{(v)} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & 0 \end{bmatrix} \quad G_x^{(s)} = \frac{1}{2} \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & -1 \\ 0 & 0 & 0 \end{bmatrix}. \quad (\text{C.4})$$

Der diskrete Gradient ist sehr rauschempfindlich.

Roberts-Operator Der Roberts-Operator [Rob63] ist ein diskreter Gradient in Diagonalrichtung. Mittels der Filterkerne

$$G_x = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad G_y = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \quad (\text{C.5})$$

lässt sich der Betrag und Winkel des Gradienten approximieren durch

$$|\nabla I| = \sqrt{G_x^2 + G_y^2} \quad \varphi_{\nabla I} = \arctan(G_y/G_x) - 3/4\pi. \quad (\text{C.6})$$

Als diskreter Gradient ist der Roberts-Operator sehr rauschempfindlich.

Prewitt-Operator Der Prewitt-Operator [Pre70] berechnet eine diskrete Approximation der Ableitung in x - und y -Richtung und glättet das Bild dabei gleichzeitig mit einem gleitenden Mittelwert in Form eines Rechteckfilters.

$$G_x = \frac{1}{6} \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix} \quad G_y = \frac{1}{6} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix} \quad (\text{C.7})$$

Dieser Operator ist weniger rauschempfindlich als der diskrete Gradient.

Sobel-Operator Der Sobel-Operator [Dav75] verbessert den Prewitt-Operator dahin gehend, dass anstatt eines Rechteckfilters ein Gaußfilter zur Glättung verwendet wird.

$$G_x = \frac{1}{8} \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \quad G_y = \frac{1}{8} \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} \quad (\text{C.8})$$

Der Sobel-Operator approximiert die erste Ableitung eines 2D-Gaußkerns nach x bzw. y .

Kirsch-Operator Der Kirsch-Operator [Kir71], auch **Kompass-Operator** genannt, verwendet acht Vorlagen (Templates) $T^{(0)}, T^{(45)}, T^{(90)}, T^{(135)}, T^{(180)}, T^{(225)}, T^{(270)}, T^{(315)}$, welche jeweils eine ideale gerade Kante in die Richtungen $0^\circ, 45^\circ, 90^\circ, 135^\circ, 180^\circ, 225^\circ, 270^\circ$ und 315° nachbilden.

$$\begin{bmatrix} 5 & 5 & 5 \\ -3 & 0 & -3 \\ -3 & -3 & -3 \end{bmatrix} \quad \begin{bmatrix} 5 & 5 & -3 \\ 5 & 0 & -3 \\ -3 & -3 & -3 \end{bmatrix} \quad \begin{bmatrix} 5 & -3 & -3 \\ 5 & 0 & -3 \\ 5 & -3 & -3 \end{bmatrix} \quad \begin{bmatrix} -3 & -3 & -3 \\ 5 & 0 & -3 \\ 5 & 5 & -3 \end{bmatrix}$$

$$\begin{bmatrix} -3 & -3 & -3 \\ -3 & 0 & -3 \\ 5 & 5 & 5 \end{bmatrix} \quad \begin{bmatrix} -3 & -3 & -3 \\ -3 & 0 & 5 \\ -3 & 5 & 5 \end{bmatrix} \quad \begin{bmatrix} -3 & -3 & 5 \\ -3 & 0 & 5 \\ -3 & -3 & 5 \end{bmatrix} \quad \begin{bmatrix} -3 & 5 & 5 \\ -3 & 0 & 5 \\ -3 & -3 & -3 \end{bmatrix}$$

Die Maximumbildung liefert die Kantenrichtung mit der besten Übereinstimmung.

$$G_{\text{kir}} = \max(|T^{(0)}|, |T^{(45)}|, \dots, |T^{(315)}|) \quad (\text{C.9})$$

Anhang D

Laufzeitverhalten aktiver Konturmodelle

D.1 Lösungsschemata für aktive Konturmodelle

Lösungsschema	Komplexität	reelle Multiplikationen
$\mathbf{v}^t = (\mathbf{I} + \tau \mathbf{K})^{-1} (\mathbf{v}^{t-1} + \tau \mathbf{F}_{\text{ext}}(\mathbf{v}^{t-1}))$ (3.71)	$\mathcal{O}(N^2)$	$2N^2$
$\mathbf{v}^t = \text{IFFT} \left\{ \frac{1}{1 + \tau \hat{\mathbf{k}}} \odot \text{FFT} \{ \mathbf{v}^{t-1} + \tau \mathbf{F}_{\text{ext}}(\mathbf{v}^{t-1}) \} \right\}$ (3.78)	$\mathcal{O}(N \log N)$	$2(N + 2N \log_2 N)$
$\mathbf{v}^t = \text{IFFT} \left\{ \frac{1}{1 + \tau \hat{\mathbf{k}}} \right\} \circledast (\mathbf{v}^{t-1} + \tau \mathbf{F}_{\text{ext}}(\mathbf{v}^{t-1}))$ (3.79)	$\mathcal{O}(MN)$	$2MN$

Tabelle D.1: Drei verschiedene Lösungsschemata für aktive Konturmodelle (vorgestellt in Abschnitt 3.3). Ihr numerischer Aufwand hängt von der Konturlänge N und der Anzahl der Samples des Stabilisierungskerns $\text{IFFT}\{1/(1 + \tau \hat{\mathbf{k}})\}$ (im Falle des Lösungsschemas mittels zyklischer Faltung) ab.

In Tabelle D.1 sind die in Abschnitt 3.3 vorgestellten Lösungsschemata für aktive Konturmodelle zusammengefasst. Des Weiteren ist sowohl die numerische Komplexität als auch die tatsächliche Anzahl der notwendigen reellen Multiplikationen angegeben¹.

D.2 Tatsächliches Laufzeitverhalten

Das Laufzeitverhalten eines Algorithmus hängt jedoch nicht nur von der Anzahl der Rechenoperationen, sondern auch von der Art der Implementierung auf der jeweiligen Hardware ab. Durch geschickte Programmierung lässt sich beispielsweise die Anzahl der Speicherzugriffe reduzieren bzw. können die Daten idealerweise vollständig im Cachespeicher gehalten werden.

Um den tatsächlichen Geschwindigkeitsvorteil der Lösung mittels schneller Fouriertransformation (*Fast Fourier Transform, FFT*) nach Gleichung (3.78) und zyklischer Faltung mit reduzierter Kerngröße nach Gleichung (3.79) gegenüber der klassischen Lösung mittels Matrix-Vektor-Multiplikation nach Gleichung (3.71) bewerten zu können, soll das reale Laufzeitverhalten der Algorithmen untersucht werden.

Unter Ausklammerung der Berechnung der externen Kräfte $F_{\text{ext}}(\mathbf{v}^{t-1})$ lassen sich die Kern-Algorithmen folgendermaßen formulieren:

¹Für die FFT wird eine Konturlänge N auf Basis einer Zweierpotenz angenommen.

- Matrix-Vektor-Multiplikation

$$\begin{aligned} \mathbf{x}^{t+1} &= \mathbf{A}\mathbf{x}^t \\ \mathbf{y}^{t+1} &= \mathbf{A}\mathbf{y}^t \end{aligned} \quad \mathbf{A} \in \mathbb{R}^{N \times N} \quad \mathbf{x}, \mathbf{y} \in \mathbb{R}^{N \times 1}$$

- Zyklische Faltung mit variabler Kernlänge M

$$\begin{aligned} \mathbf{x}^{t+1} &= \mathbf{a} \circledast \mathbf{x}^t \\ \mathbf{y}^{t+1} &= \mathbf{a} \circledast \mathbf{y}^t \end{aligned} \quad \mathbf{a} \in \mathbb{R}^{M \times 1} \quad \mathbf{x}, \mathbf{y} \in \mathbb{R}^{N \times 1}$$

- Schnelle Fouriertransformation (FFT)

$$\mathbf{z}^{t+1} = \text{IFFT}\{\mathbf{a} \odot \text{FFT}\{\mathbf{z}^t\}\} \quad \mathbf{a} \in \mathbb{R}^{N \times 1} \quad \mathbf{z} \in \mathbb{C}^{N \times 1}.$$

Vergleicht man die Anzahl der benötigten Multiplikationen der FFT-Lösung und der Lösung mittels zyklischer Faltung (siehe Tabelle D.1), so stellt sich heraus, dass die FFT-Lösung bis zu einer bestimmten Kernlänge

$$M \geq 1 + 2 \log_2 N \tag{D.1}$$

schneller sein müsste, weil weniger Multiplikationen auszuführen sind. Für eine Konturlänge von $N = 1024$ ist dies beispielsweise für Kernlängen von $M \geq 21$ Samples der Fall.

Ergebnisse

Um die tatsächliche Leistungsfähigkeit der vorgestellten Lösungsschemata zu bewerten, wurden die Kern-Algorithmen unter MATLAB² hinsichtlich ihres Laufzeitverhaltens getestet. MATLAB verwendet neben der FFTW-Bibliothek für die schnelle Fouriertransformation generell hochoptimierte Bibliotheken für Operationen im Bereich der linearen Algebra (z.B. für die Matrix-Vektor-Multiplikation) und ist daher als Werkzeug hervorragend geeignet.

In den Abbildungen D.1 und D.2 ist das Laufzeitverhalten der Kern-Algorithmen auf einem UltraSparc III Prozessor mit 1.2 GHz dargestellt. Es ist deutlich zu erkennen, dass das FFT-Lösungsschema alle anderen Lösungsschemata in puncto Geschwindigkeit überbietet. Bereits bei sehr kurzen Konturlängen von $M \approx 300$ unterscheiden sich das FFT-Lösungsschema und die klassische Lösung (Matrix-Vektor-Multiplikation) um eine Größenordnung (vgl. Abbildung D.2)! Des Weiteren ist in Abbildung D.2 Folgendes zu beobachten: Die Kurven der Matrix-Vektor-Multiplikation und der zyklischen Faltung mit Kernlänge $M = N$ zeigen zwar die gleiche Tendenz, sind jedoch um einen Offset gegeneinander verschoben. Da beide Lösungsverfahren prinzipiell die *gleichen* mathematischen Operationen durchzuführen haben, ist dieser Offset nur durch eine weniger effiziente Implementierung der zyklischen Faltung zu erklären³. Unter der Annahme, die zyklische Faltung ähnlich effizient wie die Matrix-Vektor-Multiplikation implementieren zu können, lässt sich mit den vorhandenen Daten mittels eines Korrekturfaktors die theoretisch minimale Laufzeit der zyklischen Faltung ermitteln. In den Abbildungen D.3 und D.4 ist dieser Fall dargestellt⁴. Durch diese Korrektur rückt die Laufzeit der zyklischen Faltung (mit kurzen Kernlängen $M = 21$ bzw. $M = 31$) sehr eng an die der FFT heran.

Für die praktische Nutzung aktiver Konturmodelle ist das Lösungsschema mittels FFT die erste Wahl und ermöglicht durch seine nahezu lineare Komplexität den Einsatz aktiver Konturmodelle auch für Echtzeit- und Hochdurchsatzanwendungen.

²MATLAB[®] R14, Version 7.0.4

³Für die MATLAB-Implementierung wurde die zyklische Faltung mittels eines FIR-Filters nachgebildet.

⁴Der tatsächliche Wert des Korrekturfaktors beträgt $\frac{1}{6.5}$.

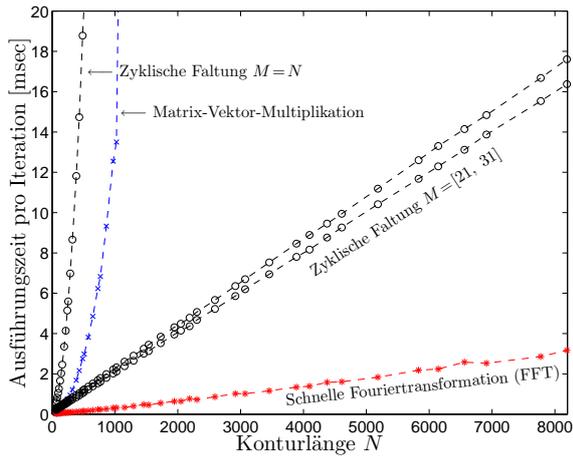


Abbildung D.1: Laufzeit der Kern-Algorithmen der vorgestellten Lösungsschemata pro Iterationsschritt der Schlange (in Millisekunden) auf einem UltraSparc III Prozessor mit 1.2 GHz. Es wurden nur Konturlängen N betrachtet, deren größter Primfaktor maximal drei ist. (Sequenzen dieser Längen lassen sich besonders gut faktorieren.) Das FFT-Lösungsschema ist das mit Abstand effizienteste und ermöglicht den Einsatz aktiver Konturmodelle für Hochdurchsatz- bzw. Echtzeitanwendungen.

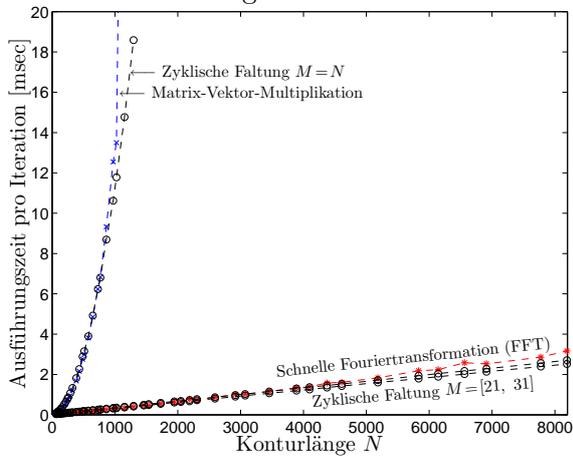


Abbildung D.3: Normierte Variante von Abbildung D.1. Die Ausführungszeiten aller zyklischen Faltungen wurden mittels eines Faktors korrigiert, so dass die Ausführungszeiten der zyklischen Faltung mit $M = N$ und der Matrix-Vektor-Multiplikation etwa gleich sind. Diese Normierung ergibt eine Abschätzung des Laufzeitverhaltens der zyklischen Faltung, wenn sie ähnlich effizient implementiert würde wie die Matrix-Vektor-Multiplikation.

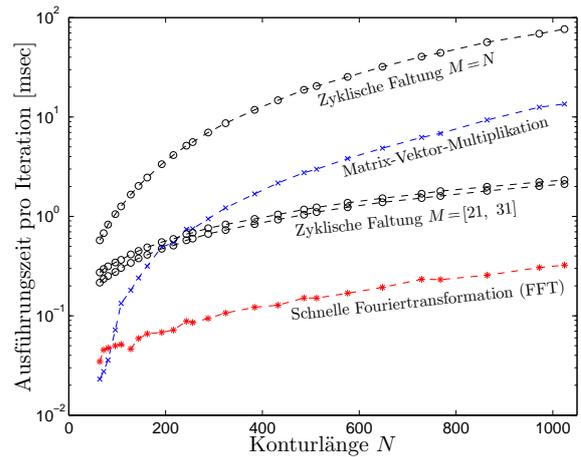


Abbildung D.2: Detailausschnitt von Abbildung D.1 zur Hervorhebung der Ergebnisse bei kurzen Konturlängen. (Die Y-Achse ist hier logarithmisch skaliert.)

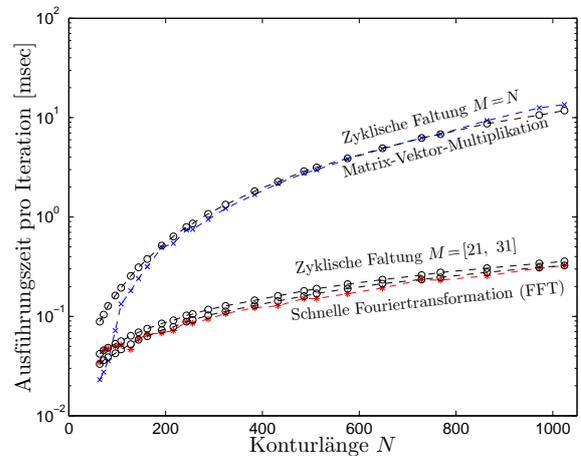


Abbildung D.4: Detailausschnitt von Abbildung D.3 zur Hervorhebung der Ergebnisse bei kurzen Konturlängen.

Anhang E

Reelle vs. komplexe Darstellung Kartesischer Fourierdeskriptoren

In der Literatur existieren zwei verschiedene Darstellungen kartesischer Fourierdeskriptoren. Beide Varianten sind gleichwertig und können ineinander umgerechnet werden. Aus Sicht des Autors ist der komplexen Variante aufgrund der kompakteren und leichter zu handhabenden Darstellung der Vorzug zu geben.

Aus Gründen der Vollständigkeit und des besseren Verständnisses soll an dieser Stelle gezeigt werden, wie sich beide Darstellungen ineinander überführen lassen. Dazu vergleichen wir die komplexe Repräsentation

$$z(t) = \sum_{\mu=-\infty}^{\infty} \hat{Z}_{\mu} e^{j2\pi\mu t/T} \quad \text{mit} \quad z(t) = x(t) + jy(t) \quad \circ\bullet \quad \hat{Z}_{\mu} = \hat{X}_{\mu} + j\hat{Y}_{\mu} \quad (\text{E.1})$$

und die reelle Repräsentation

$$\begin{bmatrix} x(t) \\ y(t) \end{bmatrix} = \begin{bmatrix} \hat{A}_0 \\ \hat{C}_0 \end{bmatrix} + \sum_{\mu=1}^{\infty} \begin{bmatrix} \hat{A}_{\mu} & \hat{B}_{\mu} \\ \hat{C}_{\mu} & \hat{D}_{\mu} \end{bmatrix} \begin{bmatrix} \cos(2\pi\mu t/T) \\ \sin(2\pi\mu t/T) \end{bmatrix}. \quad (\text{E.2})$$

Im Folgenden soll gezeigt werden, wie sich die reellwertigen Fourierdeskriptoren \hat{A}_{μ} , \hat{B}_{μ} , \hat{C}_{μ} , \hat{D}_{μ} aus den komplexen Fourierdeskriptoren \hat{Z}_{μ} berechnen lassen. Dazu betrachten wir zunächst nur den x -Anteil der Kontur und vergleichen die komplexe Darstellung

$$x(t) = \sum_{\mu=-\infty}^{\infty} \hat{X}_{\mu} e^{j2\pi\mu t/T} \quad (\text{E.3})$$

und die reelle Darstellung

$$x(t) = \hat{A}_0 + \sum_{\mu=1}^{\infty} \hat{A}_{\mu} \cos(2\pi\mu t/T) + \hat{B}_{\mu} \sin(2\pi\mu t/T). \quad (\text{E.4})$$

Mittels der EULERSchen Formel $e^{\pm jx} = \cos(x) \pm j \sin(x)$ lässt sich Gleichung (E.4) schreiben als

$$x(t) = \hat{A}_0 + \sum_{\mu=1}^{\infty} \hat{A}_{\mu} \frac{1}{2} \left[e^{j2\pi\mu t/T} + e^{-j2\pi\mu t/T} \right] + \hat{B}_{\mu} \frac{1}{2j} \left[e^{j2\pi\mu t/T} - e^{-j2\pi\mu t/T} \right]. \quad (\text{E.5})$$

Durch eine Aufspaltung der unendlichen Reihe ergibt sich

$$x(t) = \sum_{\mu=-\infty}^{-1} \frac{1}{2} \left[\hat{A}_{|\mu|} - j\hat{B}_{|\mu|} \right] e^{j2\pi\mu t/T} + \hat{A}_0 + \sum_{\mu=1}^{\infty} \frac{1}{2} \left[\hat{A}_{\mu} + j\hat{B}_{\mu} \right] e^{j2\pi\mu t/T}. \quad (\text{E.6})$$

Entstanden ist eine unendliche Reihe, die wie Gleichung (E.3) von $\mu = -\infty \dots \infty$ läuft. Durch Koeffizientenvergleich der komplexen Drehzeiger in den Gleichungen (E.3) und (E.6) lassen sich nun leicht die folgenden Zusammenhänge finden:

$$\hat{X}_{-\mu} = \frac{1}{2} [\hat{A}_{\mu} - j \hat{B}_{\mu}] \quad (\text{E.7})$$

$$\hat{X}_{\mu} = \frac{1}{2} [\hat{A}_{\mu} + j \hat{B}_{\mu}] . \quad (\text{E.8})$$

Dabei treten die Koeffizienten $\hat{X}_{-\mu}$ und \hat{X}_{μ} zueinander konjugiert komplex auf, weil die Funktion $x(t)$ reellwertig ist.

Das gleiche Schema läßt sich für das Signal $y(t)$ durchführen. Daraus ergeben sich die Zusammenhänge

$$\hat{Y}_{-\mu} = \frac{1}{2} [\hat{C}_{\mu} - j \hat{D}_{\mu}] \quad (\text{E.9})$$

$$\hat{Y}_{\mu} = \frac{1}{2} [\hat{C}_{\mu} + j \hat{D}_{\mu}] . \quad (\text{E.10})$$

Mit

$$x(t) + j y(t) = z(t) \quad \circ \bullet \quad \hat{Z}_{\mu} = \hat{X}_{\mu} + j \hat{Y}_{\mu} \quad (\text{E.11})$$

lassen sich die Beziehungen zwischen \hat{A}_{μ} , \hat{B}_{μ} , \hat{C}_{μ} , \hat{D}_{μ} und \hat{Z}_{μ} nun angeben.

$$\hat{A}_{\mu} = \text{Re}\{\hat{Z}_{\mu} + \hat{Z}_{-\mu}\} \quad (\text{E.12})$$

$$\hat{B}_{\mu} = \text{Re}\{j(\hat{Z}_{\mu} - \hat{Z}_{-\mu})\} \quad (\text{E.13})$$

$$\hat{C}_{\mu} = \text{Im}\{\hat{Z}_{\mu} + \hat{Z}_{-\mu}\} \quad (\text{E.14})$$

$$\hat{D}_{\mu} = \text{Im}\{j(\hat{Z}_{\mu} - \hat{Z}_{-\mu})\} \quad (\text{E.15})$$

Anhang F

Streudiagramme der Merkmale

Auf den folgenden Seiten sind die Streudiagramme der in Tabelle 5.1 auf Seite 82 aufgeführten Merkmale zur Haustorienerkennung dargestellt. Die Farblegende ist **rot** für die Klasse „*Haustorium*“ und **grün** für die Klasse „*kein Haustorium*“.

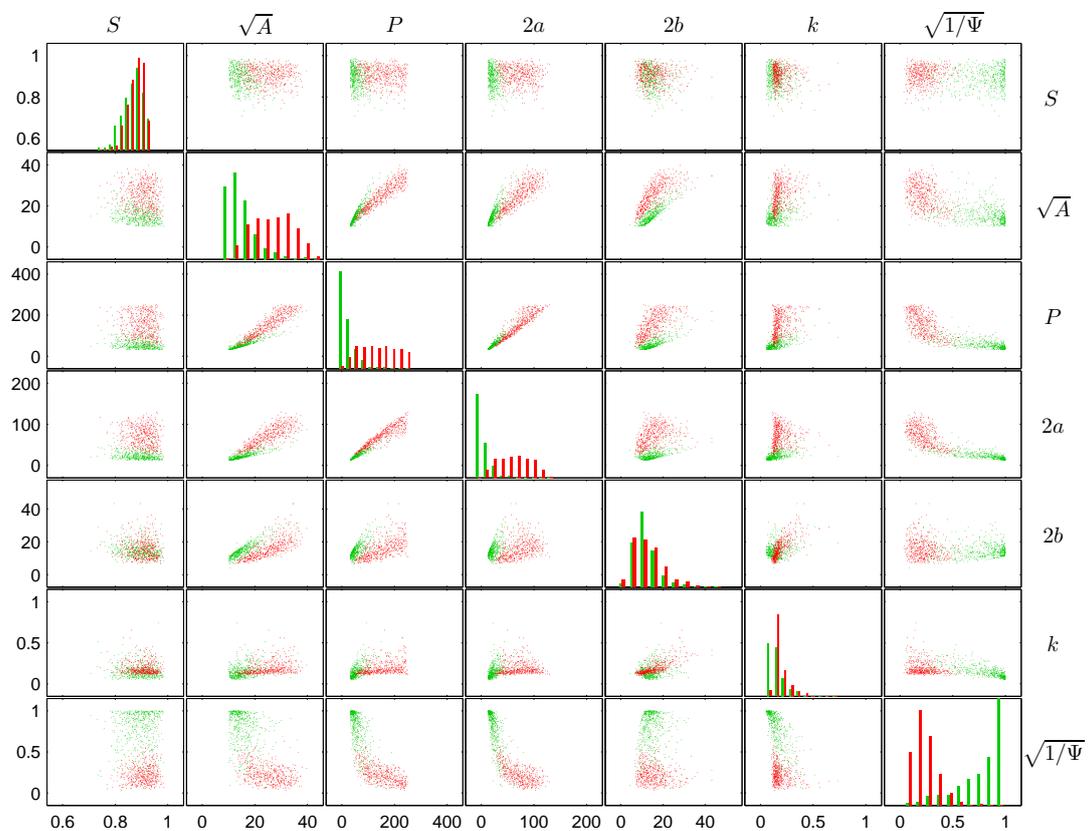


Abbildung F.1: Streudiagramme der einfachen Merkmale.

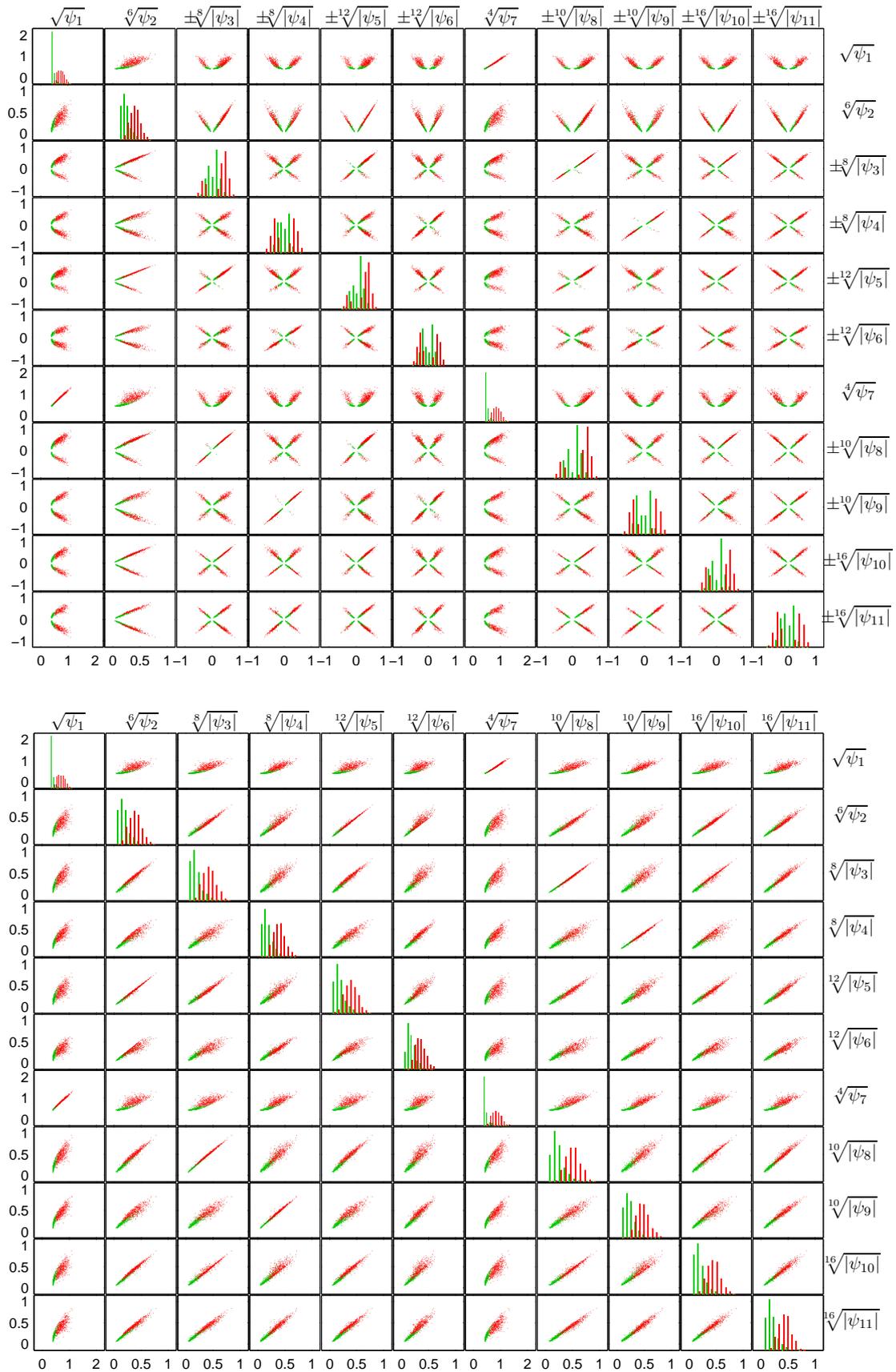


Abbildung F.2: Streudiagramme der Flussler-Momentinvarianzen $\psi_1 \dots \psi_{11}$.
Oben: Berücksichtigung des Vorzeichens. (Anmerkung: ψ_1 , ψ_2 und ψ_7 sind immer positiv.)
Unten: Keine Berücksichtigung des Vorzeichens.

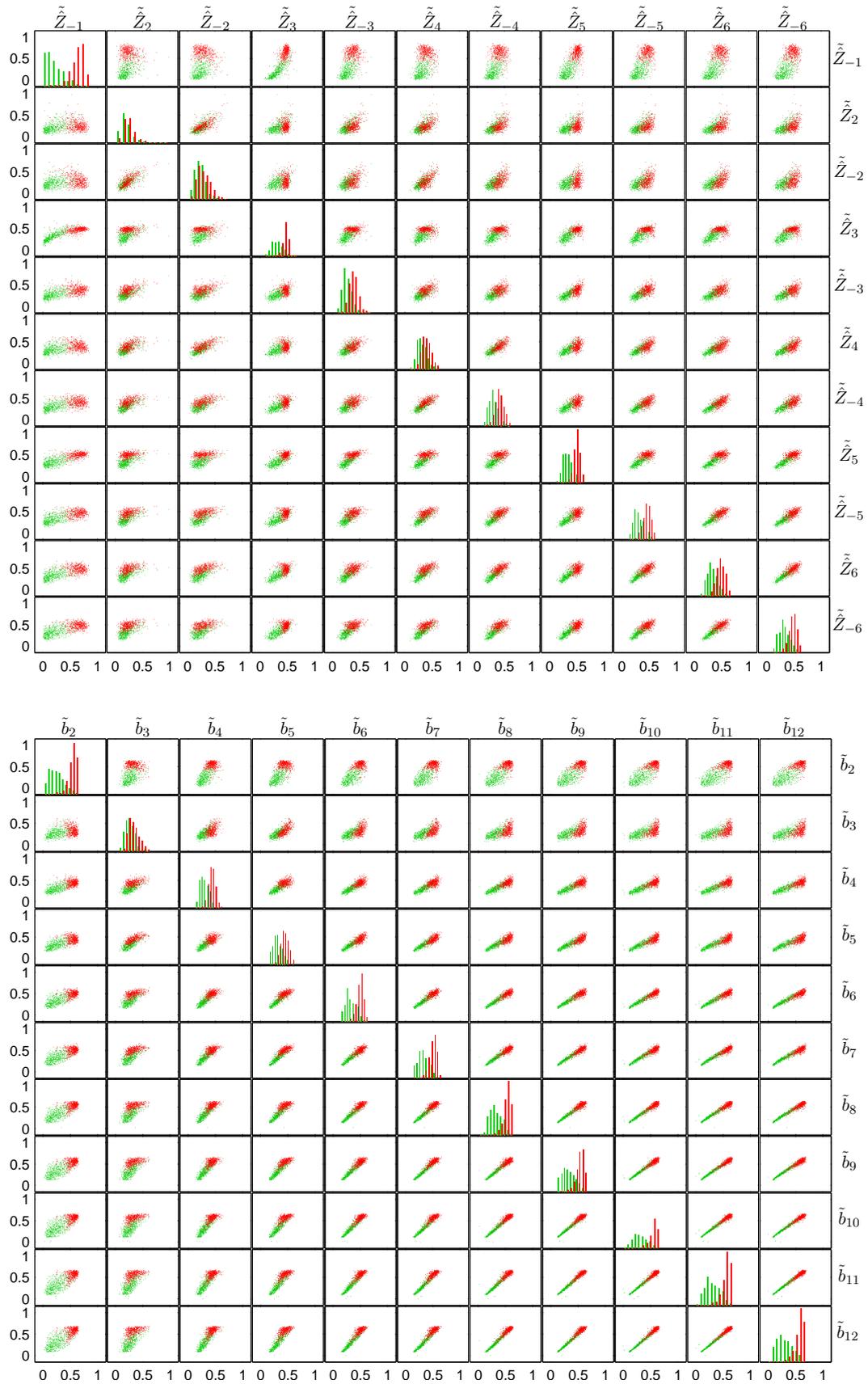


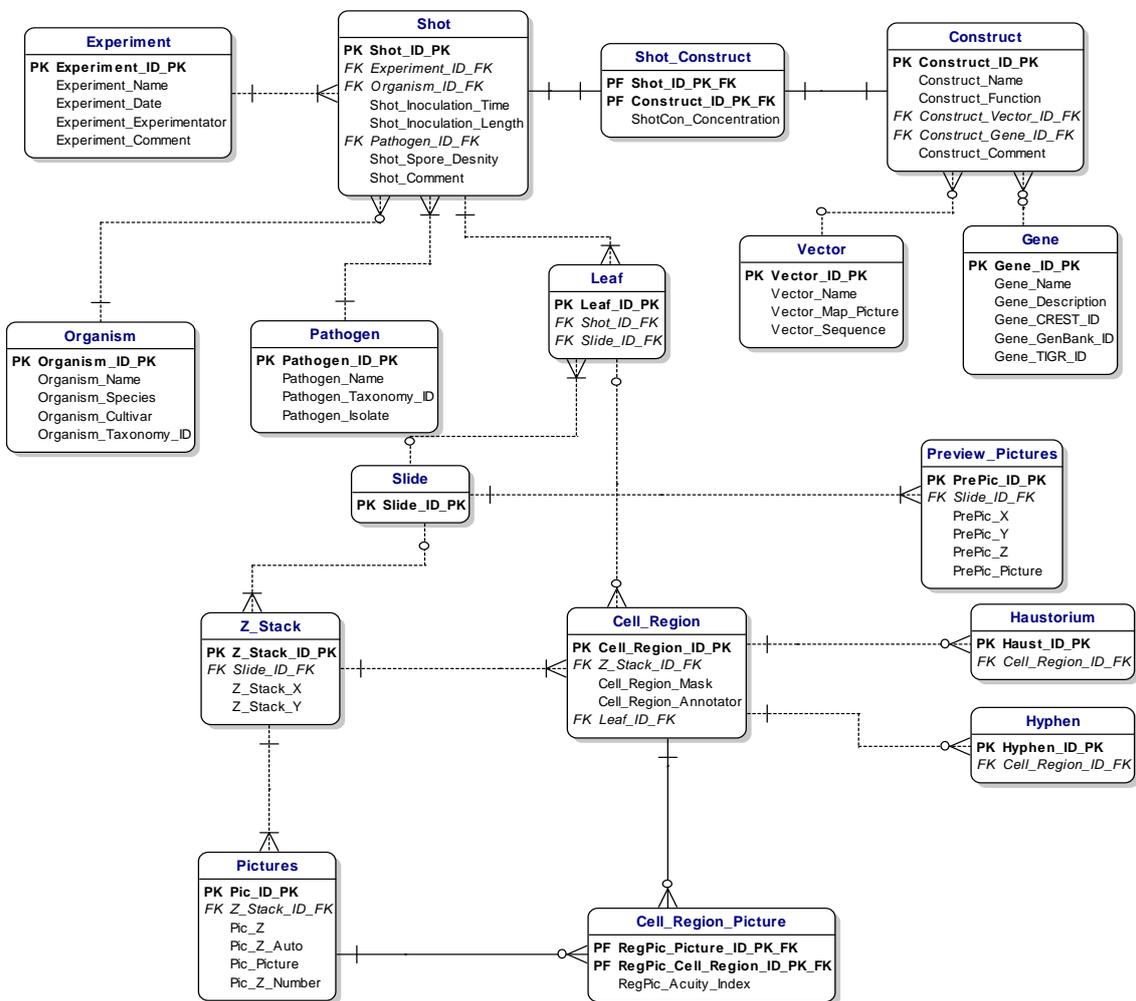
Abbildung F.3: Streudiagramme der Spektralmerkmale.

Oben: Beträge der gemäß Gleichung (5.17) normierten Fourierdeskriptoren.

Unten: Beträge der gemäß Gleichung (5.18) normierten Granlund-Spektralinvarianzen.

Anhang G

Entity-Relationship-Diagramm der Datenbank



Literaturverzeichnis

- [Bar76] BARNETT, Vic: The ordering of multivariate data. In: *Journal of the Royal Statistical Society. Series A* 139 (1976), Nr. 2, S. 331–354
- [Bay53] BAYES, Thomas: An Essay Towards Solving a Problem in the Doctrine of Chances. In: *Philosophical Transactions of the Royal Society of London* 53 (1753), S. 370–418
- [Bha43] BHATTACHARYYA, Anil K.: On a measure of divergence between two statistical populations defined by their probability distributions. In: *Bulletin of the Calcutta Mathematics Society* 35 (1943), S. 99–110
- [Bis95] BISHOP, Christopher M.: *Neural Networks for Pattern Recognition*. Oxford : Oxford University Press, 1995. – 482 S. – ISBN 0–19–853849–9
- [BSMM01] BRONSTEIN, Ilja N. ; SEMENDJAJEW, Konstantin A. ; MUSIOL, Gerhard ; MÜHLIG, Heiner: *Taschenbuch der Mathematik*. Thun : Verlag Harri Deutsch, 2001. – 1191 S. – ISBN 3–8171–2005–2
- [BW64] BROWN, Paul K. ; WALD, George: Visual pigments in single rods and cones of the human retina. In: *Science* 143 (1964), April, S. 45–51
- [Can86] CANNY, John F.: A Computational Approach to Edge Detection. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)* 8 (1986), November, Nr. 6, S. 679–698
- [Cas96] CASTLEMAN, Kenneth R.: *Digital Image Processing*. Englewood Cliffs, NJ : Prentice Hall, 1996. – 667 S. – ISBN 0–13–211467–4
- [CH67] COVER, Thomas M. ; HART, Peter E.: Nearest neighbor pattern classification. In: *IEEE Transactions on Information Theory* 13 (1967), S. 21–27
- [Che52] CHERNOFF, Herman: A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations. In: *The Annals of Mathematical Statistics* 23 (1952), S. 493–507
- [CJFC97] CESAR JR., Roberto M. ; FONTOURA COSTA, Luciano da: Application and assessment of multiscale bending energy for morphometric characterization of neural cells. In: *Review of Scientific Instruments* 68 (1997), Mai, Nr. 5, S. 2177–2186
- [CLMC92] CATTÉ, Francine ; LIONS, Pierre-Louis ; MOREL, Jean-Michel ; COLL, Tomeu: Image selective smoothing and edge detection by nonlinear diffusion. In: *SIAM Journal on Numerical Analysis* 29 (1992), Februar, Nr. 1, S. 182–193
- [CST00] CRISTIANINI, Nello ; SHAWE-TAYLOR, John: *An Introduction to Support Vector Machines*. Cambridge : Cambridge University Press, 2000. – 189 S. – ISBN 0–521–78019–5

- [CT65] COOLEY, James W. ; TUKEY, John W.: An algorithm for the machine calculation of complex Fourier series. In: *Mathematics of Computation* 19 (1965), S. 297–301
- [Dav75] DAVIS, Larry S.: A Survey of Edge Detection Techniques. In: *Computer Graphics and Image Processing* 4 (1975), September, Nr. 3, S. 248–270
- [DHS01] DUDA, Richard O. ; HART, Peter E. ; STORK, David G.: *Pattern Classification*. New York : John Wiley & Sons, 2001. – 654 S. – ISBN 0–471–05669–3
- [DL88] DUCHENE, J. ; LECLERCQ, S.: An Optimal Transformation for Discriminant and Principal Component Analysis. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)* 10 (1988), Nr. 6, S. 978–983. – ISSN 0162–8828
- [DLR77] DEMPSTER, Arthur P. ; LAIRD, Nan M. ; RUBIN, Donald B.: Maximum Likelihood from Incomplete Data via the EM Algorithm. In: *Journal of the Royal Statistical Society, Series B* 39 (1977), Nr. 1, S. 1–38
- [FBR05] FRINTROP, Simone ; BACKER, Gerriet ; ROME, Erich: Goal-directed Search with a Top-down Modulated Computational Attention System. In: KROPATSCH, Walter G. (Hrsg.) ; SABLATNIG, Robert (Hrsg.) ; HANBURY, Allan (Hrsg.): *Pattern Recognition: 27th DAGM Symposium*. Wien, Austria : Springer, August–September 2005 (LNCS 3663). – ISBN 3–540–28703–5
- [FCCJ01] FONTOURA COSTA, Luciano da ; CESAR JR., Roberto M.: *Shape Analysis and Classification*. Boca Raton : CRC Press, 2001. – 659 S. – ISBN 0–8493–3493
- [Fis36] FISHER, Ronald A.: The use of multiple measurements in taxonomic problems. In: *Annals of Eugenics* 7 Part II (1936), S. 179–188
- [FJ] FRIGO, Matteo ; JOHNSON, Steven G.: *FFTW – “Fastest Fourier Transform in the West”*. <http://www.fftw.org>. – Online-Ressource, Abruf: Jan 2006
- [FJ05] FRIGO, Matteo ; JOHNSON, Steven G.: The Design and Implementation of FFTW3. In: *Proceedings of the IEEE* 93 (2005), Nr. 2, S. 216–231
- [FJKR05] FRIEDLAND, Gerald ; JANTZ, Kristian ; KNIPPING, Lars ; ROJAS, Raul: Image Segmentation by Uniform Color Clustering – Approach and Benchmark Results / Department of Computer Science, Freie Universität Berlin. 2005. – Forschungsbericht
- [FJR05] FRIEDLAND, Gerald ; JANTZ, Kristian ; ROJAS, Raul: SIOX: Simple Interactive Object Extraction in Still Images. In: *IEEE International Symposium on Multimedia (ISM2005)*. Irvine (California), Dezember 2005
- [Flu00] FLUSSER, Jan: On the Independence of Rotation Moment Invariants. In: *Pattern Recognition* 33 (2000), Nr. 9, S. 1405–1410
- [For00] FORSTER, Malcolm R.: Key Concepts in Model Selection: Performance and Generalizability. In: *Journal of Mathematical Psychology* 44 (2000), S. 205–231
- [FPT81] FOWLER, Robert J. ; PATERSON, Mike ; TANIMOTO, Steven L.: Optimal Packing and Covering in the Plane are NP-Complete. In: *Information Processing Letters* 12 (1981), Nr. 3, S. 133–137
- [Fre88] FREY, Herbert: *Digitale Bildverarbeitung in Farbräumen*, TU München, Diss., März 1988. – 112 S

- [Fuk94] FUKUNAGA, Keinosuke: *Introduction to Statistical Pattern Recognition*. 2nd. Boston : Academic Press, 1994. – 591 S. – ISBN 0–12–269851–7
- [Gau66] GAUSS, Carl F.: Nachlass: Theoria interpolationis methodo nova tractata. In: *Königliche Gesellschaft der Wissenschaften, Göttingen* (1866), S. 265–327
- [GIM] *the GIMP*. <http://www.gimp.org>. – Online-Ressource, Abruf: Jan 2006
- [GJJ96] GOSE, Earl ; JOHNSONBAUGH, Richard ; JOST, Steve: *Pattern Recognition and Image Analysis*. Upper Saddle River, NJ : Prentice Hall, 1996. – 484 S. – ISBN 0–13–236415–8
- [Gol04] GOLDBERG, David E.: *Genetic Algorithms in Search, Optimization and Machine Learning*. Boston : Addison-Wesley, 2004. – 412 S. – ISBN 0–201–15767–5
- [Gra72] GRANLUND, Gösta H.: Fourier Preprocessing for Hand Print Character Recognition. In: *IEEE Transactions on Computers* C-21 (1972), März, Nr. 3, S. 195–201
- [Hay99] HAYKIN, Simon: *Neural networks: a comprehensive foundation*. 2nd. Upper Saddle River, NJ : Prentice Hall, 1999. – 842 S. – ISBN 0–13–273350–1
- [Hel67] HELMHOLTZ, Hermann von: *Handbuch der physiologischen Optik*. Leipzig, 1867
- [HFPK00] HÜCKELHOVEN, Ralph ; FODOR, József ; PREIS, Christine ; KOGEL, Karl-Heinz: Hypersensitive cell death and papilla formation in barley attacked by the powdery mildew fungus are associated with hydrogen peroxide but not with salicylic acid accumulation. In: *Plant Physiology* 119 (2000), Nr. 4, S. 1251–1260
- [HJB84] HEIDEMAN, Michael T. ; JOHNSON, Don H. ; BURRUS, C. S.: Gauss and the history of the FFT. In: *IEEE Acoustics, Speech, and Signal Processing Magazine* 1 (1984), Nr. 4, S. 14–21
- [Hot33] HOTELLING, Harold: Analysis of a complex of statistical variables into principal components. In: *Journal of Educational Psychology* 24 (1933), S. 417–441
- [Hu61] HU, Ming-Kuei: Pattern recognition by moment invariants. In: *Proceedings of the IRE* 49 (1961), September, S. 1428
- [Hu62] HU, Ming-Kuei: Visual pattern recognition by moment invariants. In: *IRE Transactions on Information Theory* 8 (1962), Februar, Nr. 2, S. 179–187
- [IKN98] ITTI, Laurent ; KOCH, Christof ; NIEBUR, Ernst: A Model of Saliency-based Visual Attention for Rapid Scene Analysis. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)* 41 (1998), Nr. 3, S. 523–537
- [Inc02] INC., Adobe S.: *Adobe Photoshop User Guide*. 2002
- [IP93] IVINS, James P. ; PORRILL, John: Everything you always wanted to know about snakes (but were afraid to ask) / Artificial Intelligence Vision Research Unit, University of Sheffield, England S10 2TP. 1993. – Forschungsbericht
- [IS03] IHLOW, Alexander ; SEIFFERT, Udo: Microscope Color Image Segmentation for Resistance Analysis of Barley Cells against Powdery Mildew. In: *9. Workshop "Farbbildverarbeitung"*. Ostfildern-Nellingen, Germany, Oktober 2003 (ZBS Zentrum für Bild- und Signalverarbeitung e.V. Ilmenau, Report Nr. 3/2003). – ISSN 1432–3346, S. 59–66

- [IS04a] IHLOW, Alexander ; SEIFFERT, Udo: Automating Microscope Colour Image Analysis using the Expectation Maximisation Algorithm. In: RASMUSSEN, Carl E. (Hrsg.) ; BÜLTHOFF, Heinrich H. (Hrsg.) ; GIESE, Martin A. (Hrsg.) ; SCHÖLKOPF, Bernhard (Hrsg.): *Pattern Recognition: 26th DAGM Symposium*. Tübingen, Germany : Springer, August/September 2004 (LNCS 3175). – ISBN 3-540-22945-0, S. 536-543
- [IS04b] IHLOW, Alexander ; SEIFFERT, Udo: Haustoria Segmentation in Microscope Colour Images of Barley Cells. In: DROEGE, Detlev (Hrsg.) ; PAULUS, Dietrich (Hrsg.): *10. Workshop "Farbbildverarbeitung"*. Koblenz : Der andere Verlag, Oktober 2004. – ISBN 3-89959-245-X, S. 119-126
- [IS05] IHLOW, Alexander ; SEIFFERT, Udo: Snakes Revisited – Speeding Up Active Contour Models using the Fast Fourier Transform. In: *Proceedings of the Eight IASTED International Conference on Intelligent Systems and Control (ISC 2005)*. Cambridge, MA, USA, Oktober 2005, S. 416-420
- [Jäh02] JÄHNE, Bernd: *Digitale Bildverarbeitung*. Springer, 2002. – 618 S. – ISBN 3-540-41260-3
- [JM97] JOHNSON, David S. ; MCGEOCH, Lyle A.: The Traveling Salesman Problem: A Case Study in Local Optimization. In: AARTS, Emile (Hrsg.) ; LENSTRA, Jan K. (Hrsg.): *Local Search in Combinatorial Optimization*. London : John Wiley & Sons, 1997, S. 215-310
- [Jol02] JOLLIFFE, Ian T.: *Principal Component Analysis*. 2nd. New York : Springer, 2002. – 500 S. – ISBN 0-387-95442-2
- [Jun05] JUNGnickel, Dieter: *Graphs, networks and algorithms*. 2nd. Berlin : Springer, 2005. – 611 S. – ISBN 3-540-21905-6
- [KI89] KRESS, Dieter ; IRMER, Ralf: *Angewandte Systemtheorie*. Berlin : Verlag Technik, 1989. – 336 S. – ISBN 3-341-00693-1
- [Kir71] KIRSCH, Russell A.: Computer Determination of the Constituent Structure of Biological Images. In: *Computers and Biomedical Research* 4 (1971), S. 315-328
- [KK04] KEMPken, Frank ; KEMPken, Renate: *Gentechnik bei Pflanzen*. Berlin : Springer, 2004 (Springer-Lehrbuch). – ISBN 3-540-01216-8
- [Koe84] KOENDERINK, Jan J.: The Structure of Images. In: *Biological Cybernetics* 50 (1984), S. 363-370
- [Koz93] KOZIOL, James A.: Probability plots for assessing multivariate normality. In: *The Statistician* 42 (1993), Nr. 9, S. 161-173
- [KWT88] KASS, Michael ; WITKIN, Andrew ; TERZOPOULOS, Demetri: Snakes: Active contour models. In: *International Journal of Computer Vision* 1 (1988), Nr. 4, S. 321-331
- [Len05] LENNERZ, Christian: *Distance Computation for Extended Quadratic Complexes*, Universität des Saarlandes, Diss., Dezember 2005. – 394 S
- [Lew01] LEW, Michael S. (Hrsg.): *Principles of Visual Information Retrieval*. London : Springer, 2001. – 356 S. – ISBN 1-85233-381-2

- [Loè77] LOÈVE, Michel: *Probability Theory I*. 4th. New York : Springer, 1977. – 425 S. – ISBN 3-540-90210-4
- [Loè78] LOÈVE, Michel: *Probability Theory II*. 4th. New York : Springer, 1978. – 413 S. – ISBN 3-540-90262-7
- [LS02] LENNERZ, Christian ; SCHÖMER, Elmar: Efficient Distance Computation for Quadratic Curves and Surfaces. In: *2nd IEEE Conference on Geometric Modeling and Processing (GMP 2002)*. Wako, Saitama, Japan, 2002, S. 60–69
- [MDMJ64] MARKS, W. B. ; DOBELLE, W. H. ; MACNICHOL JR., E. F.: Visual pigments of single primate cones. In: *Science* 143 (1964), März, S. 1181–1183
- [MH80] MARR, David ; HILDRETH, Ellen C.: Theory of Edge Detection. In: *Proceedings of the Royal Society of London B* 207 (1980), S. 187–217
- [Mun05] MUNSELL, Albert H.: *A Color Notation*. 1st. Boston : Geo. H. Ellis Co., 1905
- [Mun15] MUNSELL, Albert H.: *Atlas of the Munsell Color System*. Malden, Massachusetts; Wadsworth, Howland, 1915
- [NIS] NIST/SEMATECH: Quantile-Quantile Plot. <http://www.itl.nist.gov/div898/handbook/eda/section3/qqplot.htm>. In: NIST/SEMATECH (Hrsg.): *NIST/SEMATECH e-Handbook of Statistical Methods*, Kapitel 1.3.3.24
- [NKCNAJ01] NG KUANG CHERN, Nathaniel ; NEOW, Poo A. ; ANG JR., Marcelo H.: Practical Issues in Pixel-Based Autofocusing for Machine Vision. In: *Proceedings of the 2001 IEEE International Conference on Robotics and Automation*. Seoul, Korea, Mai 2001, S. 2791–2796
- [OKS80] OHTA, Yu-Ichi ; KANADE, Takeo ; SAKAI, Toshiyuki: Color Information for Region Segmentation. In: *Computer Graphics and Image Processing* 13 (1980), S. 222–241
- [Pav78] PAVLIDIS, Theo: A review of algorithms for shape analysis. In: *Computer Graphics and Image Processing* 7 (1978), Nr. 2, S. 243–258
- [Pea01] PEARSON, Karl: On lines and planes of closest fit to systems of points in space. In: *Philosophical Magazine* 2 (1901), 559-572. <http://pbil.univ-lyon1.fr/R/liens/pearson1901.pdf>
- [PM90] PERONA, Pietro ; MALIK, Jitendra: Scale-Space and Edge Detection Using Anisotropic Diffusion. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)* 12 (1990), Nr. 7, S. 629–639
- [PM96] PROAKIS, John G. ; MANOLAKIS, Dimitris G.: *Digital Signal Processing*. 3rd. Upper Saddle River : Prentice Hall, 1996. – 1016 S. – ISBN 0-13-373762-4
- [Poya] POYNTON, Charles A.: *Frequently Asked Questions about Gamma*. <http://www.poynton.com/GammaFAQ.html>. – Online-Ressource, Abruf: Jan 2006
- [Poyb] POYNTON, Charles A.: *YUV and luminance considered harmful: A plea for precise terminology in video*. http://www.poynton.com/papers/YUV_and_luminance_harmful.html. – Online-Ressource, Abruf: Jan 2006
- [Poy96] POYNTON, Charles A.: *A Technical Introduction to Digital Video*. New York : John Wiley & Sons, 1996. – ISBN 0-471-12253-X

- [Pre70] PREWITT, Judy M.: Object Enhancement and Extraction. In: LIPKIN, Bernice S. (Hrsg.) ; ROSENFELD, Azriel (Hrsg.): *Picture Processing and Psychopictorics*. New York : Academic Press, 1970, S. 75–149
- [PV00] PLATANIOTIS, Konstantinos N. ; VENETSANOPOULOS, Anastasios N.: *Color Image Processing and Applications*. Berlin : Springer, 2000. – 355 S. – ISBN 3–540–66953–1
- [Rei91] REISS, Thomas H.: The revised Fundamental Theorem of Moment Invariants. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)* 13 (1991), August, Nr. 8, 830-834. <http://csdl.computer.org/comp/trans/tp/1991/08/i0830abs.htm>
- [Rob63] ROBERTS, Lawrence G.: *Machine Perception of Three-Dimensional Solids*, MIT, Diss., Juni 1963. – 82 S
- [RW84] REDNER, Richard A. ; WALKER, Homer F.: Mixture densities, maximum likelihood, and the EM algorithm. In: *SIAM Review* 26 (1984), S. 195–239
- [SB59] STILES, W. S. ; BURCH, J. M.: NPL colour-matching investigation: Final report. In: *Optica Acta* 6 (1959), S. 1–26
- [SBS99] SÜSTRUNK, Sabine ; BUCKLEY, Robert ; SWEN, Steve: Standard RGB Color Spaces. In: *Seventh Color Imaging Conference: Color Science, Systems and Applications* Bd. 7. Scottsdale, Arizona, November 1999, S. 127–134
- [SBV95] SCHÖLKOPF, Bernhard ; BURGESS, Chris ; VAPNIK, Vladimir: Extracting Support Data for a Given Task. In: FAYYAD, Usama M. (Hrsg.) ; UTHURUSAMY, Ramasamy (Hrsg.): *Proceedings of the First International Conference on Knowledge Discovery and Data Mining*, AAAI Press, Menlo Park, California, 1995, S. 252–257
- [Sch06] SCHULZE, Christian: *Intelligente automatische Bildaufnahmesteuerung zum computergestützten Hochdurchsatz-Screening in einem biologischen System mittels Mikroskop-Farbbildanalyse*, Otto-von-Guericke-Universität Magdeburg, Diplomarbeit, Januar 2006. – 96 S
- [Sei04] SEIFFERT, Udo: Biologically Inspired Image Compression in Biomedical High-Throughput Screening. In: IJSPEERT, Auke J. (Hrsg.) ; MURATA, Masayuki (Hrsg.) ; WAKAMIYA, Naoki (Hrsg.): *Biologically Inspired Approaches to Advanced Information Technology* Bd. 3141. Heidelberg : Springer, Oktober 2004, S. 428–440. – ISBN 3–540–23339–3
- [Soi04] SOILLE, Pierre: *Morphological Image Analysis*. 2nd. Berlin : Springer, 2004. – ISBN 3–540–42988–3
- [SPSLD00] SCHWEIZER, Patrick ; POKORNY, Jana ; SCHULZE-LEFERT, Paul ; DUDLER, Robert: Double-stranded RNA interferes with gene function at the single-cell level in cereals. In: *Plant Journal* 24 (2000), Nr. 6, S. 895–903
- [SS05] SEIFFERT, Udo ; SCHWEIZER, Patrick: A Pattern Recognition Tool for Quantitative Analysis of In Planta Hyphal Growth of Powdery Mildew Fungi. In: *Molecular Plant-Microbe Interactions (MPMI)* 18 (2005), September, Nr. 9, S. 906–912

- [TA77] TIKHONOV, Andrey N. ; ARSENIN, Vasiliy Y.: *Solutions of Ill-posed Problems*. Washington, DC : Winston & Sons, 1977. – ISBN 0–470–99124–0
- [Tau04] TAUTENHAHN, Ralf: *Intelligente Algorithmen zum Finden von Haustorien in Mikroskopbildern*, Martin-Luther-Universität Halle-Wittenberg, Diplomarbeit, Dezember 2004. – 70 S
- [Ten70] TENENBAUM, Jay M.: *Accommodation in Computer Vision*, Stanford University, Diss., September 1970. – 452 S
- [Tik63] TIKHONOV, Andrey N.: Regularization of incorrectly posed problems. In: *Soviet Mathematics Doklady* 4 (1963), S. 1624–1627
- [TIS05] TAUTENHAHN, Ralf ; IHLOW, Alexander ; SEIFFERT, Udo: Adaptive Feature Selection for Classification of Microscope Images. In: *Proceedings of the International Workshop on Fuzzy Logic and Applications (WILF 2005)*. University of Milan, Crema, Italy, September 2005, S. 215–222
- [TK03] THEODORIDIS, Sergios ; KOUTROUMBAS, Konstantinos: *Pattern Recognition*. 2nd. Amsterdam : Academic Press, 2003. – 689 S. – ISBN 0–12–685875–6
- [TSV⁺98] TANG, Xiaoou ; STEWART, W. K. ; VINCENT, Luc ; HUANG, He ; MARRA, Marty ; GALLAGER, Scott M. ; DAVIS, Cabell S.: Automatic Plankton Image Recognition. In: *Artificial Intelligence Review* 12 (1998), Februar, Nr. 1-3, S. 177–199
- [Wei] WEISSTEIN, Eric W.: *Quadratic Surface*. <http://mathworld.wolfram.com/QuadraticSurface.html>. From MathWorld – A Wolfram Web Resource. – Online-Ressource, Abruf: Jan 2006
- [Wei98a] WEICKERT, Joachim: *Anisotropic Diffusion in Image Processing*. Stuttgart : Teubner, 1998 (ECMI Series). – 140 S. – ISBN 3–519–02606–6
- [Wei98b] WEICKERT, Joachim: Efficient and Reliable Schemes for Nonlinear Diffusion Filtering. In: *IEEE Transactions on Image Processing* 7 (1998), März, Nr. 3, S. 398–410
- [Wei98c] WEILER, Frank: *Modellbasierte Bildsegmentierung mit aktiven Kontourmodellen*, RWTH Aachen, Diss., März 1998. – 111 S
- [Wit83] WITKIN, Andrew P.: Scale-Space Filtering. In: *Proc. 8th International Joint Conference on Artificial Intelligence*. Karlsruhe, Germany, 1983, S. 1019–1022
- [WVM04] WERUAGA, Luis ; VERDÚ, Rafael ; MORALES, Juan: Frequency Domain Formulation of Active Parametric Deformable Models. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)* 26 (2004), Dezember, Nr. 12, S. 1568–1578
- [WW80] WALLACE, Timothy P. ; WINTZ, Paul A.: An Efficient Three-Dimensional Aircraft Recognition Algorithm Using Normalized Fourier Descriptors. In: *Computer Graphics and Image Processing* 13 (1980), S. 99–126
- [XP98] XU, Chenyang ; PRINCE, Jerry L.: Snakes, Shapes, and Gradient Vector Flow. In: *IEEE Transactions on Image Processing* 7 (1998), März, Nr. 3, S. 359–369
- [You02] YOUNG, Thomas: On the theory of light and colours. In: *Philosophical Transactions of the Royal Society of London* 92 (1802), S. 12–48

LITERATURVERZEICHNIS

- [YWB74] YOUNG, Ian T. ; WALKER, Joseph E. ; BOWIE, Jack E.: An analysis technique for biological shape I. In: *Information and Control* 25 (1974), August, Nr. 4, S. 357–370

Lebenslauf

persönliche Angaben

Name	Alexander Ihlow
Geboren	08.04.1976 in Magdeburg
Familienstand	ledig

Schulbildung

1982 – 1991	Polytechnische Oberschule in Ballenstedt
1991 – 1994	Wolterstorff-Gymnasium Ballenstedt
	Abitur 1994

Hochschulausbildung

1994 – 1999	Studium der Elektrotechnik an der TU Ilmenau, Fachrichtung Informationstechnik
April – Oktober 1999	Diplomarbeit „ <i>Einfluss realer Umgebungsbedingungen auf die Richtcharakteristik linearer Antennenarrays</i> “ Diplom 1999

Berufstätigkeit

01.11.1999 – 31.12.2002	wissenschaftlicher Mitarbeiter im Fachgebiet Nachrichtentechnik an der TU Ilmenau
seit 01.01.2003	wissenschaftlicher Mitarbeiter in der Arbeitsgruppe Mustererkennung am IPK Gatersleben

Magdeburg im Juli 2006