

Aus dem Helmholtz Zentrum München  
Deutsches Forschungszentrum für Gesundheit und Umwelt  
Institut für Biologische und Medizinische Bildgebung  
(Institutsdirektor: Prof. Dr. Vasilis Ntziachristos)

# **Inhaltsbasierte Autofokussierung in der automatisierten Mikroskopie**

Dissertation  
zum Erwerb des Doktorgrades der Humanbiologie  
an der Medizinischen Fakultät der  
Ludwig-Maximilians-Universität zu München.

Vorgelegt von:

Peter Hamm

aus Warschau

Jahr:

2011

Mit Genehmigung der Medizinischen Fakultät  
der Universität München

<b>Berichterstatter:</b>	Prof. Dr. Karl-Hans Englmeier
<b>Mitberichterstatter:</b>	Prof. Dr. Jochen Herms Prof. Dr. Christoph Schmitz
<b>Dekan:</b>	Prof. Dr. med. Dr. h. c. M. Reiser, FACR, FRCR
<b>Tag der mündlichen Prüfung:</b>	03.11.2011

---

---

Die vorliegende Dissertation entstand während meiner Tätigkeit als wissenschaftlicher Mitarbeiter am Institut für Biologische und Medizinische Bildverarbeitung (Leiter: Prof. Dr. Vasilis Ntziachristos) in der Arbeitsgruppe Bild- und Signalverarbeitung unter Herrn Prof. Dr. Karl-Hans Englmeier in Zusammenarbeit mit Carl Zeiss MicroImaging GmbH München.

The experimental work was done at the Institute for Biological and Medical Imaging of the Helmholtz Zentrum München (Head: Prof. Dr. Vasilis Ntziachristos) within the research unit 'Image and Signal Processing' under the supervision of Prof. Dr. Karl-Hans Englmeier in cooperation with Carl Zeiss MicroImaging GmbH Munich.



Meinen Eltern

# Inhaltsverzeichnis

<b>Abbildungsverzeichnis</b>	<b>8</b>
<b>Tabellenverzeichnis</b>	<b>10</b>
<b>Abkürzungsverzeichnis</b>	<b>12</b>
<b>1 Zusammenfassung und Abstract</b>	<b>14</b>
<b>2 Einleitung</b>	<b>16</b>
2.1 Motivation . . . . .	16
2.2 Problemstellung . . . . .	17
2.3 Zielsetzung . . . . .	18
<b>3 Grundlagen</b>	<b>19</b>
3.1 Geschichte des Autofokus . . . . .	19
3.2 Autofokus-kategorien . . . . .	22
3.2.1 Aktiver Autofokus . . . . .	23
3.2.2 Passiver Autofokus . . . . .	25
3.3 Herausforderungen für einen Autofokus . . . . .	29
3.3.1 Mikroskop-Backlash und Drift . . . . .	29
3.3.2 Sensor-Rauschen und Sensor-Fehler . . . . .	30
3.3.3 Artefakte . . . . .	30
3.3.4 Vielschichtige Proben und Präparate hoher Schichtdicke . . . . .	32
<b>4 Stand der Forschung</b>	<b>34</b>
4.1 Algorithmen passiver Methoden . . . . .	34
4.1.1 Funktionen im Ortsfrequenzraum . . . . .	35
4.1.2 Gradienten-Berechnung . . . . .	37
4.1.3 Funktionen der Bildstatistik . . . . .	39

4.2	Strategien der Fokussuche . . . . .	40
4.2.1	Suche mit fester und variabler Schrittweite . . . . .	40
4.2.2	Einfache iterative Suche . . . . .	41
4.2.3	Fibonacci-Suche . . . . .	41
4.2.4	Suche mit Fokuskurven-Interpolation . . . . .	42
4.3	Autofokussierung unter Berücksichtigung von Artefakten und hoher Proben- dicke . . . . .	43
<b>5</b>	<b>Material und Methoden</b>	<b>45</b>
5.1	Konzept des Inhaltsbasierten Autofokus . . . . .	45
5.1.1	Modulare Struktur . . . . .	45
5.2	Universelle Objekterkennung in mikroskopischen Proben . . . . .	47
5.2.1	Merkmalsextraktion mit Haar-Filtern . . . . .	48
5.2.2	Merkmalsselektion mit Adaptive Boosting . . . . .	51
5.3	Wahl der Schärfefunktion . . . . .	53
5.4	Konfiguration der Gesamtmethode . . . . .	56
5.5	Umsetzung des IB-AF . . . . .	57
5.5.1	Implementierung in die Mikroskop-Software . . . . .	57
5.5.2	User-Interface . . . . .	58
<b>6</b>	<b>Ergebnisse</b>	<b>63</b>
6.1	Auswertungsverfahren . . . . .	63
6.2	Ergebnisse . . . . .	64
6.2.1	Fokussierungsqualität bei verrauschten Bildern . . . . .	65
6.2.2	Fokussierungsvergleich zwischen Ganzbild- und ROI-Analyse . . . . .	65
6.2.3	Abhängigkeit der Resultate von Trainings- und Klassifikationspa- rametern . . . . .	69
6.2.4	Studie am Modell . . . . .	75
6.2.5	Resultate des IB-AF bei biologischen Präparaten . . . . .	76
6.2.6	Resultate - Material . . . . .	79

<b>7 Diskussion</b>	<b>82</b>
<b>8 Schluss und Ausblick</b>	<b>86</b>
<b>9 Anhang</b>	<b>88</b>
9.1 Zusätzliche Abbildungen und Fokuskurven . . . . .	88
9.2 Verfahren der Merkmalsextraktion . . . . .	92
9.2.1 Principle Components Analysis (PCA) . . . . .	92
9.2.2 Scale Invariant Feature Transform (SIFT) . . . . .	93
9.3 Merkmalsselektionsverfahren . . . . .	96
9.3.1 Support Vector Machine (SVM) . . . . .	96
9.4 Veröffentlichungen und Patentanmeldung . . . . .	102
9.5 Danksagung . . . . .	103
9.6 Curriculum Vitae . . . . .	104
<b>Literatur</b>	<b>106</b>
<b>Stichwortverzeichnis</b>	<b>110</b>

## Abbildungsverzeichnis

1	Hopkins automatisch fokussierende Vergrößerungsapparatur . . . . .	20
2	Kodak A3 Autographic Special und der im Benutzerhandbuch erläuterte Bildmesssucher. . . . .	21
3	Automatische Fokussierungsapparatur von Reinheimer und Leitz . . . . .	22
4	Distanzmessung per Laserstrahl und Triangulationsmessung . . . . .	24
5	Phasenvergleichsmethode zur Bestimmung der Fokusslage . . . . .	25
6	Gegenüberstellung von Bild und Transformation in den Ortsfrequenzraum	27
7	Sobel-Operator im Bild- und Ortsfrequenzraum . . . . .	28
8	Verdeutlichung des Backlashs an einer Stellschraube . . . . .	29
9	Artefakte bei Phasenkontrast- und Fluoreszenzpräparaten . . . . .	31
10	Kristall-Artefakt und Phasen-Artefakte . . . . .	32
11	HeLa-Probe mit Zellen in drei unterschiedlichen Fokusebenen . . . . .	33
12	Iterationsschritte der Fibonacci-Suche . . . . .	42
13	Module der Inhaltsbasierten Fokussuche . . . . .	46
14	Schritte der Inhaltsbasierten Autofokussmethode . . . . .	47
15	Haarfilter-Typen und ihre mögliche Positionierung im Bild bzw. Bildaus- schnitt . . . . .	49
16	Berechnung der Haar-Filter-Antwort mit Hilfe des Integralbildes . . . . .	50
17	Kaskadierung von Klassifikatoren in einzelne Stufen . . . . .	52
18	Abhängigkeit der gefundenen Fokusebene vom Schwellwert . . . . .	53
19	Fokusergebnisse bei einer stark verrauschten Fluoreszenz-Beads-Probe . .	54
20	Die zu Abbildung 21 gehörigen Fokusskurven . . . . .	55
21	Schärfemessung innerhalb von drei ROIs in einer Caco2-Fluoreszenz-Probe	55
22	Ablauf und eingesetzte Methoden und Daten der Gesamtmethode . . . .	57
23	Markierung von Bildregionen in der Entwickleroberfläche . . . . .	59
24	Paneele der Entwickler- und Benutzer-GUI . . . . .	61
25	Bildrauschen im Teststapel . . . . .	65

26	Fokuskurven bei verschiedenen Schärfefunktionen . . . . .	66
27	Fokuskurven verschiedener Schärfefunktion bei ganzheitlicher Bildbetrachtung . . . . .	66
28	Resultierende Fokuskurven für Einzelbildbereiche . . . . .	67
29	CaCo-Zellen im FITC- und DAPI-Kanal . . . . .	68
30	Cytoskelett im DsRed- und DAPI-Kanal . . . . .	69
31	Cytoskelett im Phasenkontrast und Beads-Probe . . . . .	70
32	CaCo-Zellpräparat mit einem interessierenden Objekt (hellere Intensität in der rechten Bildmitte), das DNA-Strukturen zeigt . . . . .	72
33	Anzahl der Hypothesen in Abhängigkeit vom Abstand zur Fokusebene . . . . .	73
34	Diagramm der in Tabelle 8 enthaltenen Ergebnisse bei Änderung der Trainingsbildgröße . . . . .	74
35	Einzelbilder des Modell-Stapels . . . . .	76
36	Autofokussierung an Modellobjekten . . . . .	77
37	Objekthypothesen und Verifikation mit Fokusebenenkoordinate . . . . .	77
38	Auswahl von vier Beispielen der CaCo-2-Testreihe . . . . .	79
39	Zwei Beispiele der Zellkolonien-Proben . . . . .	79
40	Repräsentative Auswahl des Material-Probensatzes . . . . .	80
41	Fokuskurven für das Kristall-Artefakt-Präparat im Phasenkontrast-Modus . . . . .	88
42	Fokuskurven für das Phasen-Artefakte-Präparat im Phasenkontrast-Modus . . . . .	89
43	Mitotische HeLaR-Zellen im Fluoreszenz-Kanal in verschiedenen Fokusebenen . . . . .	90
44	Kristall-Artefakt im Phasenkontrast-Modus . . . . .	90
45	Phasen-Artefakte im Phasenkontrast-Modus . . . . .	90
46	Fokuskurven des Modellstapels . . . . .	91
47	Bestimmung der Hauptgradientenorientierung über Gradientenbetrags-histogramm . . . . .	95
48	Gradientenhistogramme für die Merkmalsvektorberechnung bei SIFT . . . . .	96
49	Trennebene in einem linearen Merkmalsraum mit linearer Kernel-Funktion und Hard-Margin-Bedingung . . . . .	100

## Tabellenverzeichnis

1	Fibonacci-Folge bis zur 12. Stelle. . . . .	41
2	Ablauf des AdaBoost-Algorithmus mit Haar-Filtern als so genannte schwache Klassifikatoren (weak classifiers) . . . . .	52
3	Autofokusergebnisse der Analyse kompletter Ansichten #1 . . . . .	68
4	Autofokusergebnisse der Analyse kompletter Ansichten #2 . . . . .	69
5	Autofokusergebnisse der Analyse von bestimmten Bildregionen #1 (Beispielebene siehe Abb. 21) . . . . .	70
6	Autofokusergebnisse der Analyse von bestimmten Bildregionen #2 . . . . .	71
7	Autofokusergebnisse der Analyse von bestimmten Bildregionen #3 . . . . .	71
8	Erkennungsrate in Abhängigkeit von der Herunterskalierung der Trainingsbilder . . . . .	72
9	Detektionsraten in Abhängigkeit von der Trainingsdatensatzgröße . . . . .	72
10	Trainingsdauer und Detektionsraten in Abhängigkeit von der minimalen Richtig-Positiv-Rate . . . . .	73
11	Trainingsdauer und Detektionsraten in Abhängigkeit von der maximalen Falsch-Positiv-Rate . . . . .	74
12	Detektionsrate und Berechnungsdauer in Abhängigkeit vom Skalierungsfaktor . . . . .	75
13	Vergleich der Erkennungsraten des Inhaltsbasierten Autofokus mit klassischen AF-Methoden . . . . .	78
14	Klassifikationsergebnisse des IB-AF auf CaCo-2-Testreihe . . . . .	78
15	Vergleich der Erkennungsraten des Inhaltsbasierten Autofokus mit klassischen AF-Methoden bei Zellkolonie-Proben . . . . .	78
16	Klassifikationsergebnisse auf den vom Hypothesenmodul detektierten und von der Verifikation eingeordneten ROIs durch das Inhaltsbasierte AF-System bei Zellkolonie-Proben. . . . .	79
17	Vergleich der Erkennungsraten des Inhaltsbasierten Autofokus mit klassischen AF-Methoden für die Materialprobe. . . . .	81
18	Klassifikationsergebnisse auf den vom Hypothesenmodul detektierten und von der Verifikation eingeordneten ROIs durch das Inhaltsbasierte AF-System. . . . .	81

19	Ablauf der Berechnung der Transformationsmatrix auf der Basis der Hauptkomponenten (principle components) . . . . .	93
----	---	----

## Abkürzungsverzeichnis

AdaBoost Adaptive Boosting

AF Autofokus

CCD Charge Couple Device

CMOS Complementary Metal Oxide Semiconductor

DAB Discrete AdaBoost

DAPI 4,6-Diamidin-2-phenylindol

DIC Differentieller Interferenzkontrast

DIN Deutsches Institut für Normung

DLL Dynamic Link Library

DoG Difference of Gaussian

FITC Fluoreszeinisothiocyanat

FN Falsch-Negative

FP Falsch-Positive

GAB Gentle AdaBoost

GUI Graphical User Interface

HCS High Content Screening

HoG Histogram of Gradients

HTS High Throughput Screening

IB-AF Inhaltsbasierter Autofokus

IPL Image Processing Library

IR Infrarot

LED Licht Emittierende Diode

LG Logit Boost

LoG	Laplacian of Gaussian
LSM	Least Square Minimum
NA	Numerische Apertur
OpenCV	Open Source Computer Vision
OTF	Optical Transfer Function
PCA	Principle Components Analysis
PSF	Point Spread Function
RAB	Real AdaBoost
RBF	Radial Basis Function
RN	Richtig-Negative
ROI	Region Of Interest
RP	Richtig-Positive
SIFT	Scale Invariant Feature Transform
STED	Stimulated Emission Depletion
SVM	Support Vector Machine
XML	Extensible Markup Language

# 1 Zusammenfassung und Abstract

## Zusammenfassung

Die automatische Fokussierung ist ein grundlegender Arbeitsschritt für die Bildaufnahme und –auswertung mit motorgetriebenen Mikroskopen. Auch wenn die Forschungs- und Entwicklungsarbeit auf dem Gebiet des kontrastbasierten Autofokus nunmehr auf eine viele Dekaden lange Geschichte zurückblicken kann, fehlt es selbst aktuellen Methoden an Robustheit gegenüber Bildstörungen und der Handhabung komplexerer Präparatstrukturen. Diese Dissertation stellt einen neuen Autofokusansatz vor, der grundsätzlich mit jeder Mikroskopieart wie unter anderem Fluoreszenz-, Hellfeld- oder auch Phasenkontrast-Mikroskopie verwendet werden kann. Die Neuheit der Methode besteht in einer inhaltsbasierten Fokussuche, die für eine gezieltere Autofokussierung Vorwissen über das zu untersuchende Präparat verwendet. Dabei stellen die von den Trainingsdaten extrahierten und per Boosting selektierten lokalen Haar-Merkmale die Wissensbasis. Die im folgenden als Inhaltsbasierte Autofokus (IB-AF) bezeichnete Methode verfährt in drei Schritten: Zuerst werden an beliebiger z-Koordinate innerhalb des Präparats Regions-of-Interest (ROI) ermittelt, die starke Objekthypothesen enthalten. Danach wird nur auf diesen Regionen eine Kontrastmessung zur Fokusalagenfindung der jeweiligen Region entlang der z-Achse durchgeführt. Im letzten Schritt werden die Regionen in ihrer Fokusalage einer genaueren Verifikation unterzogen, um ergänzend etwaige uninteressante Objekte auszuschließen. Mit dieser Herangehensweise wendet sich der IB-AF von traditionellen Methoden ab, welche den gesamten Bildbereich einer Fokussuche unterziehen. Dadurch ist es möglich, sowohl Artefakte aus der Schärfemessung auszuschließen, als auch gezielt spezifische Objekte in den Fokus zu bringen. Die vorgestellte Methode wurde auf Präparaten mit unterschiedlichen Herausforderungen getestet und erzielte ein erfolgreiches Fokussieren, wo andere Methoden bisher scheiterten.

## Abstract

Autofocusing is a fundamental step when it comes to image acquisition and analysis with automated microscopy devices. But despite all efforts that have been put into developing a reliable autofocus system, recent methods still lack robustness towards artifacts and complex specimen structures. This thesis paper presents a novel automated focusing approach that is generally applicable to different microscope modes such as brightfield, phase contrast, Differential Interference Contrast (DIC) and fluorescence microscopy. The main innovation consists in a content-based focus search that makes use of a priori

knowledge about the observed objects by employing local object features (Wavelets) and Boosted Learning. The autofocus routine is processed in three steps. First, image regions that contain strong object hypothesis are detected on which, in the second step, image contrast measurement is conducted to find the respective in-focus layer along the Z-axis. In the third step, these areas are verified in their respective in-focus layer. Hence, this method turns away from common autofocus approaches that apply holistic image frequency measurements to obtain the focus plane. Thus, it is possible to exclude artifacts from being brought into focus calculation as well as locating the in-focus layer of specific microscopic objects. The proposed method was tested on data sets of different specimen types. The method proofed to cope with focus difficulties that cause other algorithms to fail and was able to detect the focal planes only of desired objects.

## 2 Einleitung

### 2.1 Motivation

Die Mikroskopie ist zweifelsohne das wichtigste bildgebende Instrument in Biologie und Medizin. Die Verwendung von Mikroskopen hat einen stetig wachsenden Anteil an den Forschungsfortschritten in den Bereichen der Proteomik, Genomik, Biochemie und Molekularbiologie. Die hierbei erzielten Ergebnisse sind ein entscheidender Motor sowohl für die kontinuierliche Verbesserung der medizinischen Diagnostik, als auch für die Entwicklung von neuen Medikamenten und Therapien.

In der Biomedizin hat sich durch die technische Evolution in den letzten Jahren und Jahrzehnten die Arbeitsweise mit den modernen Mikroskopen geändert. Vor allem durch die Entwicklungen in der Computertechnik und im Bereich der CCD- und CMOS-Kameras rückt die digitale Bildverarbeitung als Mittel zur Probenanalyse in den Vordergrund. Die Fluoreszenz-Mikroskopie profitiert zudem von der zunehmenden Anzahl von Fluoreszenz-Markern und den Neuerungen der Laser-Technik ([17]). Die fortschreitende Technik bietet dabei die Voraussetzung, den Untersuchungsprozess mit Mikroskopen weitestgehend zu automatisieren und um ein Vielfaches zu beschleunigen; dies sowohl in quantitativer (z.B. die Bildaufnahme) als auch in qualitativer Hinsicht wie die Analyse von Zellen ([24, 32]). In diesem Zusammenhang spricht man einerseits von High-Throughput-Screening (HTS), das die Untersuchung von größeren Probensätzen (Assays) mit hohem Durchsatz realisiert, und andererseits von High-Content-Screening (HCS), welches die automatisierte Interpretation von Datensätzen mit hohem Informationsgehalt beschreibt. Gegenwärtige kommerzielle HTS-Systeme sind in der Lage, bis zu 100 Mikrotiterplatten im 96-Well-Format vollständig automatisch zu bearbeiten. Typische HCS-Anwendungen sind die Quantifizierung von Zellen bestimmter Ausprägung oder auch die Zell-Zyklus-Analyse von markierten Proben. Letzteres Beispiel steht für eine typische 4D-Anwendung, bei der die Analyse einer dreidimensionalen Probe um die Dimension der Zeit ergänzt wird.

Über Medizin und Biologie hinaus spielt die Mikroskopie eine tragende Rolle auf dem Gebiet der Materialwissenschaften und Werkstofftechnik. Bei industriellen Produktionsprozessen werden Mikroskope für qualitätssichernde und schadenanalysierende Untersuchungen von Materialoberflächen herangezogen ([28]). In der Geologie werden Polarisationsmikroskope eingesetzt, um Eigenschaften wie Brechungsindizes und Schwingungsverhalten von Kristallen bzw. Mineralien zu bestimmen ([38]). Die Polarisationsmikroskopie findet ebenfalls Einsatz in der Eigenschaftsbestimmung von Polymeren. Weitere Einsatzgebiete reichen von Lebensmittelprüfung über kriminologische Untersuchungen zu einer

hier nicht aufzählbaren Liste von Untersuchungsfeldern.

Der initiale Arbeitsschritt in den oben beschriebenen Automatisierungen besteht darin, die zu untersuchende Probe in den Fokus zu bringen, d.h. scharf auf dem Kamera-Sensor abzubilden. Alle folgenden Prozesse, wie z.B. eine konkrete Bildanalyse, schließen sich diesem Vorgang an und sind außerordentlich vom Ergebnis der Fokussierung abhängig. In der Praxis wird das Fokussierungsproblem auf zwei Wegen zu lösen versucht: Aktive und Passive Autofokusse. Ersterer bezeichnet die Fokussierung mittels eines infrarot-nahen Lasers, der die Abstandsmessung der Linse zum Objekt ermöglicht. Passive Fokussierungstechniken umfassen die Bestimmung der Fokusebene über die Analyse des aufgenommenen Bildes. Hierbei wird vornehmlich der Kontrast unter der Annahme betrachtet, dass ein hoher Kontrast einer hohen Bildschärfe entspricht. Auf dem Markt befinden sich Systeme, die entweder passive oder aktive Fokussmethoden oder eine Kombination beider verwenden.

### 2.2 Problemstellung

Es existiert eine Vielzahl von Einflüssen, die eine exakte Fokuseinstellung während mikroskopischer Untersuchungen verhindern. Diese Einflüsse können systemischen Ursprungs sein oder in der Beschaffenheit und Präparation der zu untersuchenden Probe liegen. Durch die Probenherstellung verursachte Artefakte führen objektfremde Strukturen in das Präparat ein und bewirken eine Addition von Bildinformationen, die vor allem für die vorhandenen bildkontrastbasierten Autofokusverfahren eine erhebliche Störgröße bedeuten. Ebenso problematisch sind Proben, die Untersuchungsobjekte in mehreren Fokusebenen enthalten. Die Bestimmung der Fokusebene mittels aktiver Fokussysteme als Basis setzt eine eindeutige Oberfläche im Präparat voraus, welche die elektromagnetischen Wellen (Laser) reflektiert. In der Lichtmikroskopie ist diese Oberfläche zumeist der Objektträger selbst, so dass Objekte wie z.B. Zellen unscharf abgebildet werden, wenn sie nicht auf dem Träger fixiert sind. Die eben zuvor erwähnte signifikante Dicke mikroskopischer Präparate, die hohe Vergrößerungen erfordern, machen aktive Fokussysteme unpraktikabel. Bildbasierte passive Fokustechniken dagegen bieten das Potential, die genannten Schwierigkeiten zu bewältigen. Jedoch ist der überwiegende Teil der bestehenden Verfahren unter anderem dadurch limitiert, dass lediglich eine Fokuslage ausgegeben wird. Es ist deshalb in der Praxis gängige Vorgehensweise, diesen Problemen aus dem Weg zu gehen, indem so lange fehlerfreie Proben gesammelt werden, bis eine statistisch auswertbare Anzahl von Bildaufnahmen vorliegt. Das Problem wird damit nicht gelöst, es entsteht jedoch auch kein unmittelbarer Schaden für die wissenschaftliche Auswertung der Aufnahmen. Anders verhält es sich jedoch, wenn der Fokussierungsfehler

vom Wissenschaftler unbemerkt in der anschließenden Analyse zum Tragen kommt. Zur Veranschaulichung sei folgendes Beispiel erläutert: Die Beobachtung des Zellwachstums von immortalisierten Zellkulturen ist eine der wichtigsten Untersuchungsgegenstände in der biomedizinischen Wissenschaft. Dabei wird anhand der Zellform und -struktur eine Aussage über das Zellstadium getroffen. Die Aussage, ob es sich um sich teilende, apoptotische oder nekrotische Zellen handelt ([54]), gibt Aufschluss z.B. über die Wirksamkeit eines Zytostatikums, das das Wachstum von kanzerogenen Zellen hemmt. Solche Untersuchungen erfolgen unüberwacht mit hohem Proben-Durchsatz ([11]). Sowohl die durch einen Autofokus unterstützte Bildaufnahme, als auch die Bildanalyse, im vorliegenden Fall die Zellsegmentierung und Forminterpretation, vollziehen sich automatisch. Dem Forscher präsentieren sich lediglich die quantitativen Ergebnisse, die er im Anschluss auswertet. Nun kann allerdings eine z.B. durch Artefakte induzierte Fehlfokussierung um lediglich einen geringen Versatz zur Verwischung von Strukturen und Formen führen. Die darauf folgende Bildanalyse liefert unweigerlich falsche Ergebnisse und deren Interpretation durch den Wissenschaftler führt am Ende zu falschen Aussagen.

### 2.3 Zielsetzung

Das Ziel dieser Forschungsarbeit war die Entwicklung geeigneter Methoden und Algorithmen, die eine automatische Bestimmung der optimalen Fokusposition(en) eines Präparates bei motorisierten Mikroskopen ermöglichen. Zusätzlich soll das Verfahren generell auf verschiedene Untersuchungsobjekte und bei verschiedenen Mikroskop-Modi anwendbar sein. Das Hauptaugenmerk liegt auf schwierigen Proben, die entweder aufgrund komplexer Strukturen oder durch das Vorhandensein von irrelevanten Elementen neben den eigentlichen Untersuchungsobjekten Standard-Autofokusroutinen in ihrer Funktion beeinträchtigen. Die Ursache der Funktionseinschränkungen aufgrund von Artefakten ist nicht unbedingt bei den jeweiligen Umsetzungen der existierenden Algorithmen zu suchen, sondern liegt vielmehr im Ansatz, den alle existierenden Methoden gemein haben. Denn die Fokus-Messergebnisse beruhen allein auf der Kontrastanalyse des gesamten Bildes, ohne zwischen verschiedenen Bildobjekten zu differenzieren. Auf diese Weise liefert auch ein Staubkorn Kontrastwerte, die in die Berechnung einfließen. Der hier gewählte Ansatz soll sich auf eine Inhaltsbasierte Autofokus-Suche stützen, bei der nur relevante Objekte in der mikroskopischen Probe in die Fokusmessung eingehen dürfen. Das Vorwissen über diese relevanten Objekte soll der Fokussierungsroutine über eine vorangegangene einmalige Klassifikation bereitgestellt werden. Auf diese Weise kann die Schärfefunktion konkret auf bestimmte lokale Bereiche im Präparat eingeschränkt werden.

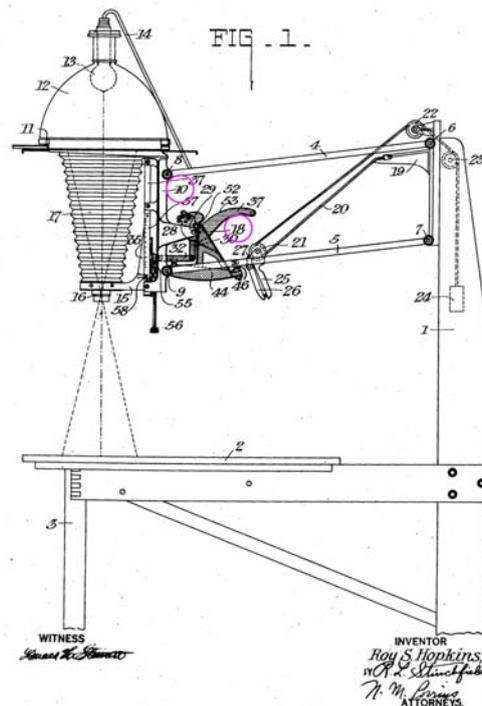
### 3 Grundlagen

Definiert man den Begriff Autofokus als das selbstständige Einstellen der Entfernung eines optischen Apparats zum Betrachtungsobjekt, dann lässt sich der Beginn seiner historischen Entwicklung bis zum Anfang des 20. Jahrhunderts zurückverfolgen. Der Abschnitt 3.1 beschäftigt sich ausführlich mit der Geschichte des AF. Diesem Abschnitt schließt sich die Erklärung der zwei grundsätzlichen Autofokusmethoden an. Man unterscheidet zwischen den aktiven und den passiven AF-Systemen, wobei der in dieser Arbeit vorgestellte bildbasierte AF-Ansatz zu den passiven Methoden zählt. Das Grundlagenkapitel endet mit den Erläuterungen zu den vielseitigen unerwünschten Einflüssen, die erhebliche Herausforderungen für alle AF-Funktionen darstellen.

#### 3.1 Geschichte des Autofokus

Ein Autofokus (AF) in der Optik bezeichnet eine Vorrichtung, die selbstständig, also automatisch auf ein Objekt scharfstellt, d.h. die Entfernung zwischen den Linsen und dem Objekt einstellt. In dieser Bedeutung wird der Begriff AF erstmals im Jahr 1921 in einer Patentschrift von Edwin Fritts ([15]), Eastman Kodak Company, verwendet. Darin wird eine Apparatur zur Vergrößerung von Negativen auf eine Photoplatte beschrieben. Durch entsprechende Mechaniken wird dabei in Abhängigkeit von der gewählten Vergrößerung der korrekte Abstand zwischen Negativ und Linsensystem beibehalten, und somit eine fokussierte Projektion garantiert. Noch früher ist die AF-Vorentwicklung aus dem Jahre 1917 von Roy S. Hopkins erwähnt, die im Jahr 1922 als Patent mit dem Titel „Automatically-Focusing Camera“ ([23]) angemeldet wurde. Auch sie beschreibt ein sich selbst fokussierendes Vergrößerungsgerät (siehe Abb. 1).

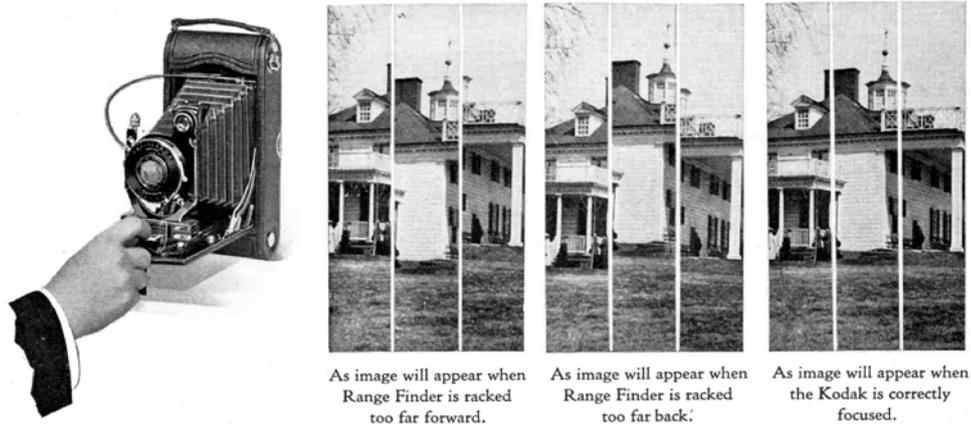
Bei diesen Projektionsgeräten war es aufgrund des mehr oder weniger bekannten Objektabstands möglich, mit einfachen Mechaniken automatisch eine scharfe Abbildung zu erzielen. Die Entwicklung automatisierter Fokuseinstellungen vor allem im Bereich der Photographie nahm den Zwischenschritt über Hilfssysteme zur Fokuseinstellung, wie sie in so genannten Messsucherkameras verbaut wurden. Bei diesen Kameras erhielt der Anwender eine optische Rückmeldung über ein Suchfenster, das ihm die korrekte Entfernungseinstellung des Objektivs erleichterte. Solche Messsucher arbeiteten mit einem Schnitt- bzw. Mischbildentfernungsmesser, der über ein Spiegelsystem unterschiedliche Strahlengänge vom Aufnahmeobjekt zusammen lenkt und bei richtiger Objektiveneinstellung ein einheitliches Bild im Sucher erscheinen ließ. Die erste in den Markt eingeführte Messsucherkamera war 1916 die Kodak A3 Autographic Special ([12]), deren Zeichnung die Abbildung 2 darstellt. Die ersten 35mm Kameras mit ausgestatteten Messsuchern



**Abbildung 1: Zeichnung aus Hopkins Patentanmeldung für eine automatisch fokussierende Negativ-Vergrößerungskamera. Die mit einem Kreis gekennzeichneten Bauteile sind in der Patentschrift als die Teile benannt, die für den korrekten Abstand sorgen.**

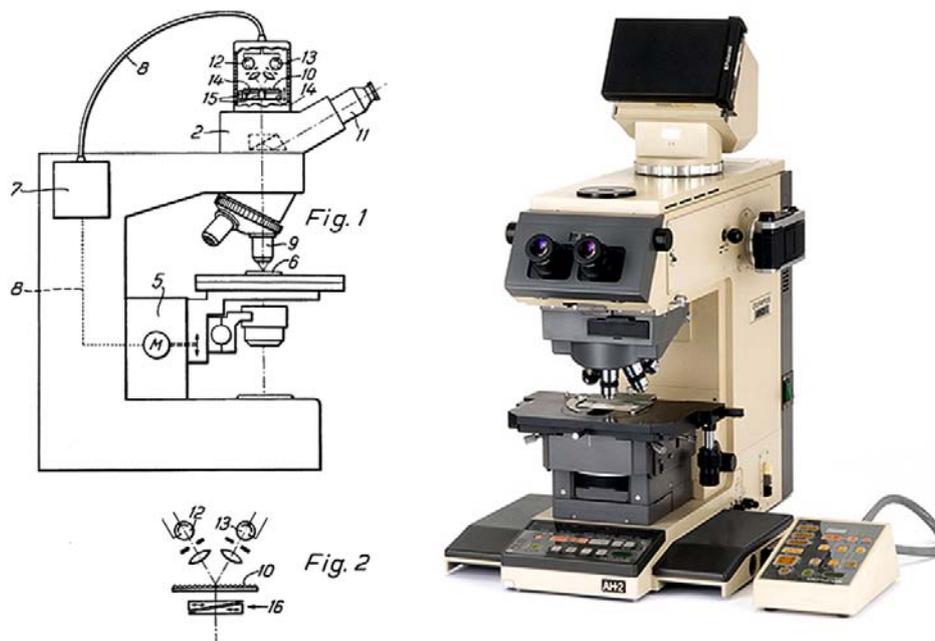
wurden ab 1932 von Leica mit der Leica II und Zeiss mit der Contax I verkauft. Der erste Autofokus ohne jegliche manuelle Unterstützung geht auf die Leitz Wetzlar GmbH und den Erfinder Ludwig Leitz im Jahr 1960 zurück ([29]). Auf diesem Patent aufbauend veröffentlichten - ebenfalls unter der Firma Leitz - die Erfinder Fromund Hoch und Dr. Gustav Weissenberg die deutsche Patentschrift „Verfahren und Vorrichtung zur Entfernungsmessung“ ([22]) und beschrieben darin ein passives Fokussystem mit zwei Strahlendetektoren, deren Signale sich bei Fokusslage gegenseitig kompensieren - und erfanden damit den ersten Autofokus auf Basis des Phasenvergleichs (siehe Abschnitt 3.2.2). Die Firma Leitz war es auch, die in den Folgejahren durch Werner Holle, Ludwig Leitz und Gunter Reinheimer die Entwicklung im Bereich der automatischen Schärfereinstellung maßgeblich vorantrieben. Im Mai 1974 erfolgte die Anmeldung eines „Automatic Microscope Focussing Device“ ([42]) beim amerikanischen Patentamt, das auf dem oben genannten Phasenvergleich beruhte, aber speziell für die Ausstattung und Funktion mit einem motorgesteuerten Objektträger eines Mikroskops vorgestellt wurde (siehe Abbildung 3).

Diesem Patent folgten wenige Jahre später auch Prototypen. Nichtsdestoweniger war es die Firma Olympus, die erst 1983 das erste serienmäßige Mikroskop mit verbautem Autofokus produzierte. Mitte der 70er Jahre spielte sich bei den Herstellern von Verbrau-



**Abbildung 2: Kodak A3 Autographic Special und der im Benutzerhandbuch erläuterte Bildmessucher.**

cherkameras eine ganz eigene Geschichte mit weitreichenden Konsequenzen ab. Zwar war die Firma Leitz zweifelsohne der Erfinder des Autofokus, doch ihre Patente waren lediglich in Deutschland angemeldet. Während Leitz Wetzlar 1976 mit dem Correfot CK-2 eine Kamera mit passivem Autofokus auf der Kölner Photokina präsentierte, entwickelte die Firma Honeywell Inc. mit ihrem Entwickler Norman L. Stauffer Autofokus-Module. Die Patentschriften „Focus Detecting Apparatus“ ([49]) und „Autofocus Camera“ ([48]) stammen aus dem Jahr 1975. Honeywells Autofokus-Methode arbeitete prinzipiell nach derselben phasenvergleichenden Methode, auch wenn sie technisch gegenüber dem Correfot anders realisiert wurde. Honeywell selbst baute ab 1975 lediglich das Autofokus-Modul, bekannt unter dem Namen „Visiotronic“; das aber aufgrund der geringen Stückzahlkosten von 10\$ großen Erfolg hatte. Dennoch verletzte das Visiotronic-System offensichtlich die Patente von Leitz. Ein sich anbahnender Patentstreit endete schließlich im Verkauf der Leitz-Patente an Honeywell im Jahre 1978. Die erste käufliche Photokamera mit Autofokus (Konica C35-AF) erschien 1977 mit eben diesem Visiotronic-System von Konica. Die Firma Minolta, welche ab 1976 mit Leitz kooperierte, brachte 1985 mit der Minolta 7000 die weltweit erste Spiegelreflexkamera mit im Gehäuse integriertem Autofokus auf den Markt. Die Kamera verwendete eine Weiterentwicklung einer von Honeywell stammenden Technik. Die Folge war ein 1987 angestrebter Prozess, der für Minolta nach vier Jahren mit einer Strafzahlung von fast 120 Millionen Dollar an Honeywell endete. Während die bisher genannten Hersteller mit dem Phasenvergleich als Fokussierungsmethode arbeiteten, entwickelten andere Hersteller mitunter auch wegen fehlender Patentrechte für den Phasen-AF neue Autofokusarten. Polaroid integrierte 1972 in die Sofortbildkamera SX-70 einen Ultraschall-Distanzsensoren, der allerdings keine große weitere Verbreitung fand. Canon stellte mit der AF35M 1979 die erste Kamera mit aktivem Fokus-Infrarotsensoren vor. Das Bildkontrastverfahren zur Fokussuche, welches den Bildinhalt auswertet, fand



**Abbildung 3: Links: Skizze des in [42] vorgestellten Mikroskop-Apparats mit automatischer Fokussierung aus dem Jahr 1974. Rechts: Das erste mit einem AF ausgestattete kommerzielle Mikroskop aus dem Jahr 1983.**

erstmalig 1981 bei der Pentax ME-F Spiegelreflexkamera Anwendung. Gleich dem Infrarotlicht, fand auch die Entwicklung von Lasersensoren zur Distanzberechnung für die Fokuseinstellung statt. Laser-Autofokussysteme hielten dann vor allem in der industriellen Materialfertigung Einzug und sind heutzutage das Mittel der Wahl, um den korrekten Abstand zu Werkstücken beizubehalten.

Sehr aktuell ist die Forschung auf dem Gebiet von Flüssiglinsen. Darauf zurückzuführende Patente (u.a.[3]) von Bruno Berge stammen aus den Jahren 1997 bis 2002. Die Firma Varioptic verkündete Ende des Jahres 2006 die Realisierung des ersten Autofokus mit Flüssiglinse ohne jegliche bewegliche Teile, die Echtzeitfokussierung verwirklicht ([4]).

## 3.2 Autofokuskategorien

In der Literatur werden die Autofokussmethoden in aktive und passive Verfahren eingeteilt. Eine Fokussierung auf Basis einer direkten Entfernungsmessung mit einem Hilfsensor gilt als aktiv und eine über die Kameraoptik und die Auswertung des Bildinhalts als passiv. Bei den passiven Methoden arbeiten die Fokussierungsansätze über Phasendetektion oder durch Bildkontrastmessung.

### 3.2.1 Aktiver Autofokus

Für die direkte Abstandsmessung sind verschiedene elektromagnetische Strahlarten denkbar. Polaroid beispielsweise stattete in den 80er Jahren ihre Sofortbildkameras mit einem Autofokus aus, der über die Laufzeitmessung emittierten Ultraschalls funktionierte. Während sich diese Technik im Kamera- und Mikroskopbereich aufgrund vieler Nachteile nicht durchsetzte, findet man sie noch in der industriellen Abstandsmessung unter definierten Arbeitsbedingungen, die ungewollte Schallstreuung und ähnliche Störfaktoren ausschließen. Eine weitere Möglichkeit ist die Entfernungsmessung mit Laserstrahlung, wie man sie in Autofokussystemen von Photokameras und Diaprojektoren aber auch in Mikroskopen finden kann. Wie bei der Ultraschallvariante ist die Laufzeitmessung zur Abstandsbestimmung möglich. Weiterhin üblich ist die Analyse des Leistungsdichteverlusts der reflektierten IR-Strahlung, die Messung der Interferometrie-Phasenverschiebung oder die Entfernungsberechnung über Triangulation. Die Phasenverschiebung wird auf Basis der Überlagerung zweier Strahlen ermittelt, die derselben Laserquelle entstammen. Hierbei findet der Vergleich statt zwischen einem Lichtbündel, das von einem Referenzobjekt reflektiert wird, und dem Reflektionsstrahl vom Messobjekt. Über die Phasendifferenz  $\omega_{\Delta}$  und die Wellenlänge des Lasers  $\lambda_{laser}$  ergibt sich die zum Referenzobjekt relative Wegdifferenz  $d$  zu:

$$d = \frac{\lambda_{laser}}{2\pi} \omega_{\Delta} \quad (1)$$

Diese Messung ist jedoch nicht eindeutig, wenn die Distanz  $d$  größer ist als die Wellenlänge, und außerdem die Anzahl der zurückgelegten Wellenperioden nicht bekannt ist. In diesem Fall wird das Lasersignal amplitudenmoduliert und der Phasenunterschied der Trägerwelle bestimmt.

Triangulationsmessgeräte bestimmen den Objektabstand über den Winkel des reflektierten Strahles. Der emittierende Laser und der Detektor sind mit einem bekanntem Abstand versetzt angeordnet. Zusammen mit dem Objekt bilden sie ein Dreieck, das mit bekannten Winkeln die Abstandsberechnung ermöglicht. Als Detektor fungiert ein CCD-Sensor. Der reflektierte Laserstrahl verändert die Position auf dem Sensor in Abhängigkeit vom Eintrittswinkel. Der Abstand des Lasersignals zu einer definierten Sensorkoordinate führt direkt zum Winkelbetrag.

Aus der Abbildung 4 werden die Winkelbeziehungen ersichtlich:

$$\tan(\alpha_{ref}) = \frac{d_{ref}}{D_h} \quad (2)$$

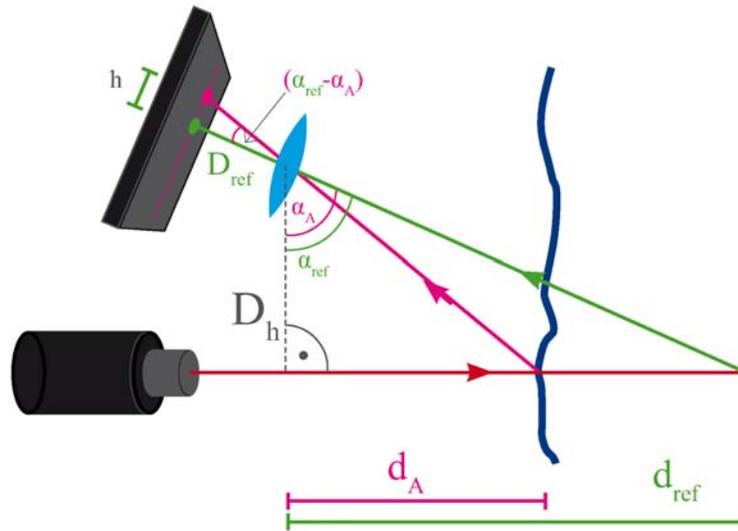


Abbildung 4: Ein weiter entferntes Objekt vergrößert den Einfallswinkel  $\alpha$  und verändert entsprechend die Position des Strahls auf dem CCD-Sensor. Mit den bekannten Größen des Referenzstrahls wird die Objektdistanz über Dreiecksbeziehungen ermittelt. (Quelle: eigene Darstellung)

$$\tan(\alpha_{ref} - \alpha_A) = \frac{h}{D_{ref}} \quad (3)$$

$$\tan(\alpha_A) = \frac{d_A}{D_h} \rightarrow d_A = \tan(\alpha_A) D_h \quad (4)$$

Um die Lösung für Gl. 4 und somit die Entfernung  $d_A$  zu erhalten, wird Gleichung 3 nach  $\tan(\alpha_A)$  umgestellt:

$$\frac{h}{D_{ref}} = \tan(\alpha_{ref} - \alpha_A) = \frac{\tan(\alpha_{ref}) - \tan(\alpha_A)}{1 + \tan(\alpha_{ref})\tan(\alpha_A)} \quad (5)$$

$$\tan(\alpha_A) = \frac{\frac{d_{ref} D_{ref}}{D_h} - h}{D_{ref} + \frac{d_{ref} h}{D_h}} \quad (6)$$

Somit ergibt sich der Objektabstand zu:

$$d_A = \left( \frac{\frac{d_{ref} D_{ref}}{D_h} - h}{D_{ref} + \frac{d_{ref} h}{D_h}} \right) D_h \quad (7)$$

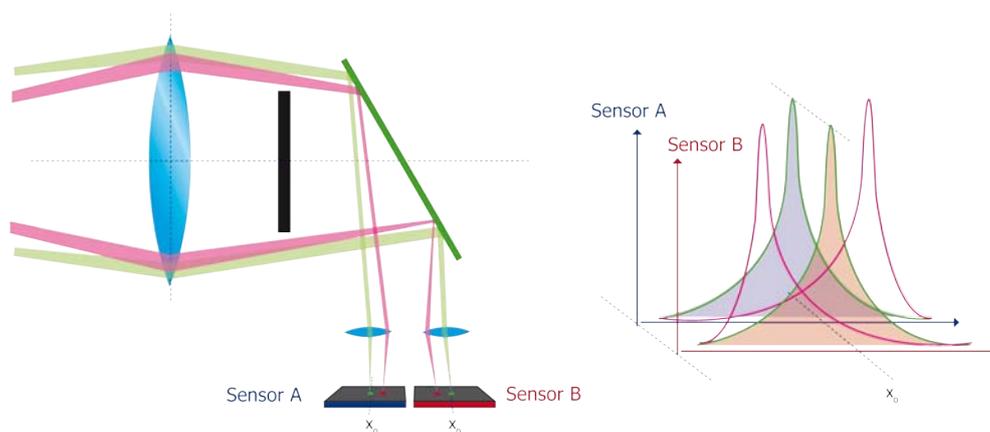
Während in der industriellen Fertigung zumeist Laser sichtbaren Lichts eingesetzt werden,

ist in der biologischen und medizinischen Mikroskopie vor allem der Infrarot-Laser von Bedeutung, weil es bei diesen Anwendungen generell erwünscht ist, die Proben einer möglichst geringen Strahlung auszusetzen. Präzise Autofokussysteme in der HTS- und HCS-Mikroskopie (siehe Abs. 2.1) arbeiten mit einer Kombination aus aktivem Laser-AF und passivem Kontrast-AF. Aktive Fokussierungssysteme spielen ihre Stärken vor allem in der Autofokusnachjustierung aus. Sie sind in der Lage, auf kleinste Positionsänderung des Deckglases zu reagieren, haben allerdings Probleme, biologische Proben scharf zu stellen. Deshalb werden der Laser-AF zur schnellen Grob- und ein passiver AF zur Feinjustierung eingesetzt.

#### 3.2.2 Passiver Autofokus

##### Phasenvergleich

Systeme, die auf dem sog. Phasenvergleich beruhen, teilen das vom Aussenrand der Objektivlinse eintreffende Licht und projizieren es auf zwei versetzte photo-elektrische Empfänger. Das Projektionsmuster gibt Aufschluss sowohl über den Fokussierungsgrad als auch die Korrekturrichtung der Linse bei erkannter Defokussierung. Abbildung 5 stellt diese Sachverhalte grafisch dar. Fallen beide Intensitätsmaxima an derselben Sensorposition zusammen, befindet sich das abzubildende Objekt im Fokus.



**Abbildung 5:** Bei der Phasenvergleichsmethode werden die Randstrahlen geteilt und auf zwei Photosensoren fokussiert. Bei Strahlen aus der Fokusebene (grüne Strahlenbündel) überlappen sich die Intensitätsmaxima der Sensoren. (Quelle: eigene Darstellung)

Der Vergleich der Intensitätskurven erfolgt über Phasenkorrelation, welche der Methode gleichfalls den Namen gibt. Bei der Phasenkorrelation wird der Versatz von zwei ähnlichen Signalverläufen bzw. Bildern über die Kreuzkorrelation der in den Ortsfrequenzraum transformierten Daten ermittelt. Im vorliegenden Fall, wo die Sensoren ein zwei-

dimensionales Intensitätsbild liefern, ist die Phasenkorrelation als das normalisierte Kreuz-Leistungsspektrum  $S(u, v)$  der Fourier-Transformierten der Eingangssignale  $A(x, y)$  und  $B(x, y)$  definiert:

$$A(x, y) \longleftrightarrow F_A(u, v) \quad (8)$$

$$B(x, y) \longleftrightarrow F_B(u, v) \quad (9)$$

Nach dem Fourier-Shift-Theorem drückt sich die örtliche Verschiebung als Phasenunterschied im Frequenzbereich aus. Wenn also B und A nur durch die Translation um  $(x_0, y_0)$  verschieden sind, findet sich dieser Offset in der Phase der Fourier-Transformierten wieder:

$$B(x, y) = A(x - x_0, y - y_0) \quad (10)$$

$$F_B(u, v) = F_A(u, v)e^{-j(ux_0 + vy_0)} \quad (11)$$

Für die konkrete Ermittlung der Verschiebungskordinaten wird das Kreuz-Leistungsspektrum aus Transformierten  $F_B(u, v)$  und komplex konjugierten  $F_A^*(u, v)$  gebildet:

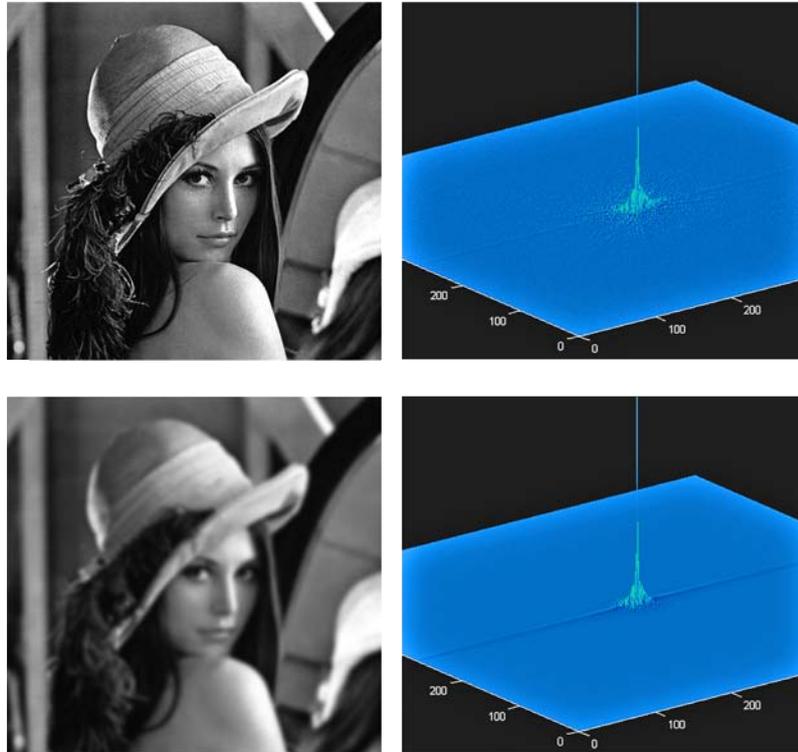
$$S(u, v) = \frac{F_B(u, v)F_A^*(u, v)}{|F_B(u, v)F_A^*(u, v)|} \quad (12)$$

Die inverse Fourier-Transformation von  $S(u, v)$  führt bei kontinuierlichen Eingangssignalen zu einer Delta-Funktion, die im Punkt  $(x_0, y_0)$  einen Dirac-Stoß aufweist. Dessen Koordinaten geben die direkte Verschiebung der beiden zu vergleichenden Bilder wieder.

### Kontrastvergleich

Für die Autofokussierung durch Kontrastmessung sind zusätzlich zum Bildsensor keine Optiken erforderlich, weil die Auswertung der Fokusslage aufgrund der Bildinformation geschieht. Diese Auswertung erfolgt heutzutage grundsätzlich in den folgenden Schritten: Digitalisierung der Sensordaten, Software-seitige Berechnung eines Schärfewertes des aufgenommenen Bildes, Interpretation des Wertes im Kontext zuvor analysierter Bilder in anderen z-Lagen und schließlich die Hardwareanpassung, um den korrekten Abstand zwischen Optik und Untersuchungsobjekt herzustellen. Der Schärfewert wird mit einem Schärfalgorithmus bzw. -funktion berechnet. Für sich genommen liefert er weder eine Aussage über den Grad der Schärfe, noch zeigt er die Richtung der Fokusslage an. Erst

durch den Vergleich der Messwerte mehrerer Bilder entlang der z-Achse ist es möglich, die Ebene mit maximalem Schärfewert als Fokusebene zu bestimmen. Das Abtragen der Messwerte bezüglich der z-Koordinaten führt zur Fokuskurve, die im Idealfall ein gut definiertes globales Maximum aufweist.



**Abbildung 6:** Gegenüberstellungen des Lena-Testbildes und ihrer jeweiligen Transformation in den Ortsfrequenzraum mittels diskreter Kosinustransformation (MatLab-Plot). Die Plot-Mitte zeigt den Betrag des Gleichanteils, während mit größerem Abstand von der Mitte die Frequenz zunimmt. Während ein fokussiertes Bild mit vielen Strukturen zu einem Spektrum mit Anteilen auch im höherfrequenten Bereich führt (obere Reihe), zeigt das defokussierte Bild kaum Anteile außerhalb der tiefen Ortsfrequenzen (untere Reihe). (Bild-Quellen - rechts: eigene Darstellung )

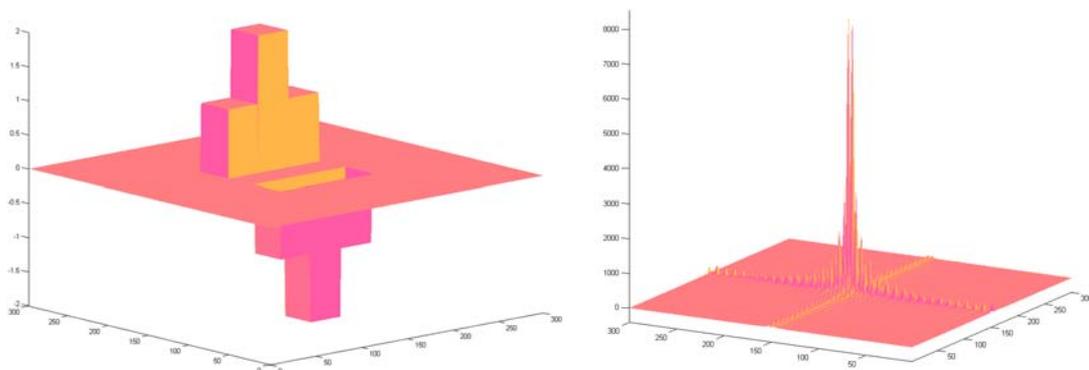
Die verwendeten Schärfefunktionen legen dabei folgende Tatsachen zugrunde. Die Darstellung eines Objektpunktes in der Bildebene  $B(x)$  durch ein optisches System folgt einer bestimmten Intensitätsverteilung. Diese Intensitätsspreizung des Bildpunktes ist das Ergebnis der Faltung des idealen Abbilds  $O(x)$  mit einer Point-Spread-Funktion (PSF, dt. Punktspreizfunktion).

$$B(x) = PSF(x) \otimes O(x) \quad (13)$$

Im Fourier-Raum wandelt sich die Faltung zu einer Multiplikation der Transformaten. Die Fourier-Transformation der Point-Spread-Funktion  $F_{PSF}$  wird als *Optical-Transfer-Function* (OTF) bezeichnet:

$$F_B(u) = F_{PSF}(u) \cdot F_O(u) = OTF(u) \cdot F_O(u) \quad (14)$$

Die Form der PSF nimmt entsprechend der Lichtbeugungseigenschaften die Form von Airy-Beugungsrings an. Die PSF-Funktion selbst enthält die Bessel-Funktion 1. Ordnung. Bei polychromatischem Licht ist es legitim, die Airy-Funktion durch eine Gauß-Funktion zu ersetzen. Die Superposition der Intensitätsverteilung bei den einzelnen Wellenlängen macht diese Approximation möglich ([40]). Wird also eine Gauß-Verteilung im Ortsraum angenommen, bedeutet das nach Gleichung 14 eine Multiplikation der transformierten Abbildungsfunktion mit einem Rechteck im Frequenzraum. Das ist nichts anderes als eine Tiefpass-Filterung im Frequenzbereich. Die Messung des Bildkontrasts ist also unmittelbar mit dem Frequenzspektrum im Ortsfrequenzraum des Bildes verbunden. Je defokussierter ein Bild ist, desto größer fällt der Radius der Gauß-Funktion aus, und ein größerer Radius bedeutet einen schmaleren Tiefpass im Frequenzbereich. Eine Abbildung im Fokus heißt also eine Beibehaltung hoher Frequenzanteile, während ein defokussiertes Bild deren Abwesenheit durch die Tiefpassfilterung bedeutet (siehe Abbildung 6 mit Lena<sup>1</sup>-Testbild).



**Abbildung 7: Links: Sobel-Operator im Ortsraum. Rechts: Entsprechung des Operators im Frequenzraum.** (Quelle: eigene Darstellung)

Eine ähnliche Entsprechung lässt sich auch für Kantenoperatoren wie zum Beispiel den Sobel-Filter finden. Abbildung 7 zeigt einen 3x3-Operator, der auf horizontale Bildkanten

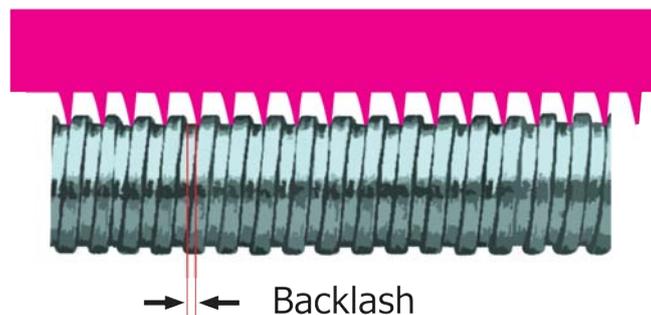
<sup>1</sup>Das sog. "Lena"-Testbild stammt aus dem Jahr 1972 und somit aus der Zeit des Beginns der rechnergestützten Bildverarbeitung. Es zeigt das Centerfold "Lenna Sjöobloom" (ihr richtiger Name ist Lena Söderberg) der Novemberausgabe. Das Bild wurde und wird weiterhin aufgrund der vielfältigen Strukturen als Standard-Testbild für bildverarbeitende Algorithmen verwendet. (Quelle: <http://www.cs.cmu.edu/~chuck/lennapg/>)

anspricht. Die Fourier-Transformierte zeigt zum einen hohe Amplitudenbeträge entlang der y-Achse, zum anderen weist das Spektrum keinen Gleichanteil auf. Diese Operation liefert demnach im Frequenzraum ein deutliches Produktresultat bei Vorhandensein von Frequenzen entlang der y-Achse des Bildfrequenzspektrums.

## 3.3 Herausforderungen für einen Autofokus

### 3.3.1 Mikroskop-Backlash und Drift

Mikroskope besitzen einen systemischen Fehler, der wiederholtes Einstellen des Objektives auf exakt dieselbe Position verhindert. Dieses als Backlash bezeichnete Spiel der Trieb-Mechaniken ist die Summe aller konstruktiven Lücken zwischen den ineinander greifenden Zahnrädern und Spindeln (siehe Abbildung 8). Vor allem bei Richtungsänderungen des Tisches kommt ein Backlash zum Tragen, welcher je nach Fertigungsqualität zwischen 0,1 Mikrometer und zweistelligen Mikrometerbeträgen liegt.



**Abbildung 8: Verdeutlichung des Backlashs an einer Stellschraube.** (Quelle: eigene Darstellung)

Der Backlash-Fehler kann zum einen durch hohe Fertigungspräzision und ein spezielles Design der Getriebeteile minimiert werden; zum anderen kann die Korrektur der Ist- zur Soll-Position Software-seitig vorgenommen werden, indem die vorherige Bewegungsrichtung gespeichert wird.

Die durch den Backlash verursachte Trieb-Abweichung lässt sich auf verschiedene Weise bestimmen. Mit einem zusätzlichen Laser-Modul kann die Abweichung exakt gemessen werden. Es ist aber ebenfalls möglich, mit vorhandener Optik und Kamera den Backlash per Software zu überprüfen. Hierbei wird mit einer klassischen Schärfefunktion (siehe Abschnitt 4.1.2) jeweils die Fokuskurve sowohl in die eine z-Richtung als auch in die entgegengesetzte Richtung für ein geeignetes Präparat aufgenommen, das bekanntermaßen eine unimodale Kurve mit definierter Spitze liefert. Werden die zwei Kurvenresultate übereinandergelegt, zeigt der Abstand der Spitzen den Backlash an.

Bei Reihenaufnahmen über einen längeren Zeitraum kommt es mit dem Fokus-Drift zu einer weiteren Beeinträchtigung. Temperaturschwankungen im Mikroskop oder der Umgebung aber auch Vibrationen führen zu einer Bewegung weg von der Fokusebene. Dieses Wegdriften kann wiederum durch laser-gestützte Abstandsmessung oder Software-Kontrastanalyse korrigiert werden.

#### **3.3.2 Sensor-Rauschen und Sensor-Fehler**

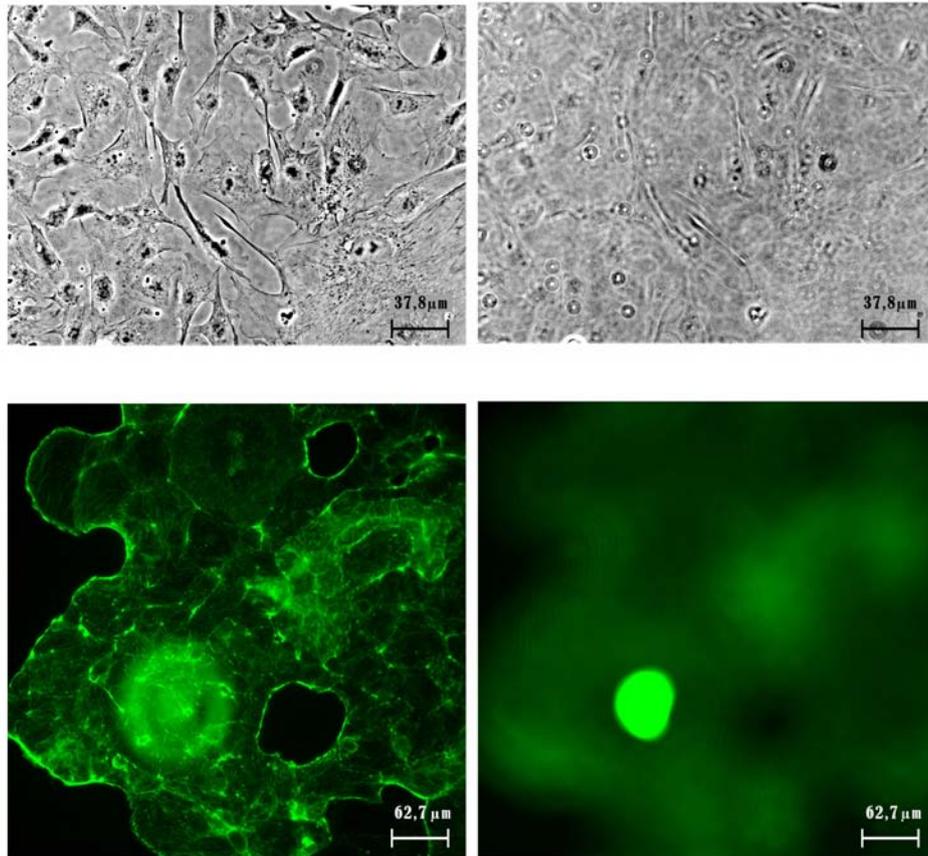
Die bildgebenden Mikroskopsensoren zeigen störende Eigenschaften, die zur Bilddegradation beitragen und somit auch die Qualität von kontrastbasierten Autofokus-Methoden beeinflussen. Das Ausmaß der negativen Auswirkung dieser Eigenschaften auf das Bildergebnis ist maßgeblich von der CCD-Charakteristik abhängig. Dazu zählen die Lichtquantenausbeute der Photodioden, die maximale Elektronenanzahl bis zur Pixel-Sättigung (Full Well Capacity), die Dunkelstromeigenschaften und die Fähigkeit, überschüssige Ladung bei Überbelichtung abzuleiten, sowie die Leistungsfähigkeit, Ladungen auszulesen, zu transportieren und zu digitalisieren ([27]).

Die Rauscharten unterscheidet man wie folgt: Schrotrauschen bzw. im speziellen Fall das Photonenrauschen bezeichnet den ungleichmäßigen Ladungsfluß aufgrund stochastisch eintreffender Photonen. Thermisches Rauschen oder auch Dunkelstromrauschen tritt als Folge der Bewegung freier Ladungsträger auf, bedingt durch ihre thermische Energie in Leitern- und Halbleitern. Das dadurch verursachte Auftreten von unerwünschten Elektronen verstärkt sich mit zunehmender Temperatur des CCD-Sensors. Neben dem als Rosa Rauschen modellierten  $1/f$ -Rauschen, das z.B. durch Halbleiterverunreinigungen oder Kristall-Defekte entsteht, existiert weiterhin das Quantisierungsrauschen. Sowohl thermisches als auch Photonenrauschen können als additives weißes Rauschen modelliert werden.

Pixel-Fehler können Autofokus-Schärfefunktionen erheblich beeinflussen, weil sie hohe real nicht existierende Gradienten erzeugen. Zu ihnen zählen die Hot Pixel, die einen sehr hohen Dunkelstrom und bei langen Belichtungszeiten und thermischer Aufheizung überdurchschnittlich helle Bildpunkte erzeugen, des Weiteren die Dead Pixel und Stuckpixel, die entweder immer dunkel oder Licht unabhängig immer hell sind.

#### **3.3.3 Artefakte**

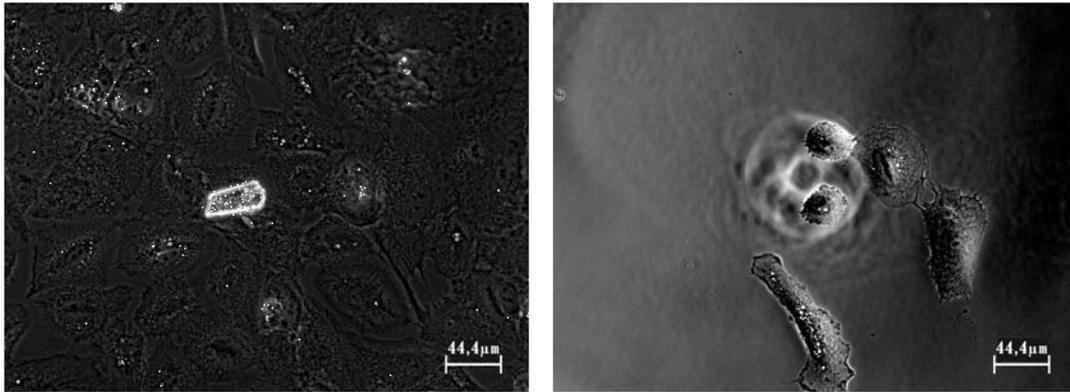
Die erläuterten Rauscheinflüsse und Pixelfehler sind im weiten Sinne als Artefakte zu bezeichnen. Artefakte bringen in die Bildaufnahme unerwünschte Strukturen ein, die nicht



**Abbildung 9:** Im Phasenkontrastverfahren weisen Objektkanten sehr helle Halos auf (links oben) und bilden Phasenringe in von der Fokusslage entfernten Ebenen (rechts oben) (hier: Cytoskelett-Probe im Phasenkontrast- Mikroskop: Zeiss Axio Observer.Z1 mit 20x/0.80NA-Objektiv). Staub und andere Verunreinigungen bringen sich störend in die Aufnahme der Caco2-Zellprobe (FITC-Färbung - Mikroskop: Zeiss Axio Imager.Z1 mit 20x/0.5NA-Objektiv) ein (links unten) und zeigen zumeist hohe Intensitäten durch stark reflektiertes Licht (rechts unten). (Quellen: eigene Darstellung; Zur Verfügung gestellt durch Zeiss MicroImaging GmbH)

direkt die Probenmerkmale darstellen. Artefakte entstehen weiterhin durch fehlerhafte beziehungsweise unreine Präparation und auch durch die Eigenschaften bestimmter Mikroskop-Modi. Zu letzteren zählen die charakteristischen Shade-Off- und Halo-Effekte (siehe Abbildung 9 oben) in der Phasenkontrast-Mikroskopie. Bei positivem Phasenkontrast werden die Konturen von stark Licht brechenden Objekten bzw. zwischen Grenzen von Objektmerkmalen von hellen Konturen (Halos) begleitet. Mit Shade-Off wird der Helligkeitsverlauf beschrieben, der innerhalb von Objektflächen mit gleichem Brechungsindex auftritt. Diese Effekte werden durch Streulicht verursacht, das den Phasenring passiert. Des Weiteren ist bei Phasenkontrast das Phänomen zu beobachten, dass Objekte mit zunehmender Defokussierung keine gewöhnliche Gaußsche Glättung erfahren, sondern Ringe bilden (Abb. 10 rechts).

Die durch Präparation induzierten Artefakte sind vielfältiger Art. Die linke Aufnahme in



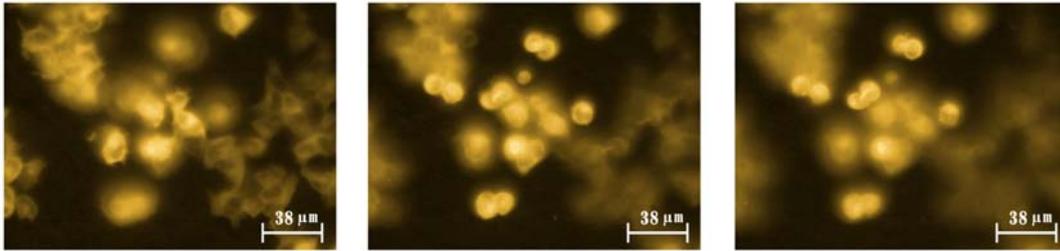
**Abbildung 10: Links: Ein nicht aufgelöstes Kristall des Nährmediums beeinträchtigt die Aufnahme durch hohen Kontrast und Helligkeit. Rechts: Phasenringe von Verunreinigung wirken bis in die Fokusebene der eigentlich zu betrachtenden Zellen.** (Quellen: eigene Darstellung; Zur Verfügung gestellt durch Zeiss MicroImaging GmbH)

Abbildung 10 zeigt zum Beispiel ein Kristallstück, das einer Nährlösung mit nicht vollständig aufgelösten Salzen entstammt. Dessen hohe Helligkeitswerte und die starken Gradienten beeinträchtigen die korrekte Autofokussuche besonders bei Untersuchungsobjekten mit geringem Kontrast. Unerwünschte Objekte können auch durch abgebrochene Teile von Membran-Filtern in die Bildaufnahme gelangen. Solche Filter bestehen beispielsweise aus Zellulose und werden unter anderem zur Filterung von Bakterien beim Einsäen von Zellen benutzt. Die Liste von Probenunregelmäßigkeiten kann um eine Vielzahl unterschiedlichster Faktoren erweitert werden: Wasser- und Lufteinschlüsse, Stärke-Kristalle von OP-Handschuhen, Verfärbungsartefakte des Probenmaterials durch das Färbemittel selbst, Pilzbefall oder bakterielle Verunreinigung und vieles mehr ([36]).

#### 3.3.4 Vielschichtige Proben und Präparate hoher Schichtdicke

Eine große Anzahl von mikroskopischen Proben besitzt die angenehme Eigenschaft, die Untersuchungsobjekte in lediglich einer optischen Ebene zu beinhalten. Zum Beispiel proliferieren diverse Tumorzelllinien in einer Schicht und verbleiben aufgrund von Kontaktinhibition als Monolayer. Bei pathologischen Untersuchungen erfolgen Gewebsschnitte in derart dünnen Schichten, dass sie auf dem Objektträger ebenfalls in einer Fokusebene aufliegen. Gleichzeitig findet man jedoch in nahezu allen mikroskopischen Untersuchungsbereichen eine Vielzahl von Präparaten, die Objekte bzw. Objektmerkmale in mehreren Fokuslagen aufweisen.

Nicht immer sind z.B. Zellen auf den Objektträger fixiert, das heißt, sie schwimmen vielmehr in der Suspension in unterschiedlichen fokalen Ebenen. Ist die Lösung aus



**Abbildung 11: HeLa-Probe mit Zellen in drei unterschiedlichen Fokusebenen.** (Quellen: eigene Darstellung; Zur Verfügung gestellt durch Zeiss MicroImaging GmbH)

unterschiedlichen Zelltypen oder Zellstadien zusammengesetzt, erhöht sich die Wahrscheinlichkeit, dass die Probe mehrere Fokuslagen besitzt (siehe Abbildung 11). Gleiches ist der Fall, wenn in Vergrößerungsbereichen gearbeitet wird, bei denen die Untersuchungsobjekte ein Vielfaches größer als die Schärfentiefe sind. Erwähnt sei auch, dass die anfangs genannten Monoschichtpräparate unter ungünstigen Bedingungen oder aufgrund schlechter Präparation zu Polylayer-Proben werden. Dazu kommt es zum Beispiel, wenn Gewebebereiche vom Objektträger abheben bzw. zerreißen oder wenn Zelllinien ohne Kontaktinhibition vor rechtzeitigem Passagieren übereinander wachsen.

An dieser Stelle soll noch einmal erwähnt werden, dass sich die Arbeit in einem Forschungsfeld bewegt, wo mikroskopische Proben erwartungsgemäß eine fokussierte und relevante Informationsebene je Objekt besitzen. Darunter fallen deshalb nicht Anwendungen, wie z.B. im Bereich der Stereologie, die quantitative Analysen in Präparaten ([19]) durchführen, wo mitunter jede Probenebene relevante Informationen enthält, die es gilt, diese zu extrahieren.

## 4 Stand der Forschung

### 4.1 Algorithmen passiver Methoden

Die im Folgenden erläuterten Schärfemessfunktionen folgen meistens den unter anderem durch Groen et al.([21]) aufgestellten Kriterien eines Autofokus-Algorithmus. Diese sind: **(a)** die Richtigkeit des Messergebnisses, **(b)** das Erzielen möglichst eines globalen Maximums, bestenfalls ohne Nebenmaxima (Unimodalität), **(c)** die Reproduzierbarkeit des Ergebnisses, **(d)** eine monoton fallende und steigende Fokuskurve, **(e)** eine gewisse Robustheit gegenüber Störungen und Bildrauschen und **(f)** eine möglichst allgemein gültige Anwendung auf verschiedene Untersuchungsbereiche. Neben den genannten Kriterien finden sich in anderen Veröffentlichungen ([32], [24]) weitere wie die Berechnungsgeschwindigkeit und der Abstand zweier lokaler Maxima in der Nähe des globalen. Während Batten ([1]) den Punkt der Unimodalität als sehr wichtige Eigenschaft hervorheben, wehren sich Geusenbroek et al. ([18]) entschieden gegen dieses Kriterium und begründen dies mit fehlender Durchsetzbarkeit in der Praxis.

In einer Veröffentlichung von Sun et al. ([50]) werden die unterschiedlichen Algorithmen in vier Gruppen eingeteilt: **(1)** ableitungsbasierte, **(2)** statistische, **(3)** histogrammbasierte und **(4)** sogenannte unmittelbare (engl. ‚intuitive‘) Algorithmen. Diese Einteilung ist in der Literatur viel zitiert und hat größtenteils ihre Berechtigung. Die Einleitung dieses Abschnitts behandelte das Kontrastmessverfahren als Synonym für alle bildverarbeitenden passiven Autofokusmethoden und setzte es mit der Analyse des Ortsfrequenzspektrums in den Hochfrequenzbereichen gleich. Dies trifft auf die eben genannten ableitungsbasierten wie auf die statistischen Methoden zu, wenn man sich z.B. vor Augen führt, dass die Autokorrelationsfunktion nach dem Wiener-Khintchine-Kriterium mit dem Leistungsspektrum im Frequenzbereich verknüpft ist ([5]). Jedoch fallen die histogramm-basierten und unmittelbaren Algorithmen nicht direkt in die Kontrastverfahren, weil sie nur elaborierte Verfahren auf der Basis der Pixelintensitäten darstellen. Damit folgen sie nicht der menschlichen Kontrastwahrnehmung, die auf der Verarbeitung von Helligkeitssprüngen beruht. In publizierten Untersuchungen ([50],[9]) haben sich allerdings diese Methoden als weniger erfolgreich erwiesen. Der Vollständigkeit halber seien sie aber erwähnt.

Es liegt nahe, für die Fokusfindung mehrere der im Folgenden besprochenen Schärfefunktionen auszuwerten. Entscheidet sich die Mehrzahl der Algorithmen für eine Ebene, kann diese mit größerer Sicherheit als die Fokusposition erkannt werden. Ein anderer Ansatz besteht in der Kombination verschiedener Methoden (eindimensionale Fourier-Transformation mit Pearson-Korrelation).

Im Folgenden finden sich die bekanntesten Algorithmen in kurzer Form erläutert. Dabei liefern alle Gleichungen als Ergebnis ein Schärfemaß  $S$ , das für jeweils ein 2-D-Bild berechnet wird. Jenes Bild, das den höchsten absoluten Wert aus einer Reihe von Bildern erzielt, wird als Fokusebene ausgegeben.

#### 4.1.1 Funktionen im Ortsfrequenzraum

Wavelet-basierte Fokusmessung – Hui et al. ([25])

Der Schärfemesswert  $S$  wird durch den Quotienten aus den Quadraten der Hochpassanteile  $M_{LH}$ ,  $M_{HL}$  und  $M_{HH}$  und aus dem Tiefpassanteil  $M_{LL}$  der diskreten Wavelet-Dekomposition gebildet. Hierbei liegt die Annahme zugrunde, dass die Summen der drei Hochpassfunktionen mit steigender Bildschärfe zunehmen während der Tiefpassanteil abnimmt.

$$S = \frac{M_H^2}{M_L^2} \quad (15)$$

mit:

$$M_H^2 = \sum_{(x,y) \in S_{LH}} M_{LH}^2(x,y) + \sum_{(x,y) \in S_{HL}} M_{HL}^2(x,y) + \sum_{(x,y) \in S_{HH}} M_{HH}^2(x,y) \quad (16)$$

und:

$$M_L^2 = \sum_{(x,y) \in S_{LL}} M_{LL}^2(x,y) \quad (17)$$

Wavelet-Algorithmus mit Mittelwert

Yang und Nelson ([55]) präsentieren zwei Messfunktionen, die ebenfalls den auf der Wavelet-Dekomposition beruhenden Algorithmus benutzen, allerdings den Tiefpassanteil außer acht lassen. Zum einen summieren sie die Beträge der Hochpassanteile auf:

$$S_1 = \sum_{(x,y) \in S} |M_{LH}(x,y)| + |M_{HL}(x,y)| + |M_{HH}(x,y)| \quad (18)$$

zum anderen subtrahieren sie die Mittelwerte  $\mu$  von den jeweiligen Anteilen, bilden die

Gesamtsumme der Quadrate und nehmen eine Normierung mit der Gesamtpixelzahl vor:

$$S_2 = \frac{1}{\text{Höhe} \cdot \text{Breite}} \sum_{(x,y) \in S} (|M_{LH}(x,y)| - \mu_{LH})^2 + (|M_{HL}(x,y)| - \mu_{HL})^2 + (|M_{HH}(x,y)| - \mu_{HH})^2 \quad (19)$$

In einer aktuelleren Veröffentlichung von Makkapati ([35]) wird die Formel in Verbindung mit einer vorangestellten Hintergrund-Subtraktion verwendet, um Rauschen nicht in die Schärferechnung einfließen zu lassen. Als Ergebnis nehmen die Fokuskurven eine geglättete Form mit leichter zu bestimmenden Maxima an.

#### Bandpass-Filter

Die Untersuchung der Abhängigkeit der Optical-Transfer-Function(OTF) vom Grad der Defokussierung führte Boddeke und Boddeke et al. ([6, 7]) zur Verwendung eines Bandpass-Filters  $h_{\theta_{Bp}}$ , wobei  $\theta_{Bp}$  die Mittenfrequenz darstellt. In dieser Betrachtung variiert die OTF vor allem in den mittleren Frequenzbändern, so dass allein die Beträge dieser Spektralanteile für die Fokussmessung herangezogen werden.

$$S = \sum_{(x,y) \in I} (I(x,y) \otimes h_{\theta_{Bp}}(x,y))^2 \quad (20)$$

Weil ein solcher Bandpass im Bildbereich zu einem entsprechend großen Operator führt, verwendet Boddeke schlussendlich den Brenner-Gradient nach Bild-Sampling mit der Nyquist-Frequenz in der Form:

$$S = \sum_{(x,y) \in I} (I(x,y) - (I(x-1,y)))^2 \quad (21)$$

#### Eindimensionale Fourier-Transformation mit Pearson-Korrelation

Bueno et al. ([10]) kombinieren eine Fourier-Transformation mit angeschlossener Korrelation, um eine akkurate und schnelle Autofokus-Routine zu implementieren. Im ersten Schritt wird ein Vektor aus den Fourier-Antworten einer Anzahl von äquidistanten vertikalen Bildzeilen gebildet. Der eigentliche Fokussmesswert  $S$  errechnet sich dann aus dem Pearson-Korrelationskoeffizienten  $r$  der Vektoren der zu untersuchenden Bildebene  $F_i$  und einem Referenzvektor  $F_{defocus}$ , der für ein äußerst defokussiertes Bild steht.

$$S = |r_{Pearson}| = \left| \frac{\sum_{(x,y) \in I} (F_i(x,y) - \overline{F_i(x,y)})(F_{defocus}(x,y) - \overline{F_{defocus}(x,y)})}{\sqrt{\sum_{(x,y) \in I} (F_i(x,y) - \overline{F_i(x,y)})^2 \sum_{(x,y) \in I} (F_{defocus}(x,y) - \overline{F_{defocus}(x,y)})^2}} \right| \quad (22)$$

Der Koeffizient nimmt Werte von 0 bis 1 an. Nun zeigt aber jene Fouriervektor-Matrix auf die Fokusebene, die den geringsten Korrelationskoeffizient aufweist, weil das schärfste Bild am wenigsten mit dem Referenzbild korreliert.

#### Beträge der Fourier-Komponenten

Die von Batten ([1]) beschriebene Schärfefunktion arbeitet ebenfalls im Frequenzraum. Nach Anwendung der Fast-Fourier-Transformation werden die Beträge der Komponenten des Spektrums  $F(u, v)$  in Abhängigkeit eines Frequenz-Schwellwertes  $\omega_t$  summiert.

$$S = \frac{1}{4\omega_t^2} \sum_{(u,v) \in F} |F(u, v)| \dots \text{für} : \omega < \omega_t \quad (23)$$

Das heißt, dass nur die Frequenzbänder addiert werden, die eine geringere Frequenz als  $\omega_t$  aufweisen. Auf den ersten Blick ist dieser Tiefpass im Frequenzraum nicht nachzuvollziehen, weil eben nicht die tiefen Spektralanteile für ein fokussiertes Bild stehen. Batten erwähnt allerdings nur, dass die Wahl des Schwellwertes nicht zu tief angesetzt werden darf, weil sonst die wichtigen hohen Anteile fehlen würden.

#### 4.1.2 Gradienten-Berechnung

##### Schwellwert-basierter Gradientenbetrag

Eine äußerst einfache Fokusfunktion findet sich in der Veröffentlichung von Santos et al. ([44]). Der Fokuswert ergibt sich durch die Summierung der quadrierten Intensitätsunterschiede direkt benachbarter horizontaler Pixel, falls der Gradientenbetrag einen definierten Schwellwert  $\theta_t$  überschreitet.

$$S = \sum_{(x,y) \in I} (I(x+1, y) - (I(x, y)))^2 \dots \text{für} : |I(x+1, y) - (I(x, y))| < \theta_t \quad (24)$$

Diese enge Nachbarschaftsbeziehung bedeutet allerdings eine hohe Sensitivität gegenüber Bildrauschen.

Eine Erweiterung, der Brenner-Gradient ([8]), unterscheidet sich von der vorherigen Funktion lediglich darin, dass die Intensitätsdifferenz mit dem übernächsten Nachbarn gebildet wird.

$$S = \sum_{(x,y) \in I} (I(x+2, y) - I(x, y))^2 \dots \text{für} : |I(x+2, y) - I(x, y)| < \theta_t \quad (25)$$

#### Tenengrad-Gradient

Das nach Tenenbaum ([51]) formulierte Fokuskriterium sieht die Faltung des Bildes mit 3x3-Sobel-Operatoren  $s_x, s_y$  sowohl in x- als auch in y-Richtung vor.

$$S = \sum_{(x,y) \in I} (I(x, y) \otimes s_x)^2 + (I(x, y) \otimes s_y)^2 \quad (26)$$

mit:

$$s_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}; s_y = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} \quad (27)$$

#### Gaussian Derivatives - Gaußsche Ableitung 1. Ordnung

Das von Geusebroek in [18] erläuterte und auch patentierte Fokusverfahren verwendet als zentrales Element die Ableitung der Gauß-Funktion 1. Ordnung (First-order Gaussian Derivatives). Der Schärfewert errechnet sich dann über die Summen der Quadrate der Faltungen des Bildes mit den 1D-Operatoren  $G_x, G_y$  in x- und in y-Richtung.

$$S = \frac{1}{\text{Höhe} \cdot \text{Breite}} \sum_{(x,y) \in I} (I(x, y) \otimes G_x(x, y, \sigma))^2 + (I(x, y) \otimes G_y(x, y, \sigma))^2 \quad (28)$$

Zur Veranschaulichung: der Prewitt-Kantenoperator ist eine Approximation der Gaußschen Ableitung.

## Laplace-Operator

In gleicher Weise wie zuvor der Sobel-Operator kann der Laplace-Operator mit verschiedenen Gewichtungen  $g$  als Kernel zur Faltung mit dem Bild verwendet werden. Er spricht zudem auf Kanten aller Orientierungen an.

$$S = \sum_{(x,y) \in I} (I(x,y) \otimes L)^2 \quad (29)$$

mit:

$$L = g \cdot \begin{bmatrix} -1 & -4 & -1 \\ -4 & 20 & -4 \\ -1 & -4 & -1 \end{bmatrix} \quad (30)$$

Eine Erweiterung stellt der Laplace-Of-Gaussian-Operator als zweite Ableitung der Gauß-Funktion dar ([47]).

### 4.1.3 Funktionen der Bildstatistik

#### Varianz-Messung

In vielen Autofokus-Untersuchungen hat sich die Verwendung der normalisierten Varianz-Funktion als sehr erfolgreich herausgestellt. Hierbei werden statt der direkten Gradientenbetrachtung die Pixel-Intensitätsunterschiede im Vergleich zum Mittelwert  $\mu$  summiert und das Maximum ermittelt.

$$S = \frac{1}{\text{Höhe} \cdot \text{Breite}} \sum_{(x,y) \in I} (I(x,y) - \mu)^2 \quad (31)$$

Die Quadrierung der Differenz soll hohe Unterschiede hervorheben.

#### Korrelationsmessung

Diese Schärfefunktionen haben einen direkten Bezug zum Leistungsbetrag des Frequenzspektrums, werden allerdings durch die Autokorrelation im Bildraum berechnet.

$$S = \sum_{(x,y) \in I} (I(x,y) \cdot I(x+1,y) - \sum_{(x,y) \in I} I(x,y) \cdot I(x+2,y)) \quad (32)$$

Vollath ([53]) führt mehrere korrelationsbasierte Funktionen ein; darunter auch die „Standard Deviation-based Correlation“, welche die Autokorrelation unter Einbeziehung des Intensitätsmittelwerts  $\mu$  berechnet:

$$S = \sum_{(x,y) \in I} I(x, y) \cdot I(x + 1, y) - (Höhe \cdot Breite \cdot \mu^2) \quad (33)$$

## 4.2 Strategien der Fokussuche

Die Suchtechnik bestimmt die Vorgehensweise, in der das Untersuchungsobjekt entlang der z-Achse durchfahren, die Fokusschärfefunktion ausgeführt und die Fokusebene prädiert wird. Die wichtigsten Kriterien der Suchstrategie sind die Schnelligkeit der Konvergenz und die Zuverlässigkeit, das globale Maximum zu finden. Die zeitliche Komponente ist besonders bei ausbleichenden Proben von großer Bedeutung.

### 4.2.1 Suche mit fester und variabler Schrittweite

Im einfachsten Fall wird das Präparat mit einer festen Schrittweite zwischen unterer und oberer Probengrenze gescannt. Die optimale Schrittweite  $z_{\Delta}$  wird durch die Abbesche Auflösungsgrenze vorgegeben:

$$z_{\Delta} = \frac{0,61\lambda}{NA} \quad (34)$$

Der Scan-Prozess kann nach der Hill-Climbing-Strategie erfolgen. Hierbei wird aus zwei Messungen die Differenz des Schärfewertes berechnet. Bei positivem Vorzeichen behält die Bewegung der Optik ihre Richtung, andernfalls wird die entgegengesetzte eingeschlagen. Dieses Vorgehen setzt jedoch strikt eine unimodale Fokuskurve voraus.

Bezüglich des zeitlichen Aufwands ist es jedoch sinnvoller, mit einer variablen Schrittweite zu arbeiten. Durch die Kombination einer initialen groben Suche mit großer Schrittweite und einer Feinsuche in kleinen Schritten wird ein erheblicher Zeitgewinn erzielt. Allerdings besteht die Gefahr, ein lokales Maximum zu ermitteln, wenn die grobe Schrittweite zu groß gewählt wird. Die Schrittweite wechselt adaptiv von grob zu fein, wenn die Schärfedifferenz zweier Messpunkte einen bestimmten Schwellwert überschreitet und somit den steilen Anstieg zum Maximum signalisiert.

### 4.2.2 Einfache iterative Suche

Bei der iterativen Strategie wird der Suchbereich sukzessiv eingegrenzt. Nach jeder Iteration bildet das ermittelte Maximum der Schärfefunktion den Mittelpunkt des neuen verkleinerten Suchintervalls. Wird die Anzahl der äquidistanten Messpositionen für jede Iteration konstant gehalten, verringert sich automatisch die Schrittgröße und wird zum globalen Maximum hin am kleinsten. Die iterative Suche ist eine Form der Teile-Und-Herrsche-Suche, die als binäre Suche verstanden werden kann. Ziel der binären Suche ist es, das gesuchte Element mit der Suchintervallmitte in Übereinstimmung zu bringen. Im konkreten Fall der Fokussuche ist die Fokusebene gefunden, wenn die Intervallmitte dem Funktionsmaximum entspricht.

### 4.2.3 Fibonacci-Suche

Eine weitere Form der iterativen Teile-und-Herrsche-Methode ist die Fibonacci-Suche, die aber auch zwingend einen unimodalen Fokuskurvenverlauf voraussetzt. Sie ist im Grunde genommen eine Suche entsprechend des Goldenen Schnitts (Golden-Section-Search), die in jeder Iteration den Messpunkt innerhalb des Intervalls so legt, dass das Verhältnis der Abstände dieses Punktes zur linken und zur rechten Intervallgrenze die Goldene Zahl 1,618 beträgt ([14]). An dieser Einteilung orientiert sich auch die Fibonacci-Suche, die jedoch mit ganzen Zahlen operiert. Diese Zahlen werden der Fibonacci-Folge entnommen, die durch die Formel:

$$F_k = \frac{1}{\sqrt{5}}(G^k - (G + \sqrt{5})^k) \quad (35)$$

mit:

$$G = \frac{1 + \sqrt{5}}{2} \quad (36)$$

gebildet wird. Dabei ist jede Zahl  $F_k$  dieser Folge die Summe der zwei vorangegangenen Werte  $F_{k-1}$ ,  $F_{k-2}$  und das Verhältnis zweier aufeinanderfolgender Zahlen ist in der Näherung die Goldene Zahl  $G$ .

<b>k</b>	$F_0$	$F_1$	$F_2$	$F_3$	$F_4$	$F_5$	$F_6$	$F_7$	$F_8$	$F_9$	$F_{10}$	$F_{11}$	$F_{12}$
<b>F<sub>k</sub></b>	0	1	1	2	3	5	8	13	21	34	55	89	144

**Tabelle 1: Fibonacci-Folge bis zur 12. Stelle.**

Der Such-Algorithmus verfährt auf der Grundlage der Fibonacci-Zahlen wie folgt: Gegeben sei das Suchintervall in der ersten Iteration durch die Grenzen  $[a, b]$  mit  $N$  möglichen

äquidistanten Messpositionen. Idealerweise ist  $N$  eine Zahl der Fibonacci-Folge, so dass  $N = F_k$  ist. Die zwei Schärfemessungen der ersten Iteration werden nun an den Positionen  $a + F_{k-2}$  und  $a + F_{k-1}$  berechnet. Für den Fall, dass der Wert an der Stelle  $a + F_{k-2} > a + F_{k-1}$  ist, wird der Suchbereich in der nächsten Iteration auf das Intervall  $[a, a + F_{k-1}]$  und im umgekehrten Fall auf das Intervall  $[F_{k-2}, b]$  eingegrenzt. Die folgenden Iterationen arbeiten dann nach demselben Schema. Dieser Algorithmus ist noch einmal grafisch durch Abbildung 12 erläutert.

Neben der Unimodalitätsbedingung besitzt die Fibonacci-Suche einen weiteren Nachteil. Durch das Vor- und Zurückfahren der Optik bzw. des Objektisches während der Suche tritt der Backlash-Fehler (siehe Abschnitt 3.3.1) entsprechend stärker in Erscheinung.

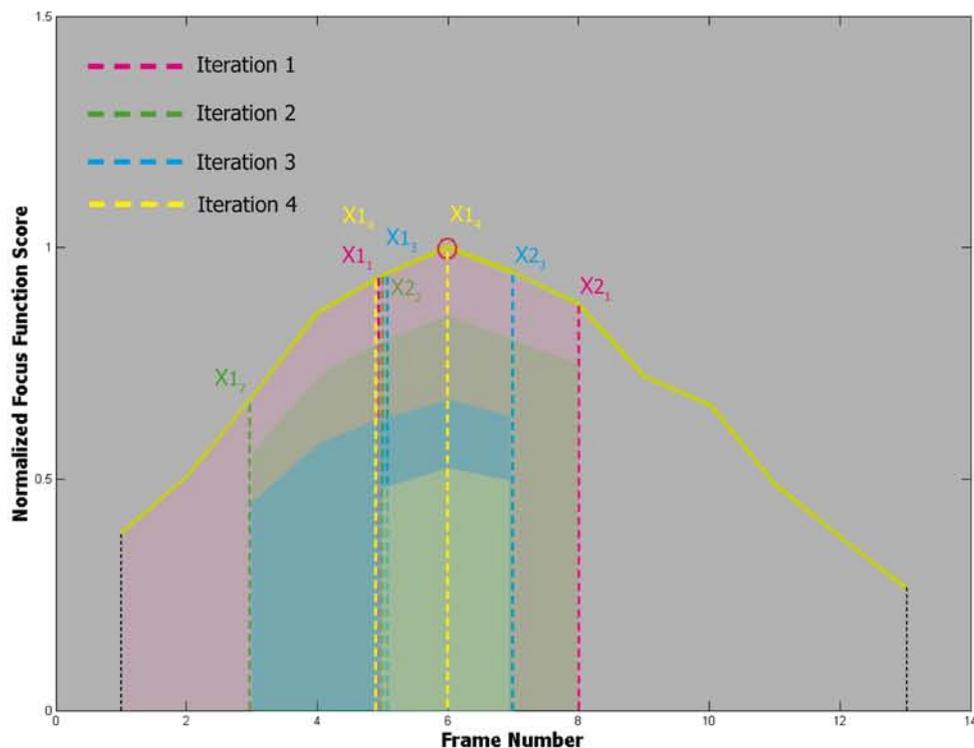


Abbildung 12: Grafische Darstellung der Iterationsstufen einer Fibonacci-Suche an einer simulierten Fokuskurve. Nach der 1. Iteration wird das Intervall auf den Bereich  $[1, X_{2_1}]$  eingegrenzt, weil  $X_{1_1} > X_{2_1}$  ist. Die ersten zwei Messwerte liegen entsprechend der Fibonacci-Folge bei den Frame-Nummern 5 und 8, da der Ursprungsbereich 13 mögliche Messpositionen umfasst. In der zweiten Iteration liegen die Messwerte bei 3 und gleichfalls 5. Nach der 4. Iteration hat sich der Suchbereich auf das Intervall  $[X_{3_1}, X_{3_2}]$  verkleinert. (Quelle: eigene Darstellung)

#### 4.2.4 Suche mit Fokuskurven-Interpolation

Die Anzahl der Messungen kann durch die Schätzung des Maximums per Interpolation der Fokuskurve mit wenigen Messwerten verringert werden. Diese Messwerte produziert

ein initialer Suchdurchlauf mit großen Schrittweiten. Im Anschluss erfolgt eine Kurvenanpassung (Curve Fitting). Mit dem Wissen, dass die Fokuskurve gleich welcher Schärfefunktion im Idealfall eine Gauß-Form annimmt, können die Parameter der Gauß-Funktion schrittweise durch die Methode der kleinsten Fehlerquadrate (Method of Least Squares) an die Messpunkte angepasst werden. Weil diese Interpolation immer nur eine Schätzung darstellt, ist es sinnvoll, die Region des Maximums der angepassten Gaußkurve einer genaueren Fokussuche zu unterziehen.

### 4.3 Autofokussierung unter Berücksichtigung von Artefakten und hoher Probendicke

Den im Unterkapitel 4.1 erläuterten passiven Autofokus-Algorithmen liegt immer die Annahme zugrunde, dass lediglich eine Fokusebene existiert oder von Interesse ist. Setzt man störungsfreie und einschichtige Präparaten voraus, ist das Gros dieser Methoden bei der Fokusfindung hinreichend erfolgreich. Es finden sich jedoch wenige Publikationen, welche die nicht idealen Fokussierungsbedingungen durch den Einfluss von Artefakten und speziellen Präparateigenschaften genauer thematisieren. Darunter fällt unter anderem die Verteilung von zu untersuchenden Objekten in mehreren Fokusebenen bei Proben großer Dicke.

So stellen Shen et al. ([47]) wie auch Price und Gough ([41]) richtig fest, dass dicke Proben mit mehreren Fokusebenen zu polymodalen Fokuskurven führen, bei denen eine Kurvenanpassung (Curve Fitting) versagt. Die daraus gezogene Konsequenz bedeutet jedoch nur eine Anpassung der Hochpass-Schärfefunktion, um die Kurve idealer zu formen und letztlich nur auf die eine Ebene scharf zu stellen, welche die größte Information enthält. Die untersuchten Proben beinhalteten jeweils ein Objekttyp und keine Artefakte, so dass bei multiplen lokalen Maxima von Fokusebenen ausgegangen werden konnte, die interessierende Objekte enthielten.

Einen Schritt weiter gehen Brázdilová und Kozubek ([9]), die gezielt mit dicken Präparaten arbeiten, um die Vorteile ihres Autofokus mit ‚Informationsgehalt-Analyse‘ bei vielschichtigen Proben zu demonstrieren. Sie benutzen mit der Normalisierten Varianz einen existierenden Schärfalgorithmus, legen ihr Augenmerk jedoch auf die Suchstrategie, die es verbietet, lokale Maxima zugunsten eines globalen Maximums auszuschließen. Damit erreichen die auf 10 getesteten Proben eine Erkennung von 85% aller möglichen 57 Fokusebenen.

Bei der Untersuchung von Zell- und Gewebeproben stießen Della Mea et al. ([13]) sporadisch auf Leerflächen im Präparat, welche die Autofokusmessung veranlasste, Dreck-

partikel auf dem Abdeckglass scharf zu stellen. Sie umgingen diesen Fehler durch Voranalyse des Histogramms und schlossen Ausschnitte mit geringer Pixelwertverteilung von vornherein aus. Somit stellten sie dem AF-Schritt eine, wenn auch unspezifische, Inhaltsuntersuchung voran.

Die Fokussierung von Blutausrichen in den Untersuchungen von Makkapati ([35]) setzt eine weitere Überlegung um. Hierbei führt die der Wavelet-basierten Fokussierung vorgestellte Extraktion des Hintergrundes im grünen Farbkanal zu einer robusteren Schärfstellung. In der Veröffentlichung wurde ebenfalls das Vorhandensein von unterschiedlichen Objekttypen in einem Präparat thematisiert. Neben den Erythrozyten wurden gleichfalls weiße Blutkörperchen segmentiert und in die Schärfmessung einbezogen. Dies wurde jedoch nicht als problematisch angesehen und sogar für die qualitative Nachuntersuchung als erwünscht dargestellt.

Weil für die vorliegende Arbeit die Fokussierung speziell auf relevante Objekte die Hauptmotivation darstellt, sind vor allem die folgenden zwei Patente von Interesse. Zum einen handelt es sich um das patentierte Verfahren „zum automatischen Fokussieren auf Objekte“ ([46]), das in einem ersten Schritt festlegt, auf welchen Bereich des Blickfeldes die Fokussierung angewendet wird. Dieses Konzept, eine adaptive ROI-Einschränkung nur auf interessierende Objekte anzuwenden, erbringt die elaborierteste aller bisher besprochenen Methoden. Allerdings ist die Umsetzung der ROI-Suche mit einer auf Farb- oder Grauwerten basierten Charakterisierung nur vage beschrieben und erscheint unausgegrenzt.

Zum anderen findet sich in dem amerikanischen Patent von Frost et al. ([16]), die ebenfalls die Fokussuche auf interessierende Objekt einengen, die konkrete Aussage, dass mit ihrem Verfahren auch Artefakte ausgeschlossen werden sollen. Die Umsetzung erfolgt per Klassifizierung eines Bildpixels auf der Grundlage des Pixelintensitätswertes. Das Ausführungsbeispiel in der Patentschrift erläutert die Erkennung anhand eines Pap-Abstrichs, indem vornehmlich die Zellen von Interesse sind. Durch einfachen Grauwertvergleich mit zwei Schwellwerten werden die Pixel in die Klassen Zelle, heller oder dunkler Hintergrund eingeteilt, wobei Artefakte die gleichen Intensitätswerte wie der Hintergrund besitzen sollen. Die Schärfmessung wird im Anschluss nur auf Arealen der Zell-Klasse durchgeführt. Es kann zusammengefasst werden, dass beide Patente den fortschrittlichen Gedanken gemein haben, den tatsächlichen Probeninhalt auf interessierende Objekte zu untersuchen und nur auf diese zu fokussieren. Die tatsächliche Umsetzung ist jedoch auf sehr konkrete Anwendungen zugeschnitten und kaum universell anwendbar. Zudem liegen zu keinem der beiden Patente dokumentierte Auswertungen hinsichtlich der Autofokussierungsqualität vor.

## 5 Material und Methoden

Die vorliegende Arbeit führt das Konzept, die Fokussmessung auf relevante Bildbereiche einzuschränken, konsequent durch, indem sie einen grundlegend neuen Ansatz in das Entwicklungsgebiet der automatisierten Fokussierung einbringt. Im Unterschied zu vorhandenen Methoden (siehe Abschnitt 4.3) soll der hier vorgestellte Inhaltsbasierte Autofokus (IB-AF) in der Lage sein, nicht nur mehrere Fokusebenen in einem Präparat zu finden, sondern auch unter unterschiedlichen Lichtmikroskop-Modi bei allen Arten von Präparaten zu arbeiten und explizit die Existenz von Artefakten und irrelevanten Objekten einzubeziehen. Hiervon ausgeschlossen sind solche Modi wie z.B. die konfokale Mikroskopie, die aufgrund ihrer Bauart bereits fokussierte Bilder liefern.

Eine wichtige Voraussetzung für die Umsetzung des IB-AF ist ein Vorwissen über die zu untersuchende Probe. Im neuen IB-AF müssen demnach Schärfemessung und Objekterkennung zusammenspielen. Es ist eine Klassifikationsaufgabe aus dem Bereich der Objekt- und Mustererkennung, die dem AF Informationen über den Inhalt der zu untersuchenden Probe liefert und es ihm ermöglicht, irrelevante Subjekte auszuschliessen. Wie im Folgenden näher erläutert wird, fußt die Bereitstellung des Vorwissen jedoch nicht auf konkretem Wissen über das Untersuchungsobjekt wie z.B. Probenart und -inhalt. Vielmehr bringt das Verfahren eine allgemeingültige Methodik hervor und ist somit eine Art Gegenentwurf zu speziellen Insellösungen.

### 5.1 Konzept des Inhaltsbasierten Autofokus

#### 5.1.1 Modulare Struktur

Das Inhaltsbasierte Autofokusverfahren gliedert sich in drei hintereinander geschaltete Teilaufgaben, die als Module angesehen werden können. Auf das erste Modul, die Erzeugung von Objekthypothesen und deren Bildorte, folgt die Fokussmessung auf den ermittelten Bildregionen. Daran anschließend werden diese Regionen in der entsprechenden Fokusslage einer Objektverifikation zur Vermeidung von Falsch-Positiv-Erkennungen unterzogen.

Jedes Modul kann einzeln betrachtet und modifiziert werden, weil es theoretisch allein über die Schnittstellen mit den Nachbarmodulen verbunden ist: Das zweite Fokussierungsmodul ist lediglich auf die im ersten Schritt ermittelten Koordinaten der Region-Of-Interest (ROI) angewiesen, während der Verifikationsteil neben diesen Koordinaten zusätzlich die jeweils erkannten Fokusslagen benötigt. Den schematischen Ablauf und die Übergabeparameter stellen die Abbildungen 13 und 14 dar.

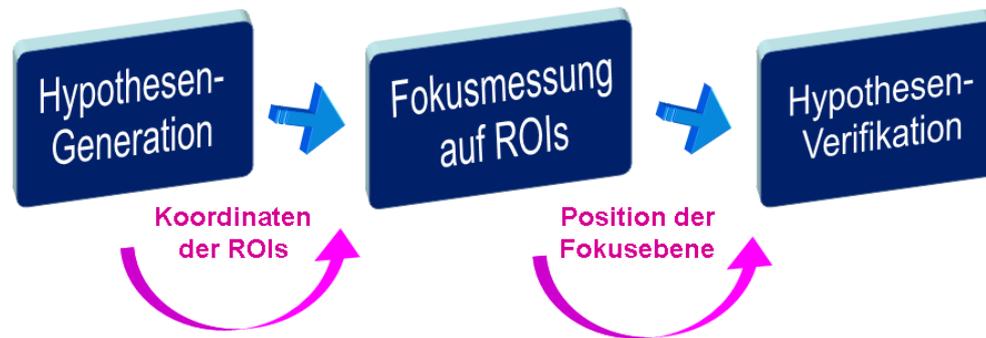


Abbildung 13: Module der inhaltsbasierten Fokussuche. (Quelle: eigene Darstellung)

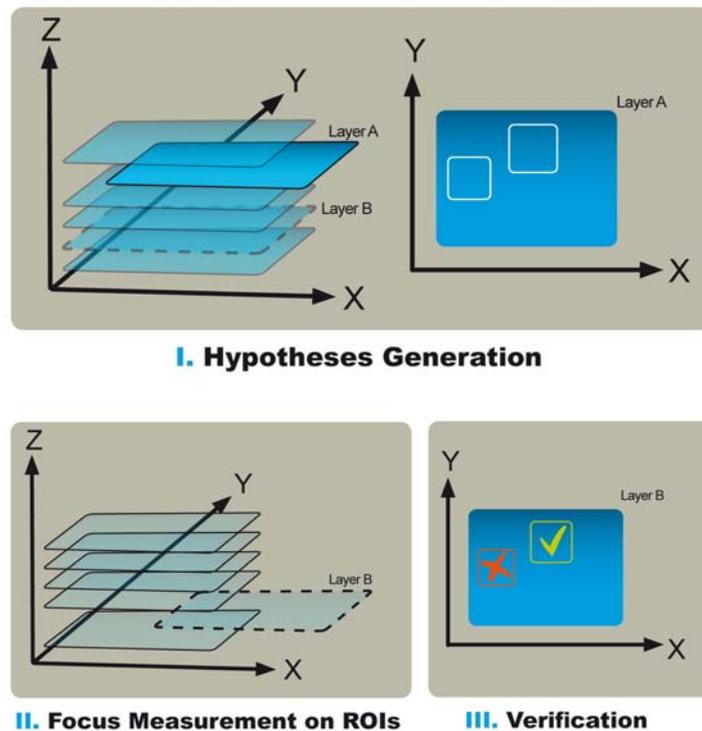
Für die Umsetzung des ersten und des dritten Moduls müssen Fragestellungen der Objektklassifikation beantwortet werden. Damit verbunden ist die Suche nach geeigneten Methoden, um Objektmerkmale zu extrahieren und mit ihnen einen Klassifikator zu trainieren. Das mittlere Modul beinhaltet dagegen lediglich eine passende Fokusschärfefunktion und die dazugehörige Suchstrategie.

**Hypothesengeneration** Hypothesen bezeichnen in dieser Arbeit durchweg Bildareale, die mit einer höheren Wahrscheinlichkeit die Position in der XY-Ebene von interessierenden Objekten angeben. Bei der Erzeugung von Hypothesen bzw. ROIs innerhalb einer Bildaufnahme kann allein davon ausgegangen werden, dass die Mikroskopoptik auf einen Bereich zwischen der oberen und unteren Probengrenze eingestellt ist.

Dabei ist immer davon auszugehen, dass die erste Bildaufnahme Objekte im defokussierten Zustand abbilden. Auch wenn die abgebildete Ebene nicht zu stark defokussiert ist, dient sie dennoch zur initialen Bestimmung von möglichen Objektpositionen. Eine sichere Objekterkennung ist bei unscharfer Abbildung nur schwer beziehungsweise überhaupt nicht zu realisieren. Deshalb wird in diesem Stadium noch von Objekthypothesen gesprochen.

**Fokussierung innerhalb der ROIs** Die Ergebnisse in Abschnitt 6.2.2 zeigen, dass die Maxima-Bestimmung von bekannten Schärfefunktionen eine zuverlässige Adressierung der Fokusebene ermöglicht, sofern die Suchregion auf die interessierenden Objekte eingeschränkt wird. In dieser Hinsicht greift die Wahl der Schärfefunktion für das Fokussmodul auf eine klassische Gradientenfunktion zurück.

**Hypothesenverifikation** Es ist die Funktion der Hypothesenverifikation, die ROIs der fokussierten Objekte in die Klassen „erwünschte“ und „unerwünschte“ Bildbereiche ein-



**Abbildung 14:** Die drei Schritte der modular aufgebauten Autofokussuche. I) Hypothesengeneration in einer unbestimmten unfokussierten Ebene II) Fokusmessung und Bestimmung der Fokalebene innerhalb der Hypothesenregionen III) Verifikation der Hypothesen in der Fokusebene. (Quelle: eigene Darstellung)

zuteilen. Dieser Schritt entspricht einer reinen Objekterkennungsaufgabe ohne Berücksichtigung von defokussierten Bildinhalten. Im Vergleich zu Hypothesengenerierung sind Falsch-Positiv-Erkennungen unerwünscht. Deshalb unterscheidet sich das Training des Klassifikators einerseits hinsichtlich der Datenbasis, die keine defokussierten Bildinhalte berücksichtigen muss, andererseits auch bezüglich der Klassifikationsgenauigkeit.

## 5.2 Universelle Objekterkennung in mikroskopischen Proben

Das Hypothesengenerierungsmodul muss zwei Aspekte bei der Ermittlung von möglichen Arealen mit interessierenden Objekten berücksichtigen. Zum einen soll ein universelles Verfahren die Bestimmung von vielen erdenklichen Objektarten ermöglichen, zum anderen muss davon ausgegangen werden, dass sich die zu erkennenden Objekte im defokussierten Zustand befinden. Der Weg führt dabei über eine Klassifikation, die auf dem vorherigen Training auf extrahierten Merkmalen basiert. Für die Auswahl der Merkmalsextraktion innerhalb der Objekterkennung existieren verschiedene Ansätze ([43]). Modellbasierte oder formbasierte Ansätze scheiden an dieser Stelle aus, denn sie setzen

ein detailliertes Modell des zu erkennenden Objekts voraus. Dies steht zum einem der in dieser Arbeit formulierten Ziele, nämlich der universellen Objekterkennung, entgegen. Als allgemein einsetzbare Methode kann dementsprechend nur ein ansichtsbasiertes Verfahren Anwendung finden, weil es nicht auf spezifische Formen oder Muster zurückgreift. Zum anderen muss wegen der variablen Mikroskopvergrößerung und der unbestimmten Lage der zu untersuchenden Objekte im Präparat die Methode unbedingt skalen- und rotationsinvariant sein.

Für die Implementierung in die Erkennungsmodule kommen deshalb nur die drei folgenden Merkmalsextraktionsverfahren und die mit ihnen verbundenen Merkmalsselektionsmethoden potentiell in Frage: Die Principle Components Analysis (PCA) ([39]) berechnet Merkmalsvektoren, die durch Transformation des ursprünglichen Datensatzes auf der Grundlage der Datenvariation entstehen. Die Skaleninvariante Merkmalstransformation (SIFT - Scale Invariant Feature Transform) ist eine Interesting-Points-Detektion, während die Merkmalsextraktion mit Haar-Filtern im Grunde genommen auf einer einfachen Gradienten-Betrachtung basiert.

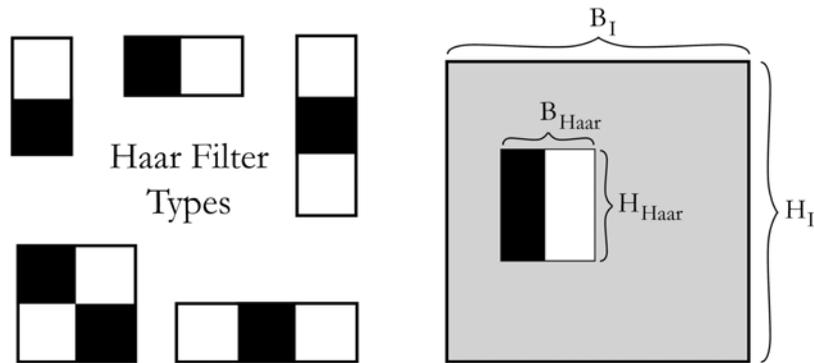
Auf Grundlage der Arbeit von Papageorgiou et al. ([37]) finden haar-ähnliche Filter zur Berechnung der Merkmale Verwendung, die ebenso wie die Haar-Transformation eine Wavelet-Transformation darstellen. Im Konkreten ist das Merkmal der Wert, den die Faltung der jeweiligen Haar-Filtermaske mit dem Bildausschnitt an einer bestimmten Bildkoordinate liefert. Diese Faltung kann auch als Differenz der Pixelintensitätssummen benachbarter Bildregionen beschrieben werden.

Die Bezeichnung "Haar-ähnlich" ist darauf zurückzuführen, dass sich der hier verwendete Satz an Filterfunktionen von den ursprünglichen Basis-Haar-Wavelets unterscheidet. Beide Transformationen liefern allerdings Filterantworten, die als Differenzen der Pixelwerte der zu filternden Bildregionen angesehen werden können.

### 5.2.1 Merkmalsextraktion mit Haar-Filtern

Abbildung 15 stellt die Basis-Wavelets der Haar-ähnlichen Filter samt unterschiedlicher Orientierungen dar. Es ergeben sich drei grundsätzliche Typen:

- **Zwei-Rechtecke-Filter**, das die Differenz der Summen der Pixelwerte zweier rechteckiger Regionen bildet (TYP 1)
- **Drei-Rechtecke-Filter**, das die Differenz der Pixelwert-Summen von zwei außen liegenden rechteckigen Regionen mit der Summe der innen liegenden Region berechnet (TYP 2)



**Abbildung 15: Haarfilter-Typen und ihre mögliche Positionierung im Bild bzw. Bildausschnitt.** (Quelle: eigene Darstellung)

- **Vier-Rechtecke-Filter**, das die Differenzen der Diagonal-Paare von rechteckigen Regionen ermittelt (TYP 3)

Aus diesen Filtergrundtypen kann nun ein Satz von Filtermasken gebildet werden, die sich untereinander in Skalierung und Orientierung, sowie Position relativ zum Ursprung des Bildausschnitts unterscheiden. Grundsätzlich hat jeder der drei Filtergrundtypen eine Mindestgröße und eine maximale Ausdehnung. Letztere ist auf die Größe des Bildausschnitts begrenzt. Im Falle der Verwendung des Typs 1 wäre die Mindestgröße  $2 \times 1$  Pixel, die maximale bei einem Bildbereich der Dimension  $30 \times 30$  Pixel entsprechend  $30 \times 30$ .

Die Anzahl  $N$  der möglichen Filtermasken eines Typs ergeben sich nach Lienhart und Maydt ([31]):

$$N = \frac{B_I}{B_{Haar}} * \frac{H_I}{H_{Haar}} \left( B_I + 1 - B_{Haar} * \frac{\frac{B_I}{B_{Haar}} + 1}{2} \right) * \left( H_I + 1 - H_{Haar} * \frac{\frac{H_I}{H_{Haar}} + 1}{2} \right) \quad (37)$$

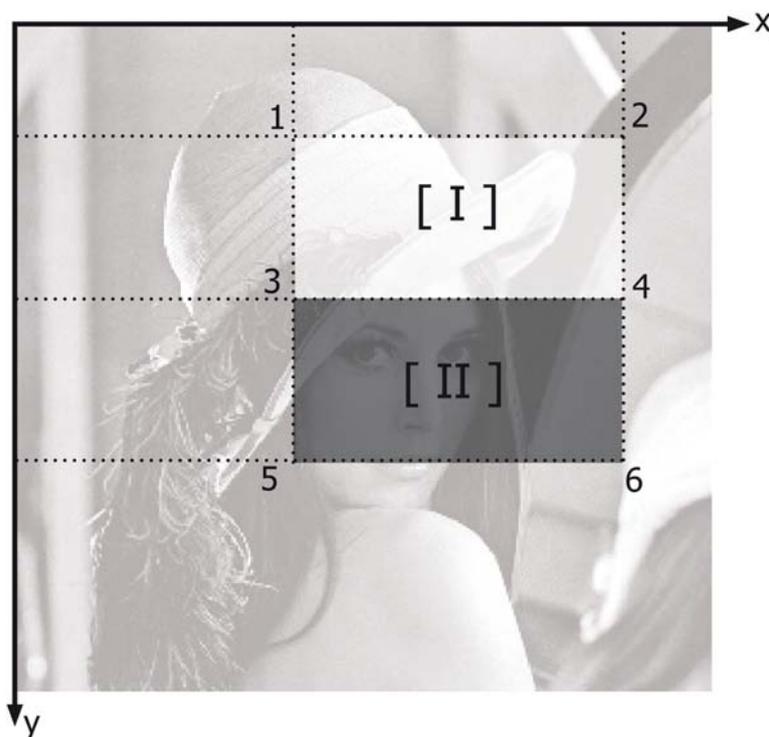
( $B_{Haar}$  und  $H_{Haar}$  stehen für die kleinstmögliche Breite und Höhe des Haarfilter in Pixel, wobei  $B_I$  und  $H_I$  für die Dimension des Bildausschnitts  $I$  stehen.)

Für das oben genannte Beispiel ergeben sich somit allein für einen Filtermaskentyp bereits über 100.000 mögliche Formen und folglich ebenso viele Merkmalswerte, die diesen Bildausschnitt beschreiben. Diese Anzahl der Merkmale übersteigt bei weitem die Pixelanzahl des Ausschnittes selbst und stellt somit eine Übererfüllung der Anforderung nach ausreichender Anzahl von Merkmalen dar. Des Weiteren führte die Berechnung all dieser Merkmale zu einem unverhältnismäßig hohen Rechenaufwand.

**Integralbild** Viola und Jones ([52]) verwendeten das Integralbild, um die Berechnung der Haar-Filter-Antworten erheblich zu beschleunigen. Das Integralbild  $B_{int}$  besitzt die selbe Dimension wie das Ausgangsbild  $B$ , enthält an den jeweiligen XY-Positionen jedoch nicht den Intensitätswert, sondern den Summenwert aller Intensitätswerte der Bildpunkte, die sich links und über dieser XY-Koordinate befinden.

$$B_{int}(x, y) = \sum_{(x,y) \in B} B(x, y) \quad (38)$$

Die Berechnung der Differenzwerte von summierten Intensitätssummen, also die hier besprochene Faltungsantwort eines Bildausschnitts mit einem Haar-Filter, vereinfacht sich durch die Verwendung des Integralbildes zu einer Summgleichung mit wenigen Summanden. Das wiederholte Auslesen einzelner Pixelwerte ist nicht mehr notwendig. Die Funktionsweise soll hier am Bildbeispiel 16 veranschaulicht werden. Die Filterantwort der dargestellten Typ-2-Maske über der Bildregion errechnet sich aus der Differenz der Pixelsummen der Areale [I] und [II].



**Abbildung 16: Berechnung der Haar-Filter-Antwort mit Hilfe des Integralbildes.** (Quelle: eigene Darstellung)

Mit Hilfe des Integralbildes errechnen sich die einzelnen Summen zu:

$$B[I] = B_{int}(4) - B_{int}(2) - B_{int}(3) + B_{int}(1) \quad (39)$$

$$B[II] = B_{int}(6) - B_{int}(4) - B_{int}(5) + B_{int}(3) \quad (40)$$

so dass die Filterantwort  $D$  einfach durch die Differenz beider Summen bzw. mit der Addition von sechs Summanden berechnet wird:

$$D = B[I] - B[II] = B_{int}(1) - B_{int}(2) - 2B_{int}(3) + 2B_{int}(4) + B_{int}(5) - B_{int}(6) \quad (41)$$

### 5.2.2 Merkmalssektion mit Adaptive Boosting

Es ist die Aufgabe des Trainings, aus der immensen Anzahl der möglichen Haar-ähnlichen Merkmale jene mit besonders diskriminierenden Eigenschaften zu ermitteln. Der modifizierte Boosting-Ansatz von Schapire und Singer ([45]) löst diese Aufgabe nicht nur, sondern verringert durch die Kaskadierung der Klassifikatoren die Menge an Rechenoperation um ein Vielfaches. Die Boosting-Kategorie steht für eine Reihe von maschinellen überwachten Lernverfahren, die auf Grundlage eines Trainingsdatensatzes iterativ einen Klassifikator erstellen. Dieser Klassifikator entsteht durch schrittweises Aneinanderfügen von mehreren schwachen Klassifikatoren.

Ein schwacher Klassifikator besitzt für sich eine schlechte Klassentrennungseigenschaft und erwirkt damit kaum mehr als eine Zufallsentscheidung. Die Gesamtperformance wird allerdings durch die Linearkombination vieler schwacher Klassifikatoren erheblich gesteigert. Der AdaBoost-Algorithmus (Adaptive Boosting) von Schapire unterscheidet sich vom klassischen Boosting in der Hinsicht, dass dem starken Klassifikator schrittweise beliebig viele schwache Klassifikatoren hinzugefügt werden, bis eine vorgegebene Fehlerrate erreicht ist. Der Algorithmus wird deshalb als adaptiv bezeichnet, weil sich die Methode an die Trainingsfehler der schwachen Klassifikatoren in jeder Iteration anpasst. Diese Adaption ist eine Neugewichtung der Elemente des Trainingsdatensatzes. Dabei werden Elemente, die von einem vorherigen schwachen Klassifikator falsch eingeordnet wurden, eine höhere Gewichtung zugeordnet. Der vereinfachte Ablauf des Algorithmus ist in der Tabelle 2 dargestellt.

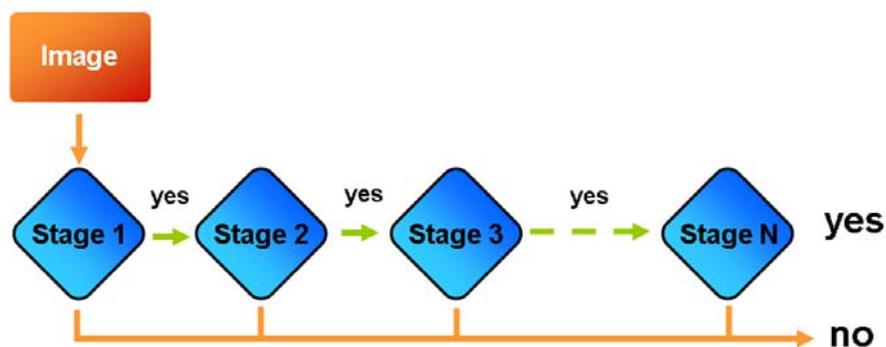
**Kaskadierung der Klassifikatoren** Viola und Jones präsentierten in ihrer maßgeblichen Veröffentlichung ([52]) über Gesichtserkennung nicht nur das oben beschriebene

**FOR t = 1 : T Iterationen**

<b>I.) Bestimmung des Weak Classifiers</b>	Finde für jeden Haar-Filter $h_i \in H$ aus dem Gesamtpool aller möglichen Haar-Filter, die noch nicht ausgewählt worden, den Schwellwert $s_i$ , der das beste Klassifikationsergebnis auf dem gelabelten Trainingsdatensatz $I(x)$ unter Berücksichtigung der Gewichtung $D_t$ der einzelnen Trainingsbilder erzielt. Der Startwert aller Gewichtungen ist die reziproke Bilderanzahl: $D_1(x) = NbrOfTrainingImages^{-1}$
<b>II.) Bestimmung des Schwellwert</b>	Bestimme für diese Haar-Filter mit ermittelten Schwellwert $s_i$ den einen Filter, der im Vergleich das beste Klassifikationsergebnis auf dem Trainingsdatensatz erzielt. Das heißt: Wähle den Haar-Filter $h_t$ , der zum geringsten Klassifikationsfehler führt.
<b>III.) Berechnung des Beitragsfaktors</b>	Berechne den Beitragsfaktor $\alpha_t = \frac{1}{2} * \log(\frac{1-E_t}{E_t})$ dieses Haar-Filters zum Gesamtklassifikator $C = \alpha_t * h_t + \sum_{u=1..t-1} \alpha * h_u$ mit: $E_t = \frac{W_a}{W_e}$ , wobei $W_a$ die Summe aller Gewichte und $W_e$ die Summe aller Gewichte der falsch-klassifizierten Trainingsbilder darstellen.
<b>IV.) Aktualisierung der Gewichte</b>	Aktualisiere die Gewichte der Trainingsbilder für die nächste Iteration: $D_{t+1}(x) = D_t(x) * e^{\alpha_t}$ wenn: $I(x)$ durch den Haar-Filter $h_t$ korrekt klassifiziert wurde $D_{t+1}(x) = D_t(x) * e^{-\alpha_t}$ wenn: $I(x)$ durch den Haar-Filter $h_t$ falsch klassifiziert wurde

**Tabelle 2: Ablauf des AdaBoost-Algorithmus mit Haar-Filtern als so genannte schwache Klassifikatoren (weak classifiers)**

Integralbild, sondern auch einen effektiven Weg, den Erkennungsablauf über eine Kaskadierung der Klassifikatoren zu beschleunigen (siehe Abb. 17).



**Abbildung 17: Kaskadierung von Klassifikatoren in einzelne Stufen.** (Quelle: eigene Darstellung)

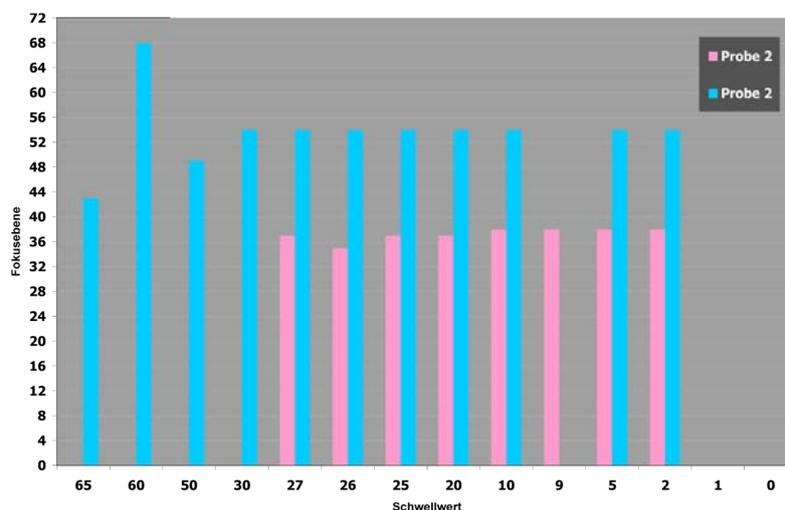
Ein zu untersuchender Bildausschnitt muss alle Kaskadenstufen mit positivem Ergebnis durchlaufen, um als Positiverkennung klassifiziert zu werden. Liefert eine Stufe ein negatives Ergebnis, wird der Durchlauf beendet.

Die grundsätzliche Idee besteht darin, dass auf einem zu untersuchenden Objekt nicht ka-

tegorisch alle Klassifikatoren angewendet werden. Vielmehr kann der Erkennungsprozess dadurch abgekürzt werden, dass die Klassifikatoren sukzessiv angewendet werden und die Prozedur an der Stelle endet, wo ein Klassifikator das Untersuchungsobjekt ablehnt. Das Training der einzelnen Kaskadenstufen erfolgt auf der Basis einer vorgegebenen Erkennungsrate und einer Falsch-Positiv-Rate (FP-Rate). Bei fortschreitender Stufe werden die Erkennungsrate erhöht und die FP-Rate verringert, bis die Gesamtklassifikation die Zielraten erreicht. Im Ergebnis wird bereits in der ersten Stufe eine Vielzahl von Negativen verworfen.

### 5.3 Wahl der Schärfefunktion

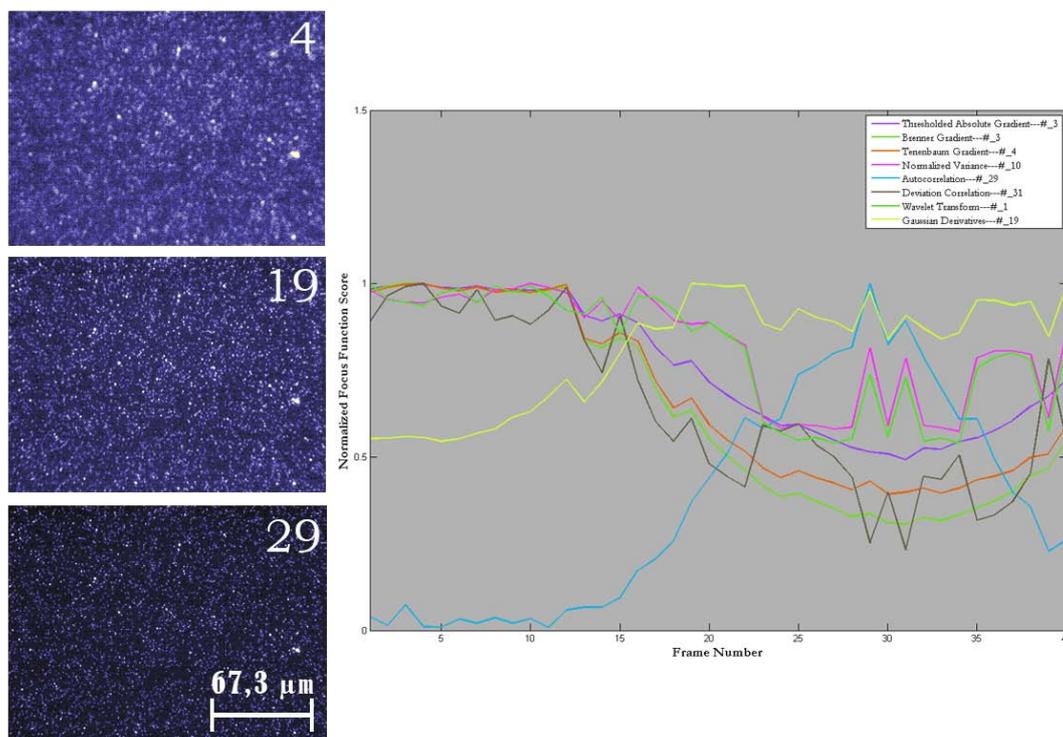
Im Gesamtsystem des IB-AF spielt die eigentliche Schärfefunktion zwar eine untergeordnete Rolle. Eine unzuverlässige Methode kann dennoch die Gesamtfunktionalität erheblich degradieren. Die Mehrheit der bisherigen Veröffentlichungen zeigt übereinstimmend schlechte Ergebnisse im Bezug auf schwellwertabhängige Verfahren wie dem schwellwertbasierten Gradientenbetrag ([44]) und dem Brenner-Gradienten (siehe Abschnitt 4.1.2), die sich vollständig mit den eigenen Untersuchungen decken. Vor allem die Abhängigkeit von einem Schwellwert hat zum Ergebnis, dass besonders in Proben mit Polyfokalebenen bei unterschiedlichen Schwellwerten auch jeweils unterschiedliche Fokusebenen angezeigt werden (siehe Abbildung 18).



**Abbildung 18: In Abhängigkeit vom Schwellwert ermittelt die schwellwertbasierte Gradientenmethode unterschiedliche Fokusebenen - hier dargestellt für zwei CaCo-Zellproben im Fluoreszenzkanal.** (Quelle: eigene Darstellung)

Weiterhin hat Bildrauschen einen gewichtigen Einfluss auf die Fokussierungsgenauigkeit. Diesen Umstand teilen sich die schwellwertbasierten Verfahren nicht nur mit der

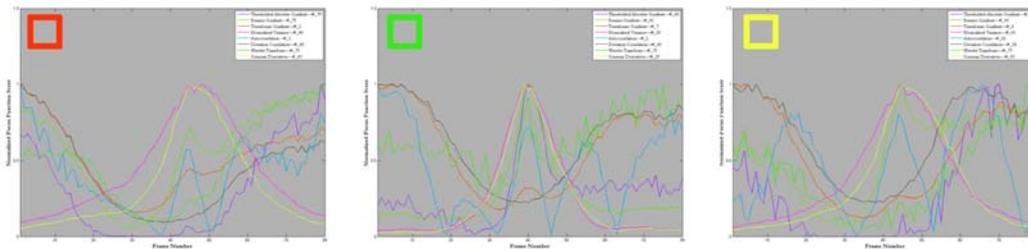
Tenengrad-Methode. Wie die Abbildung 19 verdeutlicht, leiden sowohl der Wavelet-basierte Algorithmus als auch die Normalisierte Varianzanalyse unter hohem Rauschen. Dahingegen kann sowohl das Verfahren der mittelwertbasierten Korrelation (siehe Gleichung 33) als auch Gaussian Derivatives (Gaußschen Ableitung) eine gewisse Robustheit aufweisen, die bei letzteren mit der entsprechenden Wahl des Sigamwertes der Gauß-Funktion erzielt wird. In der Gesamtbetrachtung, das heißt unter Berücksichtigung von normalem Hellfeld-, Fluoreszenz- und Phasenkontrast-Proben, lieferten vor allem die Autokorrelation, Gaussian Derivatives und die ableitungsbasierte Korrelation die besten Ergebnisse. Autokorrelation- und Gauß-Methode zeigen - genauso wie die Varianzmethode - am häufigsten unimodale Fokuskurven, die zudem noch steile Flanken besitzen. Ganz im Gegenteil dazu stehen die sehr flachen Kurven der Deviation-based-Correlation-Methode. Zur Normalisierten Varianzmethode ist noch hinzuzufügen, dass sie alles in allem hinter den Erwartungen bezüglich der Fokussierungsqualität aus anderen Veröffentlichungen ([50, 1, 9]) zurückblieb und nur in Fluoreszenzproben überzeugen konnte.



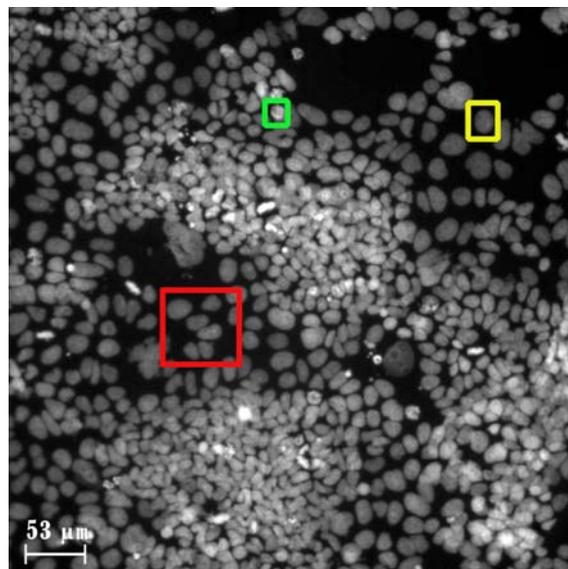
**Abbildung 19: Fokusergebnisse bei einer stark verrauschten Fluoreszenz-Beads-Probe. Die Nummern in den Bildern bezeichnen die Ebenenindizes im digitalisierten Bildstapel. Diese Indizes sind die x-Achsenbasis im Fokuskurvenplot (rechte Grafik).** (Quelle: Präparat zur Verfügung gestellt von Zeiss MicroImaging GmbH; rechter Plot: eigene Darstellung)

Neben den Ergebnissen der Ganzbildanalysen ist vor allem die Fokussierungsqualität der Schärfefunktionen bei ROI-Messungen, bei denen die Hintergrundinformationen ausgeschlossen werden, für die Algorithmenwahl relevant. Abbildung 21 zeigt die ROI-Mess-

ungen an drei Arealen, die entweder nur ein Objekt oder ein Gebiet mit gleichen Objekten in derselben Schärfeebene enthalten und frei von Bildstörungen sind. Die dazugehörigen Fokuskurven sind im Plot der Abbildung 20 dargestellt.



**Abbildung 20: Die zu 21 gehörigen Fokuskurven.** (Quelle: eigene Darstellung)



**Abbildung 21: Schärfemessung innerhalb von drei ROIs in einer Caco2-Fluoreszenz-Probe.**

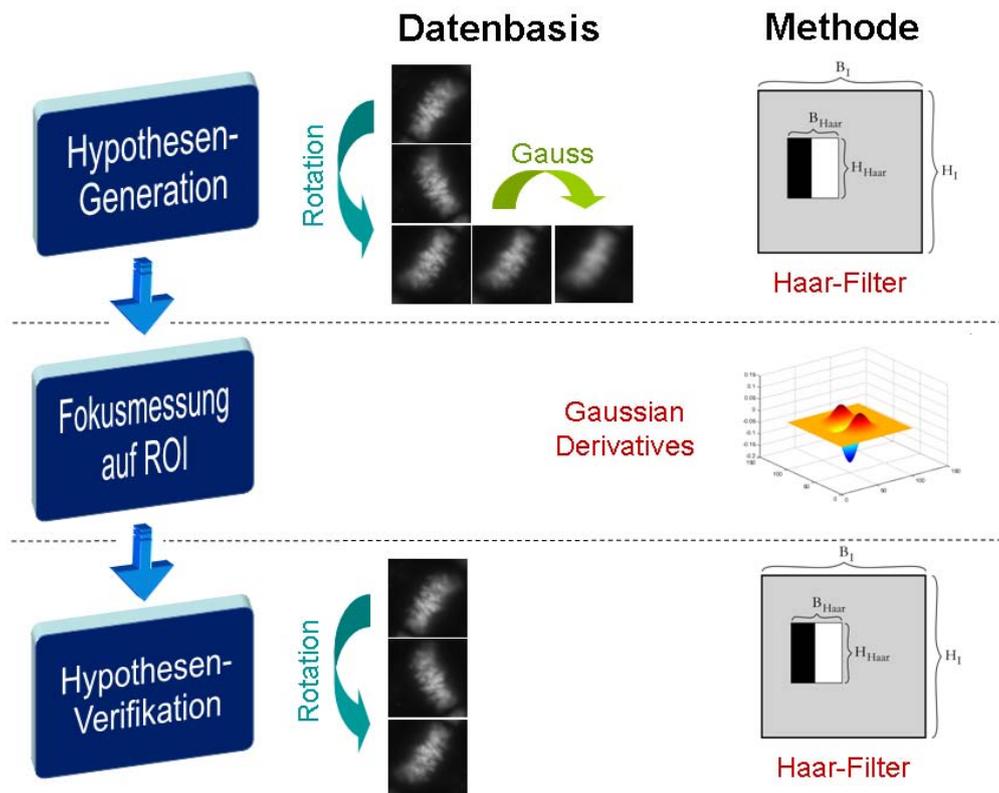
(Quelle: Präparat zur Verfügung gestellt von Zeiss MicroImaging GmbH)

Die Ergebnisse der Evaluation auf den eingeschränkten Bildbereichen decken sich dabei nicht gänzlich mit den Resultaten der Gesamtbildanalysen. Die Gradienten-Methoden können in gleicher Weise nicht überzeugen, während die Varianzmethode bei Fluoreszenzpräparaten sehr genaue Ergebnisse liefert, aber im Phasenkontrast versagt. Die Autokorrelation zeigt in der ROI-Untersuchung entgegengesetzten Charakter: sie funktioniert präzise im Phasenkontrast und scheitert in der Fluoreszenz. Letztlich gehören der Gaussian-Derivatives-Algorithmus und die standardabweichungsbasierte Korrelation in der Betrachtung auf eingeschränkten Bildbereichen zu den zuverlässigsten Funktionen.

## 5.4 Konfiguration der Gesamtmethode

Die in 5.1.1 beschriebenen Module verlangen nach drei separaten Methoden, welche die in gleichem Abschnitt erläuterten Kriterien erfüllen müssen. Die Wahl der Hypothesengenerationsmethode ist dabei die interessanteste. Die Klassifizierung von Bildregionen geschieht hier auf leicht bis stark unscharf abgebildeten Objekten und muss eine hohe Sensitivität aufweisen. Für diese Aufgabe bietet sich die SIFT-Methode (die Erläuterung der Methode findet sich im Anhang unter Abschnitt 9.2.2) besonders an, weil der Verwendung von Gauß-Pyramiden bereits eine Repräsentation von Unschärfe innewohnt. Die SIFT-Methode hat allerdings zwei entscheidende Nachteile im Bezug auf die Implementierung in die Autofokusanwendung. Denn zum einen stellt die Zuordnung ermittelter Schlüsselpunktmerkmale zu bekannten Trainingsmerkmalen die eigentliche Herausforderung bei der SIFT-Methode dar. Die Keypoint-Suche auf Bildregionen liefert grundsätzlich eine variable Anzahl von Schlüsselpunkten und damit auch Merkmale unterschiedlicher Größe. Die von Lowe ([34, 33]) vorgeschlagene erweiterte Nearest-Neighbour-Suche mit einem Best-Bin-First-Algorithmus zur Beschleunigung der Suche konnte in Testversuchen nicht die gewünschten Ergebnisse liefern. Die Kombination der SIFT-Merkmale mit Trainingsmethoden wie der SVM (Erläuterung in Abschnitt 9.3.1) scheitert im ersten Schritt an der variablen Anzahl von Merkmalen. Um die Trainingsmerkmale auf eine feste Länge zu bringen, können die Merkmale in Histogramm-Pyramiden ([20]) überführt werden. Allerdings ließ sich mit diesem aufwendigen Umweg kein Vorteil gegenüber der Haar-Methode belegen. Der zweite Nachteil der SIFT-Methode liegt in der Patentierung des Verfahrens, die bei einem kommerziellen Einsatz bzw. bei einem möglichen Patentantrag des Inhaltsbasierten Autofokus zu beachten wäre.

Die Kombination aus PCA (siehe Abschnitt 9.2.1) und SVM zeichnet sich in erster Linie durch ein vergleichsweise schnelles Training aus, das zudem gute Ergebnisse erbringt. Dennoch sind es die Haar-Filter mit AdaBoost-Training, die für die Objekterkennung letztlich verwendet werden. Neben dem hohen Maß an Einstellungen ist es vor allem die sehr schnelle Klassifikation von Bildregionen im Erkennungsprozess, welche durch die Kaskadierung der Klassifikatoren erreicht wird. Die Haar-Merkmale werden nicht nur für die Objekthypothesen, sondern auch für die Verifizierung im dritten Modul verwendet. Dabei unterscheidet sich das Training in der Datenbasis und den Trainingsparametern. Während das Training für die Objekthypothesen mit großen Zugeständnissen an die Spezifität, aber mit hoher Sensitivität durchgeführt wird, zielt das Training der Klassifikatoren für die Verifizierung auf eine geringe Falsch-Positiv-Rate ab. Für das Hypothesenmodul liegen sowohl rotierte als auch mit Gauß-Filtern unterschiedlicher Sigma-Werte gefilterte Instanzen der Trainingsbilder vor. Damit wird zum einen eine Rotationsinvari-



**Abbildung 22: Ablauf und eingesetzte Methoden und Daten der Gesamtmethode.** (Quelle: eigene Darstellung)

anz erzielt, zum anderen simulieren die weichgezeichneten Bildkopien Unschärfe. Für das Verifizierungsmodul besteht die Trainingsbasis nur aus scharf aufgenommenen Bildern. Zusätzlich kommen auch hier gedrehte Objektinstanzen hinzu. Für die Schärfefunktion wird die Funktion der Gaussian Derivatives gewählt. Sie hat sowohl auf den Ganzbildanalysen, als auch auf den Einzelobjekttests sehr gute Ergebnisse erzielt und ist auch bei verschiedenen Mikroskopmodi universell einsetzbar. Damit ergibt sich das Ablaufschema in Abbildung 22.

## 5.5 Umsetzung des IB-AF

### 5.5.1 Implementierung in die Mikroskop-Software

Für das Zusammenspiel aus Merkmalsextraktion mit Haar-Filtern und dem Boosting-Training konnte auf die Funktionalität der OpenCV-Bibliothek ([26]) zurückgegriffen werden, die in die Zeiss-eigene Mikroskop-Software integriert wurde. Sie liegt in der verwendeten Version als C-Code vor und zeichnet sich zum einen durch den großen Umfang an Bildverarbeitungs- und Interpretationsalgorithmen als auch durch die extrem effiziente Implementierung für die Intel-Prozessorstruktur aus. OpenCV basiert auf der

Intel Image Processing Library (IPL), die Funktion beinhaltet, die auf Assembler-Ebene optimiert wurden. Im Detail werden folgende Funktionen explizit verwendet:

- **cvCreateCascadedClassifier** - AdaBoost-Training mit den Eingangsparametern: Trainingsbilder Klasse 1, Trainingsbilder Klasse 2, Anzahl der Kaskaden, minimum RP-Rate je Kaskade, maximale FP-Rate je Kaskade, Höhe und Breite auf die die Trainingsbilder einheitlich skaliert werden
- **cvHaarDetectObjects** - Klassifikation mit einem Schiebefenster mit den Eingangsparametern: Bildausschnitt, Klassifikator-Kaskade, Iterationsfaktor mit dem sich das Schiebefenster bis hin zur Bildausschnittsgröße sukzessive höher skaliert, minimale Anzahl von Nachbardetektionen (um eine Region als positiv zu klassifizieren), Startgröße des Schiebefensters, Boosting-Methode

Alle variablen Werte wie u.a. die Dimension der Trainingsbildskalierung (gewöhnlich 24x24 Pixel), die Richtig- und Falschdetektionsraten, die initiale Schiebefenstergröße und der Iterationsfaktor der Schiebefensterskalierung können über das User-Interface verändert werden. Weiterhin geschieht die Auswahl der Objektausschnitte, die der Ausgangspunkt für die Generierung der Trainingsbilddatenbasis ist, ebenfalls über dieses User-Interface.

Die gesamte Funktionalität wird über ein GUI gesteuert und kann sowohl offline auf Bildstapeln als auch online mit angeschlossenem Mikroskop operieren.

### 5.5.2 User-Interface

Innerhalb der IB-AF-Programmoberfläche stehen zwei Paneele zur Verfügung: zum einen die Entwickleroberfläche (siehe Abbildung 23), welche die Einstellung aller Parameter ermöglicht, und zum anderen die Benutzeroberfläche. Diese ist als Entwurf für ein mögliches Endbenutzer-Interface zu verstehen und versucht, den Funktionsumfang auf ein Minimum an Interaktionsmöglichkeiten zu begrenzen. Beide Paneele sind in die drei separaten Teilfunktionen untergliedert: Generierung von Trainingsbildern, Training der kaskadierten Klassifikatoren (`cvCreateCascadedClassifier`) und die Autofokussierung mit integrierter Bildausschnittsklassifikation (`cvHaarDetectObjects`).

#### Entwickleroberfläche

##### “Generate Training Images” - Generierung der Trainingsbilder (Abb. 23)

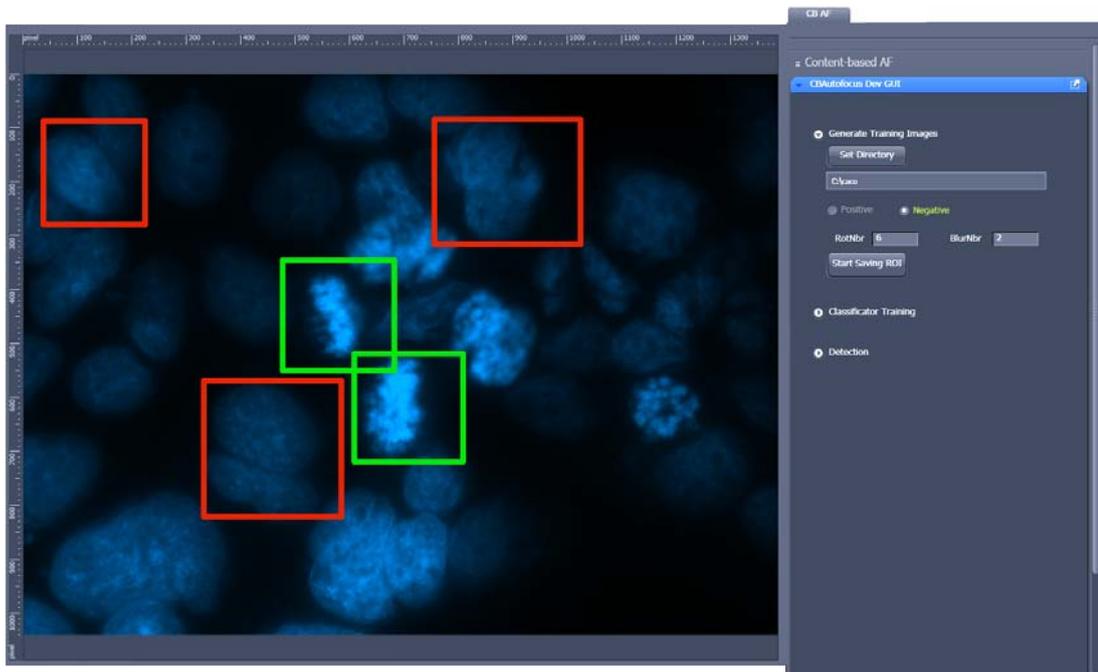


Abbildung 23: Programmoberfläche, in deren Hauptfenster Bildregionen mit positiven (grün = sichtbare DNA) und negativen (rot) Untersuchungsobjekten durch ROIs markiert wurden.

- “Set Directory” - Auswahl des Dateordners für die Trainingsbilder auf der lokalen Festplatte. Es wird automatisch eine entsprechende Ordnerunterstruktur für die positiven und negativen Trainingsbilder angelegt. Dies geschieht sowohl für das Hypothesentraining als auch für das Training des Verifikationsklassifikators.
- “Positive / Negative-Radiobutton” - Mit der Computer-Maus werden Rechtecke als Overlay-Elemente in das Hauptfenster gezeichnet. Entsprechend dem aktivierten Button werden die Bildausschnitte der positiven oder negativen Trainingsklasse zugeordnet.
- “RotNbr” - Eintrag bestimmt die Anzahl der rotierten Instanzen des jeweiligen Bildausschnitts, d.h. 0 = keine Rotation, 6 = Bildkopien unterscheiden sich durch eine Rotation von 30° zueinander.
- “BlurNbr” - Eintrag bestimmt die Anzahl der unscharf gefilterten Instanzen je rotierten Bildausschnitts.
- “Start Saving ROI” - Speichert die eingezeichneten Bildausschnitte und ihre rotierten und geblurten Kopien als Bilddateien entsprechend der Vorgaben.

### “Classifier Training” - Klassifikatortraining (Abb. 24 1. v.l.)

- “Set Train Path” - Angabe des Dateipfad, in dem sich die Trainingsbilder befinden.
- “SPECIFICITY maxfalse” - Maximal erlaubte Falsch-Positiv-Rate (in %) des Klassifikators auf den Trainingssatz je Kaskade.
- “SENSITIVITY minpos” - Minimal zu erreichende Richtig-Positiv-Rate (in %) des Klassifikators auf den Trainingssatz je Kaskade.
- “Type TrainClassifierName” - Eingabefeld für den selbst zu wählenden Klassifikatornamen.
- “CLASSIFIER WINDOW SIZE” - Höhe und Breite in Pixel, auf welche die Bildregionen skaliert werden.
- “NUMBER OF CASCADE STAGES” - Maximale Anzahl der Kaskaden.
- “Haarfilter Mode” - Auswahl des zugrunde liegenden Haar-Filtersatzes.
- “Boosting Mode” - Auswahl der Boosting-Methode. “GAB” Gentle AdaBoost, “LB” LogitBoost, “RAB” Real AdaBoost, “DAB” Discrete AdaBoost.
- “Start Training” - Startet das Training entsprechend der Vorgaben. Die Klassifikatoren werden in einer Ordnerstruktur angelegt, wobei jeweils ein Ordner für eine Kaskade steht und die Klassifikatorparameter (Position, Haar-Filter-Typ, Größe, Schwellwert) innerhalb einer Textdatei in dem jeweiligen Ordner gespeichert werden. Der Gesamtklassifikator wird zudem als xml-File unter dem in “Type TrainClassifierName” angegebenen Namen gespeichert. Bei Bedarf wird dieser Klassifikator mit dieser Datei geladen. Es werden zwei Klassifikatoren trainiert: einen für die Hypothesengenerierung und ein zweiter für die Verifikation.

### “Detection” - Klassifikation und Fokussierung (Abb. 24 2. v.l.)

- “Load Classifier” - Auswahl des Klassifikators über die xml-Datei.
- “TakeSharpClassifier”-Checkbox - Wenn aktiv, wird explizit der Verifikationsklassifikator geladen.
- “Scale Iteration Factor” - Iterationsfaktor, um den das Schiebefenster sukzessive vergrößert wird.
- “SEARCH WINDOW - init size” - Startgröße des Schiebefensters.

- “SEARCH WINDOW - end size” - Maximale Größe des Schiebefensters.
- “TrainWinSize” - Dimension der skalierten Trainingsbilder.
- “Minimum of Neighbours” - Minimale Anzahl von Nachbardetektionen, damit eine Bildregion als Hypothese anerkannt wird.
- “Start Detect” - Hypothesengenerierung, d.h. Suche nach positiven Bildregionen entsprechend der Vorgaben ohne Fokussierungsmessung. Alle erkannten Bildausschnitte werden als Rechteck-Overlay-Objekte im Bild eingezeichnet.
- “Multi ROI AF” - Bestimmung der Fokusebene für jedes Overlay-Objekt und Anzeige der jeweiligen Ebene als Text-Objekt innerhalb der ROI.
- “AF” - Vollzieht die Schritte der Hypothesengenerierung und anschließender Autofokussierung in einem Vorgang.
- “Sim Mode”-Checkbox - Wenn aktiviert, erfolgt die Berechnung auf dem geöffneten Bildstapel und nicht im Live-Bild des Mikroskops.
- “Processing Time” - Berechnungszeit

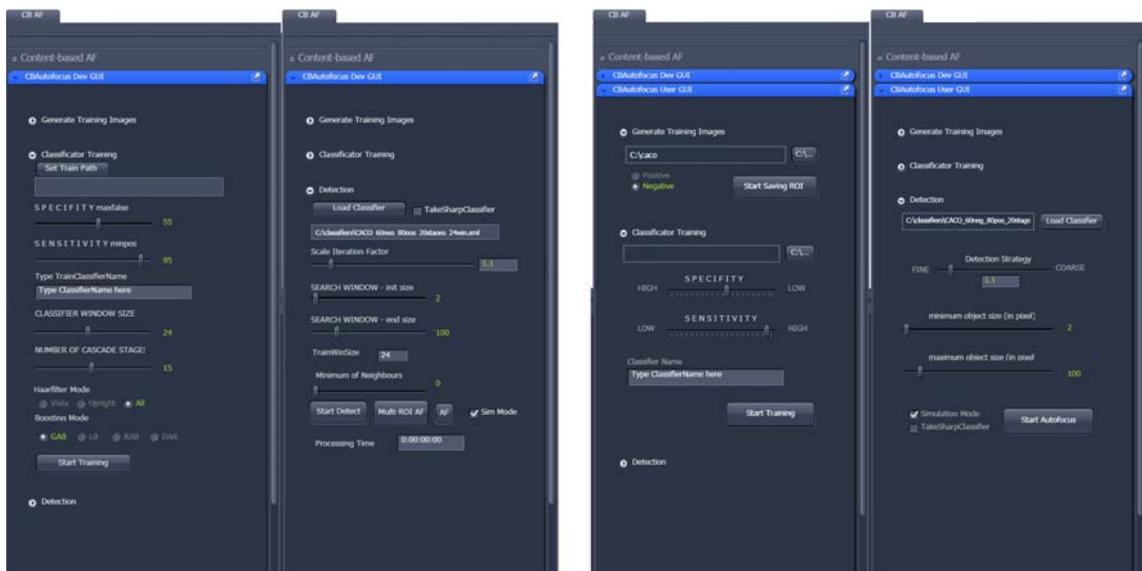


Abbildung 24: Aufgeklappte Paneele der Entwickler-GUI (1. und 2. v.l.) und der Benutzer-Oberfläche (3. und 4. v.l.)

## Benutzeroberfläche

“Generate Training Images” - Generierung der Trainingsbilder (Abb. 24 3. v.l.)  
 Analog zur Entwickleroberfläche muss der Datenort auf der Festplatte und die Klasse

der zu speichernden Bildregionen angegeben werden. Die Anzahl der Rotationen und der Unschärfe-Instanzen ist jedoch auf 6 und 2 fixiert.

**“Classifier Training” - Klassifikatortraining (Abb. 24 3. v.l.)** Hier muss nur der Ort der Trainingsbilder auf der Festplatte und der Name des Klassifikators angegeben werden. Die Richtig-Positiv-Rate und die maximale Falsch-Positiv-Rate wird über den “SPECIFICITY”- bzw. entsprechend über den “SENSITIVITY”-Slider geregelt. Alle weiteren Parameter sind unveränderlich: Haarfilter Mode - ALL, Boosting Mode - DAB, Classifier Window Size = 24, Number of Stages = 12.

**“Detection” - Klassifikation und Fokussierung (Abb. 24 4. v.l.)** Entsprechend muss auch hier die xml-Datei aus dem Klassifikator-Ordner angegeben werden. Der Skalierungsfaktor des Schiebefensters ist als “Detection Strategy”-Slider umgesetzt. “Start Autofocus” setzt analog zum “AF”-Button der Entwickleroberfläche die Hypothesengenerierung und anschließende Autofokussierung auf den ermittelten ROIs in Gang.

## 6 Ergebnisse

### 6.1 Auswertungsverfahren

Die Evaluierung des inhaltsbasierten Autofokusverfahrens gestaltet sich komplexer als die Auswertung von bisher publizierten und angewandten Autofokusmethoden. War bislang lediglich die Frage zu beantworten, ob ein Algorithmus eine gewünschte Fokusebene ermittelt oder nicht, soll das hier beschriebene Verfahren eine Mehrzahl von Fokusebenen finden und auch Fokusebenen mit nicht interessierenden Objekten ausschließen. Weil die Gesamtfokussierungsmethode zum großen Teil auf Objekterkennung aufbaut, muss auch eine Auswertung der Klassifikationsgenauigkeit stattfinden. Die in dem Rahmen dieser Arbeit verwendeten Proben zeichneten sich immer durch polyfokale Präparate aus. Dabei können die Proben sowohl vollständig gleiche Objekte jedoch in unterschiedlichen Ebenen, als auch verschiedene Objektarten in unterschiedlichen Lagen beinhalten. Wobei dann nur eine bestimmte Objektklasse jeweils für die Untersuchung von Interesse ist. Die klassischen Fokusmethoden können bei diesen Untersuchungsbeispielen aufgrund der Annahme einer Fokusebene niemals vollständig richtige Ergebnisse erzielen. Denn entweder ermitteln sie eine falsche Ebene oder eben nur eine einzige der gesuchten Fokuslagen. Im Gegenzug ist es jedoch möglich, dass der Inhaltsbasierte AF nicht nur eine, sondern mehrere falsche Fokusebenen präsentiert. Für die konkrete Auswertung wurde deshalb wie folgt verfahren.

Im ersten Schritt wurden die Proben mit dem Mikroskop und mit großem Offset in z-Richtung um die Fokusebene(n) herum aufgenommen. Dabei wurde die Schrittweite optimal zur Mikroskopvergrößerung gewählt. Mit diesem Offset stehen ausreichend nicht-fokussierte Lagen je Aufnahme zur Verfügung, um auch die Hypothesengenerierung in stark unscharfen Ebenen zu testen. Die Aufnahmen liegen dann als digitale Bildstapel vor. Weil die Ergebnisse reproduzierbar sein sollen und weil wiederholtes Testen bei Fluoreszenzproben zu fortschreitendem Ausbleichen führte, wurden alle Auswertungen offline auf diesen z-Stapeln vollzogen. Für jede Aufnahme liegt eine Expertenentscheidung bezüglich der Fokusebene(n) vor, die für alle Auswertungen den Goldstandard definierten. Diese Entscheidung bestimmt auch den Ausgangspunkt für die Messungen mit dem Autofokusmodul. Insgesamt wurden für jeden z-Stapel zwei Fokusedurchläufe durchgeführt, wobei die erste Messung im Abstand von 5 Ebenen oberhalb und die zweite im Abstand von 10 Fokuslagen unterhalb der mittleren Fokuslage gestartet wurden. In diesen zwei Ebenen sind die Objekte unterschiedlich stark defokussiert. Diese Vorgehensweise soll eine nicht zielgerichtete initiale Abstandseinstellung des Objektivs bezüglich des Präparats simulieren.

## 6.2 Ergebnisse

Im Folgenden bewerte ich zunächst die klassischen Fokusfunktionen, die der Wahl der Schärfemethode für den IB-AF zugrunde lagen. Dabei werden die Fokusanalysen von konkreten Bildregionen und von kompletten Aufnahmen gegenüber gestellt. Anschließend stelle ich die Ergebnisse des IB-AF vor, wie sie mit der Software-Implementierung erzielt wurden. Dabei erörtere ich auch den Einfluss der verschiedenen Parameter sowohl beim Training als auch bei der Generierung und Verifizierung von den Hypothesen.

Für die Untersuchungen wurden folgende Mikroskop-Setups verwendet:

- **Zeiss Axio Observer.Z1** mit **AxioCam HR3**-Kamera mit einer Auflösung von 1388x1040 Pixel mit einer Bitgröße von 14 bit je Kanal (**Z1/HR3**)
- **Zeiss Axio Imager.Z1** mit **AxioCam MR3**-Kamera mit einer Auflösung von 1388x1040 Pixel mit einer Bitgröße von 12 bit je Farbkanal (**Imager/MR3**)
- **Zeiss Axio Observer.Z1** mit **AxioCam MR3**-Kamera mit einer Auflösung von 1388x1040 Pixel mit einer Bitgröße von 12 bit je Farbkanal (**Z1/MR3**)

Die Probensammlung (bereitgestellt durch Zeiss) setzte sich wie folgt zusammen:

- **Kol1-5/7-8:** CaCo mit einzelnen Zellen und Zellkolonien - Fluoreszenz (Hoechst 33342)  
- EC Plan-Neofluar 40x/0.75NA-Objektiv (Imager/MR3)
- **Kristall:** Kristall-Artefakt im Phasenkontrast mit adherenten CaCo-Zellen -  
20x/0.80NA-Objektiv (Z1/HR3)
- **Hela2:** HeLa-Zellen im Fluoreszenzkanal (DsRed) - Plan-Apochromat 40x/0.95NA-Objektiv (mit Z1/HR3)
- **Ph.-A.:** Artefakte im Phasenkontrast - Plan-Apochromat 20x/0.80NA-Objektiv (Z1/HR3)
- **Beads:** Fluoreszenz-Probe mit starkem Rauschen  
- EC Plan-Neofluar 40x/1,30NA (Imager/MR3)
- **CaCo DAPI:** Zellkern-Färbung (DAPI) / Fluoreszenz - 20x/0.50NA-Objektiv (Imager/MR3)
- **CaCo FITC:** Immunglobulin-Färbung (FITC) / Fluoreszenz  
- 20x/0.50NA-Objektiv (Imager/MR3)

- **Cyt. DAPI:** Cytoskelett mit Zellkern-Färbung (DAPI) / Fluoreszenz - Plan-Apochromat 20x/0,8NA-Öl-Immersionsobjektiv (Z1/HR3)
- **Cyt. Ph.:** Cytoskelett im Hellfeld-Phasenkontrast - Plan-Apochromat 20x/0,8NA-Öl-Immersionsobjektiv (Z1/HR3)
- **Cyt. Red:** Cytoskelett mit Tubulin-Färbung (DsRed) / Fluoreszenz - Plan-Apochromat 20x/0,8NA-Öl-Immersionsobjektiv (Z1/HR3)
- **Mitotische CaCo-2:** Zellkern-Färbung (DAPI) / Fluoreszenz - 63x/1.30NA-Objektiv (Imager/MR3)
- **Material:** Silizium - EC Epiplan-Neofluar 10x/0,25NA (Z1/MR3)

### 6.2.1 Fokussierungsqualität bei verrauschten Bildern

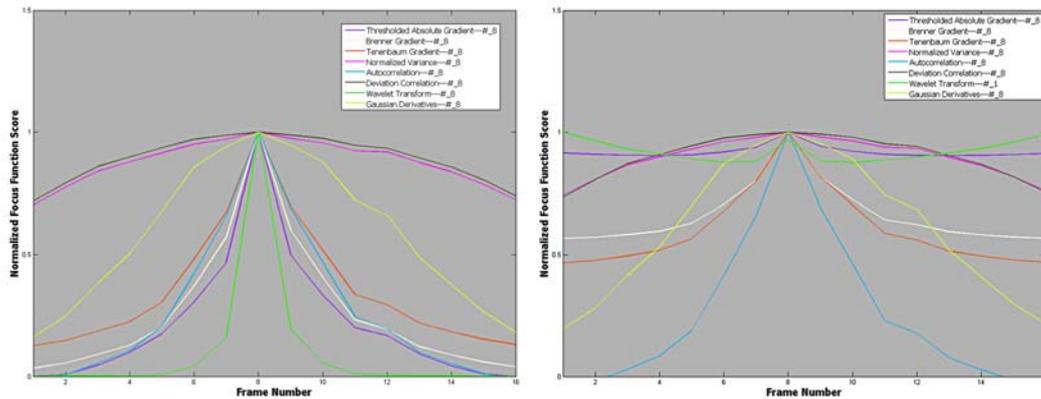
Die Auswirkungen von Rauschen (siehe Abschnitt 3.3.2) auf eine Schärfemessung treten vor allem bei Gradienten-basierten Schärfefunktionen deutlich zu Tage. In einem simulierten Test wurde einem Bildstapel Gauß-verteilter Rauschen hinzugefügt (Abb. 25). Während alle verwendeten Fokussierungsmethoden (siehe Legende in Abb. 26) die richtige Fokussierungslage, nämlich Frame #8, im Bildstapel ohne Rauschen ermittelte, versagt der Wavelet-basierte Algorithmus im mit Rauschen addierten Bildstapel. Weiterhin ist aus den Fokussierungskurven (Abb. 26) abzulesen, dass neben der Wavelet-Methode vor allem die Gradientenmethoden Schwellwert-basierter Gradientenbetrag (Thresholded Absolute Gradient), Brenner- und Tenenbaum-Gradient eine stark veränderte Kurve mit flacheren Flanken aufweisen.



**Abbildung 25:** Links: Fokussierte und defokussierte Darstellung des Lena-Testbildes aus dem simulierten rauschfreien Bildstapel. Rechts: Fokussierte und defokussierte Darstellung des Lena-Testbildes aus dem simulierten Stapel mit Rauschen Gaußscher Verteilung.

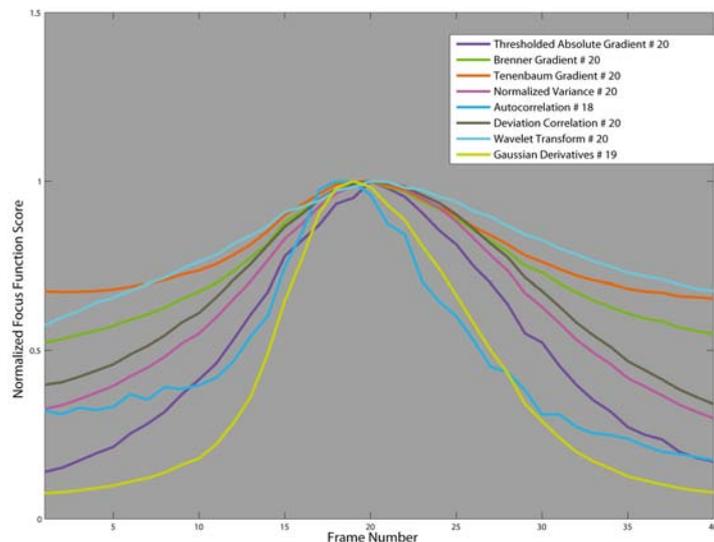
### 6.2.2 Fokussierungsvergleich zwischen Ganzbild- und ROI-Analyse

Dass klassische Fokussierungsmethoden bei Vorhandensein von polyfokalen Proben nicht in der Lage sind, alle relevanten Fokusebenen zu detektieren (siehe Abschnitt 4.3), geht aus



**Abbildung 26:** Die Fokuskurven für eine Auswahl an Fokusschärfefunktionen anhand des simulierten Stapels in Bild 25 ohne Rauschen (links) und für den mit Gauß-verteilterm Rauschen versetzten Stapel (rechts). (Quelle: eigene Berechnung und Darstellung per MatLab-Plot)

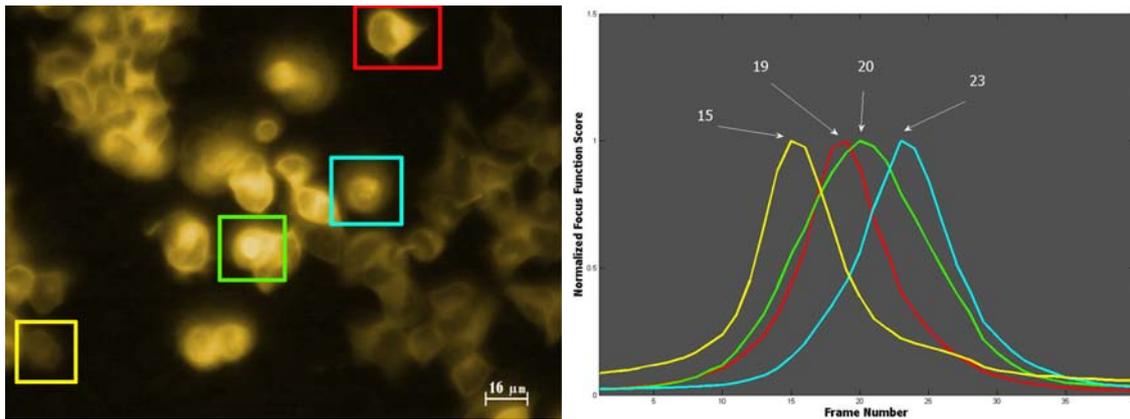
dem Versuch mit einer Fluoreszenz-Hela-Probe (Linkes Bild in Abb. 28) hervor. Dort befinden sich Zellen in vier verschiedenen Fokusebenen. Die Fokusergebnisse der bekannten Schärfefunktionen sind in Abbildung 27 dargestellt. Die Fokusergebnisse (rechter Plot in Abb. 28) mit der Varianz-basierten Schärfefunktion auf den einzelnen ROIs zeigen allerdings die korrekten Fokusebenen der einzelnen Regionen.



**Abbildung 27:** Resultierende Fokuskurven bei Betrachtung des gesamten Blickfeldes der HeLa-Zellenprobe in Abb. 28. (Quelle: eigene Darstellung)

In einer Testreihe mit unterschiedlichen Präparaten war zu untersuchen, wie erfolgreich einzelne Autofokusmethoden bei unterschiedlichen Modi operieren und ob die Eingrenzung des Suchfeldes auf interessierende Objekte eine Steigerung der Fokussierungsqualität erzielt.

Die ebenengenaue Bestimmung der Fokusslage war oft nicht möglich, weil die Objekte



**Abbildung 28: Resultierende Fokuskurven für einzelne Bildbereiche, welche die korrekten Fokusebenen für die jeweiligen Objekte liefern. Links: Bild eines aufgenommenen Bildstapels eines Fluoreszenz-HeLa-Zellenpräparats. Die farbige umrandeten ROIs markieren die Bereiche in der XY-Dimension, in denen die Fokusbewertung mit einer varianzbasierten Funktion durchgeführt wurde. Die Farben korrespondieren mit den Farben der Fokuskurven im rechten Plot. Die jeweiligen Zahlen zeigen die Nummer des Frames (Stapelbildes) mit dem jeweils höchsten Fokuswert. (Quellen: eigene Darstellung; Zur Verfügung gestellt durch Zeiss MicroImaging GmbH)**

nicht nur in einer Ebene scharf abgebildet wurden. Aus diesem Grund stuften wir die Nachbarlagen der durch den Experten festgelegten Fokuscoordinate (in den Tabellen in dunkelgrüner Schrift) ebenfalls als korrekt ermittelt ein. Sie sind in den Ergebnistabellen mit hellgrüner Schriftfarbe eingetragen.

Tabelle 3 zeigt die Ergebnisse bei der Probe mit Zellkolonien und vereinzelt Zellen im normalen Fluoreszenz-Kanal. Einzelzellen und Kolonien befinden sich in unterschiedlichen Fokuslagen im Abstand von bis zu drei Ebenen. Hier schneidet die Gaussian-Derivatives-Methode am besten ab, weil sie entweder eine der zwei Fokusebenen ermittelt oder mit nur einer Ebene von einer der vom Fachmann<sup>2</sup> bestimmten Fokusebene abweicht und immer zwischen den zwei Schärfeebenen liegt. Die Autokorrelationsfunktion erzielt ebenfalls sehr gute Ergebnisse, während alle anderen Schärfemaße mittelmäßige (Normalisierte Varianz, Brenner-Gradient) bis komplett falsche Fokusebenen (Wavelet und Ableitungsbasierte Korrelation) ermitteln.

Das Kristallartefakt-Präparat, das in allen Ebenen besonders aber in der Ebene #39 starke Kontraste liefert, enthält am Objektträger anhaftende Zellen, deren Ausdehnung sich über mehrere Ebenen (25-31) erstreckt. Hier haben die Methoden Normalisierte Varianz, Wavelet und Absoluter Gradient mit den Phasenartefakten in defokussierten Ebenen Probleme und ermitteln eine inkorrekte Fokusposition. In der HeLa2-Probe (Abb. 28) im Fluoreszenz-Kanal errechnen alle Funktionen eine richtige bzw. fast richtige Fokusebene (Tabelle 4) genauso wie im DsRed-Fluoreszenzbild der Cytoskelett-Probe. In den Phasenkontrast-Proben mit Phasen-Artefakten (Ph.-A.) und Cytoskelett (Cyt.-Ph.)

<sup>2</sup>Dr. Andreas Murr - Institut for Biological and Medical Imaging - Doktor der Biologie

		Probe						
		Kol1	Kol2	Kol3	Kol4	Kol5	Kol7	Kol8
Fokusebene		40/37	45/42	40/37	50/48	49/46	47/45	46/45
Meth.	Abs. Grad.	37	78	37	41	49	46	43
	Brenner	37	74	38	75	45	47	45
	Tenengrad	1	63	38	75	48	47	47
	Varianz	20	41	37	44	48	49	51
	Autokorr.	1	43	38	49	46	47	45
	Abl. Korr.	80	80	71	64	54	70	99
	Wavelet	78	80	79	81	100	97	99
	Gauss. Der.	38	43	38	49	47	47	46

Tabelle 3: Autofokusergebnisse der Analyse kompletter Ansichten #1

zeigt sich die Autokorrelationsfunktion als zuverlässig, während alle anderen nur teilweise oder gar nicht richtige Ergebnisse erzeugen. Allein die Fluoreszenz-Beads-Probe stellt wegen des hohen Bildrauschens für alle Methoden eine Hürde dar.

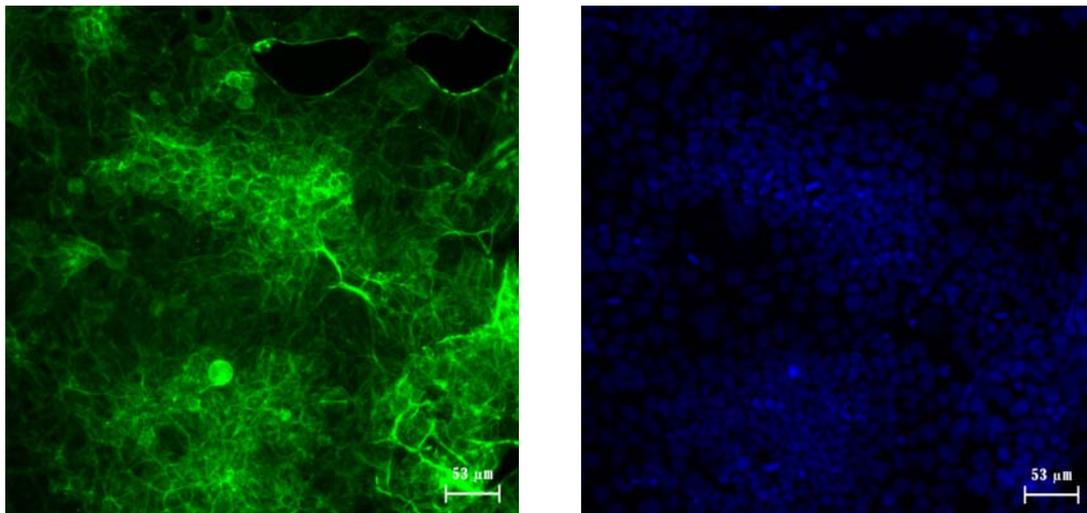


Abbildung 29: CaCo-Zellen in derselben Probe jeweils im FITC- und im DAPI-Kanal. (Quellen: eigene Darstellung; Zur Verfügung gestellt durch Zeiss MicroImaging GmbH)

Tabellen 5 und 6 listen die Ergebnisse der Fokusfunktion für bestimmte Bildbereiche auf. Es wurden im Voraus die ROI-Koordinaten von einzelnen Objekten bzw. einheitlichen Objektbereichen manuell so bestimmt, dass keine störenden Hintergrundinformationen innerhalb dieser Bildregionen liegen und jeder ROI-Stapel nur eine Fokusebene besitzt. Die jeweiligen Bildregionen sind u.a. in den Abbildung 30 und 31 als weiße Rahmen markiert. Die Auswertung der Fokussierungsergebnisse auf bestimmten Bildarealen entlang der z-Ebene bringt das überraschende Ergebnis, dass die Einschränkung des Untersuchungsbereichs für die wenigsten Schärfefunktionen eine Verbesserung des Fokusergebnisses bewirkt.

		Probe								
		Kristall	HeLa2	Ph.-A.	Beads	CaCo FITC	CaCo DAPI	Cyt. Ph	Cyt. RED	Cyt. DAPI
<b>Fokusebene</b>		25-31	15/19/20/23	60	31	38/43	40/44	21	21	23
Meth.	<b>Abs. Grad.</b>	39	20	57	3	79	79	7	21	40
	<b>Brenner</b>	28	20	60	3	79	44	6	21	40
	<b>Tenengrad</b>	28	20	60	4	42	44	6	21	40
	<b>Varianz</b>	1	20	54	10	42	62	8	21	23
	<b>Autokorr.</b>	28	18	60	29	42	44	21	21	23
	<b>Abl. Korr.</b>	29	20	54	29	43	38	18	22	23
	<b>Wavelet</b>	17	20	25	1	43	80	5	21	40
	<b>Gauss. Der.</b>	29	19	5	19	42	45	20	21	23

Tabelle 4: Autofokusergebnisse der Analyse kompletter Ansichten #2

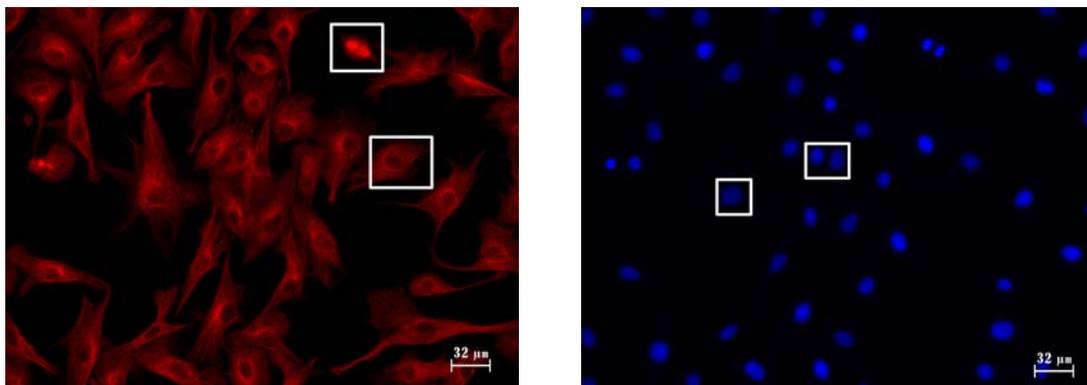


Abbildung 30: Cytoskelett im DsRed- und DAPI-Kanal. (Quellen: eigene Darstellung; Zur Verfügung gestellt durch Zeiss MicroImaging GmbH)

### 6.2.3 Abhängigkeit der Resultate von Trainings- und Klassifikationsparametern

#### Hypothesengenerierung in Abhängigkeit von der Startposition

Auch wenn die Klassifikatoren mit einer gewissen Information an Unschärfe der interessierenden Objekte versorgt werden, ist es dennoch nachvollziehbar, dass ab einem bestimmten Grad der Defokussierung eine Generierung von Hypothesen nicht möglich ist. Die Abbildung 33 verdeutlicht, dass mit zunehmendem Abstand von der Fokusebene die Anzahl der Hypothesen abnimmt. Die zugrunde liegende Bildquelle (siehe Abb. 32) des Diagramms beinhaltet genau ein interessierendes Objekt (mitotische Zelle), das von der Hypothesengenerierung bei einer optimalen Schrittweite von  $0,03 \mu\text{m}$  in z-Richtung im Intervall von  $-0,42 \mu\text{m}$  bis  $0,54 \mu\text{m}$  noch detektiert wird.

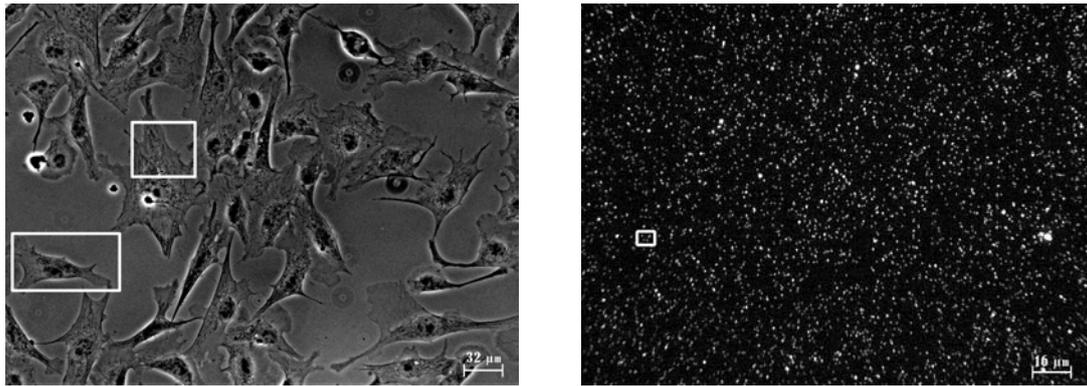


Abbildung 31: Cytoskelett im Phasenkontrast-Modus (links) und Beads-Probe (rechts). (Quellen: eigene Darstellung; Zur Verfügung gestellt durch Zeiss MicroImaging GmbH)

		Probe					
		CaCo DAPI-Kanal			CaCo FITC-Kanal		
		ROI 1	ROI 2	ROI 3	ROI 1	ROI 2	ROI 3
	<b>Fokusebene</b>	<b>39</b>	<b>44</b>	<b>44</b>	<b>40</b>	<b>40</b>	<b>39</b>
Meth.	<b>Abs. Grad.</b>	40	70	79	75	79	1
	<b>Brenner</b>	40	44	79	28	79	42
	<b>Tenengrad</b>	7	4	1	2	1	79
	<b>Varianz</b>	39	45	48	30	39	29
	<b>Autokorr.</b>	1	66	1	7	78	11
	<b>Abl. Korr.</b>	40	38	40	43	43	43
	<b>Wavelet</b>	75	77	79	80	43	42
	<b>Gauss. Der.</b>	39	45	45	30	39	42

Tabelle 5: Autofokusergebnisse der Analyse von bestimmten Bildregionen #1 (Beispielebene siehe Abb. 21)

Fokussierungsqualität und Trainingsdauer aufgrund der Trainingsparameter

**Boosting-Modus und Trainingsbildgröße** Die Distanz der Startebene von der tatsächlichen Fokusebene spielt für das Klassifikationsergebnis eine wichtige Rolle. Darüber hinaus sind es die Einstellungsparameter des Klassifikationstrainings, die den hauptsächlichen Einfluss auf die Klassifikationsgenauigkeit ausüben. Einige der Parameter blieben in dieser Arbeit allerdings unangetastet. Darunter fiel der Modus des AdaBoost-Trainings, welcher in allen Kalkulationen auf Gentle AdaBoost (GAB) eingestellt war. Diese Maßnahme stützt sich auf die ausführlichen Untersuchungen von Lienhart et al. ([30]), der in der Wahl von GAB das schnellste Training mit gleichzeitig bester Detektionsrate nachweist. Lienhart et al. gehen zudem auf die Größe der Trainingsbilder ein und zeigen, dass sowohl die Skalierung auf  $<20 \times 20$  Pixel als auch die Vergrößerung auf mehr als  $24 \times 24$  Pixel hinsichtlich der Detektionsrate nicht sinnvoll ist.

Meine eigenen Untersuchungen bestätigen, dass die Skalierung größer als 18 Pixel aus-

		Probe					
		Cytoskelett DAPI		Cytoskelett DsRed		Cytoskelett Phasenk.	
		ROI 1	ROI 2	ROI 1	ROI 2	ROI 1	ROI 2
	<b>Fokusebene</b>	<b>23</b>	<b>23</b>	<b>21</b>	<b>21</b>	<b>21</b>	<b>21</b>
Meth.	<b>Abs. Grad.</b>	40	40	21	21	8	8
	<b>Brenner</b>	40	40	21	21	8	8
	<b>Tenengrad</b>	1	1	21	21	7	7
	<b>Varianz</b>	24	24	22	21	8	8
	<b>Autokorr.</b>	23	1	21	21	21	21
	<b>Abl. Korr.</b>	23	23	22	40	19	18
	<b>Wavelet</b>	40	40	21	21	3	4
	<b>Gauss. Der.</b>	23	23	22	21	20	20

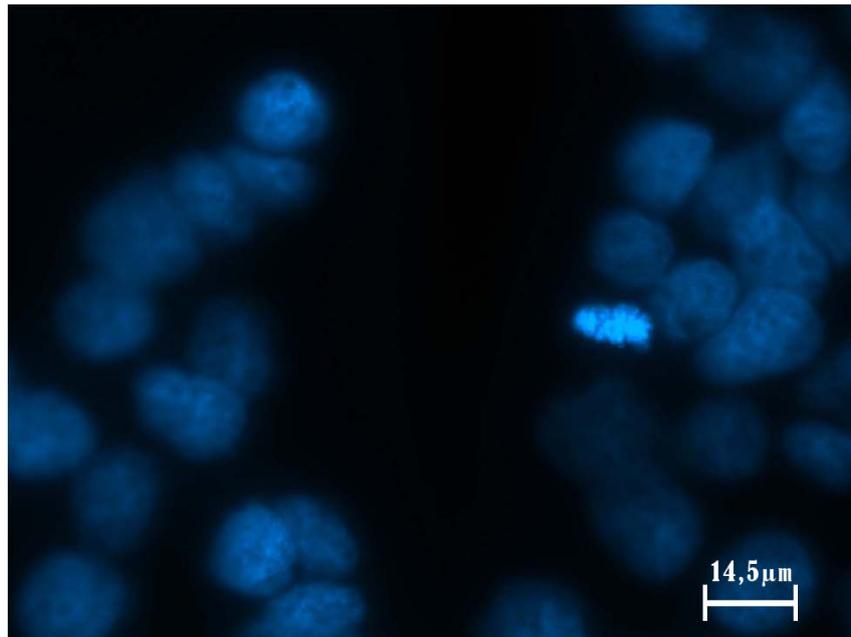
Tabelle 6: Autofokusergebnisse der Analyse von bestimmten Bildregionen #2

		Probe				
		Beads	KOL1		Kristall	Ph.-A.
		ROI 1	ROI 1	ROI 2	ROI 1	ROI 1
	<b>Fokusebene</b>	<b>31</b>	<b>38</b>	<b>40</b>	<b>29</b>	<b>60</b>
Meth.	<b>Abs. Grad.</b>	10	36	42	28	59
	<b>Brenner</b>	10	37	41	29	60
	<b>Tenengrad</b>	10	1	1	29	60
	<b>Varianz</b>	16	39	32	29	60
	<b>Autokorr.</b>	7	37	41	29	60
	<b>Abl. Korr.</b>	31	40	80	12	2
	<b>Wavelet</b>	2	37	78	15	2
	<b>Gauss. Der.</b>	29	37	40	30	60

Tabelle 7: Autofokusergebnisse der Analyse von bestimmten Bildregionen #3

fallen sollte. Die Ergebnisse in Tabelle 8 und deren zugeordnete Darstellung in Abbildung 34 zeigen eine geringere Erkennungsrate ( $< 41\%$ ) bei Skalierung kleiner 20 Pixel als bei der Wahl ab 20 Pixel. Ausreißer hierbei ist allerdings die 22x22 px Größe mit 36,4%. Aus den Auswertungen ergibt sich gleichfalls, dass sehr kleine Skalierungen (12 und 14 px) eine höhere Falsch-Positiv-Rate mit sich bringen und damit die Sensitivität der Methode verringern. Grundsätzlich gilt die Aussage, dass mit hohen Skalierungen bessere Ergebnisse erzielt werden. Es steigt jedoch auch der Rechenaufwand exponentiell. Beim Training mit 12x12 px Merkmalen mussten  $\sim 15.600$  Merkmale, mit 24x24 px  $\sim 261.600$  Merkmale, bei 30x30 px  $\sim 642.600$  und bei 35x35 px gar über 1.1 Millionen Merkmale berechnet werden.

**Größe der Trainingsdatenbasis** Nahezu alle Veröffentlichungen zum AdaBoost-Algorithmus betonen die Wichtigkeit der Anzahl der Trainingsdatenbasis. Dabei soll die Anzahl der Trainingsbilder so groß wie möglich sein und meist mehrere Tausend je Klassen



**Abbildung 32: CaCo-Zellpräparat mit einem interessierenden Objekt (hellere Intensität in der rechten Bildmitte), das DNA-Strukturen zeigt.** (Quellen: eigene Darstellung; Zur Verfügung gestellt durch Zeiss MicroImaging GmbH)

Größe	12x12	14x14	18x18	20x20	22x22	24x24	30x30
Objekte	22						
Hypothesen (% bzgl. Objekte)	76 (345%)	49 (223%)	25 (114%)	56 (255%)	32 (145%)	49 (223%)	53 (241%)
RP (RP/Objekte)	7 (31,8%)	9 (40,9%)	8 (36,4%)	12 (54,5%)	8 (36,4%)	12 (54,4%)	12 (54,4%)
FP	3	4	7	8	2	6	9
RN	49	26	7	28	17	28	24
FN (bzgl. Hypothesen)	17 (22,4%)	10 (20,4%)	3 (12%)	8 (14,3%)	5 (15,6%)	3 (6,1%)	8 (15,1%)

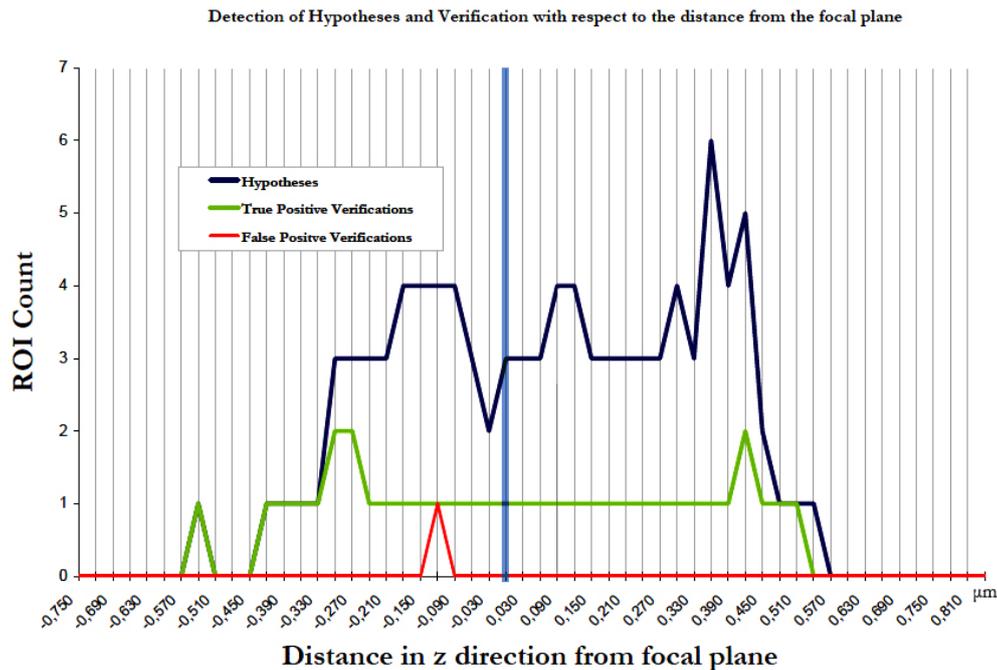
**Tabelle 8: Erkennungsrate in Abhängigkeit von der Herunterskalierung der Trainingsbilder auf 12x12 px bis 30x30px**

betragen. Dies steht im Widerspruch zu den für den Inhaltsbasierten Autofokus festgelegten Bedingungen, die u.a. eine vergleichsweise geringe Menge von Input-Bildern fordern, um die Anwenderaktionen zu begrenzen. Die Auswirkung der Anzahl der Trainingsbilder für eine Testreihe geringeren Umfangs ist in Tabelle 9 festgehalten.

Trainingsatz	72 Pos./108 Neg.	378 Pos./ 424 Neg.	450 Pos./ 810 Neg.
Erkennungsrate	66,7 %	94,4 %	100 %

**Tabelle 9: Detektionsraten in Abhängigkeit von der Trainingsdatensatzgröße**

Für diese Auswertung wurde ein Klassifikator mit einer Trainingsbildgröße von 20x20 px, einer Maximum-FP-Rate von 50% pro Kaskade und einer Minimum-RP-Rate von 99% trainiert. Die Trainingsbilder beinhalten in diesem Fall auch zwei Gauß-geglättete und um schrittweise 30° gedrehte Bildinstanzen, so dass dem Trainingsatz mit 76 Positiv-



**Abbildung 33: Anzahl der Hypothesen in Abhängigkeit vom Abstand zur Fokusebene ( $z = 0\mu m$ ), wie sie in Abbildung 32 dargestellt ist. (Quelle: eigene Darstellung)**

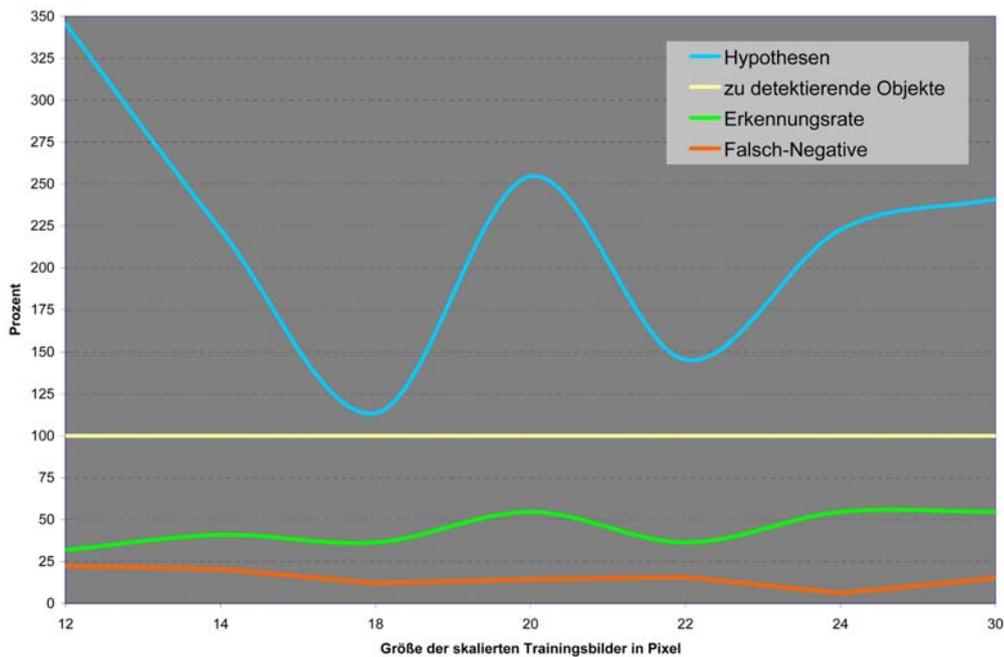
Bildern ursprünglich nur 4 Inputsamples zu Grunde liegen, die aber zweimal mit steigendem Sigma Tiefpass-gefaltet und jeweils 6-mal gedreht wurden. Entsprechend bestanden die Datensätze 387 Pos./ 324 Neg. aus 21 Positiv- bzw. 24 Negativ-Samples und 450 Pos./810 Neg. aus 25 Positiv- bzw. 45 Negativ-Samples. Nur der umfangreichste Datensatz schafft es, alle gewünschten Objekte (mitotische Zellen in CaCo-2-Zellproben) zu ermitteln. Meine Ergebnisse bestätigen damit die bisherigen Publikationen.

**RP-Rate und FP-Rate des Klassifikatortrainings bei kleiner Datenbasis** Die Veränderungen der Trainingsklassifikationsparameter (RP- und FP-Rate) beeinflussen zwei Leistungsvariablen: die Trainingsdauer und die Erkennungsrate. Im Zuge des Tests mit CaCo-2-Zellkolonie-Proben wurden zum einen die minimale RP-Rate (Tab. 10), zum anderen die maximale FP-Rate (Tab. 11) des Trainings variiert.

Richtig-Positiv-Rate	99,5%	99%	95%	90%
Zeit in Minuten	14	14	14	17
RP (RP/Objekte)	18 (90%)	18 (90%)	18 (90%)	16 (80%)
FP	29	29	20	25

**Tabelle 10: Trainingsdauer und Detektionsraten in Abhängigkeit von der minimalen Richtig-Positiv-Rate bei einer festen FP-Rate von 50% je Kaskade**

Dabei bestand die Trainingsbildbasis aus 100 Positiv-Samples mit Zellkolonien und 150



**Abbildung 34: Diagramm der in Tabelle 8 enthaltenen Ergebnisse bei Änderung der Trainingsbildgröße.** (Quelle: eigene Darstellung)

Negativ-Beispielen mit vereinzelt Zellen geringerer Fluoreszenz. Für diesen kleinen Datensatz hat die Änderung der RP-Rate von 99,5% zu 90%, was der Gesamt-RP-Rate bei 10 Kaskaden von 95,11% auf 34,87%) entspricht, einen Erkennungsratenunterschied von 10%. Lediglich die Trainingszeit steigt mit Verringerung der FP-Rate exponentiell. Während das Training mit 60%-FP-Rate je Kaskade 4 Minuten dauert, verlängert sich die Trainingsdauer auf 171 Minuten bei einer gewünschten FP-Rate von 40% pro Kaskade. Die FP-Rate selbst hat keinen signifikanten Einfluss. Die 60% bzw. 40% entsprechen bei einem Training mit 10 Kaskaden einer Gesamt-FP-Rate von  $(0.60^{10}) * 100\% = 0,605\%$  bzw. 0,01%, so dass der geringe Unterschied durch die unwesentlichen Schwankungen der RP-Rate erklärt ist.

Falsch-Positiv-Rate	40%	45%	50%	60%
Zeit in Minuten	171	95	14	4
RP (RP/Objekte)	13 (59,1%)	15 (68,2%)	15 (68,2%)	15 (68,2%)
FP	16	12	20	11

**Tabelle 11: Trainingsdauer und Detektionsraten in Abhängigkeit von der maximalen Falsch-Positiv-Rate bei einer festen RP-Rate von 95% je Kaskade**

Fokussierungsqualität aufgrund der Detektionsparameter

**Iteration des Skalierungsfaktors des Schiebefensters** Während des Detektionsprozesses kann auf zwei Variable Einfluss genommen werden: die durch einen Faktor bestimmte iterative Skalierung des Schiebefensters und die Anzahl notwendiger überlappender Detektionen, um einen Bildbereich als korrekt zu klassifizieren. Im Hinblick auf die Berechnungsgeschwindigkeit ist es sinnvoll, die Skalierungsschritte in Abhängigkeit von der anzunehmenden Objektbildgröße innerhalb der Gesamtaufnahme zu wählen. Jedoch hat die Erfahrung gezeigt, dass die Skalierung um wenige Pixel ein anderes Klassifikationsergebnis erbringt. Große Schritte erhöhen die Wahrscheinlichkeit, dass positive Bildregionen übergangen werden, während kleine Schrittweiten zu einer höheren Anzahl von Falsch-Positiv-Detektionen führen. Für letzteren Fall findet sich der Parameter „Minimale Anzahl an Nachbarn“ (Min. Nachbarn in Tab. 12), der die Anzahl von sich überschneidenden Detektionen angibt, bevor eine Bildregion als positiv klassifiziert wird.

Skalierungsfaktor Min. Nachbarn	1,001		1,005		1,01		1,05		1,1		1,5		2	
	1	2	1	2	1	2	1	2	1	2	1	2	1	2
Zeit in sek.	13,22	13,27	2,61	2,60	1,48	1,44	0,60	0,67	0,45	0,46	0,42	0,39	0,38	0,38
Detektionen	26	22	22	17	18	12	14	7	9	4	3	1	3	2
TP	4	4	4	3	3	2	3	3	3	3	1	1	2	2

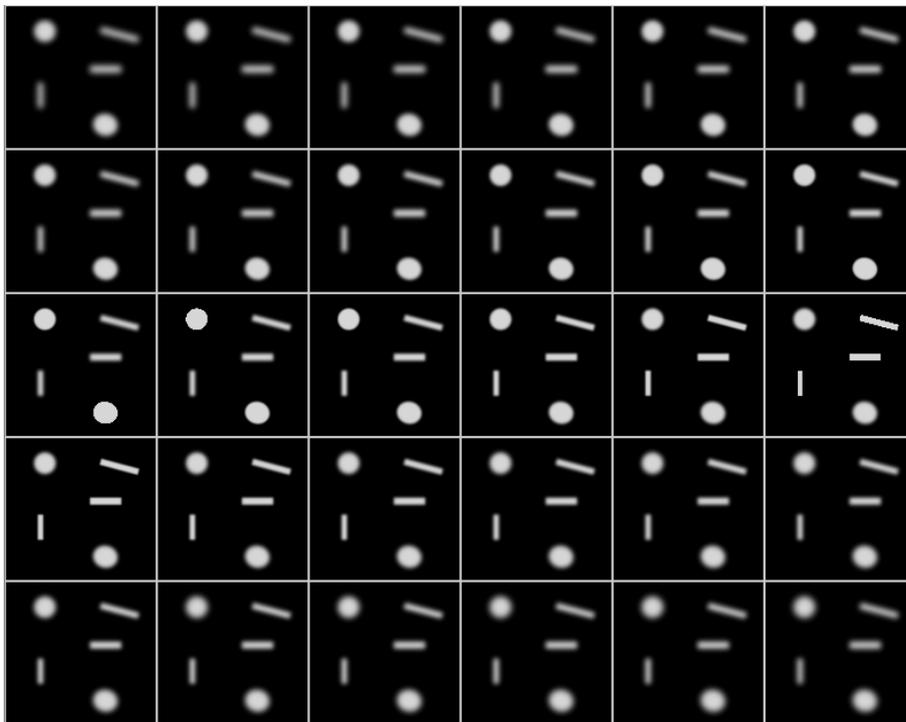
**Tabelle 12: Detektionsrate und Berechnungsdauer in Abhängigkeit von der Schrittgröße der Skalierung bei vier zu erkennenden Objekten.**

Die Werte der Tabelle 12 entstammen einem Versuch mit einem 1388 x 1040 px großen CaCo-Zellbild in der Fokusebene. Der Klassifikator wurde mit einer Merkmalskalierungsgröße von 20 x 20 Pixel und einem Kaskadentrainingswerten von 95% Detektionsrate und 50% maximaler Falsch-Positiv-Rate trainiert. Die Werte belegen, dass kleine Schrittweiten (alias Skalierungsfaktoren) bis zu einem Iterationsfaktor von 1,005 einerseits alle interessierenden Objekte finden, andererseits aber auch viele Falsch-Positiv-Detektionen hervorbringen. Eine hohe Erkennungsrate bei kleinen Schrittweiten geht allerdings mit einer hohen Berechnungszeit einher, während sich die Rechenzeit bei Iterationsfaktoren größer 1,05 im Millisekundenbereich bewegt. Es ist hinzuzufügen, dass die Kalkulationszeit die Darstellung der Detektion einschließt, so dass die eigentliche Klassifikationszeit ein Bruchteil der in Tabelle 12 aufgeführten Werte beträgt.

#### 6.2.4 Studie am Modell

In einem Vortest wurde ein Z-Stapel mit 30 Bildern (Abb. 35) modelliert, der rechteckige und runde Objektmodelle in unterschiedlichen Fokusebenen beinhaltet. Der Bilddaten-

satz für das Training entstand auf der Basis von 5 Positiv- und 5 Negativ-Beispielen. Der Klassifikator wurde mit einer Skalierungsfenstergröße von 24 Pixeln, einer Minimum-RP-Rate von  $0.99^{15} \cdot 100\%$  und einer maximalen FP-Rate von  $0.50 \cdot 100\%$  auf die Erkennung von Kreis-Objekten trainiert. Das Training selbst benötigte 11 Minuten. Die Autofokussierung mit dem IB-AF startete in einer defokussierten Ebene (Abb. 36links) und ermittelte dort die mit rot umrandeten Bildregionen, ohne Falsch-Hypothesen zu liefern. Die Fokuslagen der Kreise in den Ebenen 13 und 14 wurden in der anschließenden Fokusberechnung ermittelt und die scharfen Bildregionen in der Verifizierungsphase als positiv bestätigt (Abb. 36 rechts). Klassische Verfahren ermittelten auf diesem simulierten Stapel die Ebene 18 als Fokusslage (siehe Anhang Abb. 46).

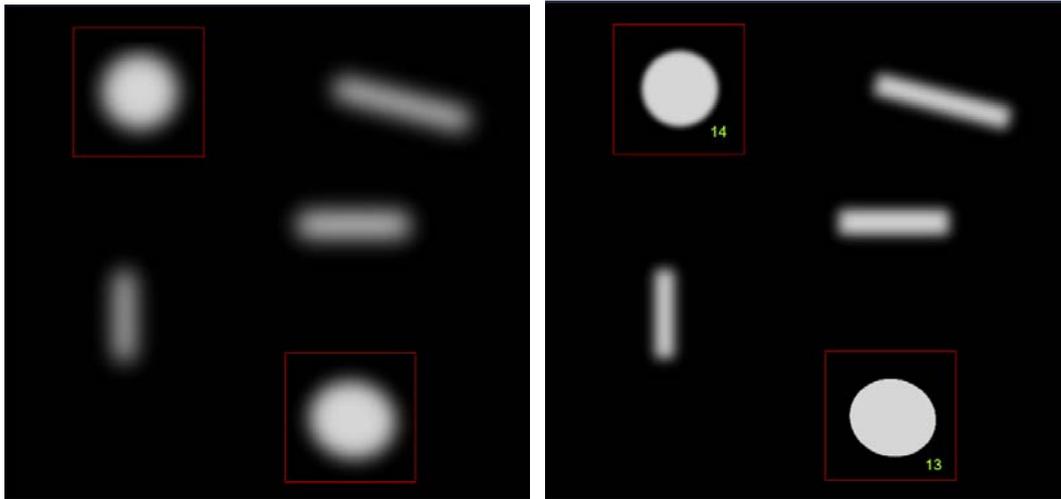


**Abbildung 35: Einzelbilder des Modell-Stapels.** (Quelle: eigene Darstellung)

### 6.2.5 Resultate des IB-AF bei biologischen Präparaten

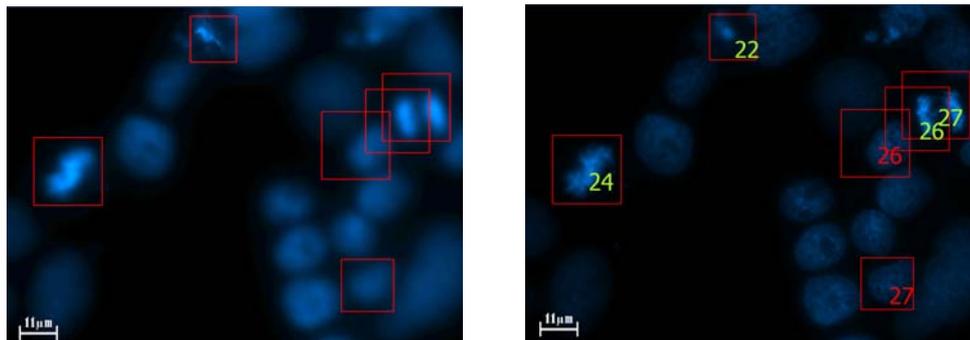
#### Mitotische CaCo-2-Zellen

Die umfangreichsten Auswertungen erfolgten an CaCo-2-Zellpräparaten (siehe Abb. 38), die Zellen in unterschiedlichen Zellstadien enthalten. Der Klassifikator des IB-AF wurde bei dieser Probenreihe auf Zellen trainiert, bei denen die DNA-Strukturen deutlich sichtbar sind. Abbildung 37 zeigt eine repräsentative Aufnahme aus dieser Reihe. Dabei stellen die roten ROI im linken Bild die in der defokussierten Ebene erzeugten Objekthypothesen



**Abbildung 36:** Links: Ausgangsebene und die darauf ermittelten Hypothesen. Rechts: Ermittelte Ebene 13 mit verifiziertem Kreisobjekt unten rechts. Der Kreis links oben befindet sich scharf abgebildet in Ebene 14. (Quelle: eigene Darstellung)

dar. Die grünen Zahlen visualisieren die im Verifikationsschritt als korrekt klassifizierten Bildausschnitte. Jede einer ROI zugehörigen Zahl bedeutet die z-Koordinate der jeweils zugeordneten Fokusebene (rechte Bild von Abb. 37).



**Abbildung 37:** Links: Hypothesen in einer nicht-fokussierten Probe. Rechts: Verifizierte ROI erhalten grüne Z-Werte, welche für die Fokusebene stehen. (Quellen: eigene Darstellung; Zur Verfügung gestellt durch Zeiss MicroImaging GmbH)

Insgesamt konnten 89 Bildstapel aufgenommen und ausgewertet werden. Sie enthielten insgesamt 150 zu erkennende mitotische Zellen. Im Schritt der Hypothesengenerierung ermittelte der IB-AF 309 Bildreale, von denen insgesamt 155 Objekte als positiv verifiziert wurden. Diese unterteilten sich in 92 RP- und 63 FP-Erkennung. Meine AF-Methode konnte trotz der hohen FP-Zahl 71,31% der vorhandenen Ebenen finden (siehe Tab. 14). Im Vergleich absolvierten die ausgewählten klassischen AF-Methoden die Fokussierungsaufgabe deutlich schlechter. Der Gaussian Derivatives Algorithmus ermittelte 56,43%, Autokorrelation 46,21% und die Methode der Normalisierten Varianz nur 12,80% der gewünschten Fokussierungen (Tab. 13).

<b>CACO-2 Serie</b>	<b>Erkannte Fokusebene</b>
IB-AF	71,31 %
Autokorrelation	46,21 %
Normalisierte Varianz	12,80 %
Gaussian Derivatives	56,43 %

**Tabelle 13: Vergleich der Erkennungsraten des Inhaltsbasierten Autofokus mit klassischen AF-Methoden**

	<b>Absolute Anzahl</b>	<b>Prozentsatz</b>
Mitotische Zellen	150	100 %
Hypothesen	309	206 %
RP-Verifikationen	92	61,34 %
FP-Verifikationen	63	/
RN-Verifikationen	136	/
FN-Verifikationen	18	12 %
von der Hypothesenfindung nicht erkannte mitotische Zellen	40	26,67 %

**Tabelle 14: Klassifikationsergebnisse auf den vom Hypothesenmodul detektierten und von der Verifikation eingeordneten ROIs durch das Inhaltsbasierte AF-System.**

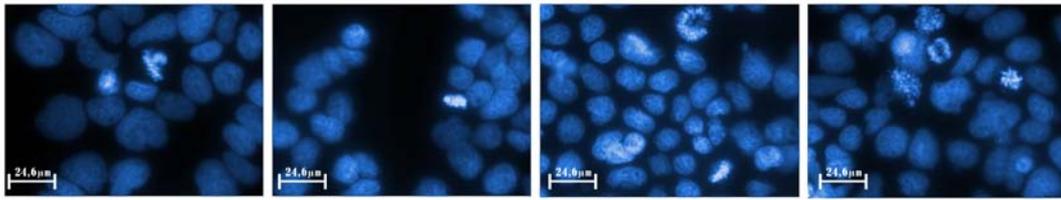
#### CaCo-2-Zellkolonien

CaCo-Zellen waren ebenfalls Untersuchungsobjekt in dieser Testreihe (Abb. 39). Die mit Hoechst 33342 markierten Proben bestanden entweder aus losen Zellen oder aus Zellkolonien mit erhöhter Fluoreszenzintensität. Dabei treten beide Zellformen in unterschiedlichen Fokusebenen auf (siehe Tabelle 3 "Fokusebenen"). Die Fokusebenen liegen ein bis drei Bildlagen auseinander, was bereits zu merklichen Unschärfen führt.

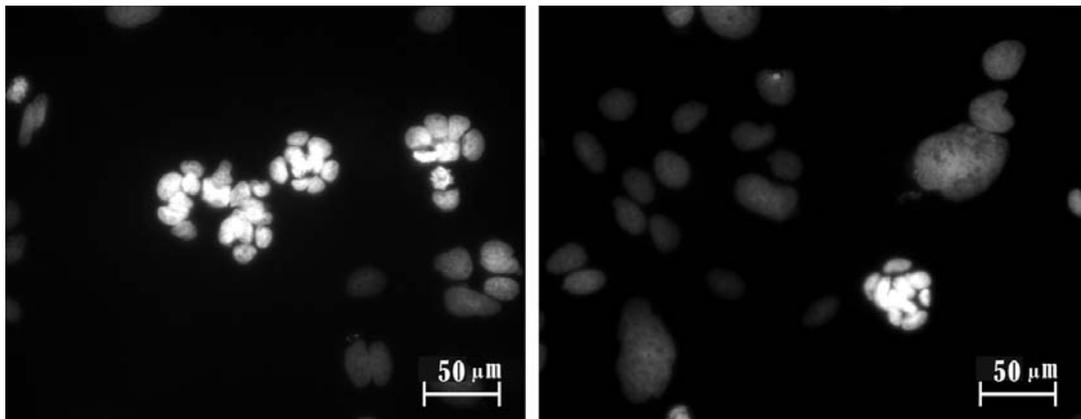
<b>Zellkolonie-Serie</b>	<b>Erkannte Fokusebene</b>
IB-AF	71,43 %
Autokorrelation	57,14 %
Normalisierte Varianz	42,86 %
Gaussian Derivatives	42,86 %

**Tabelle 15: Vergleich der Erkennungsraten des Inhaltsbasierten Autofokus mit klassischen AF-Methoden bei Zellkolonie-Proben**

Von insgesamt 8 Stapeln wurde eine Aufnahme für die Generierung der Trainingsbilder verwendet. Der Klassifikator wurde in diesem Test mit lediglich drei Positiv- und sechs Negativ-Beispielen und bezüglich einer RP-Rate von 99% und einer maximalen FP-Rate von 50% je Kaskade auf die Detektion von Kolonien hin trainiert. Die sieben Teststapel beinhalteten insgesamt 20 zu detektierende Zellhaufen in 7 Fokuslagen. Der Inhaltsbasierte AF war wiederum den klassischen Ganzbildfokusmethoden hinsichtlich



**Abbildung 38: Auswahl von vier Beispielen der CaCo-2-Testreihe.** (Quellen: eigene Darstellung; Zur Verfügung gestellt durch Zeiss MicroImaging GmbH)



**Abbildung 39: Zwei Beispiele der Zellkolonien-Proben.** (Quellen: eigene Darstellung; Zur Verfügung gestellt durch Zeiss MicroImaging GmbH)

der Erkennungsrate der gewünschten Fokusebenen überlegen (siehe Tab. 15). Bei einer Detektionsrate von 71,43 % wurden je Stapel durchschnittlich 0,5 falsche Fokusebenen ermittelt.

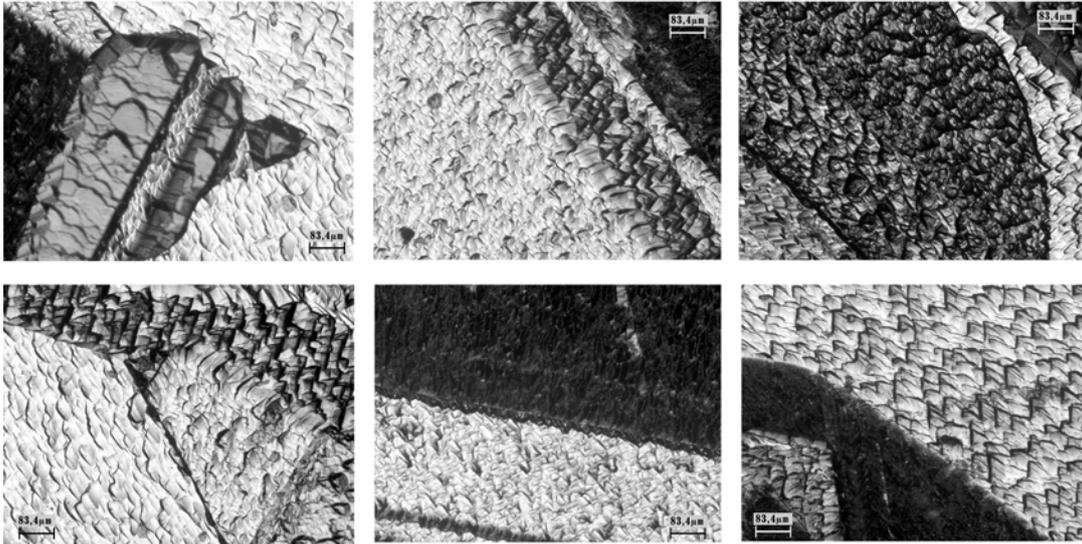
	<b>Absolute Anzahl</b>	<b>Prozentsatz</b>
Zellkolonien	20	100 %
Hypothesen	33	165 %
RP-Verifikationen	12	60 %
FP-Verifikationen	8	/
RN-Verifikationen	5	/
FN-Verifikationen	5	25 %
von der Hypothesenfindung nicht erkannte mitotische Zellen	3	15 %

**Tabelle 16: Klassifikationsergebnisse auf den vom Hypothesenmodul detektierten und von der Verifikation eingeordneten ROIs durch das Inhaltsbasierte AF-System bei Zellkolonie-Proben.**

### 6.2.6 Resultate - Material

Für den Test mit einer Materialprobe lag eine Silizium-Platte vor, deren Oberfläche mit einem DIC-Auflichtmikroskop und einem 10x/0,25NA-Objekt aufgenommen wurde. Die

Probe zeigt angrenzende Flächen unterschiedlicher Struktur (siehe Abbildung 40), die sich in unterschiedlichen Fokusebenen befinden. Die mit einer Schrittweite in Z-Richtung von  $5,88 \mu\text{m}$  aufgenommenen Stapel, zeigten Abstände zwischen den Fokuslagen von bis zu fünf Ebenen und somit einem Abstand von maximal  $19,4 \mu\text{m}$ .



**Abbildung 40: Repräsentative Auswahl des Material-Probensatzes.** (Quellen: eigene Darstellung; Zur Verfügung gestellt durch Zeiss Microlmaging GmbH)

Der Klassifikator wurde mit den Parametern 99% minimale RP-Rate und 50% maximale FP-Rate je Kaskade auf der Grundlage von 10 Kaskaden und auf der Basis von 16 Positiv- und 55 Negativ-Bildern trainiert. Durch die Rotation um schrittweise  $45^\circ$  entstanden für den Verifikationsklassifikator ein Trainingsatz von 128 positiven und 440 negativen Bildern. Für den Hypothesenklassifikator kamen zusätzlich die ein- und zweimal Tiefpassgefalteten Instanzen hinzu. Das Ziel des Trainings war die Erkennung der hellen gröber strukturierten Flächen gegenüber den dunklen Flächen und Regionenkanten.

Die Auswertung erfolgte auf 19 Stapeln und jeweils zwei Durchläufen je Stapel, bei denen die Hypothesensuche zum einen fünf Ebene unter, zum anderen 10 Ebene über der Fokusebene startete. Insgesamt ermittelte der IB-AF mit Gaussian Derivatives als Schärfefunktion 35 der gewünschten 38 Ebenen (und damit 92,11 %), die insgesamt 40 zu detektieren zusammenhängende Flächen beinhalteten (siehe Tabelle 17). Damit findet der IB-AF deutlich mehr korrekte Fokusebenen als klassische Fokusmethoden.

Allerdings zeigt die Auswertung die geringe Spezifität des IB-AF bei diesem Präparat, die durch die Falsch-Positiv-Zahl von 50 Erkennungen und letztlich einer Detektion von 29 falschen Fokusebene entsteht (siehe Tabelle 18). Damit wurden pro Stapel 0,76 falsche Ebenen ausgegeben.

<b>Material Serie</b>	<b>Erkannte Fokusebene</b>
IB-AF	92,11 %
Autokorrelation	73,68 %
Normalisierte Varianz	57,89 %
Gaussian Derivatives	63,16 %

**Tabelle 17: Vergleich der Erkennungsraten des Inhaltsbasierten Autofokus mit klassischen AF-Methoden für die Materialprobe.**

	<b>Absolute Anzahl</b>
Hypothesen	109
RP-Verifikationen	46
FP-Verifikationen	50
RN-Verifikationen	11
FN-Verifikationen	2
von der Hypothesenfindung nicht erkannte Strukturen	3
Falsch detektierte Fokusebene	29

**Tabelle 18: Klassifikationsergebnisse auf den vom Hypothesenmodul detektierten und von der Verifikation eingeordneten ROIs durch das Inhaltsbasierte AF-System.**

## 7 Diskussion

Es war das Hauptziel meiner Arbeit, über das begrenzte Leistungsvermögen der klassischen AF-Methoden bei Untersuchungen von vielschichtigen Präparaten, Proben mit irrelevanten Objekten und Proben, die durch artifizielle Gegenstände störende Bildinformationen enthalten, durch eine neue Methode hinauszugehen.

Diese neue Autofokussierungsmethode sollte in der Lage sein,

- *mehrere Fokusebenen zu ermitteln, falls sich Objekte in mehreren Lagen befinden.*
- *auf interessierende Objekte zu fokussieren und irrelevante Inhalte und Artefakte von der Fokussuche auszuschließen.*
- *flexibel bei unterschiedlichen Lichtmikroskop-Modi zu arbeiten.*
- *die Fokussierung für ein breites Spektrum von Probenarten (biologische, medizinische und industrielle) zu ermöglichen.*
- *die oben aufgeführten Bedingungen mit lediglich einer generellen Methode zu erfüllen.*

Bei der Entwicklung der neuen Methode ließ ich mich von der Annahme leiten, dass eine erfolgreiche Fokussierung bei problematischen Proben den Weg über eine Einschränkung auf relevante Bildbereiche gehen muss und allein eine neue Schärfefunktion oder eine neue Suchstrategie nicht zum Ziel führt. In meiner Arbeit kombinierte ich meines Wissens erstmals eine Fokussierungsmethode mit einem Objektklassifikationsverfahren zur Autofokussierung.

### Vorausgehende Untersuchungen

In der Vortestphase (Abschnitt 6.2.2) untersuchte ich unter anderem die Fokussierungsqualität von klassischen Schärfefunktionen, wenn der Messbereich in der XY-Ebene auf ROIs beschränkt wird. Dabei zeigten sich unterschiedliche Ergebnisse. Innerhalb der Hellfeld-Phasenartefaktprobe ermittelten alle Funktionen, ausgenommen die ableitungsbasierte Korrelation und die Wavelet-Transformation, nach der Eingrenzung der Suchregion die korrekte Fokusebene. Und auch im Phasenkontrast der Kristall-Artefaktprobe zeigte sich eine Verbesserung nach der Ausschnittsbeschränkung. Jedoch konnten im Phasenkontrastkanal der Cytoskelett-Probe nur die beiden Methoden Autokorrelation und Gaussian Derivatives überzeugen.

Die Auswertung der anderen meist im Fluoreszenzkanal aufgenommenen Präparate führt ebenfalls zu der Schlussfolgerung, dass das Eingrenzen des Bildbereichs nicht grundsätzlich mit allen Schärfefunktionen und bei allen Mikroskop-Modi Erfolg verspricht. Dieses Ergebnis widerspricht der Erwartung, dass die Wahl der Schärfefunktion bei einer hinreichend guten ROI-Detektion in den Hintergrund tritt. Dennoch ist mit der richtigen Methode das Zusammenspiel von Eingrenzung auf bestimmte ROIs mit Schärfemessung erfolgreich, um mehrere bzw. bestimmte Fokusebenen in schwierigen Präparaten zu finden. Der Vergleich aller Fokusschärfefunktionen belegt, dass die richtige Methode der Algorithmus der Gaußschen Ableitungen ist. Er schneidet durchweg am besten ab, weil mit ihm in 14 von 17 Fällen die korrekte Fokusebene ermittelt wurde. Zudem besteht bei der Gaussian Derivatives Methode keine Präferenz bezüglich eines Mikroskop-Modus, wie sie beispielsweise bei der Normalisierten Varianz bezüglich Fluoreszenzproben ersichtlich ist. In der konkreten Anwendung des IB-AF auf die Teststapel konnte ich auch konstatieren, dass die Gaussian Derivatives Funktion jeweils immer die Ebene mit der schärfsten Abbildung des Objekts innerhalb der jeweiligen ROI findet.

#### Gesamtleistungsvermögen des IB-AF

Bei Gaussian Derivatives als funktionierender Schärfefunktion lag die größte Herausforderung zweifelsohne in der korrekten Detektion und Verifizierung von interessierenden Bildarealen. Die Studie am Modellstapel mit zwei Objektkategorien bewies die grundsätzliche Funktionsfähigkeit des Systems, mehrere Fokusebenen von interessierenden Untersuchungsobjekten zu bestimmen. Das Training wurde an sehr wenigen Bildausschnitten (jeweils 5 positive und negative) durchgeführt und brachte einen Klassifikator hervor, der sowohl die Hypothesendetektion mit höchstmöglicher Sensitivität in defokussierten Ebenen, d.h. Erkennung aller relevanten Objekte, als auch die vollständig richtige Verifikation der Bildregionen ermöglichte. Im Modellpräparat wurden folglich alle zwei Kreisobjekte in ihren jeweiligen Fokuslagen ermittelt.

Die Beispielprobe beinhaltete einfache Elemente, die für Objekterkennungsalgorithmen keine Hürde darstellen dürfen. Anders verhielt es sich mit realen Proben. Im umfangreichsten Versuch wurde das Inhaltsbasierte AF-System für das Scharfstellen auf sich teilende Zellen trainiert und verwendet. Die Erscheinungsunterschiede zwischen diesen mitotischen Zellen und inaktiven Zellen fielen besonders in unfokussierten Lagen aufgrund der verschwommenen bzw. nicht mehr vorhandenen Strukturen sehr gering aus. Deshalb fand der IB-AF im ersten Schritt eine Menge von hypothetischen Objekten, die der doppelten Anzahl der tatsächlich vorhandenen teilenden Zellen entsprach. Von diesen Hypothesen wurden 61,34% aller interessierenden Objekte korrekt verifiziert. Diese

Erkennungsrate befähigte den IB-AF schließlich, 71,31% aller möglichen Fokusebene auszugeben. Damit liegt die Leistung deutlich über denen der Methoden der Ganzbild-fokusmessung ([18, 53, 21]). Während die Gaussian Derivatives Funktion 56,43% aller Ebenen ermittelte, erzielte die Autokorrelation eine Rate von 46,21% und die Normalisierte Varianz nur 12,8%. Der um einiges besseren Erkennungsleistung des IB-AF steht jedoch die nicht geringe Anzahl von Falsch-Positiv-Erkennungen während der Verifikationsphase gegenüber. Diese führte zu einer Ausgabe von durchschnittlich 0,76 falschen Fokusebenen je Probe.

Die Auswertung der Zellkolonie-Proben ergaben ähnlich gute Ergebnisse. Hierbei wurden 90% aller Kolonien während der Hypothesenfindung detektiert. Weil insgesamt jedoch 25% als negativ von der Verifikation klassifiziert wurden, sank die RP-Rate auf 60%. Mit der letztendlichen Fokusebenenrate von 71,43% lag der IB-AF wiederum deutlich über den Erkennungsraten der drei klassischen AF-Methoden - Autokorrelation 57,14%, Normalisierte Varianz 42,86%, Gaussian Derivatives 42,86%. Erneut ging die hohe Erkennungsrate mit einer Detektion von durchschnittlich 0,5 Fokusebenen unerwünschter Objekte einher.

Die Ergebnisse der Materialprobenreihe zeigten eine nochmalige Steigung der Erkennungsraten. Hier konnte der IB-AF 92,11% der gewünschten Fokusebenen ermitteln und erneut eine bessere Erkennungsleistung als die Standardmethoden liefern. Wiederum mussten mit durchschnittlich 0,76 falsch detektierten Ebenen pro Stapel eine wesentliche Einschränkung in Kauf genommen werden.

Die gesamte Auswertung hätte durch eine zusätzliche Untersuchung von Gewebsschnitten abgerundet werden können. Solche - bei guter Präparierung monofokalen Schnitte- wären für Test mit dem IB-AF aber nur dann sinnvoll gewesen, wenn sie mehrere Fokusebene enthielten (z. B. aufgrund von Probenrissen und -falten). Diese lagen für diese Arbeit jedoch nicht vor. Allerdings kann ein gewisser Vergleich mit der Materialprobe gezogen werden. Denn Gewebsschnitte enthalten ebenfalls meist zusammenhängende Strukturen (von Zellen), als dass diskrete Objekte unterschieden werden. Es ist somit zu erwarten, dass auch bei histologischen Präparaten mit einer höheren Falsch-Positiv-Erkennung zu rechnen ist.

### Haar-Merkmale und AdaBoost-Training

Für die Klassifikation sowohl bei der Hypothesengenerierung als auch bei der Verifizierung der Bildareale kombinierte ich Haar-Merkmale und die AdaBoost-Merkmalss Selektion. Dabei zeigte die Methode zwei Seiten. Zum einen erfüllte sie die Forderung nach hoher Sensitivität bei der Hypothesenfindung in unscharfen Abbildungen. Zum anderen kann die

abschließende Klassifizierung der scharf eingestellten ROIs aufgrund der zu großen Zahl von Falsch-Positiv-Erkennungen im Verifikationsschritt nicht überzeugen. Sie führte zu der Ausgabe von durchschnittlich  $<1$  Fokusebenen je Probe, die irrelevante Objekte abbildeten. Die Ursache dafür sah ich bei zwei Faktoren. Einerseits spielt die geringe Anzahl von Trainingsbildern eine Rolle. Wie bereits in Abschnitt 6.2.3 festgestellt, ist mit besseren Klassifikationsraten zu rechnen, wenn die Trainingsdatenbasis einen größeren Umfang und auch eine größere Interklassenvarianz besitzt. Letztere hängt unweigerlich von der Datengröße ab. Andererseits ist es die doch vorhandene Limitierung der Haar-Merkmale durch ihre Eigenschaft als simple Kantenfilter. Sie können mitunter entscheidende Klassenmerkmale wie Farb- und Helligkeitsunterschiede nur indirekt durch Helligkeitsdifferenzen extrahieren. Dies zeigte sich zum Beispiel bei den Materialproben. Dort führten Flächen mit anderer Helligkeit aber ähnlicher Struktur wie die zu detektierenden Arealen zu FP-Erkennungen.

Mit dem AdaBoost-Training in Kombination mit der Kaskadierung der Einzelklassifikatoren konnte ich sowohl bei der Hypothesendetektion als auch in der Verifikationsphase einen sehr schnellen Erkennungsprozess erreichen. Dem stand die Dauer des Trainings gegenüber. Es kann in Abhängigkeit von der Größe der Interklassenvarianz der Datenbasis und der gewünschten Klassifikationsgenauigkeit wenige Sekunden oder aber auch Stunden dauern. Ein derart langes Training ist aber, wenn überhaupt, nur bei einer sehr hohen Zuverlässigkeit des Gesamtsystems zu vertreten.

## 8 Schluss und Ausblick

Die in dieser Arbeit vorgestellte Kopplung aus Detektion von relevanten, inhaltspezifischen ROIs und zwischengeschalteter Schärfemessung ist ein naheliegender aber bis dato nicht umgesetzter Evolutionsschritt auf dem Gebiet der bildbasierten Autofokussierung.

Der Inhaltsbasierte AF kann mehrere Fokusebenen finden und zwischen relevanten und irrelevanten Untersuchungsobjekten unterscheiden. Die Entscheidungsgrundlage ist dabei nicht das Auftreten mehrerer lokaler Maxima in der Fokuskurve, sondern die Auswertung von allen potentiellen Bildregionen, die jeweils für sich eigene Fokuslagen besitzen dürfen. Der IB-AF grenzt sich durch diese Vorgehensweise deutlich von existierenden Methoden ab und erzielte in meinen Untersuchungen eine Steigerung der Erkennungsleistung von relevanten Fokusebenen um 15 - 20% gegenüber der nächstbesten AF-Methode. Damit habe ich die zwei wichtigsten Zielsetzungen erfüllt. Weiterhin operierte der IB-AF auf unterschiedlichen Präparaten bei unterschiedlichen Mikroskop-Modi und bewies die grundsätzlich gute Funktionalität bei System- und Probeninhaltübergreifenden Untersuchungen.

Die Umsetzung des Inhaltsbasierten Autofokus zeigt das große Potential dieses neuen Autofokusansatzes. Sie konnte jedoch deren Leistungsfähigkeit nicht ausschöpfen. Der IB-AF bietet als ein umfassendes Konzept einen großen Spielraum für Weiterentwicklung, den es zu nutzen gilt. Sein Leistungsvermögen im Bezug auf Zuverlässigkeit kann durch die Qualifizierung der einzelnen Teilmethoden erhöht werden. Die wohl entscheidende Rolle spielen dabei die Klassifikationsalgorithmen. Zwar kann die Objekterkennung auf der Basis der Haar-Merkmale und des AdaBoost-Trainings eine gewünscht schnelle Detektion im Online-Modus aufweisen, so dass die Dauer des Gesamt-AF fast allein von der Mikroskopmechanik abhängt. Die Detektionsrate im Verifikationsschritt liefert jedoch zu viele Falsch-Positiv-Erkennungen. Es hat sich herausgestellt, dass mit den Haar-Merkmalen keine gleich guten Klassifikationsergebnisse auf allen Probenvarianten erzielt werden können. Soll es weiterhin darum gehen, eine generell anwendbare AF-Methode zu entwickeln, können zwei Ansätze verfolgt werden. Man sucht nach einem Ersatz für die Haar-Merkmale oder man erweitert die Haar-Merkmale um ergänzende Features. Dabei können die in der Auswertung angesprochenen Intensitäts- und Farbeigenschaften von Präparaten eine hilfreicher Zusatz sein.

Nimmt man davon Abstand, eine allgemein anwendbare Inhaltsbasierte AF-Applikation zu verwenden, lässt sich mit Sicherheit die Fokussierungsqualität durch Anpassung von Klassifikations- und Schärfelgorithmen an die jeweils verwendete mikroskopische Probe erhöhen. Dann können auch spezifische morphologische Eigenschaften und Farbinforma-

tionen einbezogen werden.

In der Umsetzung des IB-AF konnte ich zwei Aspekte noch nicht berücksichtigen. Der Anwender muss die Datenbasis über die Zeit mit weiteren Trainingsbildern erweitern können. Zudem sollten Fehlklassifikationen über eine Rückmeldung den Klassifikator verbessern. Beides erfordert ein erneutes, im Hintergrund ablaufendes Klassifikatortraining. Zum anderen kann das Anwendungsspektrum des AF vergrößert werden, wenn das Training auf multiple Klassen erweitert wird.

Darüber hinaus müssen umfangreichere Verwendungsstudien durchgeführt werden. Beispielsweise blieb die Frage unbeantwortet, ob der Anwender nach abgeschlossener Autofokussierung nur eine Fokusebene (mit der meisten Information) angezeigt haben möchte oder aber alle gefundenen Lagen bzw. eine Fusion aller ermittelten Schärfeebene zu einem Bild (Erweiterte Tiefenschärfe).

## 9 Anhang

### 9.1 Zusätzliche Abbildungen und Fokuskurven

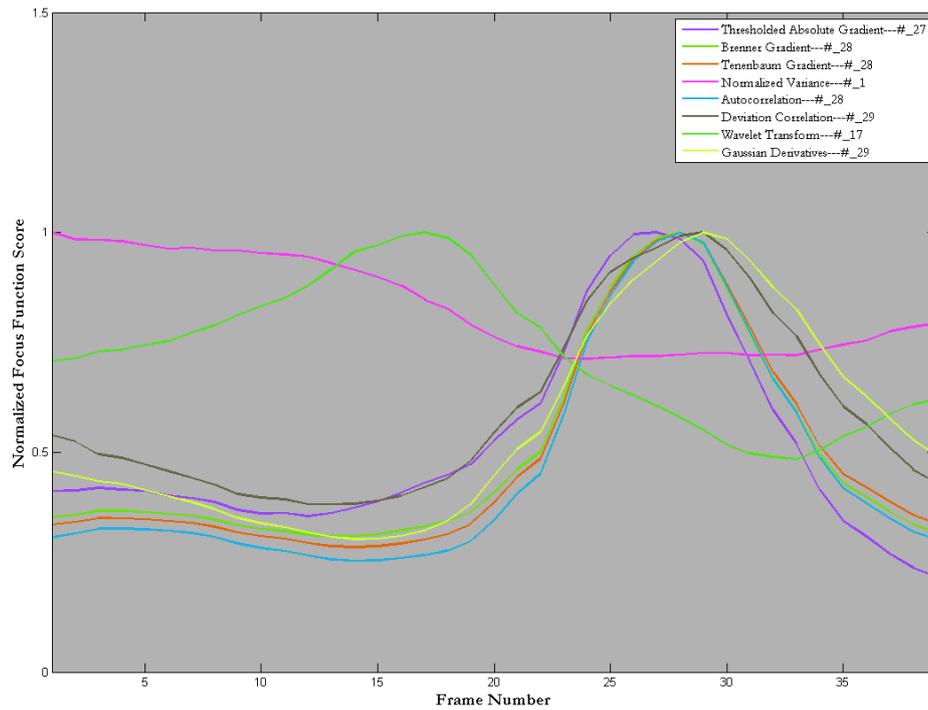


Abbildung 41: Fokuskurven für das Kristall-Artefakt-Präparat im Phasenkontrast-Modus im linken Bild der Abb. 10 und Abb. 44 . (Quelle: eigene Darstellung)

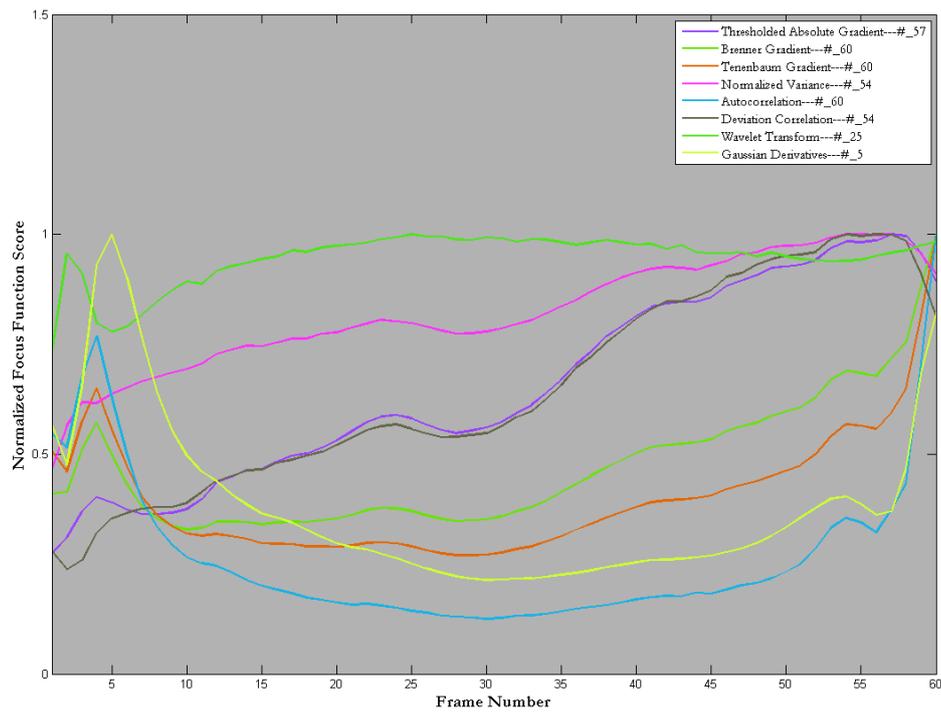
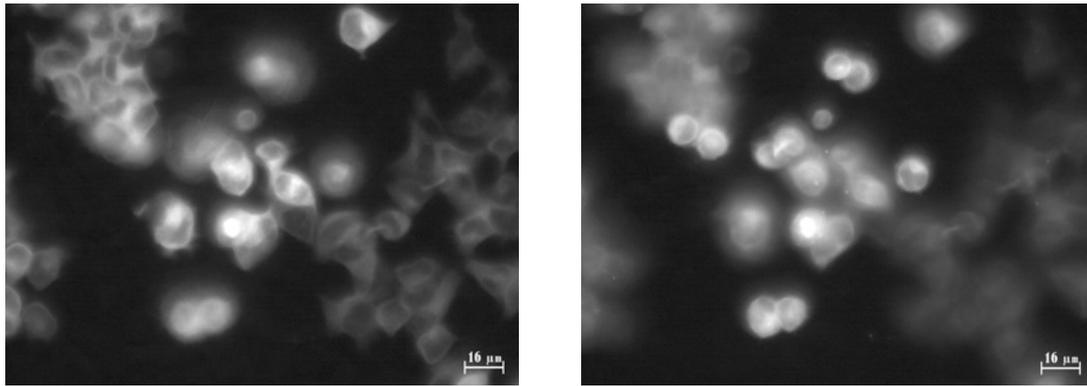
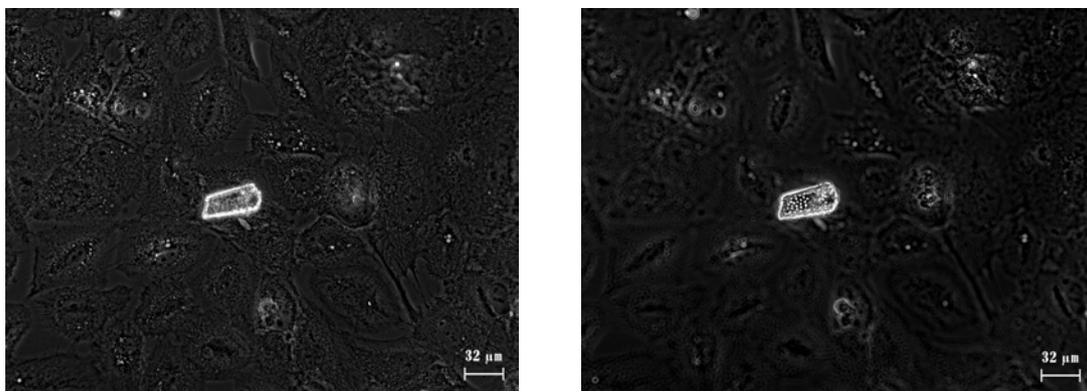


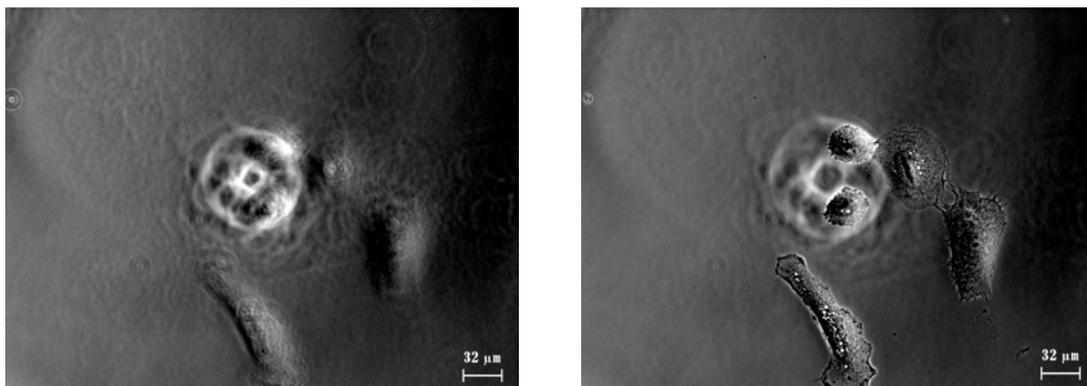
Abbildung 42: Fokuskurven für das Phasen-Artefakte-Präparat im Phasenkontrast-Modus im rechten Bild der Abb. 10 und Abb. 45. (Quelle: eigene Darstellung)



**Abbildung 43: HeLaR-Zellen im Fluoreszenz-Kanal in verschiedenen Fokusebenen.** (Quellen: eigene Darstellung; Zur Verfügung gestellt durch Zeiss Microlmaging GmbH)



**Abbildung 44: Kristall-Artefakt im Phasenkontrast-Modus.** Mikroskop: Zeiss Axio Observer.Z1 mit 20x/0.80NA-Objektiv. Aufgenommen mit einer Zeiss AxioCam HR3 mit einer Auflösung von 1388x1040 px bei 14 bit-Tiefe. (Quellen: eigene Darstellung; Zur Verfügung gestellt durch Zeiss Microlmaging GmbH)



**Abbildung 45: Phasen-Artefakte im Phasenkontrast-Modus.** Mikroskop: Zeiss Axio Observer.Z1 mit 20x/0.80NA-Objektiv. Aufgenommen mit einer Zeiss AxioCam HR3 mit einer Auflösung von 1388x1040 px bei 14 bit-Tiefe. (Quellen: eigene Darstellung; Zur Verfügung gestellt durch Zeiss Microlmaging GmbH)

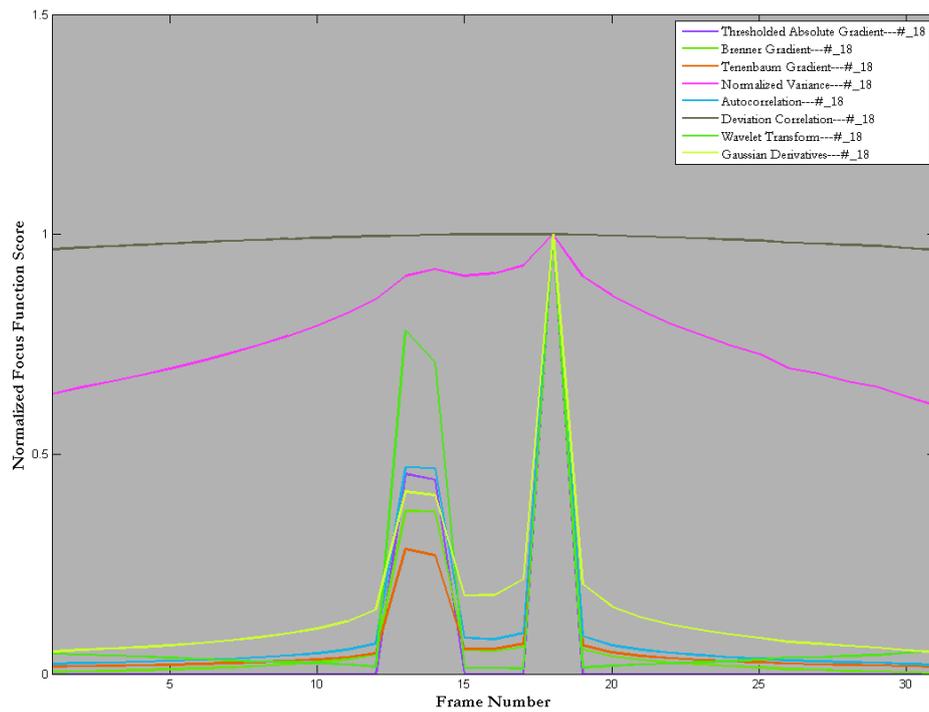


Abbildung 46: Fokuskurven des Modellstapels aus Abb. 35. (Quelle: eigene Darstellung)

## 9.2 Verfahren der Merkmalsextraktion

### 9.2.1 Principle Components Analysis (PCA)

Die "Analyse der Hauptkomponenten" ist nicht nur in der Lage, die Dimension der Ausgangsdaten zu reduzieren, sondern führt auch eine Transformation aus, die diskriminierende Merkmale auf der Basis der Datenvarianz hervorbringt. Stellt man sich einen Trainingsbildsatz als Punkte in einem mehrdimensionalen Raum vor, so lässt sich die PCA-Transformation als eine Drehung der Punktwolke verstehen, so dass die neuen Koordinatenachsen der Richtungen der größten Varianz entsprechen. Diese Richtungsvektoren stellen die gesuchten und untereinander unkorrelierten Principle Components dar.

In der Praxis der Bildverarbeitung ist die PCA eine Matrixmultiplikation (der ausführliche Berechnungsablauf ist in Tabelle 19 dargestellt) aus der Transponierten der Transformationsmatrix  $W$  und dem Datensatz  $X$ , so dass sich der Merkmalsvektor  $y$  zu:

$$y = W^T X \quad (42)$$

ergibt. Die ursprünglich zweidimensionalen Bilddaten liegen in  $X$  als in eine Matrix aneinander gefügte 1-D-Vektoren der Dimension Bildhöhe x Bildbreite z.B. durch Aneinanderreihung der Bildzeilen vor. Zudem ist jeder Bildvektor vom Bildmittelwertvektor bereinigt.

Die Berechnung der Transformationsmatrix führt über die Berechnung der Eigenvektoren  $E$  und -werte  $\lambda$  der Kovarianzmatrix  $C$  des Datensatzes. Die Lösung des Eigensystems führt zu einer Anzahl von Eigenvektoren, die der Menge der Bildsamples im Datensatz entspricht. Die Eigenvektormatrix wird entsprechend der Beträge der korrespondierenden Eigenwerte sortiert, weil ein hoher Eigenwert auf einen für die Variationsbeschreibung signifikanten Vektor zeigt. Schließlich stellt diese nach den Eigenwerten geordnete Eigenvektormatrix die in Gleichung 42 beschriebene Transformationsmatrix dar. Große Bilddatensätze führen entsprechend zu einer Vielzahl von Eigenvektoren. Praktische Untersuchungen (2) haben aber gezeigt, dass nur eine geringe Anzahl dieser Vektoren für eine Klassifikationsaufgabe sinnvoll ist. Im eigentlichen Klassifikationsschritt wird das unbekannte Bild wie zuvor auch der gesamte Trainingsdatensatz mit der Matrix  $W$  transformiert. Im einfachsten Fall kann der so berechnete transformierte Bildvektor per Minimum-Distance-Methode einer der Vektoren im transformierten Trainingssatz  $y$  und somit einer Klasse zugeordnet werden. Genauso ist es aber möglich, die PCA-reduzierten Merkmale einer erweiterten Merkmalsselektionsmethode wie der Support-Vector-Machine bereitzu-

stellen.

<b>I.) Anordnung der Trainingsbilder als Vektoren zur Matrix <math>I</math>, wobei jeder Bildvektor einer Klasse <math>K_M</math> zugehört</b>	$I = [I_1, I_2, I_3 \dots I_N]$ wobei $I_N \rightarrow 1xHöhe_{I_N} \times Breite_{I_N}$ und $I_N \in [K_1 \dots K_M]$
<b>II.) Bereinigung der Matrix durch den Trainingsbildmittelwert <math>\psi</math></b>	$X = [I_1 - \psi, I_2 - \psi, I_3 - \psi \dots I_N - \psi]$ wobei $\psi = \frac{1}{N} \sum_{n=1}^N I_n$
<b>III.) Berechnung der Kovarianzmatrix <math>C</math></b>	$C = \Gamma \Gamma^T$ mit: $X = [X_1, X_2, X_3 \dots X_N]$
<b>IV.) Berechnung der Eigenvektoren <math>V_E</math> und zugehörigen Eigenwerte <math>\lambda_E</math> über das Eigensystem</b>	$C X V_E = \lambda_E X V_E$
<b>V.) Bildung der Transformationsmatrix <math>W</math> aus den Eigenvektoren <math>V_{max}</math>, die zu den betragsmäßig größten Eigenwerten korrespondieren</b>	$W = [V_{max}]$ mit: $V_{max} \in V_E$

**Tabelle 19: Ablauf der Berechnung der Transformationsmatrix auf der Basis der Hauptkomponenten (principle components)**

### 9.2.2 Scale Invariant Feature Transform (SIFT)

Das von Lowe ([34]) entwickelte Verfahren der skalierungsinvarianten Merkmalsextraktion (scale-invariante feature transform, kurz: SIFT) ist eine pixelbasierte Methode zur Lokalisation von besonders diskriminierenden Bildregionen. Die vom SIFT-Algorithmus ermittelten Merkmale sind daraufhin optimiert, robust gegenüber Rotation, Skalierung und teilweise gegenüber Helligkeitsschwankungen zu sein. Zentrales Element der Methode ist die Ermittlung von "Schlüsselpunkten" also markanten Bildregionen. Das Gesamtverfahren läuft in vier Schritten ab: 1) Extrema-Detektion im Skalenraum, 2) Lokalisation von Schlüsselpunkten, 3) Zuweisung der Orientierung über Gradienten und 4) Berechnung des Merkmalsvektors.

**Schlüsselpunkt-Suche** Die Suche nach potentiellen so genannten Bildschlüsselpunkten (Keypoints), die für die Eignung als diskriminierendes Merkmal in Frage kommen, berücksichtigt Skalierungs- und Auflösungsinvarianz zum einen durch die Transformation

des Bildes in eine Bildpyramide. Das heißt, dass das Ausgangsbild bilinear bis zu einer sinnvollen Größe um einen konstanten Faktor herunter skaliert wird. Zum anderen wird jedem Bild in der Pyramide ein Skalenraum zugeordnet. Der Skalenraum  $L$  bezeichnet hier die sukzessive Faltung des Eingangsbildes  $I$  mit der Gauß-Funktion  $G$  mit um den Faktor  $k$  steigendem Sigma  $\sigma$ , wobei mindestens vier Skalenraumbilder pro Pyramidenlevel (auch Oktave genannt) entstehen müssen:

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (43)$$

mit:

$$\sigma = k\sqrt{2} \quad (44)$$

so dass der Gauß-Skalenraum als:

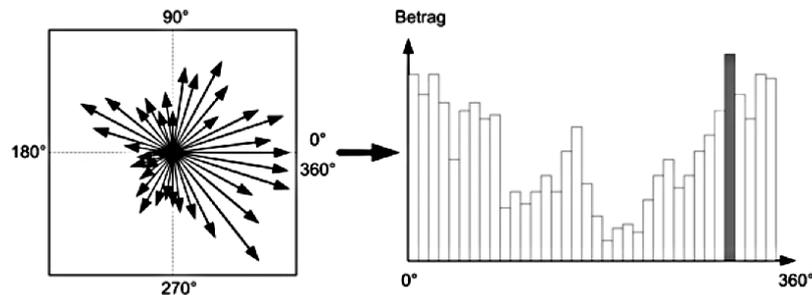
$$L(x, y, \sigma) = G(x, y, \sigma) \otimes I(x, y) \quad (45)$$

definiert ist.

Am Ende entsteht der letztlich zugrunde liegende Skalenraum aus der Subtraktion jeder Gaußinstanz mit der nächst zugeordneten. Das Ergebnis ist ein Difference-Of-Gaussian-Bild (DoG), das eine Approximation der Laplacian-Of-Gaussian-Faltung (LoG) darstellt. In der konkreten Realisierung entspricht der erste Eintrag der Pyramide dem um Faktor 2 hochskalierte Ausgangsbild und nicht dem Ursprungsbild selbst. Die Pyramidenanordnung der DoG-Skalenräume ist nur der Ausgangspunkt für die Keypoint-Suche. Ein potentieller Keypoint ist ein Pixel, dessen Intensitätswert ein Maximum oder Minimum bezüglich der 8 Pixelnachbarn im selben Skalenraumbild sowie zu den jeweils 9 Nachbarn im oberen und unteren Bild darstellt. Diese 26-Pixel-Nachbarschaft schließt somit die Schlüsselpunktsuche im obersten als auch im untersten Bild des Skalenraums aufgrund der fehlenden Nachbarebene aus. An dieser Stelle wird verständlich, warum mindestens vier Gauß-gefilterte Bilder je Skalenraum vorliegen müssen: vier Gauß-Instanzen bilden drei DoG-Bilder, welche eine Ebene zur Keypoint-Suche zulässt. Schließlich muss ein Schlüsselpunkt noch zwei weitere Bedingungen erfüllen. Einerseits werden Extrema mit geringem Kontrast und somit großer Empfindlichkeit gegenüber Rauschen durch eine Schwellenwertoperation ausgeschlossen, andererseits finden Kandidaten, die auf Bildkanten liegen, keine Berücksichtigung für die weitere Merkmalsberechnung.

**Berechnung des Merkmalsvektors** Jedem Schlüsselpunkt wird ein Merkmalsvektor zugewiesen, der sich aus dem Gradientenhistogramm (HoG) an der Position des Key-

points in der entsprechenden Ebene des Skalenraums zusammensetzt. Um eine Rotationsinvarianz zu erreichen, werden die HoG relativ zur Orientierung des Gradienten am Schlüsselpunkt berechnet.



**Abbildung 47: Bestimmung der Hauptgradientenorientierung über Gradientenbetrags-histogramm.** (Quelle: eigene Darstellung)

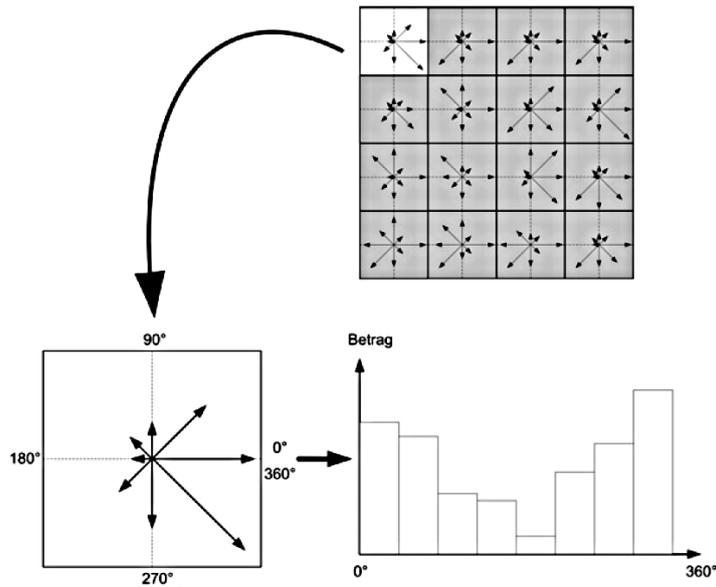
Ein Fenster wird zentral über den Bildpunkt gelegt und mit einer Gauß-Funktion gewichtet. Innerhalb dieses Fensters werden in dem Bild die Gradientenbeträge  $m$  nach der Formel:

$$m(x, y) = \sqrt{(L(x + 1, y) - L(x - 1, y))^2 + (L(x, y + 1) - L(x, y - 1))^2} \quad (46)$$

berechnet. Anschließend wird jeder Betrag, sofern er einen bestimmten Schwellwert überschreitet, in ein Histogramm übertragen, dessen x-Achse für die Orientierung im Bild steht (siehe Abbildung 47). Der zugehörige Winkel  $\rho$  des Gradientenwerts wird entsprechend berechnet durch:

$$\rho(x, y) = \tan^{-1}\left(\frac{L(x + 1, y) - L(x - 1, y)}{L(x, y + 1) - L(x, y - 1)}\right) \quad (47)$$

Das Histogramm besitzt einen Winkelabstand von  $10^\circ$ , so dass 36 Richtungen unterschieden werden. Besitzt ein Keypoint mehrere Gradientenmaxima in eine Richtung, werden entsprechend mehrere Keypoints mit diesen Orientierungen angelegt. Die Werte des Merkmalsvektor werden nun analog zur Orientierungsbestimmung aus den Gradientenbeträgen (entsprechend den Gleichungen 46 und 47) im Zusammenhang zur Gradientenrichtung ermittelt. Das Fenster besteht hierbei jedoch aus  $4 \times 4$  Unterfenstern mit jeweils einem Gradientenhistogramm, in dem Betragsmessungen je  $45^\circ$  abgetragen sind (siehe Abbildung 48). Damit ergeben sich insgesamt 128 Werte, die in Spaltenform den Merkmalsvektor je Schlüsselpunkt ergeben.



**Abbildung 48:** Sechzehn Gradienten-Unterhistogramme, die jeweils auf der Basis von 8 Orientierungen entstehen, erzeugen insgesamt 128 Gradientenbeträge. (Quelle: eigene Darstellung)

## 9.3 Merkmalsselektionsverfahren

### 9.3.1 Support Vector Machine (SVM)

Die Support-Vector-Machine ist eine spezielle Form der Kernel-Methoden zur klassenbezogenen Datentransformation. Die Theorie der Kernel-Methoden wiederum beschreibt erst einmal nur den Ansatz, Daten in einen neuen Merkmalsraum zu überführen, in dem lineare Beziehungen zwischen den Datenklassen hergestellt werden können. Dabei liegt zuallererst der von der PCA (siehe Formel 42) bekannte grundlegende Zusammenhang vor:

$$y = W^T X = \langle W, X \rangle \quad (48)$$

Im vorliegende Klassifikationskontext kann diese Formel in der Weise formuliert werden, dass die Transformations- bzw. Gewichtungsmatrix  $W$  durch Linearkombination mit den Eingangsdaten  $X = [x_1, x_2, \dots, x_n]$  die jeweils zugehörige Klasse  $y$  ermittelt. Wendet man  $W$  auf unbekannte Daten an, ergibt die Subtraktion aus der Datentransformation und der tatsächlichen Klasse im Idealfall 0.

$$y - W^T X = 0 \quad (49)$$

Für die Bestimmung von  $W$  hält dieser Idealfall auch den Lösungsansatz bereit, der in

der Minimierung des Fehlers  $\xi$  besteht, welcher sich aus der Differenz zwischen vorhergesagter und tatsächlicher Klasse ergibt. Mit der bekannten Methode der Minimierung der Fehlerquadrate (Least-Square-Minimum - LSM):

$$\min(|\xi|^2) = \min((y - W^T X)^T (y - W^T X)) \quad (50)$$

erhält man über die erste Ableitung und das Nullsetzen dieser die Lösung für die Gewichtungsmatrix:

$$\frac{d|\xi|^2}{dw} = 0 = -2X^T y + 2X^T X W \quad (51)$$

Und schließlich nach Umstellung:

$$W = (X^T X)^{-1} X^T y \quad (52)$$

Die Formel 52 zeigt zum einen bereits das zentrale Element der Kernel-Methodik, nämlich die Berechnung von  $W$  über das Skalarprodukt (Inner Product Matrix)  $X^T X$  der Input-Werte bzw., wie im Folgenden erläutert wird, über das Skalarprodukt der in einen Merkmalsraum transformierten Inputs. Zum anderen ist aus der Formel ebenfalls ersichtlich, dass die Inverse des Skalarprodukts existieren muss. Durch die Addition mit einer um  $\lambda > 0$  gewichteten Einheitsmatrix  $I$  wird letzterem aus dem Weg gegangen, weil dann eine Invertierung immer möglich ist.

$$W = (X^T X + \lambda I)^{-1} X^T y \quad (53)$$

Setzt man  $W$  in Gleichung 48 ein und nimmt eine kleine Umformung vor, erhält man die vollständige Berechnung der Schätzung  $\tilde{y}$  von  $y$  für den neuen Eingangswert  $x$ :

$$\tilde{y} = x^T (X^T X + \lambda I)^{-1} X^T y \quad (54)$$

Die direkte Berechnung der Gewichtsmatrix in Gl. 53 wird als Erst- bzw. Primallösung bezeichnet.

Es ist jedoch von grundlegender Wichtigkeit, dass die Berechnung der Schätzung  $y$  allein aus einem Skalarprodukt also nicht direkt aus den Input-Merkmalen selbst erfolgt. Dahin gelangt man, wenn die Gewichtungsmatrix als Linearkombination der Eingangswerte mit einer Gewichtung  $\alpha$  verstanden wird:

$$W = X^T \alpha \quad (55)$$

Des Weiteren ist es möglich, die Primallösung 52 umzuformulieren:

$$(X^T X + \lambda I)W = X^T y \quad (56)$$

$$X^T X W + \lambda W = X^T y \quad (57)$$

$$\lambda W = X^T (y - X W) \quad (58)$$

Mit Gleichung 55 folgt:

$$W = X^T \alpha = \lambda^{-1} X^T (y - X W) \rightarrow \alpha = \lambda^{-1} (y - X X^T \alpha) \quad (59)$$

$$\alpha \lambda = y - X X^T \alpha \quad (60)$$

$$\alpha = (X X^T + \lambda I)^{-1} y \quad (61)$$

Diese so genannte Dual-Lösung führt schließlich zur Vorhersage  $\tilde{y}$ :

$$\tilde{y} = W^T x = X^T \alpha x = x^T X^T \alpha = \alpha \langle x, X \rangle = (X X^T + \lambda I)^{-1} y \langle x, X \rangle \quad (62)$$

Diese Lösung erscheint zunächst komplex, hat jedoch folgende angenehme Eigenschaften:  $X X^T$  ist die so genannte Gram-Matrix, die lediglich das äußere Produkt der Trainingseingangsdaten darstellt.  $\langle x, X \rangle$  wiederum ist das innere Produkt aus einem neuen Eingangswert  $x$  mit unbekannter Klasse und der Input-Matrix  $X$ . Diese wird als Kernel-Matrix bezeichnet. Ein Unterschied zur Primallösung ist auch die Dimension der Matrizen. Während in Gleichung 52 das innere Produkt  $X^T X$  die Dimension  $l \times l$  besitzt, wobei  $l$  für die Länge der Merkmalsvektoren steht, hat die Gram-Matrix in Gleichung 62 nur die Dimension  $n \times n$  mit  $n$  gleich der Anzahl der Trainingssamples.

Der gesamte bisherige Abschnitt beschäftigte sich mit der Klassifizierung von Daten, die bereits lineare Beziehungen im Ursprungsraum aufwiesen, so dass auch die Vorhersage durch lineare Matrixoperationen gefunden werden konnte. Im Fall der nichtlinearen

Merkmale müssen die Trainingsdaten in einen neuen, mitunter vieldimensionalen Raum  $\phi(X)$  abgebildet werden, um lineare Beziehungen herzustellen:

$$X \rightarrow \phi(X) \quad (63)$$

Damit ändert sich die Schreibweise der Formel 62 zu:

$$\tilde{y} = (\phi(X)\phi(X)^T + \lambda I)^{-1}y \langle \phi(x), \phi(X) \rangle = \alpha \langle \phi(x), \phi(X) \rangle \quad (64)$$

Spätestens an dieser Stelle muss der Begriff der Kernel-Funktion  $\kappa$  eingeführt werden:

$$\kappa(x, X) = \langle \phi(x), \phi(X) \rangle = \kappa(x, X) = \phi(x)^T \phi(X) \quad (65)$$

Im linearen Fall (siehe Gleichung 62) bestand die Kernel-Funktion lediglich aus der Berechnung des Skalarvektorprodukts  $\langle x, X \rangle$  der Eingangsmatrizen. In nicht-linearen Fällen berechnet die Kernel-Funktion das Skalarvektorprodukt der Abbildungen  $\phi(x)$  und  $\phi(X)$  der Ursprungsdaten. Es existiert eine Vielzahl von Kernel-Funktionen, wobei die gebräuchlichsten im Folgenden aufgelistet sind:

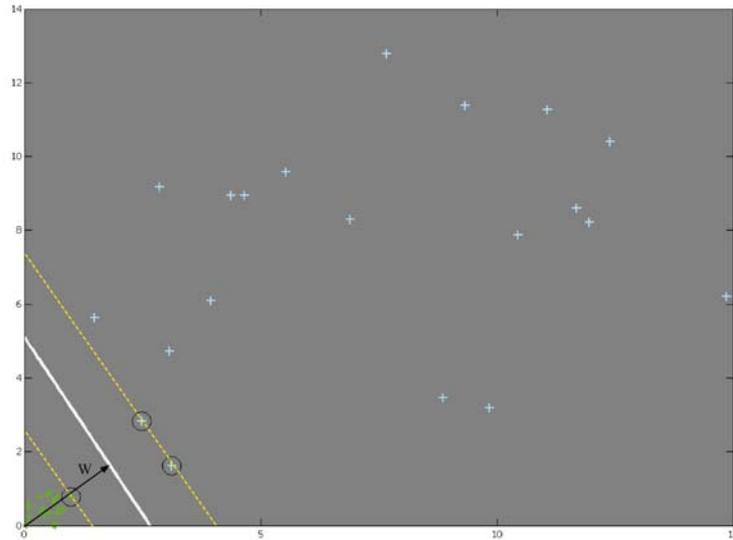
- $\kappa_{poly}(x, X) = (\phi(x)^T \phi(X))^d$  - Polynomial-Kernel mit der Potenz  $d$
- $\kappa_{rb}(x, X) = e^{-\frac{\|\phi(x) - \phi(X)\|^2}{2\sigma^2}}$  - Radial Basis bzw. Gauß-Kernel mit veränderlichem Sigma  $\sigma$
- $\kappa_{sign}(x, X) = \tan^{-1}(a\phi(x)^T \phi(X) + c)$  - Sigmoidaler Kernel mit dem veränderlichen Skalierungsparameter  $a$  und dem Shift-Parameter  $c$

Erweitert man nunmehr die Klassifizierung mit Kernel-Funktionen um den Aspekt, dass die Entscheidungsgrenze (Trennebene) den größtmöglichen Abstand zu den nächstliegende Trainingsmerkmalen beider Klassen besitzen sollen, gelangt man zu der Support-Vector-Machine-Methode.

Dabei ist die Trennebene mit dem Verschiebungsvektor  $b$  so definiert, dass auf ihr liegende Punkte  $X$  folgende Bedingung erfüllen:

$$\langle W, X \rangle + b = 0 \quad (66)$$

Ein Merkmal  $x_i$ , das ein Gleichungsergebnis größer als 0 ergibt, wird der einen Klasse zugeordnet und entsprechend der anderen Klasse, wenn das Ergebnis negativ ist:



**Abbildung 49: Trennebene in einem linearen Merkmalsraum mit linearer Kernel-Funktion und Hard-Margin-Bedingung.** (Quelle: eigene Darstellung)

$$\langle W, x_i \rangle + b > 0 \rightarrow \text{Klasse A} \quad (67)$$

$$\langle W, x_i \rangle + b < 0 \rightarrow \text{Klasse B} \quad (68)$$

Weil in diesem Fall jedoch eine unendliche Anzahl von Trennebenen möglich ist, muss ihre Position durch eine Skalierung folgende Bedingung erfüllen:

$$\min |\langle W, x_i \rangle + b| = 1 \quad (69)$$

Punkte, die auf den Randebenen (gelbe Linien in Abbildung 49) liegen, ergeben dementsprechend einem Betrag von 1. Mit dieser Bedingung lassen sich die Gleichungen 67 und 68 zusammenfassen zu ( $y_i$  ist die zu  $x_i$  zugeordnete Klasse):

$$y_i(\langle W, x_i \rangle + b) - 1 \geq 0 \quad (70)$$

Mit dem Wissen, dass  $|\langle W, X \rangle + b| = 1$  ist, berechnet sich der Abstand  $d$  der Randebenen der in der Mitte liegenden Trennebene zu:

$$d = \frac{1}{2} \|W\| \quad (71)$$

wobei  $\|W\|$  die Norm von  $W$  darstellt und sich durch  $\|W\| = \sqrt{\langle W, W \rangle}$  berechnet.

Und damit ergibt sich schließlich die Optimierungsvorschrift, um den Abstand von Trennebene zu den Randebenen zu maximieren, indem  $W$  minimiert wird:

$$\begin{aligned} &\text{minimiere : } \frac{1}{2} \|W\|^2 \\ &\text{unter Berücksichtigung von : } y_i(\langle W, x_i \rangle + b) - 1 \geq 1 \end{aligned} \quad (72)$$

Diese Optimierungsvorschrift mit der strikten Definition, dass zwischen den Randebenen keine Trainingspunkte liegen dürfen, wird Hard-Margin-Classifer genannt. Die Trainingswerte, die für die Position der Randebenen verantwortlich sind, heißen Stützvektoren bzw. Support-Vectors und sind in Abbildung 49 mit einem Kreis markiert. Die Trennung mit harter Trennschleife ist jedoch bei nicht vollständig trennbaren Klassen (beispielsweise bei Vorhandensein von Ausreißern) unmöglich, so dass es sinnvoll ist, die Regeln für die Randebenen weniger streng zu gestalten. Die Soft-Margin-Optimierung führt dafür eine Schlupfvariable  $\varsigma$  ein:

$$\begin{aligned} &\text{minimiere : } \frac{1}{2} \|W\|^2 + C \sum \varsigma_i \\ &\text{unter Berücksichtigung von : } y_i(\langle W, x_i \rangle + b) \geq 1 - \varsigma_i \end{aligned} \quad (73)$$

Die Variable  $\varsigma$  ist die Differenz aus dem Abstand des Datenpunktes und der entsprechenden Randebene und misst damit den Grad der Fehlklassifikation, die durch die Konstanten  $C$  zudem gewichtet wird. Betragsmäßig große  $C$  führen zu strikteren und enger zusammen liegenden Randebenen.

## 9.4 Veröffentlichungen und Patentanmeldung

### **„Content-based Autofocusing in Automated Microscopy“**

Hamm, P., Schulz, J., Englmeier, K.-H.

*Image Analysis & Stereology 2010; 29: S. 101-108*

### **„Automatisierte quantitative Analyse der Zellzusammensetzung von bronchoalveolaren Spülungen (BAL)“**

Wörmann, J., Braun, A., Mempel, M., Englmeier, K.-H., Hamm, P.

*Bildverarbeitung für die Medizin 2010 - Algorithmen - Systeme - Anwendungen - „Informatik Aktuell“; S.196-200*

### **„Verfahren zum automatischen Fokussieren eines Mikroskops auf ein vorbestimmtes Objekt sowie Mikroskop zum automatischen Fokussieren“**

*International Patent Application – PCT/EP2010/062185*

Inhaber: Helmholtz Zentrum München/Carl Zeiss AG - Erfinder: Hamm, P, Englmeier, K.-H., Waschulzig, Th., Wolff, H. - Veröffentlichung im März 2011

## 9.5 Danksagung

Ich bedanke mich an dieser Stelle in erster Linie bei meinem Doktorvater Prof. Dr. Karl-Hans Englmeier, auf dessen Unterstützung ich mich in jeglicher Hinsicht und zu jedem Zeitpunkt verlassen konnte. Neben ihm danke ich Dr. Thomas Waschulzik, der diese Arbeit in Kooperation mit der Carl Zeiss Microlmaging GmbH in die Wege leitete.

Ein sehr großes Dankeschön spreche ich an Janina Schulz von Carl Zeiss aus. Sie hat mich nicht nur in wissenschaftlicher und technischer Hinsicht äußerst unterstützt, sondern auch durch ihre Rückmeldungen einen motivierenden Einfluss auf meine Arbeit gehabt. Auf Seiten der Carl Zeiss standen mir stets Dr. Horst Wolff, Jasmin Hirche, Johann Kugler und viele mehr hilfreich zur Seite.

Am Helmholtz-Zentrum geht mein Dank für den wissenschaftliche Beitrag, die Hilfestellungen bei Problemlösungen und die großartige Arbeitsatmosphäre an meine geliebten Kollegen Harry Höllig, Thomas Jetzfellner, Markus Freyer, Julian Wörmann, Saskia Björn und Dr. Andreas Murr. Weiterhin möchte ich mich für die wichtige Unterstützung rund um organisatorische Fragen sehr herzlich bei Silvia Weinzierl, Susanne Stern, Richard Kumpfmüller und Ines Baumgartner bedanken.

Für das indirekte Gelingen dieser Arbeit sind aber auch und vor allem meine Eltern Gudrun und Heinz Hamm mitverantwortlich. Ich bin glücklich, dass sie mich - trotz zeitweiser abweichender Meinung - in allen meinen Lebensphasen in jeder Hinsicht unterstützt haben. Weiterhin ist es schön, eine große Schwester zu haben, die mich bewusst und unbewusst in vielen Dingen beeinflusst, und deswegen: vielen Dank Claudia.

Ich möchte weiterhin auf keinen Fall meine besten Freunde unerwähnt lassen. Denn auch Alexander Grimm, Martin Landgraf-Dietz und Marco Steinberg (in alphabetischer Folge), haben und hatten einen Einfluss auf meine Berufswahl und mein Leben.

Der letzte Absatz ist meiner Freundin Andrea gewidmet. Sie hat mir zwar auch mit ihrem Fachwissen in vielen Fragen geholfen und somit einen Anteil an dieser Arbeit, doch vielmehr möchte ich mich für die bisherige Zeit mit ihr bedanken. Und ich weiß, dass die Zukunft, die vor uns liegt, noch viel besser wird...

## Literatur

- [1] C. F. Batten. *Autofocusing and Astigmatism Correction in the Scanning Electron Microscope*. PhD thesis, University of Cambridge, 2000.
- [2] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman. Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19, No. 7:711–720, 1997.
- [3] B. Berge. Lens with variable focus - Pat. WO 99/18456, 2002.
- [4] B. Berge. Liquid lens technology for adaptive optics: principle, physical limitations and applications. In *31th Kogaku Symposium*, 2006.
- [5] W. Böge and W. Plaßmann. *Vieweg Handbuch Elektrotechnik. Grundlagen und Anwendungen für Elektrotechniker*. Vieweg + Teubner, 2007.
- [6] F. R. Boddeke. *Quantitative Fluorescence Microscopy*. PhD thesis, Technische Universiteit Delft, 1999.
- [7] F. R. Boddeke, L. J. van Vliet, H. Netten, and I. T. Young. Autofocusing in microscopy based on the OTF and sampling. *Bioimaging*, 2:193–203, 1994.
- [8] J.F. Brenner, B.S. Dew, J.B. Horton, J.B. King, P.W. Neirath, and W.D. Sellers. An automated microscope for cytologic research. *Journal of Histochemistry and Cytochemistry*, 24:100.111, 1971.
- [9] S. L. Brázdilová and M. Kozubek. Information content analysis in automated microscopy imaging using an adaptive autofocus algorithm for multimodal functions. *Journal of Microscopy*, 236:194–202, 2009.
- [10] M. A. Bueno, J. Álvarez Borrego, and L. Acho. Autofocus algorithm using one-dimensional fourier transform and Pearson correlation. In *SPIE, The International Society for Optical Engineering, 5th Iberoamerican Meeting on Optics*, 2004.
- [11] P. D. Caie, R. E. Walls, A. Ingleston-Orme, S. Daya, T. Houslay, R. Eagle, M. E. Roberts, and N. O. Carragher. High-content phenotypic profiling of drug response signatures across distinct cancer cells. *Molecular Cancer Therapeutics*, 9, Num. 6:1913–1926, 2010.
- [12] Eastman Kodak Company. *Picture taking with the No. 3A Autographic Kodak Special*. Eastman Kodak Company, January 1921.

- [13] V. Della Mea, F. Viel, and C. A. Beltrami. A pixel-based autofocusing technique for digital histologic and cytologic slides. *Computerized Medical Imaging and Graphics*, 29:333–341, 2005.
- [14] R. A. Dunlap. *The golden ratio and Fibonacci numbers*. World Scientific Pub Co Inc, 1998.
- [15] E. C. Fritts. Enlarging camera - Pat. 1.460.714, 1921.
- [16] K. L. Frost, J. W. Hayenga, J. A. Stephanick, and R. C. Schmidt. Automatic focusing of biomedical specimens apparatus - Pat. 5.647.025, 1997.
- [17] H. Gaethje. Neue Technologien und Anwendungen in der Fluoreszenzmikroskopie. *Biospektrum*, 11:688–690, 2005.
- [18] J. Geusebroek, F. Cornelissen, A. Smeulders, and H. Geerts. Robust focusing in microscopy. *Cytometry*, 39:1–9, 2000.
- [19] J. Glaser, G. Greene, and S. Hendricks. *Stereology for biological research with a focus on neuroscience*. MBF Press, 2007.
- [20] K. Graumann and T. Darrell. The pyramid match kernel: Efficient learning with sets of features. *Journal of Machine Learning Research*, 8:725–76, 2007.
- [21] F. Groen, I. T. Young, and G. Ligthart. A comparison of different focus functions for use in autofocus algorithms. *Cytometry*, 12:81–91, 1985.
- [22] F. Hoch and G. Weissenberg. Verfahren und Vorrichtung zur Entfernungsmessung - Pat. DE 1423655, 1973.
- [23] R. Hopkins. Automatically-focusing camera - Pat. 1.623.538, 1922.
- [24] W. Huan and Z. Jing. Evaluation of focus measures in multi-focus image fusion. *Pattern Recognition Letters*, 28:493–500, 2007.
- [25] X. Hui, T. Weibin, and S. Lining. Construction and evaluation of a wavelet-based focus measure for microscopy imaging. *Microscopy Research and Technique*, 70:987–995, 2007.
- [26] Intel-Corporation. Open source computer vision library - <http://opencv.willowgarage.com/wiki/>, 2006.
- [27] B. Jähne, H. Haussecker, and P. Geissler. *Handbook of computer vision and applications - Volume 1: Sensors and imaging*. Academic Press, 1999.

- [28] M. Kern and J. Trempler. *Beobachtende und messende Mikroskopie in der Materialkunde*. Brünne-Verlag Berlin, 2007.
- [29] L. Leitz. Focussing plate for optical instruments, 1960.
- [30] R. Lienhart, A. Kuranov, and V. Pisarevsky. Empirical analysis of detection cascades of boosted classifiers for rapid object detection. Technical report, Intel Corporation, Santa Clara, 2002.
- [31] R. Lienhart and J. Maydt. An extended set of haar-like features for rapid object detection. In *International Conference on In Image Processing*, 2002.
- [32] X. Y. Liu, W. H. Wang, and Y. Sun. Dynamic evaluation of autofocusing for automated microscopic analysis of blood smear and pap smear. *Journal of Microscopy*, 227:15–23, 2007.
- [33] D. G. Lowe. Object recognition from local scale-invariant features. In *International Conference on Computer Vision*, pages 1150–1157, 1999.
- [34] D. G. Lowe. Distictive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60, 2:91–110, 2004.
- [35] V. V. Makkapati. Improved wavelet-based microscope autofocusing for blood smears by using segmentation. In *5th Annual IEEE Conference on Automation Science and Engineering Bangalore, India*, 2009.
- [36] E. McInnes. Artefacts in histopathology. *Comparative Clinical Pathology*, 13:100–108, 2005.
- [37] C. P. Papageorgiou, M. Oren, and T. Poggio. A general framework for object detection. In *Sixth International Conference on Computer Vision*, pages 555–562., 1998.
- [38] W. Patzelt. *Polarized light microscopy - Principles, instruments, applications*. Ernst Leitz Wetzlar GmbH, 1985.
- [39] K. Pearson. On lines and planes of closest fit to systems of points in space. *Philosophical Magazine*, 2:559–572, 1901.
- [40] A. Pentland. A new sense for depth of field. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-9, Issue 4:523–531, 1987.
- [41] J. H. Price and D. A. Gough. Comparison of phase-contrast and fluorescence digital autofocus for scanning microscopy. *Cytometry*, 16:283–297, 1994.

- [42] G. Reinheimer and L. Leitz. Automatic microscope focussing device - Pat. 3.798.449, 1974.
- [43] P. M. Roth and M. Winter. Survey of appearance-based methods for object recognition. Technical report, Institut for Computer Graphics and Vison, Graz Univ. of Technology, 2008.
- [44] A. Santos, C. Ortiz de Solórzano, J. J. Vaquero, J. M. Pena, N. Malpica, and F. Del Pozo. Evaluation of autofocus functions in molecular cytogenetic analysis. *Microscopy*, 188:264–272, 1997.
- [45] R. E. Schapire and Y Singer. Improved boosting algorithms using coincidence-rated prediction. *Machine Learning*, 37(3):297–336, 1999.
- [46] W. Schlipf and J. Fischer. Verfahren und Vorrichtung zum automatischen Fokussieren auf Objekte - Pat. DE 4226523 B4, 2005.
- [47] F. Shen, L. Hodgson, and K. Hahn. Digital autofocus methods for automated microscopy. *Methods in Enzymology*, 414:620–632, 2006.
- [48] N. L. Stauffer. Auto focus camera - Pat. 3.860.935, 1975.
- [49] N. L. Stauffer. Focus detecting apparatus - Pat. 3.875.401, 1975.
- [50] Y. Sun, S. Duthaler, and B. J. Nelson. Autofocusing in computer microscopy: Selecting the optimal focus algorithm. *Microscopy Research and Technique*, 65:139–149, 2004.
- [51] J. M. Tenenbaum. *Accomodation in computer vision*. PhD thesis, Stanford University, USA, 1970.
- [52] P. Viola and M. Jones. Rapid object detection using boosted cascade of simple features. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2001.
- [53] D. Vollath. Automatic focusing by correlative methods. *Microscopy*, 151:279–288, 1987.
- [54] B. Xi, N. Yu, X. Wang, X. Xu, and Y. A. Abassi. The application of cell-based label-free technology in drug discovery. *Biotechnology Journal*, 3:484–495, 2008.
- [55] G. Yang and B.J. Nelson. Wavelet-based auto-focusing and unsupervised segmentation of microscopic images. In *IEEE/RSJ International Conference on Intelligent Robots and Systems. Proceedings.*, 2003.

## Index

- Fokuskurve, 27
- 1/f-Rauschen, 30
  
- Abbesche Auflösungsgrenze, 40
- AdaBoost, 51, 52, 56
- Airy-Beugungsringen, 28
- Autokorrelation, 39, 40, 54
- Autokorrelationsfunktion, 34
  
- Backlash, 29, 42
- Binäre Suche, 41
- Boosting, 51
- Brenner-Gradient, 38, 53
  
- Curve Fitting, 43
  
- Dead Pixel, 30
- Deviation-based Correlation, 40
- Difference-Of-Gaussian, 94
- Discrete AdaBoost, 60
- Drift, 30
- Dual-Lösung, 98
  
- Eigensystem, 92
- Eigenvektor, 92
  
- Fibonacci-Folge, 41
- Full Well Capacity, 30
  
- Gauß-Kernel, 99
- Gaussian Derivatives, 38, 57
- Gentle AdaBoost, 60, 70
- Golden-Section-Search, 41
- Goldener Schnitt, 41
- Gradientenhistogramm, 95
- Gram-Matrix, 98
  
- Haar-Filter, 48, 50, 56
  
- Halo-Artefakt, 31
- Hard-Margin-Klassifizierung, 100
- Hauptkomponentenanalyse, 92
- Hill-Climbing-Strategie, 40
- Hot Pixel, 30
  
- Inner Product Matrix, 97
- Integralbild, 50
- Interferometrie, 23
- Interpolation, 42
  
- Kernel-Funktion, 96, 99
- Kernel-Matrix, 98
- Keypoint (SIFT), 93
- Korrelationskoeffizient, 36
- Kreuz-Leistungsspektrum, 26
  
- Laplace-Of-Gaussian, 39
- Laplace-Operator, 39
- Laplacian-Of-Gaussian, 94
- Least-Square-Minimum, 97
- LogitBoost, 60
  
- Method of Least Squares, 43
- Methode der kleinsten Fehlerquadrate, 43
  
- Nyquist-Frequenz, 36
  
- Oktave (SIFT), 94
- Optical-Transfer-Function, 28, 36
  
- PCA, 56
- Phasenkontrast, 31, 89
- Phasenkorrelation, 25
- Phasenvergleich, 25
- Phasenverschiebung, 23
- Photonenrauschen, 30
- Point-Spread-Function, 27

Polynomial-Kernel, 99  
Primallösung, 97  
Principle Components Analysis, 92  
  
Quantenausbeute, 30  
  
Radial-Basis-Kernel, 99  
Real AdaBoost, 60  
Region-Of-Interest, 45  
  
Scale Invariant Feature Transformation, 48  
Schärfemaß, 35  
Schlupfvariable, 101  
Schrotrauschen, 30  
Sensitivität, 62  
Sensor-Rauschen, 30  
Shade-Off-Artefakt, 31  
Sigmoidal-Kernel, 99  
Skalarprodukt, 97  
Skalenraum, 93, 94  
Sobel-Operator, 28, 38  
Soft-Margin-Klassifizierung, 101  
Spezifität, 62  
Stuckpixel, 30  
Support Vector Machine, 56, 92, 96  
  
Teile-und-Herrsche-Suche, 41  
Tenengrad-Gradient, 38, 54  
Thermisches Rauschen, 30  
Triangulation, 23  
  
Unimodalität, 34  
  
Wavelet-Dekomposition, 35  
Wavelet-Transformation, 48  
weak classifier, 52  
Wiener-Khintchine-Kriterium, 34