TAMPEREEN TEKNILLINEN YLIOPISTO
TAMPERE UNIVERSITY OF TECHNOLOGY
Julkaisu 786 • Publication 786

Jyrki Selinummi

On Algorithms for Two and Three Dimensional High Throughput Light Microscopy

Tampere 2008

Tampereen teknillinen yliopisto. Julkaisu 786
Tampere University of Technology. Publication 786

Jyrki Selinummi

On Algorithms for Two and Three Dimensional High Throughput Light Microscopy

Thesis for the degree of Doctor of Technology to be presented with due permission for public examination and criticism in Tietotalo Building, Auditorium TB109, at Tampere University of Technology, on the 19th of December 2008, at 12 noon.

Tampereen teknillinen yliopisto - Tampere University of Technology
Tampere 2008

ISBN 978-952-15-2087-7 (printed)
ISBN 978-952-15-2192-8 (PDF)
ISSN 1459-2045

Abstract

In biomedical research, it is often necessary to study cell population characteristics, and quantify changes in cell phenotypes on a cell-by-cell basis. Traditionally, this work has been performed by interactive manual use of a microscope. In disciplines like systems biology, studying topologies and dynamics of complex functional networks of cells, massive systematical screens for phenotypic changes in cell populations are required. Also in drug discovery, effects of pharmacological agents on the populations must be tested automatically in a high throughput fashion.

The development of robotic arrayers and automated microscopes, together with increasing computing power and storage space have enabled the automated screening of cell populations, resulting in a revolution of microscopy imaging. Currently, imaging of hundreds of populations in parallel is common practice in a single experiment. During the screen, images of each of the cell populations are stored for subsequent analysis. The amount of image data renders manual visual analysis impossible, requiring automated image analysis systems, and software.

Current procedures of automated analysis in high throughput microscopy, however, have several drawbacks. Standard practices exist for a number of analysis approaches, but especially three dimensional studies are generally performed manually, or semi-automatically. Furthermore, greater care must be taken on practical issues, such as low computational cost and easy implementation to advance routine high throughput screening studies by bioscientist. This thesis considers fully automated methods ranging from cell enumeration, to subcellular analysis in two and three dimensions, concentrating on the applicability of the algorithms for high throughput microscopy.

Acknowledgments

The research for this thesis was carried out under the supervision of Professor Olli Yli-Harja and Lecturer Heikki Huttunen at the Department of Signal Processing of Faculty of Computing and Electrical Engineering at Tampere University of Technology during the years 2005–2008. Part of the work was also performed while at the Institute for Systems Biology, Seattle, USA, where my supervisor was Professor Ilya Shmulevich.

First of all, I want to thank Professor Olli Yli-Harja for the possibility to work in the Computational Systems Biology research group at TUT, as well as for his vision, encouragement, and support. Lecturer Heikki Huttunen I want to thank for excellent guidance and advice. My gratitude is also due to Professor Ilya Shmulevich, for great inspiration and motivation, as well as for introducing me to the interdisciplinary research environment of the Institute for Systems Biology in Seattle.

I wish to thank all my co-authors for inspiring conversations and for creating a motivating work environment. Especially I wish to mention the support by Pekka Ruusuvuori, M.Sc., Antti Lehmussola, M.Sc., Antti Niemistö, Dr.Tech., Jenni Seppälä, M.Sc., and Marja-Leena Linne, Dr.Tech. Interdisciplinary studies can only be enabled by productive co-operation with researchers from other fields of science, with background in medicine and biology. Thank you Riina Sarkanen, M.Sc., Tuula Jalonen, Ph.D., and Riitta Miettinen, Ph.D.

The financial support of Graduate School of the Tampere University of Technology, Tampere Graduate School in Information Science and Engineering (TISE), Emil Aaltosen säätiö, and Tampereen teknillisen yliopiston tukisäätiö is gratefully acknowledged.

Finally, I wish to give my greatest gratitude to my wife Ansku and my daughter Senni, as well as to my parents and friends, for always being there supporting me!

Tampere, December 2008

Jyrki Selinummi

Supervisor

Professor Olli Yli-Harja
Department of Signal Processing
Tampere University of Technology

Instructor

Lecturer Heikki Huttunen, PhD
Department of Signal Processing
Tampere University of Technology

Pre-examiners

Assistant Professor Carolina Wählby
Centre for Image Analysis
Uppsala University

Lecturer Arto Kaarna, PhD
Department of Information Technology
Lappeenranta University of Technology

Opponent

Professor Jussi Parkkinen
Department of Computer Science and Statistics
University of Joensuu

Contents

List of Publications

I J. Selinummi, J. Seppälä, O. Yli-Harja, and J. A. Puhakka, "Software for quantification of labeled bacteria from digital microscope images by automated image analysis," *BioTechniques*, vol. 39, no. 6, pp. 859–863, December 2005.

II J. Selinummi, R. Sarkanen, A. Niemistö, M.-L. Linne, T. Ylikomi, O. Yli-Harja, and T. Jalonen, "Quantification of vesicles in differentiating human SH-SY5Y neuroblastoma cells by automated image analysis," *Neuroscience Letters*, vol. 396, no. 2, pp. 102–107, March 2006.

III A. Niemistö, J. Selinummi, R. Saleem, I. Shmulevich, J. Aitchison, and O. Yli-Harja, "Extraction of the number of peroxisomes in yeast cells by automated image analysis," in *Proceedings of the 28th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC'06)*, New York, USA, August 30 – September 3, 2006, pp. 2353–2356.

IV J. Selinummi, A. Lehmussola, J.-R. Sarkanen, J. Nykky, T. O. Jalonen, and O. Yli-Harja, "Automated analysis of Golgi Apparatus dispersion in neuronal cell images," in *Proceedings of the 4th TICSP Workshop on Computational Systems Biology (WCSB'06)*, Tampere, Finland, June 12–13, 2006, pp. 89–92.

V J. Selinummi, A. Niemistö, R. Saleem, C. W. Carter, J. Aitchison, O. Yli-Harja, I. Shmulevich, and J. Boyle, "A case study on 3-D reconstruction and shape description of peroxisomes in yeast," in *Proceedings of the IEEE International Conference on Signal Processing and Communications (ICSPC'07)*, Dubai, United Arab Emirates, November 24–27, 2007, pp. 672–675.

VI J. Selinummi, P. Ruusuvuori, A. Lehmussola, H. Huttunen, O. Yli-Harja, and R. Miettinen, "Three-dimensional digital image analysis of immunostained neurons in thick tissue sections," in *Proceedings of the 28th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC'06)*, New York, USA, August 30 – September 3, 2006, pp. 4783–4786.

The author of this thesis contributed in the Publications as follows:

Each Publication is interdisciplinary in nature. In most of the Publications, J. Selinummi has been responsible for the general structure and writing, and for the design and implementations of the image analysis and processing methods. The biological experiments, as well as writing of the details of cell biological methods, were conducted by the co-authors in all the Publications.

In Publications I, II and V, J. Selinummi planned and implemented all the image processing methods, and was the main responsible for writing the manuscripts. As the second author of Publication III, J. Selinummi designed and implemented the overlapping object separation and participated in writing of the image analysis methods section. J. Selinummi designed the algorithms and was responsible for the structure and writing of Publication IV, except for the rank filtering approach in nuclei segmentation. Publication VI was jointly authored by J. Selinummi and P. Ruusuvuori, where J. Selinummi was solely responsible for the algorithms and writing related to the analysis of cell bodies.

The Publications are arranged starting from whole cell quantification in Publication I, to subcellular structures in two dimensions from Publications II to IV, and ending in three dimensional analysis in Publications V and VI. The Publication III is also a part of the Ph.D. thesis of A. Niemistö [69].

Symbols and Abbreviations

2-D Two dimensional
3-D Three dimensional
a Coefficient of a polynome
b Image background; Structuring element
c Cluster center
C Covariance matrix
cDNA Complementary DNA
CT Computed tomography
d Distance
D Domain of a structuring element
DAPI 4',6-diamidino-2-phenylindole
DIC Differential interference contrast
DNA Deoxyribonucleic acid
e Error
f Ideal image
FCM Fuzzy clustering
GA Golgi apparatus
GFP Green fluorescent protein
$\mathscr{H}$ Pattern set
h Point spread function
k Number of clusters or classes
λ Eigenvalue
l Class label
L Clustering; Maximum intensity of an image
μ_T Mean intensity of an image
μ First-order cumulative moment
m Image recorded by a fluorescent microscope
n Number of feature vectors or patterns; Number of pixels

N Noise function
ω Zeroth-order cumulative moment
o Combination of rank order filters
p Probability of an intensity
P Percentile
PCA Principal component analysis
PSF Point spread function
RNAi Ribonucleic acid Interference
σ_B^2 Between class variance
t Threshold level
t^* Optimal threshold
u Membership value
$\mathbf{U}$ Membership matrix
$\mathbf{v}$ Voxel coordinate vector
$\mathbf{w}$ Eigenvector
W Pixel neighborhood
$\mathbf{x}$ Feature vector
ζ Rank order morphological filter

Chapter 1

Introduction

The more complex the studied system, the more throughput is required from measurement techniques. This is also true in the field of systems biology, where the cell is studied as a system consisting of simple components forming an extensive network, working together resulting in complex behavior. Only by understanding the structure, dynamics, and methods for control and design, can we completely describe the system [48], enabling, for example, predictive and preventative medicine [40].

In modern cell biology, measurements such as screen expression studies [88], have been automated transforming manual instruments into high throughput systems, where hundreds or thousands of experiments are performed in parallel. In microscopy, similar transformation is converting manual microscopes into high throughput automated platforms. Especially in drug screening [35, 105] and in studies of gene knock-downs [68], robotic microscopes are routinely imaging hundreds of cell populations without user intervention. After the imaging, changes in the populations must be quantitatively measured. In such high throughput microscopy systems, automated image analysis plays a critical role, overcoming the main disadvantages of manual methods: firstly, manual analysis of cell images is subjective and error prone, inducing variance in results between researchers and between analysis instances [122, 16]. Secondly, the analysis of image data from high throughput experiments is overly laborious especially in cell-by-cell based studies, and thirdly, manual analysis generally does not yield quantitative data. Moreover, although the automated image analysis does not guarantee perfect results, the analysis is repeatable and the errors are therefore largely systematic.

The enormous diversity of cellular organisms, however, sets considerable challenges for the automated image analysis systems [132, 76], resulting in studies where manual image analysis is also applied in high throughput measurements [19]. There is great need for further algorithm development in digital image analysis, not only to enable wider range of cell image quantification applications to be automatized, but also to aim at standardized methods, enabling more practical and straightforward implementations. For example, the quantification of subcellular structures in three dimensions has not been widely applied in high throughput fashion, because of expensive equip-

ment, but also because of the challenges in image analysis. Our hypothesis is that the traditional visual measurements of cells and cell population properties can be replaced by automated image processing algorithms with suitable design, precision matching the manual procedures. In addition to cell counting, this also applies to a variety of different cellular measurements including colocalization, quantification of subcellular structures, and three dimensional feature description. The automated approach will not only remove the need for tedious manual analysis work, but also enable biologists to measure cellular features not feasible by the standard manual techniques.

This study presents fully automated digital image analysis procedures for two and three dimensional image analysis, solving a number of challenges risen in the aforementioned areas of visual analysis of cell populations. Wherever applicable, we combined well known, previously introduced methods into pipelines, also adopting algorithms not previously utilized in microscopy of cells, into the domain of high throughput microscopy. The focus being on automated image analysis, we have left certain aspects of high throughput microscopy, such as the control of the imaging device and the microscope stage, outside of this thesis. To enable result validation, the features extracted from the cells were selected to resemble the attributes a manual analyst would aim at detecting, or if the main objective has been to present a high throughput technique, a proof of principle type of approach has been given to assure the reader of the validity and applicability of the algorithm.

Publication III for this thesis was written in collaboration with the Shmulevich Laboratory, lead by Prof. Ilya Shmulevich at the Institute for Systems Biology (ISB), Seattle, USA. Publication V was composed while at the ISB as a visiting researcher. All other Publications were written while working in the Computational Systems Biology research group led by Prof. Olli Yli-Harja at the Department of Signal Processing, Tampere University of Technology. The aim has been in writing the thesis for the interdisciplinary audience of computer scientists and cell biologists, omitting jargon and including clarifications, with remarks on practical aspects of high throughput microscopy.

In Chapter 2, we briefly introduce the biological background and discuss different imaging modalities for two and three dimensional microscopy. Next in Chapter 3, we present automated image analysis procedures for microscopy of cells, concentrating on the methods utilized in the Publications. Chapter 4 discusses the main challenges and future possibilities in the field of high throughput microscopy. Finally, Chapter 5 summarizes the results presented in the Publications.

Chapter 2

Light Microscopy in Cell Biology

In biomedical research, there is a continuing trend towards more automated methods, enabling high throughput studies. The objectives in systems biology [48], as described in Chapter 1, and in *cytomics* [109], where molecular disease networks are recognized through cell phenotype studies, can only be reached by large scale exhaustive approaches integrating different data sources. *Cytometry*, the study of features and morphologies of cells, is one such source.

Traditionally, cytometry has been divided into two categories, namely *flow and image cytometry*. In flow cytometry [72, 95], a beam of light is focused into a stream of cells, and the properties of the cells are derived from scattering characteristics of the light. In image cytometry, on the other hand, the cell populations are imaged, and the phenotype data is acquired through manual or automated image analysis. The focus of this work is on image cytometry: starting from standard microscopes, different technical improvements have enabled measurements of cell properties and cellular dynamics in subcellular level, in three dimensions, and over time. Also in imaging, the measurements are constantly evolving in the direction of high throughput microscopy and high throughput screening [77], where thousands of cell populations can be imaged automatically in a sequential fashion. Naturally, the increased amount of image data has required automated image analysis techniques [132], combining biomedical research with digital signal and image processing.

In addition to light microscopy, other imaging modalities such as electron microscopy [25] and atomic force microscopy [38], have been successfully applied in cell biology, especially in structural determination of specimens and molecules respectively. These techniques are not currently feasible in high throughput studies, and therefore, are outside the scope of this thesis.

2.1 Biological Samples

In the context of cell and tissue biology, samples imaged with light microscopes can range anything from tissue, to populations of cells, to single cells, reaching subcellu-

lar structures and molecular interactions. The diversity of sample types utilized in the Publications I-VI prevents giving extensive review on each, but we briefly introduce the different sample types. While the samples differ greatly from publication to publication, the analysis methodology stays relatively constant in the context of this thesis: regardless of the exact biological aim of the study, a number of images are taken using a microscope system, requiring automated interpretation and analysis. In Chapter 3 we further show that from the image processing point of view, the algorithms used in automated analysis of different sample types share several similarities.

Starting from simpler organisms, studies of bacterial cells represent an application where the growth of different cell populations has to be monitored, studying the proportions of specific bacteria types in numerous applications utilizing different staining methods (Section 2.2.1) [47, 29]. Yeast *Saccharomyces cerevisiae* cells, on the other hand, being easy to incubate and study, are common model organisms for more complex cell types, the importance further emphasized after completing the whole genome sequence [8]. The DNA sequence of yeast combined with high throughput measurement techniques have enabled systematic analysis of, for example, gene interaction networks and functional pathways in the field of systems biology [40, 41], aiming at generalizing the results for more complex organisms. Proceeding into mammalian cells, *neuroblastoma* cells represent an example of human eucaryotes, where cancerous cells that are again easy to grow and study, can be used to model normal, healthy cell populations through differentiation. The neuroblastoma cells have been shown to share properties with normal neuronal cells after the differentiation [24, 87], forming network like structures and showing signaling activity. Continuing from cell populations into tissue biology, thick tissue samples utilized by us serve as an example of routine pathology, where tissue samples are screened for malignant tissue by medical doctors. The complexity of tissue, with interactions not only between cells of the same population, but also with the surrounding environment can mask details of intracellular activity, but on the other hand more closely resemble the cells' behavior in a larger context.

2.2 Imaging

In a standard microscope, a lens system produces a magnified view of an illuminated specimen, the schematics and principle presented, for example, in [69]. In *bright field* microscopy [18], however, cells and cell populations appear relatively transparent and colorless, with low contrast. Thus, special microscopy configurations must often be applied in the imaging procedure, or the contrast between the areas of interest (foreground) and the background must be increased by *labeling* with dyes or fluorescent proteins as described in the next Section.

Without labeling, the most common contrast enhancement methods are *phase contrast* microscopy and *differential interference contrast* (DIC) microscopy [101], where differences of the refractive index of the target and the surrounding medium cause a

phase shift in the illumination light relative to the thickness and properties of the target [82]. This shift can then be seen as intensity differences through the microscope eyepiece. More advanced contrast enhancement techniques have received rather low interest in practical applications, mainly because of the need of expensive special equipment [101].

In the next Sections, the most important labeling approaches are highlighted, followed by an introduction to three dimensional microscopy and to high throughput systems.

2.2.1 Labeling

As noted in the previous Section, the contrast of rather transparent cell samples can be enhanced using labeling, a technique in which structures or molecules of interest are highlighted with a dye. Examples including DNA staining with 4',6-*diamidino*-2-*phenylindole* (DAPI), *fluorescent in situ hybridization* where the stain is hybridized into a specific DNA strand, *immunostaining* with a specific antibody tagged fluorochrome, and fluorescent proteins [18]. While visible light microscopy combined with staining with permanent dyes is standard practice in tissue processing, in the context of this thesis, techniques of fluorescent labeling [28] and fluorescence microscopy are of more interest.

In fluorescence, a molecule embodying the fluorescent property absorbs a photon of short wavelength, transferring the system into an exited state, and subsequently releases this state, emitting a photon of longer wavelength. This difference in wavelengths, *stokes shift*, enables the target to be illuminated with high energy excitation light and visualized with emission light of lower energy. Suitable optical filters allow only the emission light to pass through to the eyepiece. Although, in practice, the non-stained areas also express some fluorescence in the form of *autofluorescence* [18], this technique greatly increases the contrast between fluorescence labeled and non-labeled structures or molecules, as areas without the labeling appear essentially black.

In addition to contrast enhancement, the specificity of different dyes and labeling techniques enables the fluorescent molecules very specifically to tag and highlight only at the structures of interest [28]. At first, it was possible to attach small organic dyes by immunostaining to specific proteins inside fixed cells, marking certain subcellular structures impossible to be distinguished without appropriate labeling. Later, fluorophores were designed to directly localize molecules and organelles of interest inside living cells. Today, the *green fluorescent protein* (GFP) originally extracted from jellyfish *Aequorea victoria*, along with it's derivatives [128], have dramatically increased the applications of fluorescence microscopy. The DNA sequence encoding GFP can be inserted into the target cells genome, resulting in fluorescence after the specific gene tagged with the GFP has been transcribed and translated into a protein. Due to the relatively small size of the GFP protein, it is generally not hindering the normal behavior of the tagged protein [14]. Some applications of the GFP include indication

of gene expression and protein dynamics inside living cells [28, 14]. Especially *fluorescence recovery after photobleaching* [98], *fluorescence resonance energy transfer*, and *fluorescence lifetime imaging* [121] have enabled studies of protein dynamics and interactions [3], utilizing GFP and other labeling methods in a nanometer scale.

The main challenges in fluorescent microscopy originate from the toxicity of the excitation light (*phototoxisity*), and the bleaching of fluorophores during exitation, forming free radicals damaging the cells (*photobleaching*) [101]. Fortunately, the development of nontoxic labeling methods such as the GFP, the increasing sensitivity of detectors, and locally adjusted lighting methods are rapidly advancing live cell fluorescence imaging [39]. Furthermore, advanced labeling dyes, such as *semiconductor quantum dots* [65, 101] increase staining efficiency, resulting in shorter excitation periods, longer lifetime, larger variety of spectral properties, and extensive brightness.

2.2.2 3-D Microscopy

Three dimensional (3-D) techniques, such as *magnetic resonance imaging* and *positron emission tomography*, are common practise in medical imaging. In cell biology, 3-D microscope systems have resulted in resolution increase, and enabled more detailed structural studies of cells. In standard two dimensional (2-D) microscopy, all the changes in illumination by the whole specimen are displayed on the microscope eyepiece. 3-D microscopy, in contrast, aims at suppressing the light emerging from *z-planes* out of focus. That is, with a very shallow depth of field, only the specific z-plane in focus is made visible.

According to [110] the image m recorded by a fluorescent microscope can be modeled as:

$$m\left(x,y,z\right) = N\left(h\left(x,y,z\right) \otimes f\left(x,y,z\right) + b\left(x,y,z\right)\right) \tag{2.1}$$

where $h\left(x,y,z\right)$ is the *point spread function* (PSF) of the microscope causing blur, $f\left(x,y,z\right)$ is the ideal image signal, $b\left(x,y,z\right)$ the background, $\otimes$ is the convolution operator, and $N(\cdot)$ is a function defining the noise characteristics, for example location invariant additive noise. In this Section we concentrate of the PSF, and its effects on 3-D imaging, analysis of the noise and background can be found in Chapter 3.

As the term implies, the PSF defines the shape of spread a single illuminated point renders when passing through an imaging system, such as the microscope. Assuming we only have a minuscule fluorescent ball, or a bead we are imaging, the shape of the point spread function is what is actually seen through the microscope eyepiece. Therefore, by reducing the size and power of the PSF in z-dimension, the out of focus blur from other z-planes is suppressed. This is referred as *optical sectioning* because information is acquired from separate z-planes without physical sectioning of the specimen.

The *confocal microscope* [75] is an imaging setup inherently suppressing out of focus light in accordance with the smaller PSF compared to 2-D systems [15]. In a standard confocal microscope, this is achieved by a two pinhole setting [75], only illumi-

nating (and measuring) the fluorescence of one small volume (cube) inside the sample, blocking most of the light emerging from outside this volume. Scanning through the whole sample results in $m(x, y, z)$ of Equation 2.1, that is, the intensity values throughout the target object. Lateral resolution increases accordingly, due to the smaller PSF. In practice, when taking the images, the microscope user interface includes zooming functions increasing the magnification, but the PSF increases simultaneously, resulting in no increase in the real resolution, but just a magnified view of the target.

Diminishing the PSF further has resulted in *multiphoton* and *4Pi microscopy* [21]. In multiphoton microscopy, the fluorophores are excited in a localized fashion using a high intensity pulsed laser with such a long wavelength (low energy) that is not able to excite the target molecule into a higher energy level, unless two low energy photons hit the fluorochrome exactly simultaneously, causing the combined energy of the two photons to be absorbed exciting the fluorescent molecule. The ability to use lower energy excitation enables deeper penetration into samples, and removes effects of toxic photobleaching outside the focal plane. The 4Pi technique [37] relies on special illumination and imaging setup of two objectives simultaneously. By setting the phases of excitation lights in a suitable fashion, and combining the emission into one detector, the fluorescence emission can be localized into an area much smaller than possible in conventional confocal equipment through interference [34]. Though these more advanced techniques have proven increase in resolution, the cost and rarity of equipment delays implementation.

Without special optical hardware, the blurring effect of the PSF can be diminished with *deconvolution* [86]. First, in non-blind deconvolution, the PSF is acquired from measurements or derived mathematically [61, 27]. Second, the convolution operation $h(x, y, z) \otimes f(x, y, z)$ in Equation 2.1 can be reversed to some extent with filtering [113]. Successful deconvolution in practice, however, is a result of numerous processing steps. Different imaging errors such as nonuniform illumination of the target specimen lead to artifacts, and although these errors can be diminished with preprocessing, a variety of defects may arise during the process [64]. Furthermore, the PSF is often assumed spatially invariant because of computational complexity, resulting in nonoptimal deconvolution. Finally, the result is greatly dependent on the choice of the filtering algorithm [113, 64], characteristics of the image affecting the performance of different approaches. Despite the disadvantages, deconvolution has been shown to result in improved measurement accuracy both with simulated and real data [86].

2.2.3 High Throughput Microscopy

As introduced in Chapter 1, the more complex the studied system, the more throughput is in general needed from the measurement instruments. Starting from *complementary DNA* (cDNA) microarrays [88], cell biological measurement systems have evolved towards massive parallelism. In microscopy, although the first automated cell analysis systems with motorized microscopy date back to the 70's [81], only during the past

years there has been a swift move from manual studies into five dimensional [4] high throughput screening [11, 105], that is, three dimensional microscopy combined with time lapse imaging of different wavelengths (colors). This process has been enabled by digital imaging sensors, by advances in automation (robotic microscope systems), and by the increase in computing power and storage space.

There are numerous example applications for high throughput microscopy, utilizing the basic microscopy schemes presented in the previous Sections. In *well plate* experiments, populations of cells are incubated in a regular grid of tiny wells. Each of the populations can be studied in a different fashion, for example mammalian cells having different genes "shut down" with *ribonucleic acid interference* (RNAi) [45] in the individual wells. In *cell arrays* [124, 134], a monolayer of cells is grown on a glass slide with printed target molecules, such as the aforementioned RNAi, also resulting in an array of cell population spots. This transition into high throughput imaging can also be seen in tissue processing, started by the introduction of the *tissue array* [50], where one single experiment is composed of hundreds of small tissue samples with different staining.

After incubation, the populations are imaged for changes in cell phenotype. Automated microscopes or robotic arrayers can be used in performing the whole chain from imaging to data storage without user intervention. The images of spot like cell populations are stored in a database system, and analyzed with the methods of digital image processing, enabling different features to be extracted from the populations, as presented in Chapter 3. The requirements for the processing chain are naturally very different from the established methods of manual microscopy. Therefore, to fully utilize the potential of these high throughput screens, the whole pipeline from incubation to data and setting storage, and result interpretation must be rigorously standardized [104], as discussed further in Chapter 4.

Chapter 3

Digital Image Analysis in Microscopy of Cells

Digital image analysis aims at automating procedures traditionally requiring manual image interpretation, enabling high throughput and perfect repeatability. In the context of light microscopy of cells, methods of signal and image processing are applied to extract information from biological cells and populations (*digital image cytometry* [119]). Despite the long history of digital image cytometry, the number and diversity of different microscopy measurement techniques challenges the robustness of any automated image analysis method [115]; no single algorithm is capable of adjusting to the vast variety of microscopy images [133], and automated image analysis has been described as "one of the greatest remaining challenges in screening" [13]. Simultaneously, from the viewpoint of biomedical research, high throughput cell measurement techniques are becoming commonplace, increasing the need, and continuously setting new requirements for automated analysis.

The image analysis procedure can be partitioned in numerous different ways. Here, we adopt the following, also representing the organization of this Chapter (from [31], modified by the author):

1. Image acquisition and restoration
2. Segmentation
3. Visualization
4. Feature extraction and measurements

This Chapter introduces the specific algorithms implemented in the Publications with descriptive examples, but the results, comparisons to manual analysis, and more extensive case studies can be found in the Publications. As discussed in Chapter 1, we mainly concentrate in quantification of subcellular structures with methods of relatively low computational cost, to enable practical implementation in high throughput microscopy applications. Although the solutions provided in the Publications are here

presented as application specific, they will most likely be applicable in a wider spectrum of image analysis problems.

3.1 Preprocessing

Images taken with microscopes always contain defects of various types [120], caused for example by a low quality imaging device or by theoretical limitations from optics and quantum effects. Because of the long history of manual microscopy, the origin of some imperfections are also due to the different requirements for image quality in manual analysis compared to the requirements in automated image processing. As an example, traditionally, when taking an image for manual analysis, the aim in general is to produce results that are as easy to interpret and as visually appealing as possible, while consistent imaging with fixed settings from image to image is critical in automated analysis to enable quantitative and repeatable results. The adaptation of our visual system [117] can also lead to images of varying quality if the result images are observed from automated image analysis point of view.

As presented in the Equation 2.1, the noise and imperfections in the recorded image can be modeled with convolution and addition operations to the underlying optimal signal. Since we were able to validate the segmentation schemes in all the Publications without deconvolution, the effect of $h\left(x,y,z\right)$ is assumed negligible, reducing the Equation 2.1 into

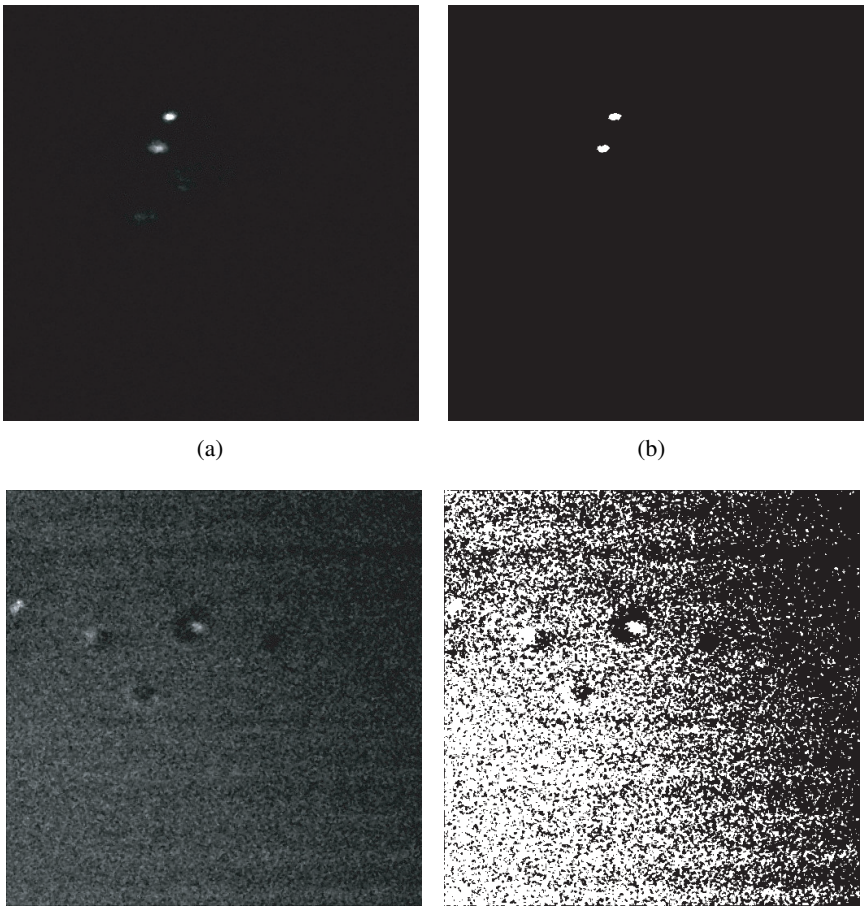
$$m\left(x,y,z\right) = N\left(f\left(x,y,z\right) + b\left(x,y,z\right)\right) \tag{3.1}$$

To obtain the original signal from the distorted output, some noise reduction scheme must be applied to diminish the effect of $N(\cdot)$, and second, the background component $b\left(x,y,z\right)$ must be estimated and removed.

Due to the error sources described above, and largely non standard high throughput microscopy setups, each batch of images is unique, differing between laboratories, equipment, and individuals performing the imaging. In all the Publications, independently designed preprocessing steps were required in order to obtain adequate estimations of $f\left(x,y,z\right)$ for segmentation, by methods described in the rest of this Section. In short, we introduce a quality control scheme for discarding images of unacceptable quality, implement a mathematical morphology based noise suppression algorithm, and reduce the effect of nonuniform background by polynomial fitting.

3.1.1 Noise Suppression

In Publication V we found that in practical applications, manual imaging may result in images with a non-acceptable signal to noise ratio due to human error in exposure control, see Figure 3.1(c) for an example. Automated imaging can also introduce similar images corrupted by noise because a batch imaging device adjusts exposure and gain settings maximizing contrast and subjective image quality. If no objects can be found



(a) (b)

(c) (d)

Figure 3.1: Image quality control by thresholding, see Publication V for details of the input data. (a) High contrast fluorescent microscopy image with two clearly visible subcellular objects. (b) Result of automated thresholding. (c) Noisy low contrast image (scaled between $[0, 1]$ for illustration purposes). (d) Thresholding of (c) displays erroneously increased number of foreground pixels, and therefore results in discarding the image.

from the field of view of the device, no signal can be detected even with high gain, yielding empty images with high level of noise. We implemented a quality control scheme for detection of these erroneous images.

The procedure is illustrated in Figure 3.1. First we *threshold* (see Section 3.2 for an extended definition of thresholding) the input images using a histogram based *Otsu's* thresholding method [74] dividing the pixel intensities into two classes, namely foreground and background. Variation between the two classes can be given as:

$$\sigma_B^2(t) = \frac{[\mu_T \omega(t) - \mu(t)]^2}{\omega(t)[1 - \omega(t)]}, \tag{3.2}$$

where μ_T is the mean intensity of the image, and $\omega(t) = \sum_{i=1}^{t} p_i$ and $\mu(t) = \sum_{i=1}^{t} i p_i$ are the zeroth- and the first-order cumulative moments of the image histogram up to the t^{th} intensity level, and p_i is the probability of intensity i in the image. By maximizing the between-class variance $\sigma_B^2(t)$ of the two classes (and minimizing within-class variance), the optimal threshold level t^* can be obtained:

$$t^* = \arg\max_{t \in [1,L)} \sigma_B^2(t), \tag{3.3}$$

where L is the maximum intensity level of the image. As the result of thresholding, pixels of the image with intensity higher than t^* are labeled as foreground objects, and the rest as background.

Figure 3.1(a) presents an image consisting of objects of interest on a darker background. Here, the Otsu's algorithm correctly separates the foreground from the background, as can be seen in Figure 3.1(b). On the other hand, applying the algorithm to an image with excessive noise (Figure 3.1(c)) leads to a clearly erroneous result (Figure 3.1(d)). This is due to the different pixel intensity distributions in the two examples. In the first case, the histogram of the image is bimodal with one large peak around the background intensities, and a small peak consisting of the foreground. In the noisy image, with a more uniform intensity distribution, the optimal threshold is closer to the mean of the intensities, resulting in virtually equal number of pixels to be assigned as background and foreground. Since, in the context of Publication V, we know *a priori* that the foreground objects comprise only $10 - 15\%$ of the total number of pixels, this information can be used in discarding images with excessive number of pixels considered foreground.

Images not discarded by the previous procedure, are impaired with noise from several sources, most relevant component in fluorescence microscopy being Poissonian [111]. To enable robust and reliable detection objects of interest, the effect of this noise must be suppressed, simultaneously preserving the object details. A standard approach for this "speckle" type noise suppression are nonlinear filters, especially the median filter [5], excessively applied also in fluorescent microscopy [36, 126, 110]. In two

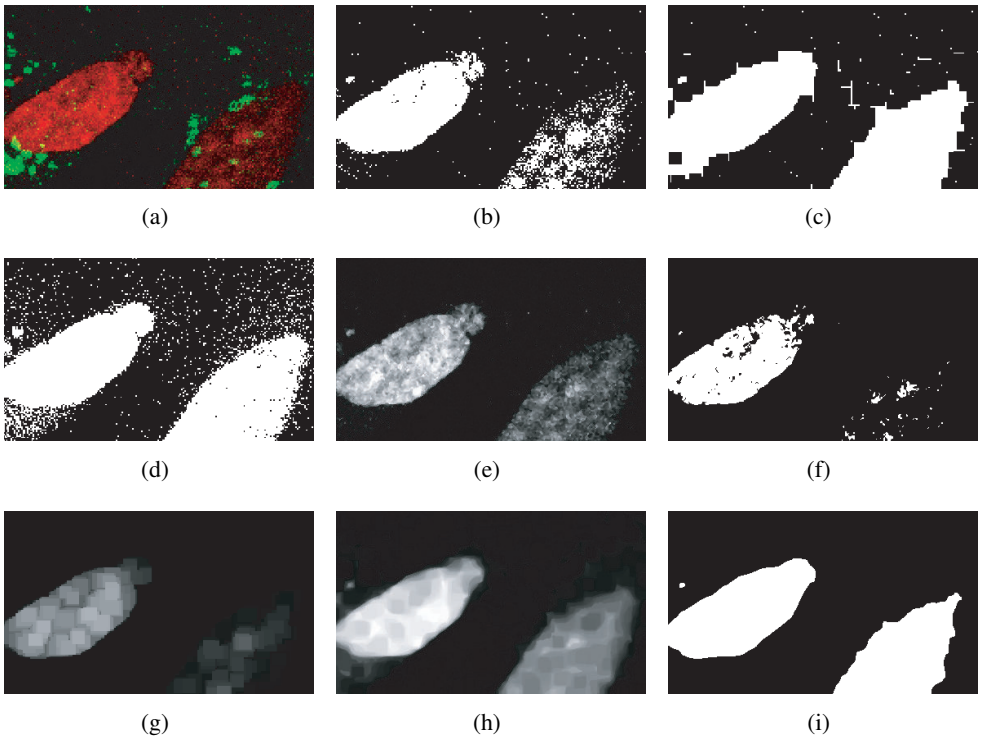


Figure 3.2: Rank order morphological filter applied to fluorescent microscopy of subcellular structures, see Publication IV for details of the data. (a) Original image with two nuclei (red). (b) Result of automated thresholding of the red channel (Otsu method [74]). (c) Postprocessing of (b) using morphological closing, with the aim of improving the segmentation result. (d) Manual thresholding of the red channel resulting in a noisy outcome, especially in areas where the green spots can be seen in (a), suggesting bleed of information between the red and green color channels. (e) The original image after filtering with the standard median filter. (f) Thresholding of (e) resulting in clear underestimation of the nuclei areas. (g) Morphological opening of the red channel also underestimating the nuclei areas. (h) Red channel after filtering with the rank ordered filter. $p = 5$, structuring element b of size 15×15. (i) Thresholding of (h) displaying two clearly visible nuclei ready for postprocessing.

dimensional median filtering, the original value of each pixel $m(x, y)$ is replaced by the median value of the pixel and it's neighborhood W:

$$m^*(x, y) = \text{median}\{m(x - s, y - t) \mid (s, t) \in W\} \tag{3.4}$$

The median filter is robust in suppressing outliers, and has excellent edge preservation properties (step response of a median filter is a step), a necessity for the forthcoming segmentation phase. Furthermore, the procedure is completely parameter free, and has a very low computational cost. We found the standard median filter to result in adequate noise suppression in all the confocal studies (Publications III and V).

In Publication IV, double staining with partly overlapping emission spectra and JPEG compression caused crosstalk between color channels, as visualized in 3.2(d). Although the automated analysis should always be applied to raw unprocessed data, lossy compression such as the JPEG is still occasionally used in practical microscopy applications due to storage space restrictions. The crossatalk artifacts rendered the median filter unfeasible. As a solution, we implemented a rank filter based modification of a gray scale morphological filter [97] $\zeta_{b,j}$:

$$[\zeta_{b,j}(m)](x,y) = P_j\{m(x-s,y-t)|(s,t) \in D_b\}, \tag{3.5}$$

where D_b is the domain of the structuring element b, and operator P_j returns j^{th} intensity percentile of the image inside structuring element. The images were filtered with the following combination:

$$o_{b,p}(m) = \zeta_{b,100-p}[\zeta_{b,p}(m)]. \tag{3.6}$$

The effectiveness of the rank ordered morphological filters stems from the possibility to adjust the "softness" of the filtering operation, enabling tuning of standard morphological operations such as opening and closing, making them more robust against outliers, still preserving shapes of the objects of interest [96]. With $p = 0$ the filter in Equation 3.6 is equivalent to the grayscale morphological opening. Increasing p will gradually filter the image with increasing smoothness, by discarding a selected percentile (rank) of the intensities inside the structuring element. We found $p = 5$, with structuring element b of size 15×15 to yield adequate results. One example of the properties of this filter, and comparisons to other filtering approaches like the median filter are given in Figure 3.2. The Figures 3.2(a), (b), (c), (d), (e), (f), and (g) present the original image, and results of different preprocessing techniques. After filtering with the rank ordered morphological filter (Figure 3.2(h)), the details of the object borders are preserved while greatly suppressing noise. Furthermore, the two objects of different intensities are filtered correctly: the one error can be corrected in a straightforward manner with a subsequent removal of undersized objects, for example. With this design, it is possible to avoid the property of the standard morphological opening (3.2(g)) to excessively emphasize the shape of structuring element, and the morphological closing (3.2(c)) to join erroneous pixels of noise by the object border to the object itself. Finally, the differences between results after the standard median filter and the proposed approach are evident by comparing the Figure 3.2(f) with Figure 3.2(i).

3.1.2 Background Correction

A varying bias field (shading, uneven background, flat field, vignetting) is often present in microscopy images. This phenomenon can be caused, for example, by thickness variation of the sample, using off-axis lighting techniques, background autofluorescence, or uneven sensitivity of the imaging detector [106]. The smoothly varying intensities of the shading do not hinder manual analysis, but automated image analysis is often

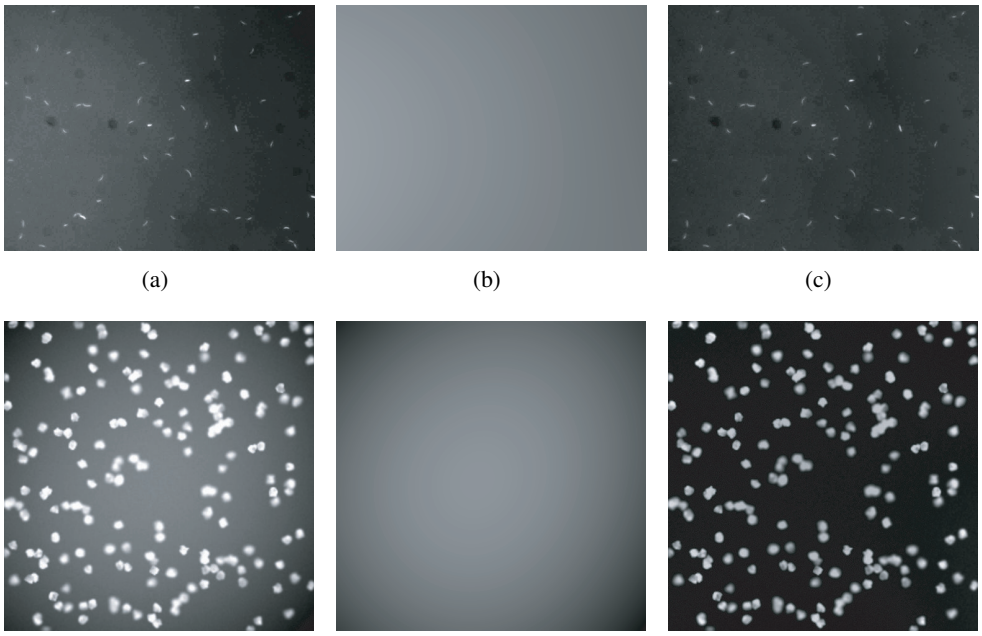


Figure 3.3: Correction of background shading. (a) Bacteria image from a water hydraulic system, see Publication I for details. (b) Two dimensional second degree polynomial fitted to (a). (c) Result after background subtraction. (d) Simulated image of DAPI stained nuclei with strong background shading, obtained using SIMCEP cell image simulator [51]. (e) Fitted surface. (f) Result after subtraction.

based on the absolute intensity values, global thresholding being one example, making the correction of the background component compulsory step.

In the literature, two main approaches have been presented for the bias, namely additive (see Equation 2.1) and multiplicative models, or a combination of the two [84, 57, 106]. Several methods have also been proposed to reduce the effect: in [108] the authors used morphological closing operator to estimate the background, [55] presents a method where background is estimated with entropy minimization and in [54] the background was estimated using gaussian low pass filtering; see [106] for an extensive review on the subject. Traditionally, to minimize computational complexity, the correction has also been applied by first imaging an empty calibration field without any objects, and subsequently subtracting this calibration image from all the images under study. This procedure adds and extra step to the imaging, and requires the calibration image to be distributed with the unprocessed data. Aiming at a more general correction scheme, the background bias should be estimated and removed given only the unprocessed images.

In Publications I and III we found that in order to enable reliable segmentation,

the background correction was a necessary preprocessing step, and that the additive model in accordance with Equation 3.1 was sufficient in modeling the bias. For two dimensional images, the light distribution in the bias field can be assumed quadratic [84, 49]:

$$b(x, y) = a_0 + a_1 x + a_2 y + a_3 xy + a_4 x^2 + a_5 y^2, \tag{3.7}$$

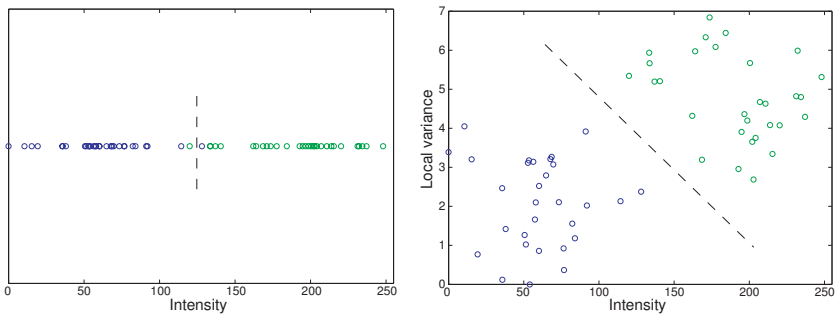
where a_n are the coefficients of the polynomial describing a two dimensional surface. In Publication I, the shading correction was performed by fitting this surface to the original image in least squares sense, resulting in estimates for the coefficients a_n of Equation 3.7. Thereafter, the estimated polynomial surface was subtracted from the image as illustrated in Figure 3.3. Figures 3.3(a), 3.3(b), and 3.3(c) show an example of DAPI stained bacteria (see Chapter 2) taken from a bioreactor for enumeration, the fitted surface, and the resulting image after subtraction for segmentation, respectively. Another example for visual evaluation of the subtraction method is given in Figures 3.3(d), 3.3(e), and 3.3(f), where the same procedure is applied to a simulated image created with the SIMCEP microscopy image simulator [53, 52, 51].

In the approach described above, because the polynomial surface is fitted to all the image pixels, the pixels not belonging to the background also affect the estimated coefficients. These effects can be avoided, for example, by selecting beforehand the pixels that will be used for the fit [84], or by performing the fit in a more robust manner discarding the outlier (foreground) pixels automatically [57]. In Publication III, we applied an M-estimator based background estimation. In short, the M-estimator approach is similar to the aforementioned polynomial fit, except that the standard sum of squares cost function in the least squares approach is replaced by a more general form, such as the Tukey cost function. This robust approach enhances the performance of the fit by suppressing the effect of outliers as demonstrated in [69], where implementation details, result examples, and discussion can also be found.

3.2 Segmentation

In segmentation, the pixels of the image are partitioned into different groups or classes. Typically, the objects of interest are separated from the background as already illustrated in Figure 3.1, aiming at a result where specific labels are assigned for specific types of objects [66], enabling automated interpretation of the image. The segmentation is often considered as the most difficult part of automated image analysis [31], and great emphasis must be put on the segmentation accuracy and robustness, not least because all the measurements of the detected objects are based on the segmentation result. Previously, numerous different segmentation schemes have been presented for microscopy images of cells (see, for example, [13] and [125]), but no single method is applicable to all the diverse image types [132].

One approach to segmentation is to consider the process as a pattern recognition problem [43, 20], where *features* calculated for pixels of the image are grouped according to some similarity measure. Formally, following the notations from [42],



(a) (b)

Figure 3.4: Feature based image segmentation. Each circle represents a measurement result (feature vector) for a pixel in an image, blue color for background and green for foreground. The data is for illustration purposes only, drawn from two normal distributions. (a) Intensities of the pixels. An adequate threshold separating background and foreground would lie somewhere around $110 - 130$ in the intensity scale. (b) The feature vectors after a second measurement, resulting in improved class separation, and therefore, depending on the segmentation algorithm, enabling improved segmentation.

we have *feature vectors* $\mathbf{x} = (x_1, \ldots x_d)$ of d measurements, forming a *pattern set* $\mathscr{H} = \{\mathbf{x}_1, \ldots \mathbf{x}_n\}$, where n is the number of feature vectors, in image analysis often the number of pixels. We want to find a method, that would assign a *class label* l_i for each of the feature vectors $\mathbf{x}_i$, where $l_i \in \{1, \ldots, k\}$ and k is the number of different groups or clusters. That is, we measure some properties of the image pixels, such as the intensities, concatenate the different measurements into vectors $\mathbf{x}_i$ and implement an algorithm labeling the vectors into predetermined number of groups. Furthermore, since we assume there is one feature vector for each pixel, the procedure results in assigning each pixel a specific label l_i. These pixel labels can refer to the background and foreground (cells) in the image, for example.

Figure 3.4 presents measurements for a small image consisting of foreground and background pixels, visualized for clarity in blue and green, respectively. The data is for illustration only, drawn from two different normal distributions. First, in Figure 3.4(a), we have measured one feature for the pixels, the intensity. In Figure 3.4(b), a second measurement has been added into the feature vectors, here the local intensity variation around the pixel. The dashed lines in both of the Figures represent the segmentation outcome, separating the groups based on the measured features, and assigning the pixels with background and foreground labels.

Preprocessing in the previous Section compensated for the distortion modeled by $b(x, y, z)$ and $N(\cdot)$ in Equation 3.1, yielding an estimate of the original image $f(x, y, z)$ given the input image $m(x, y, z)$. While this estimate may be far from perfect, it can be assumed to be normalized to fulfill the requirements set by a specific segmentation technique. Therefore, also the segmentation procedures introduced next

were designed independently for most of the Publications. Since staining is assumed, however, the objects of interest can be located by searching for areas of different intensity compared to the background. Two main approaches include the detection of objects with different pixel intensities inside the object than outside, and the detection of borders (rapid changes in intensities).

3.2.1 Detection of Foreground Objects

Generalizing the definition of thresholding given in Section 3.1.1, in thresholding [92], feature vectors $\mathbf{x}_i$ of one feature, are labeled in two classes by the magnitude of the vector. This is commonly applied with the pixel intensity as the only feature, resulting in labeling the darker areas of lower intensity values as background, and the brighter areas as foreground, as in the example with Otsu's thresholding method (Equations 3.2 and 3.3, Figures 3.2(b) and 3.4(a)). In Publications I, III, IV and V, after successful preprocessing, we found the Otsu's thresholding method with pixel intensities as features, to be the only operation needed for separating the background from the objects of interest in the images with staining.

In Publication VI, the imaging was performed by focusing through a semitransparent tissue section block of immunostained neurons, with the aim of locating and visualizing the 3-D structure of stained cells in the block. The focusing produced a series of images (referred as *z-slices*) of different z-planes, where the objects of interest appear both in and out of focus, and the segmentation procedure was required to label the *in focus* pixels of each 2-D z-slice as foreground. There are several focus level estimation algorithms presented in the literature, most of them implemented for autofocus applications selecting the best focus image from an image stack [102], for estimating 3-D shapes from focus levels [67], or for generating *all-in-focus* images through multiple image fusion [73, 129]. Our approach, resembling the proven *Tenengrad* focus estimation method [9], proceeds as follows:

1. Convolve the images with the standard Sobel mask [1].
2. Select the Sobel gradient magnitude as the only feature for each pixel.
3. Apply thresholding to label the feature vectors either as in focus candidate pixels, or out of focus.
4. Post process the candidate pixels using morphological operations.

Since the in focus pixels differentiate by rapid intensity changes from the more uniform background, and since the gradient magnitude is a measure of local changes in intensities, the pixels considered to form the in focus objects can be detected by thresholding the gradient magnitudes, separating the strongest local intensity transitions from the rest of the image. Finally, with the post filtering step, local irregularities are removed as described in more detail in the Publication VI. Figures 3.5(a) and (c) present two examples of input z-slices with two different focus levels. Results of the

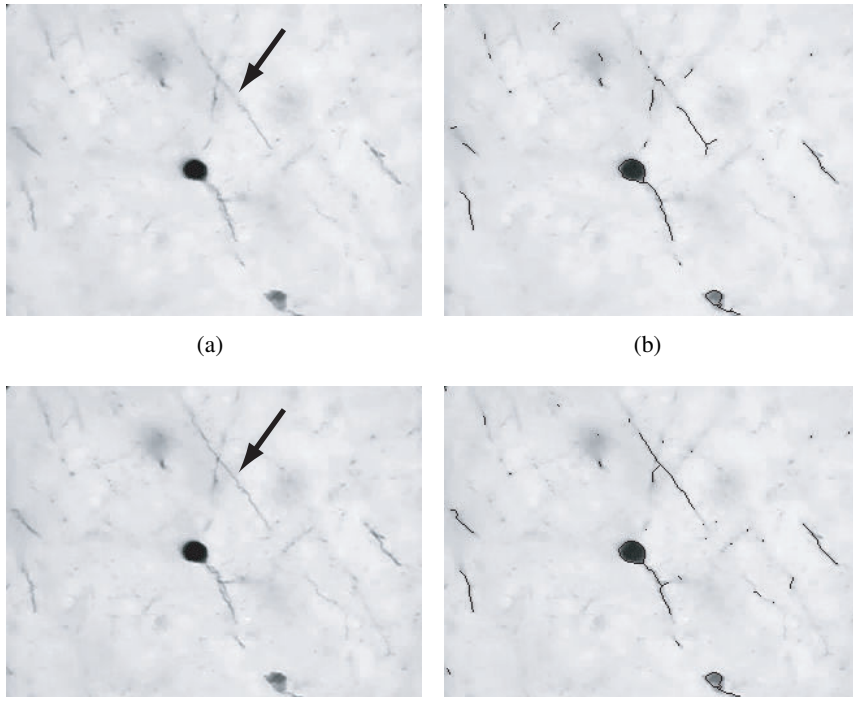


Figure 3.5: Detection of structures in focus. The images are from thick tissue sections of immunostained neurons, see Publication VI for details. (a) and (c) Images of a semitransparent tissue block with two different focus levels. (b) and (d) Result of focus detection by the Sobel-based method. The arrow highlights a structure with a change in focus from (a) to (c).

in focus detection algorithms are given in 3.5(b) and (d). By processing through all the z-slices, the method yields the 3-D structure of the object.

In the images for Publication II, a rather nonspecific fluorescent staining resulted in detection of bright spots, or *puncta*, over a heavily autofluorescent background. The rapidly altering level of background rendered background correction and thresholding methods, introduced in the previous Sections, inadequate, but while the absolute intensity values of the puncta differ, the intensities are higher than their surroundings (Figures 3.6(a) and 3.6(d)). Therefore, the pixel intensity and local intensity variation were selected as features for segmentation, the principle schemed in Figure 3.4(b). One class of algorithms for grouping two or higher dimensional feature vectors is unsupervised classification, *clustering* [42], grouping the feature vectors $\mathbf{x}_i$ of pattern set $\mathscr{H}$ around cluster centers by minimizing a certain error measure. Previously, *fuzzy clustering* (FCM) [42] has been proven effective in segmentation in biomedical image processing [78, 131].

We applied the FCM method as follows:

1. Create a random membership matrix $\mathbf{U}$ of size $n \times k$, where $u_{ij} \in [0, 1]$ representing fuzzy membership values for each of the n feature vectors in each of the k clusters.
2. Calculate a fuzzy criterion function $e^2(\mathscr{H}, \mathbf{U}) = \sum_{i=1}^{n} \sum_{j=1}^{k} u_{ij}^2 \|\mathbf{x}_i - \mathbf{c}_j\|^2$, where $\mathbf{c}_j = \frac{\sum_{i=1}^{n} u_{ij}^2 \mathbf{x}_i}{\sum_{i=1}^{n} u_{ij}^2}$ is the j^{th} fuzzy cluster center. Reassign the feature vectors into clusters decreasing the value of the criterion function. Recompute $\mathbf{U}$.
3. Repeat the Step 2 until the elements of $\mathbf{U}$ change by less than a given threshold, or a specified number of iterations has been reached.
4. Assign each feature vector to the cluster with the highest membership value in the matrix $\mathbf{U}$.

The FCM procedure labels each pixel of the original image to one of the k clusters. By utilizing the *a priori* knowledge of small fluorescent puncta on darker and larger background areas, the correct pixels are obtained by combining the clusters with the smallest number of pixels, until the next operation would force the total number of objects over a pre-estimated limit. Because of the random initialization of the matrix $\mathbf{U}$, the clustering may not always converge to exactly the same minimum (the same pixels may not always be assigned the same cluster labels). To minimize this variation, the clustering is repeated nine times, and the final results is selected to be the one with the median number of detected objects. Example results are presented in Figure 3.6, with comparisons to the Otsu thresholding method.

Despite the large number of processing phases, the clustering approach has several advantages. First, only a very limited number of parameters need to be set prior to the processing. Second, as opposed to supervised classification algorithms, clustering does not require training samples. Third, the segmentation procedure is shown to be robust to background noise such as autofluorescence. Finally, the FCM has previously been shown to encompass better convergence properties overcoming local minima in comparison to the popular k-means algorithm (Section 3.4), making the initialization of the matrix $\mathbf{U}$ less critical [42].

In Publication III, the segmentation of bright field yeast cells was also performed with a method utilizing two features, the local mean and variance, but the approach was not designed by the author of this thesis, and has already been described in detail in [69] and [71].

3.2.2 Separation of Overlapping Objects

After assigning the pixels into foreground and background, aggregated or overlapping objects must be separated to obtain object by object level data. There exists several different algorithms in the literature for the separation of overlapping objects in biomedical microscopy [91, 33, 127, 2, 58], but again, no single approach has been shown

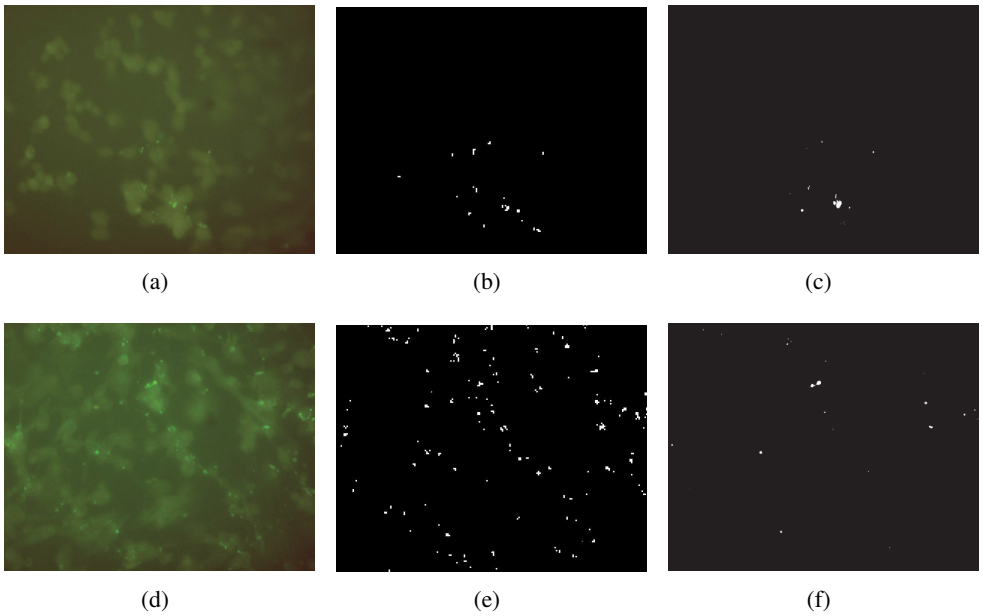


Figure 3.6: Detection of fluorescent puncta in neuroblastoma cells, details of the images given in Publication II. (a) and (d) Input images expressing heavy autofluorescence. (b) and (e) The result of the fuzzy clustering approach, displaying the detected puncta. (c) and (f) Results of Otsu thresholding for comparison, clearly underestimating the number of vesicles. If a lower threshold was set manually, the segmentation would detect the regions with strongest autofluorescence, not puncta.

to be superior in all the studies. One of the most general methods is the *watershed segmentation* [114, 97], with numerous different applications.

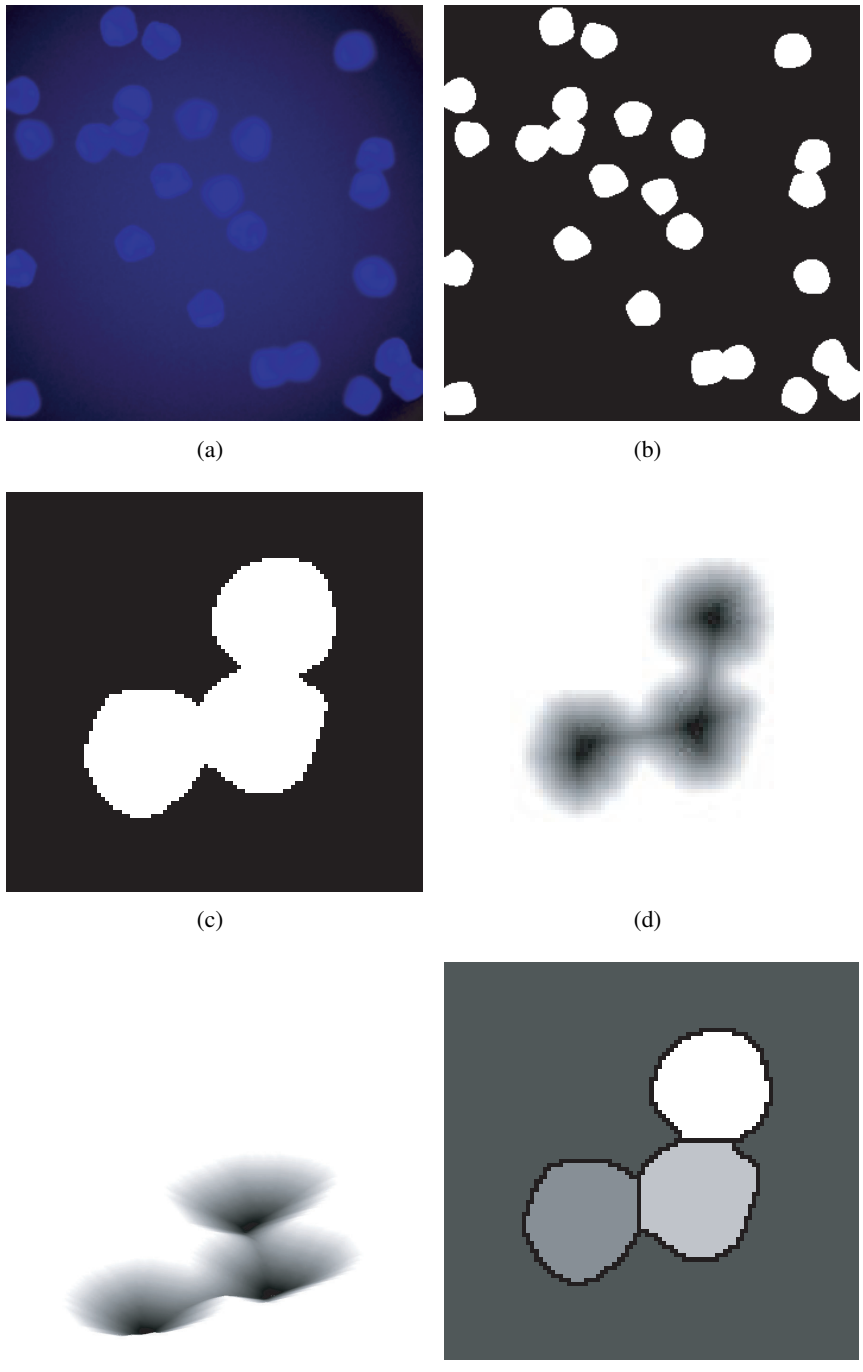
In the watershed segmentation, intensity changes of an image are considered as a topographic surface, see Figure 3.7 for an example. In this surface, the different objects are visible as pits, separated by a dividing ridge as displayed in Figure 3.7(e). The watershed transform begins to flood the surface from each regional minimum (pits or *catchment basins*), placing watershed lines where water from different pits meet (on the ridges), separating the overlapping objects. In practice, however, small irregularities in the object shapes or intensities result in several regional minima to be found for each of the objects, resulting in oversegmentation, where a single object is divided into multiple segments. A standard method of correcting the oversegmentation is *h-minima imposition* [97], where minima not deeper than a predefined threshold are suppressed. An extensive description of the watershed algorithm can be found in [114].

For the CellC software [90] introduced in Publication I, we implemented two different variations of the watershed algorithm. In the first version, the watershed lines are calculated directly from the intensity values of the input image. This procedure effectively separates overlap if the borders of the objects are of different intensity than the centers [60]. In the second version, the image is first segmented for a binary result (Figure 3.7(b) and (c)), followed by the Euclidean distance transform [7] (Figure 3.7(d)) and finally by the watershed algorithm separating the objects (Figure 3.7(f)). The distance transform in the second approach introduces information of the object shapes, resulting in successful separation of convex bodies. The second approach was also utilized in Publications III and IV. In the Publication V we applied a 3-D extension of the algorithm, separating three dimensional ball-like objects as described in Section 3.3.

3.3 3-D Reconstruction

In Publications V and VI, two different microscope configurations were used to obtain 3-D data, namely a standard 2-D light microscope (bright field) and a confocal setup. As described in the previous Chapter, segmentation of the bright field data was enabled through detection of in focus objects. In the confocal configuration of Publication V, most of the out of focus light is inherently suppressed, simplifying the segmentation step into thresholding. From a stack of z-slices from the microscope, the segmentation results in a three dimensional binary matrix, each layer presenting a section of the target object from the current focal plane. See Figure 3.8(a) for an example of sections of a spherical object. Similarly as the 2-D images consist of squares (pixels) with different intensities (0's and 1's in this binary case), this three dimensional data set can be perceived as 3-D space consisting of cubes (voxels) of different intensities. With reconstruction, this data can be visualized through rendering.

In confocal microscopy, the z resolution does not equal the resolution in x and y dimensions. For the data in Publication V, we used standard nearest neighbor interpo-



(a) (b) (c) (d) (e) (f)

Figure 3.7: Overlapping objects as topographic surface. (a) Original image of fluorescent nuclei, created with the SIMCEP cell image simulator [53]. (b) Thresholding result of (a). (c) One of the cell clusters enlarged. (d) Distance transform of (c). The lower the intensity, the further the specific pixel is from background. (e) The objects in (d) visualized as a topographic surface, ready for separation of overlapping objects. (f) The result after locating the ridges separating the three circular objects with the watershed method.

lation to cancel this anisotropy by introducing additional z-slices to the original data set. For visualization, two main approaches are commonly applied to 3-D data [23], namely (direct) volume rendering [10] and surface rendering [123], with numerous extensions [22]. Briefly, in volume rendering, each foreground voxel is represented as a semi-transparent cube in a 3-D space (Figure 3.8(d)). In surface rendering, the surface of the object is located and modeled with polygonal meshes as presented in Figure 3.8(b). Next, shading [80] is applied to take into account the objects orientation and surface characteristics with respect to the scene lighting, yielding a brightness value for different elements of the surface. Finally, the result is projected on a 2-D plane perpendicular to the observer for display. Depending on the application, several other steps can be included, such as texture mapping or alpha mapping for transparency effects, and furthermore, the rendering itself can be performed with numerous different algorithms outside the scope of this thesis. Figure 3.8 illustrates the reconstruction from a stack of binary images to the rendering of a spherical object.

The direct volume rendering was applied to display the located neurons for manual validation in Publication VI, and the both rendering approaches were compared in Publication V. The surface rendering provides a better overview of the scene in question, being computationally less expensive, and providing a more visually appealing result. On the other hand, since only the object surfaces are located and displayed, this approach is often unsuitable for detailed analysis of objects' internal structures. In contrast, the volume rendering represents the actual locations and intensities of the detected voxels, enabling a more detailed study, but the overall geometry of the shapes is difficult to visualize. This is illustrated in the Figures 3.8(c) and 3.8(d) with surface and volume rendered spherical object with inner structure, respectively.

2-D imaging, being a projection of a 3-D scene in the z dimension, can introduce heavy overlap to objects only slightly in contact in the z direction. The less overlap, the more robust the separation, implying the separation should take place in 3-D. In Publication V, we applied a 3-D extension of the watershed method, previously implemented for nuclei separation in [56, 118]. An example of the procedure is given in Figure 3.9. The Figure 3.9(a) presents an object most likely consisting of two separate structures overlapping by a few voxels. Minima of the distance transform of this object are displayed in Figure 3.9(b), defining the catchment basins for watershed algorithm as explained in Section 3.2.2. Since there are two basins, the original object will be divided in half, resulting in the two separate objects of Figure 3.9(c). For clarity, the Figures 3.9(d) and 3.9(e) display the same procedure with surface rendering. If the selection of the catchment basins can be performed in a robust way, the object separation will increase analysis reliability, especially object enumeration.

3.4 Measurements

The methods presented in this Chapter aim at extracting information from biological cells; transforming the input data (images) into quantitative results, such as the number

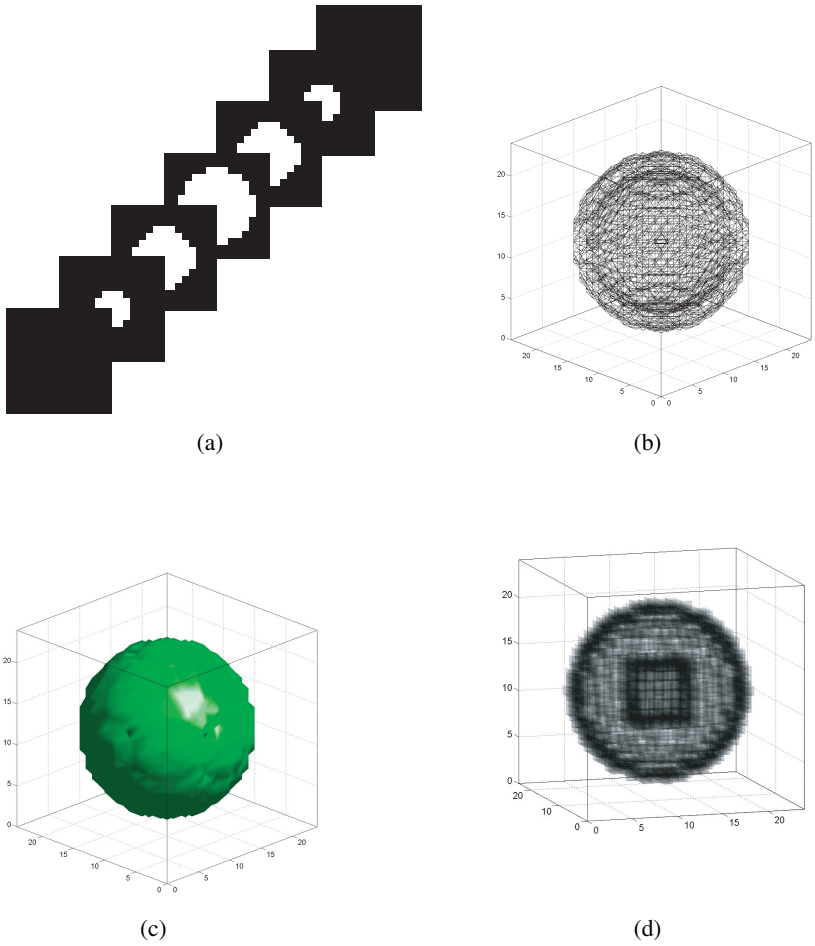


Figure 3.8: 3-D reconstruction. (a) Schema of a stack of images forming a sphere-shaped object. (b) A wire frame presentation of surface polygons of the sphere. Resolution is increased compared to the schema in (a) for illustration purposes. (c) Surface rendering of the sphere in (b). (d) Direct volume rendering, inner structure displayed to highlight differences with surface rendering. Here, a small semitransparent cube is representing each voxel of the object.

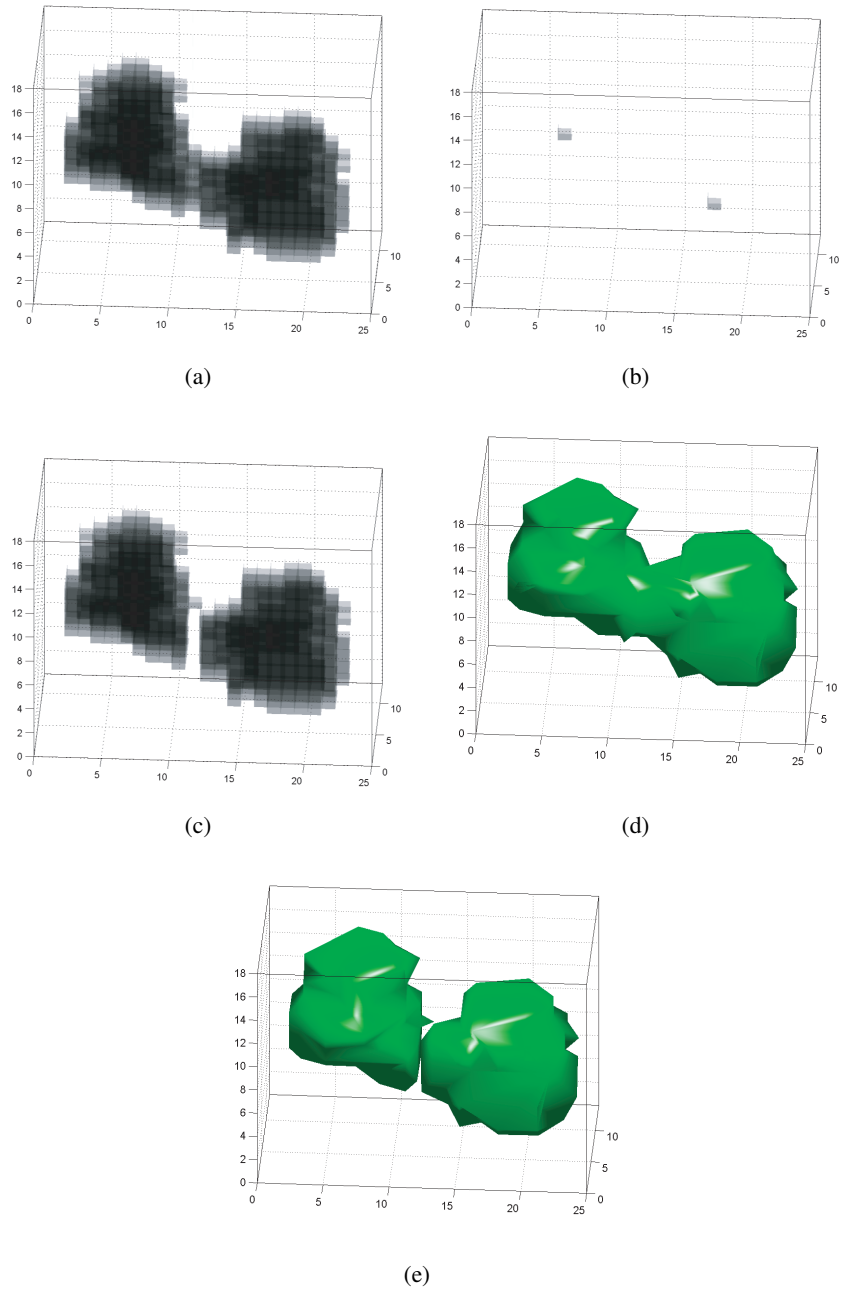


Figure 3.9: Watershed transform in 3-D applied for confocal data of subcellular structures, see Publication V for details. (a) Two partly overlapping objects. (b) Minima of the distance transform of (a). (c) Result after the watershed algorithm, initialized from the minima in (b). (d) Surface rendering of the overlapping objects. (e) Surface rendering of the result.

of objects.

In the literature, hundreds or even thousands of features quantitatively describing cell phenotype have been presented [83]. Since the segmentation extracted the detected objects from the background [66], different properties can often be measured in a very straightforward manner after applying a specific label for all the objects. For example, the number of objects results directly from the number of different labels, the object size is a sum of pixels with specific labeling, and the length of the object border can be measured by calculating the number of border pixels and assigning different weights for different border patterns [112]. Obviously, the importance of the segmentation phase is greatly emphasized during the feature extraction since all the procedures are fundamentally bound to the results of the segmentation [133]. In Publications I, II and III, all the required measurements consisted of particle enumeration, with area and intensity features that can be derived straightforwardly from the labeled segmentation result.

To analyze subcellular components on a cell-by-cell level, it is also required to label the objects according to the cell each object belongs to. This can be achieved by an additional whole cell fluorescent staining, or by whole cell segmentation from images without staining (Publication III, [71]), encompassing the subcellular structures. In Publication IV, such data was not available, but the segmentation result consisted of two channels: red channel pixels labeled as background or cell nuclei, and green channel pixels labeled as background or *golgi apparatus* (GA), a certain subcellular component. Figures 3.2(a) and 3.10(a) present examples of the input data before segmentation, and Figures 3.2(i) and 3.10(c) the segmentation results for the nuclei and golgi channels, respectively. To label the GA pixels for different nuclei, we construct feature vectors or patterns $\mathbf{x}$ from the x and y coordinates of the GA pixels, and by k-means clustering algorithm [42], cluster the vectors into k groups, where k is the number of cells (nuclei).

Similarly as the fuzzy c means clustering, the k-means clustering finds the clustering L of pattern set $\mathscr{H}$ by minimizing the error function e, here the standard mean squared error defined as:

$$e^2\left(\mathscr{H}, L\right) = \sum_{j=1}^{k} \sum_{i=1}^{n_j} \left\| \mathbf{x}_i^{(j)} - \mathbf{c}_j \right\|^2 \tag{3.8}$$

where $\mathbf{x}_i^{(j)}$ are the i^{th} pattern of j^{th} cluster, n_j is the number of patterns in the j^{th} cluster, and $\mathbf{c}_j$ are the cluster centers.

The clustering procedure is the following: First, the cluster centers $\mathbf{c}_j$ are initialized as the coordinates of the centers of mass (Figure 3.10(b)) of each of the k nuclei (Figure 3.10(a)). Second, the feature vectors $\mathbf{x}$ for each of the golgi pixels (3.10(c)) are assigned to the closest cluster center. Third, the centroids of the feature vectors assigned for the specific clusters are set as the new cluster centers $\mathbf{c}_j$. The two latter steps are repeated until the error e changes less than a given threshold, or a predefined number of iterations is reached. Since the center of mass of each nuclei was selected as

the initial cluster center, the k-means clustering procedure results in the labeling, where each of the golgi pixels are assigned to one of the nuclei. This result is illustrated in the Figure 3.10(d).

Next, the dispersion of each GA cluster (see Figure 3.10(d)) can be approximated by the median distance d_i of GA pixel coordinates to the corresponding GA center of mass:

$$d_j = \text{median}(\left\|\mathbf{x}_i^{(j)} - \mathbf{c}_j\right\|^2) \tag{3.9}$$

where $i \in \{1, \ldots, n_j\}$. In Publication IV we calculated the distances d_j for multiple image sets, resulting in measures of the golgi apparatus dispersion after several different treatments for cell populations.

In Publication V, shape descripition of subcellular components (*peroxisomes*) was studied for future detection of peroxisomal changes between different biological treatments. There exists numerous approaches for shape description in the literature, ranging from simpler and more intuitive parameters such as compactness [84] to computationally complex ones, such as spherical harmonic descriptors [94]. The selection of descriptors depends on the application, but also on the resolution and the physical sizes of the objects to be described. Even with confocal microscopy, subcellular objects often consist of only 5 to 10 voxels in diameter, hindered by imaging defects modeled by the Equation 2.1. Higher magnification would result in more pixels per millimeter, but no more real resolution after the diffraction limit [100]. With such a low number of pixels, the errors in segmentation become dominant, only enabling differentiation of basic features such as object elongation. In the Publication V, we mainly limited to features such as the lengths and widths of objects, resembling shape description applicable in manual analysis for validation.

The measurements were performed after normalizing the orientations of the segmented and reconstructed peroxisomes as visualized in Figure 3.11. In the normalization procedure, the *principal component analysis* (PCA) [44] is first applied to the covariance matrix C of the voxel coordinates $\mathbf{v}_i$ of the detected object:

$$\mathbf{C} = \frac{1}{n}\sum_{i=1}^{n} \mathbf{v}_i \mathbf{v}_i^{\mathbf{T}} \tag{3.10}$$

where $^{\mathbf{T}}$ is the transpose operator, and n is the number of voxels in an object. The principal components can be found as the eigenvectors $\mathbf{w}_i$ corresponding to the eigenvalues λ_i of

$$\mathbf{C}\mathbf{w}_i = \lambda_i \mathbf{w}_i \tag{3.11}$$

In short, this procedure can be described as finding the direction (eigenvector) of the largest variance in the input data set, followed by the direction of the second largest variance orthogonal to the previous one, and so forth, as illustrated in Figure 3.11(a).

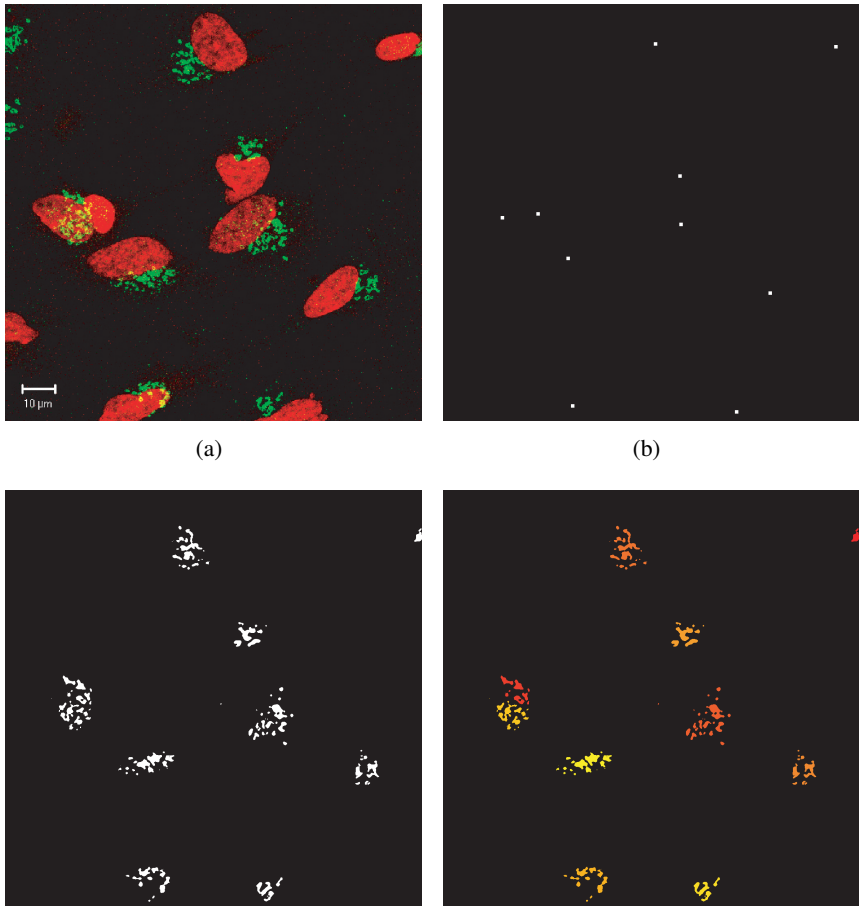
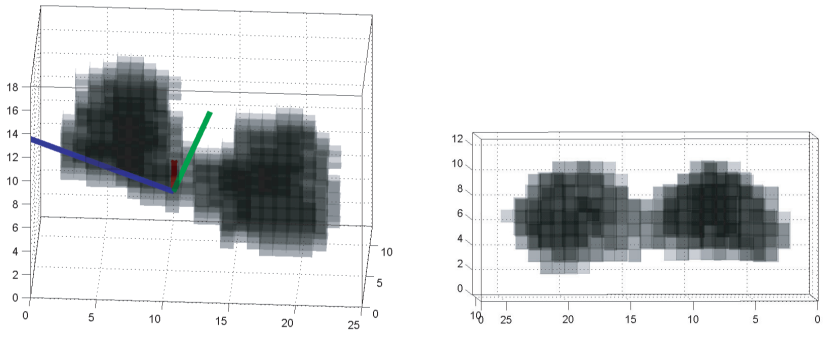


Figure 3.10: Distributing subcellular structures (GA) among several nuclei. Details of input data given in Publication IV. (a) Original two channel data of stained nuclei and GA. (b) Centers of mass of segmented nuclei. (c) Segmentation result of the GA channel. (d) Result of k-means clustering of the pixels in (c), the nuclei centers in (b) as initial cluster centers. Different colors illustrate different clusters.



(a) (b)

Figure 3.11: Orientation normalization with principal component analysis. The data is from confocal microscopy, details given in Publication V. (a) Object with eigenvectors corresponding to the three largest eigenvalues displayed in blue, green, and red. (b) Object after normalization, the eigenvectors aligned with the main x, y, and z axes. The length, width, and thickness of the object can now be measured along these axes.

Thereafter, the detected peroxisomes were rotated with a base change operation, aligning the eigenvectors corresponding to the three largest eigenvalues with the main x, y, and ,z axes shown in Figure 3.11(b). This alignment has been shown to provide an adequate rotation invariance [44, 116], and allows us to calculate the object length, width, and thickness by measuring the objects dimensions along the main axes.

Chapter 4

Discussion

Despite the recent advances, some of which presented in this work, challenges remain slowing down the change from tedious manual microscopy to fully automated high throughput processes. Here we briefly discuss some of the obstacles that have become apparent during the research performed for this thesis, and describe future aims on the area of high throughput microscopy.

4.1 Result Validation

Manual visual analysis procedures are the gold standard or ground truth against which virtually all automated image analysis methods are compared in cellular imaging. Commonly, however, direct comparisons to manually obtained results are not feasible. This is especially true in high throughput microscopy, where validation of automated image analysis results for large image data sets is laborious, increasing the possibility for human error in the manual analysis [122, 16]. In cell counting, to give an example, the results acquired by an automated enumeration algorithm can be validated by performing manual counts for a few descriptive images representing the whole batch. Unfortunately it is impossible to be sure the images chosen for the manual analysis actually represent the whole stack, and therefore, we have no way to quantitatively measure the reliability of different algorithms.

This validation problem has been addressed in other fields utilizing image processing by standard image sets, where manual analysis representing the ground truth has been done with great care [99, 62, 79]. Currently, there are few attempts in collecting similar standard image databases in the field of cell image analysis [26]. In the databases, although the manual analysis has been performed carefully, features such as intensities of cells are virtually impossible to quantify by visual analysis. A human observer is capable of sorting the cell populations according to intensity with a rough classification, but is unable to provide a quantitative intensities for each cell. Furthermore, the previously introduced weaknesses of manual analysis affect the results.

In computed tomography (CT), for example, the validation is performed by simulation [93]. In a simulated image, the ground truth is known *a priori*, enabling straightforward validation of different CT image segmentation methods. The SIMCEP simulator [51, 52] is a start for similar research in cellular imaging, enabling simulation of biomedical microscopy, and the construction of benchmark databases for validation consisting of the simulated images [85]. In these artificial images the exact shapes, locations, and other properties of the cells in the image are precisely known, enabling systematic validation of analysis methods. It is an interesting question, however, how should the simulator itself be validated.

4.2 Data Management

Data management and metadata handling are integral parts of high throughput imaging systems. Currently, the tools for data storage, processing, analysis, and integration in these systems are rather immature [133], hindering straightforward utilization of high throughput microscopy. For example in image analysis, manual work is still be needed also in high throughput applications [19]. As another example, it is intuitive to store a 2-D view of the DIC channel (described in Chapter 2.2) if locations of cells are to be detected. Recently, however, automated analysis of 3-D focus through stacks of the DIC channel has been shown to result in improved accuracy for the two dimensional detection [70], implying the requirement always to store all the data available for future research. This lack of well specified and standard approaches lead to ineffective result reproduction and algorithm reusage.

Projects such as the Cell Centered Database [63], the Open Microscopy Environment [30, 89], and WebMicroscope [59] aim at standardizing the data storage and metadata handling, combined with a web-based dissemination of images and results. CellProfiler [13], on the other hand, is one of the first attempts to design an open source platform, on which different image analysis procedures for cell biology can be implemented in a well specified manner, with a common interface [12]. For three dimensional image analysis, implementations include *daime* [17] and BioimageXD [46].

Commercial high throughput microscopy platforms are available from a number of providers, but the cost and often closed, black box type of image analysis and data handling of the platforms hinder the usage from the viewpoint of basic research. Various laboratory specific complete high throughput screening analysis systems have also been constructed [130, 32], but despite the great motivation [103], standardized general frameworks under which the whole processing chain from imaging to data analysis could be implemented are still rare, the Bisque system [107] being an exception. Only by such a freely available integrative framework, and with a well defined handling of data and metadata, the repeatability, re-usage, and integration between different data sources can be assured.

Chapter 5

Summary of Publications

This study presents digital image analysis algorithms for two and three dimensional whole cell and subcellular quantification, replacing manual visual analysis of cells and cell populations. We have introduced and validated automated measurements not previously presented in the literature, and applied algorithms that have not been used in the context of high throughput microscopy. Practically applicable solutions are introduced for all the case studies of the Publications, but no single method has been proven superior in all of them, leaving room for future research. Since there undoubtedly is an increasing need for automated methods enabling high throughput microscopy, new challenges will constantly arise in the algorithm development. The biological rationale and benefits of the measurement systems are best understood by biologists, and the technical improvements for these systems are facilitated by engineers, implying there should be even more emphasis on interdisciplinary research with real collaboration between computer scientists and biologists.

Main results of the Publications are the following:

I **"Software for quantification of labeled bacteria from digital microscope images by automated image analysis"** We created a platform for cell enumeration. At the time of the publication, no software with similar capabilities existed [6], enabling microbiologists to enumerate objects in images with different characteristics in a straightforward manner. The software was shown to be in excellent accordance with manually acquired results, is freely distributed and completely modifiable, and has since been used in several publications in microbiology, by us and others [90].

II **"Quantification of vesicles in differentiating human SH-SY5Y neuroblastoma cells by automated image analysis"** We designed a fuzzy-clustering based segmentation algorithm for quantification of small fluorescent puncta in neuroblastoma cells. The algorithm is shown to be robust against extensive background nonuniformities and autofluorescence, and to correlate well with manual

analysis.

III **"Extraction of the number of peroxisomes in yeast cells by automated image analysis"** We implemented algorithms for segmentation, and quantification of subcellular structures inside yeast cells. The fully automated method incorporates bright field microscopy data with fluorescence microscopy, and enumerates the structures on a cell-by-cell basis, separating any overlapping cells. Validation shows the algorithms to yield comparable results to manual analysis in both cell and organelle enumeration.

IV **"Automated analysis of Golgi Apparatus dispersion in neuronal cell images"** We designed a rank filtering based segmentation method for noisy fluorescent microscopy images of nuclei. Furthermore, we applied k-means clustering algorithm to distribute located subcellular structures for each of the detected nuclei. This allowed us to quantify the dispersion of the subcellular structures, a result not previously presented in the literature by automated methods. We compared the dispersion after different cell treatments, and concluded that the results are consistent with manual results from previous studies.

V **"A case study on 3-D reconstruction and shape description of peroxisomes in yeast"** In this study, we applied noise detection, segmentation, visualization, and quantification of subcellular structures in 3-D. Even with algorithms of low computational cost, we show that the 3-D approach can increase enumeration accuracy, motivating high throughput screening studies to move into three dimensional techniques, instead of current 2-D methods.

VI **"Three-dimensional digital image analysis of immunostained neurons in thick tissue sections"** We implemented an algorithm detecting in-focus areas of standard bright field microscopy images taken with different focal planes of a neuronal tissue section. We then reconstructed the data to form a 3-D view of the target. The paper shows that in contrast to commonly utilized fluorescence microscopy and expensive confocal techniques, standard bright field microscopy enables 3-D studies applicable in routine pathology, for example.

Bibliography

[1] I. Abdou and W. Pratt, "Quantitative design and evaluation of enhancement/thresholding edge detectors," *Proceedings of the IEEE*, vol. 67, no. 5, pp. 753–763, May 1979.

[2] G. Agam and I. Dinstein, "Geometric separation of partially overlapping nonrigid objects applied to automatic chromosome classification," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 11, pp. 1212–1222, 1997.

[3] H. Ai, K. L. Hazelwood, M. W. Davidson, and R. E. Campbell, "Fluorescent protein FRET pairs for ratiometric imaging of dual biosensors." *Nat Methods*, vol. 5, no. 5, pp. 401–403, May 2008. [Online]. Available: http://dx.doi.org/10.1038/nmeth.1207

[4] P. D. Andrews, I. S. Harper, and J. R. Swedlow, "To 5D and beyond: quantitative fluorescence microscopy in the postgenomic era." *Traffic*, vol. 3, no. 1, pp. 29–36, Jan 2002.

[5] J. Astola and P. Kuosmanen, *Fundamentals of Nonlinear Digital Filtering*, 1st ed. CRC-Press, Oct 1997.

[6] BioTechniques Editors, "Biospotlight: Software that really hits the spot," *BioTechniques*, vol. 39, no. 6, pp. 787–789, 2005.

[7] G. Borgefors, "Distance transformations in digital images," *Computer Vision, Graphics, and Image Processing*, vol. 34, no. 3, pp. 344–371, 1986.

[8] D. Botstein, S. A. Chervitz, and J. M. Cherry, "Yeast as a model organism." *Science*, vol. 277, no. 5330, pp. 1259–1260, Aug 1997.

[9] M. A. Bueno-Ibarra, J. Álvarez Borrego, L. Acho, and M. C. Chávez-Sánchez, "Fast autofocus algorithm for automated microscopes," *Optical Engineering*, vol. 44, no. 6, p. 063601, 2005. [Online]. Available: http://link.aip.org/link/?JOE/44/063601/1

[10] S. Callahan, J. Callahan, C. Scheidegger, and C. Silva, "Direct volume rendering: A 3D plotting technique for scientific data," *Computing in Science & Engineering*, vol. 10, no. 1, pp. 88–92, Jan.–Feb. 2008.

[11] A. E. Carpenter, "Image-based chemical screening." *Nat Chem Biol*, vol. 3, no. 8, pp. 461–465, Aug 2007. [Online]. Available: http://dx.doi.org/10.1038/nchembio.2007.15

[12] ——, "Software opens the door to quantitative imaging." *Nat Methods*, vol. 4, no. 2, pp. 120–121, Feb 2007. [Online]. Available: http://dx.doi.org/10.1038/nmeth0207-120

[13] A. E. Carpenter, T. R. Jones, M. R. Lamprecht, C. Clarke, I. H. Kang, O. Friman, D. A. Guertin, J. H. Chang, R. A. Lindquist, J. Moffat, P. Golland, and D. M. Sabatini, "CellProfiler: image analysis software for identifying and quantifying cell phenotypes." *Genome Biol*, vol. 7, no. 10, p. R100, 2006. [Online]. Available: http://dx.doi.org/10.1186/gb-2006-7-10-r100

[14] D. M. Chudakov, S. Lukyanov, and K. A. Lukyanov, "Fluorescent proteins as a toolkit for *in vivo* imaging." *Trends Biotechnol*, vol. 23, no. 12, pp. 605–613, Dec 2005. [Online]. Available: http://dx.doi.org/10.1016/j.tibtech.2005.10.005

[15] J.-A. Conchello and J. W. Lichtman, "Optical sectioning microscopy." *Nat Methods*, vol. 2, no. 12, pp. 920–931, Dec 2005. [Online]. Available: http://dx.doi.org/10.1038/nmeth815

[16] J. E. L. Corry, B. Jarvis, S. Passmore, and A. Hedges, "A critical review of measurement uncertainty in the enumeration of food micro-organisms." *Food Microbiol*, vol. 24, no. 3, pp. 230–253, May 2007. [Online]. Available: http://dx.doi.org/10.1016/j.fm.2006.05.003

[17] H. Daims, S. Lücker, and M. Wagner, "daime, a novel image analysis program for microbial ecology and biofilm research." *Environ Microbiol*, vol. 8, no. 2, pp. 200–213, Feb 2006. [Online]. Available: http://dx.doi.org/10.1111/j.1462-2920.2005.00880.x

[18] M. W. Davidson and M. Abramowitz, "Molecular expressions: Optical microscopy primer," 2008, http://micro.magnet.fsu.edu/primer/index.html.

[19] T. N. Davis, "Protein localization in proteomics." *Curr Opin Chem Biol*, vol. 8, no. 1, pp. 49–53, Feb 2004. [Online]. Available: http://dx.doi.org/10.1016/j.cbpa.2003.11.003

[20] R. Duda, P. Hart, and D. Stork, *Pattern Classification*. John Wiley & Sons, 2001.

[21] A. Egner and S. W. Hell, "Fluorescence microscopy with super-resolved optical sections." *Trends Cell Biol*, vol. 15, no. 4, pp. 207–215, Apr 2005. [Online]. Available: http://dx.doi.org/10.1016/j.tcb.2005.02.003

[22] R. Eils and C. Athale, "Computational imaging in cell biology." *J Cell Biol*, vol. 161, no. 3, pp. 477–481, May 2003. [Online]. Available: http://dx.doi.org/10.1083/jcb.200302097

[23] T. T. Elvins, "A survey of algorithms for volume visualization," *SIGGRAPH Comput. Graph.*, vol. 26, no. 3, pp. 194–201, 1992.

[24] M. Encinas, M. Iglesias, Y. Liu, H. Wang, A. Muhaisen, V. Ceña, C. Gallego, and J. X. Comella, "Sequential treatment of SH-SY5Y cells with retinoic acid and brain-derived neurotrophic factor gives rise to fully differentiated, neurotrophic factor-dependent, human neuron-like cells." *J Neurochem*, vol. 75, no. 3, pp. 991–1003, Sep 2000.

[25] J.-J. Fernandez, C. Sorzano, R. Marabini, and J.-M. Carazo, "Image processing and 3-D reconstruction in electron microscopy," *IEEE Signal Processing Magazine*, vol. 23, no. 3, pp. 84–94, 2006.

[26] E. D. Gelasca, J. Byun, B. Obara, and B. Manjunath, "Evaluation and benchmark for biological image segmentation," in *IEEE International Conference on Image Processing*, Oct 2008. [Online]. Available: http://vision.ece.ucsb.edu/publications/elisa_ICIP08.pdf

[27] S. F. Gibson and F. Lanni, "Experimental test of an analytical model of aberration in an oil-immersion objective lens used in three-dimensional light microscopy." *J Opt Soc Am A*, vol. 9, no. 1, pp. 154–166, Jan 1992.

[28] B. N. G. Giepmans, S. R. Adams, M. H. Ellisman, and R. Y. Tsien, "The fluorescent toolbox for assessing protein location and function." *Science*, vol. 312, no. 5771, pp. 217–224, Apr 2006. [Online]. Available: http://dx.doi.org/10.1126/science.1124618

[29] F. O. Glöckner, B. M. Fuchs, and R. Amann, "Bacterioplankton compositions of lakes and oceans: a first comparison based on fluorescence in situ hybridization." *Appl Environ Microbiol*, vol. 65, no. 8, pp. 3721–3726, Aug 1999.

[30] I. G. Goldberg, C. Allan, J.-M. Burel, D. Creager, A. Falconi, H. Hochheiser, J. Johnston, J. Mellen, P. K. Sorger, and J. R. Swedlow, "The Open Microscopy Environment (OME) Data Model and XML file: open tools for informatics and quantitative analysis in biological imaging." *Genome Biol*, vol. 6, no. 5, p. R47, 2005. [Online]. Available: http://dx.doi.org/10.1186/gb-2005-6-5-r47

[31] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 3rd ed. Upper Saddle River, NJ: Prentice Hall, 2008.

[32] A. Gordon, A. Colman-Lerner, T. E. Chin, K. R. Benjamin, R. C. Yu, and R. Brent, "Single-cell quantification of molecules and rates using open-source microscope-based cytometry." *Nat Methods*, vol. 4, no. 2, pp. 175–181, Feb 2007. [Online]. Available: http://dx.doi.org/10.1038/nmeth1008

[33] P. R. Gudla, K. Nandy, J. Collins, K. J. Meaburn, T. Misteli, and S. J. Lockett, "A high-throughput system for segmenting nuclei using multiscale techniques." *Cytometry A*, vol. 73, no. 5, pp. 451–66, Mar 2008. [Online]. Available: http://dx.doi.org/10.1002/cyto.a.20550

[34] M. G. Gustafsson, "Extended resolution fluorescence microscopy." *Curr Opin Struct Biol*, vol. 9, no. 5, pp. 627–634, Oct 1999.

[35] S. A. Haney, P. LaPan, J. Pan, and J. Zhang, "High-content screening moves to the front of the line." *Drug Discov Today*, vol. 11, no. 19-20, pp. 889–894, Oct 2006. [Online]. Available: http://dx.doi.org/10.1016/j.drudis.2006.08.015

[36] P. Hänninen, E. H. K. Stelzer, and J. Salo, "Nonlinear filtering in improving the image quality of confocal fluorescent images," *Mach. Vision Appl.*, vol. 4, no. 4, pp. 243–253, 1991.

[37] S. W. Hell, M. Schrader, and H. T. van der Voort, "Far-field fluorescence microscopy with three-dimensional resolution in the 100-nm range." *J Microsc*, vol. 187, no. Pt 1, pp. 1–7, Jul 1997.

[38] P. Hinterdorfer and Y. F. Dufrêne, "Detection and localization of single molecular recognition events using atomic force microscopy." *Nat Methods*, vol. 3, no. 5, pp. 347–355, May 2006. [Online]. Available: http://dx.doi.org/10.1038/nmeth871

[39] R. A. Hoebe, H. T. M. V. der Voort, J. Stap, C. J. F. V. Noorden, and E. M. M. Manders, "Quantitative determination of the reduction of phototoxicity and photobleaching by controlled light exposure microscopy." *J Microsc*, vol. 231, no. Pt 1, pp. 9–20, Jul 2008. [Online]. Available: http://dx.doi.org/10.1111/j.1365-2818.2008.02009.x

[40] L. Hood, J. R. Heath, M. E. Phelps, and B. Lin, "Systems biology and new technologies enable predictive and preventative medicine." *Science*, vol. 306, no. 5696, pp. 640–643, Oct 2004. [Online]. Available: http://dx.doi.org/10.1126/science.1104635

[41] T. Ideker, V. Thorsson, J. A. Ranish, R. Christmas, J. Buhler, J. K. Eng, R. Bumgarner, D. R. Goodlett, R. Aebersold, and L. Hood, "Integrated genomic and proteomic analyses of a systematically perturbed metabolic network." *Science*, vol. 292, no. 5518, pp. 929–934, May 2001. [Online]. Available: http://dx.doi.org/10.1126/science.292.5518.929

[42] A. K. Jain, M. N. Murty, and P. J. Flynn, "Data clustering: a review," *ACM Computing Surveys*, vol. 31, no. 3, pp. 264–323, 1999.

[43] A. Jain, R. Duin, and J. Mao, "Statistical pattern recognition: a review," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 1, pp. 4–37, Jan. 2000.

[44] I. T. Jolliffe, *Principal Component Analysis*, 2nd ed. New York, NY: Springer-Verlag New York, Inc., 2002.

[45] R. S. Kamath, A. G. Fraser, Y. Dong, G. Poulin, R. Durbin, M. Gotta, A. Kanapin, N. L. Bot, S. Moreno, M. Sohrmann, D. P. Welchman, P. Zipperlen, and J. Ahringer, "Systematic functional analysis of the *Caenorhabditis elegans* genome using RNAi." *Nature*, vol. 421, no. 6920, pp. 231–237, Jan 2003. [Online]. Available: http://dx.doi.org/10.1038/nature01278

[46] P. Kankaanpää, K. Pahajoki, V. Marjomäki, J. Heino, and D. White, "BioImageXD," 2006, http://www.bioimagexd.net.

[47] R. L. Kepner and J. R. Pratt, "Use of fluorochromes for direct enumeration of total bacteria in environmental samples: past and present." *Microbiol Rev*, vol. 58, no. 4, pp. 603–615, Dec 1994.

[48] H. Kitano, "Systems biology: a brief overview." *Science*, vol. 295, no. 5560, pp. 1662–1664, Mar 2002. [Online]. Available: http://dx.doi.org/10.1126/science.1069492

[49] A. Klein, R. van den Doel, I. Young, S. Ellenberger, and L. van Vliet, "Quantitative evaluation and comparison of light microscopes," in *Optical Investigation of Cells In Vitro and In Vivo*, ser. Proc. SPIE, Progress in Biomedical Optics, D. Farkas, R. Leif, and B. Tromberg, Eds., vol. 3260, 1998, pp. 162–173.

[50] J. Kononen, L. Bubendorf, A. Kallioniemi, M. Bärlund, P. Schraml, S. Leighton, J. Torhorst, M. J. Mihatsch, G. Sauter, and O. P. Kallioniemi, "Tissue microarrays for high-throughput molecular profiling of tumor specimens." *Nat Med*, vol. 4, no. 7, pp. 844–847, Jul 1998.

[51] A. Lehmussola, P. Ruusuvuori, J. Selinummi, T. Rajala, and O. Yli-Harja, "Synthetic images of high-throughput microscopy for validation of image analysis methods," *Proceedings of the IEEE*, vol. 96, no. 8, pp. 1348–1360, 2008.

[52] A. Lehmussola, P. Ruusuvuori, J. Selinummi, H. Huttunen, and O. Yli-Harja, "Computational framework for simulating fluorescence microscope images with cell populations." *IEEE Trans Med Imaging*, vol. 26, no. 7, pp. 1010–1016, Jul 2007.

[53] A. Lehmussola, J. Selinummi, P. Ruusuvuori, A. Niemisto, and O. Yli-Harja, "Simulating fluorescent microscope images of cell populations." *Conf Proc IEEE Eng Med Biol Soc*, vol. 3, pp. 3153–3156, 2005. [Online]. Available: http://dx.doi.org/10.1109/IEMBS.2005.1617144

[54] F. J. W.-M. Leong, M. Brady, and J. O. McGee, "Correction of uneven illumination (vignetting) in digital microscopy images." *J Clin Pathol*, vol. 56, no. 8, pp. 619–621, Aug 2003.

[55] B. Likar, J. B. Maintz, M. A. Viergever, and F. Pernus, "Retrospective shading correction based on entropy minimization." *J Microsc*, vol. 197, no. Pt 3, pp. 285–295, Mar 2000.

[56] G. Lin, U. Adiga, K. Olson, J. F. Guzowski, C. A. Barnes, and B. Roysam, "A hybrid 3D watershed algorithm incorporating gradient cues and object models for automatic segmentation of nuclei in confocal image stacks." *Cytometry A*, vol. 56, no. 1, pp. 23–36, Nov 2003. [Online]. Available: http://dx.doi.org/10.1002/cyto.a.10079

[57] J. Lindblad and E. Bengtsson, "A comparison of methods for estimation of intensity nonuniformities in 2D and 3D microscope images of fluorescence stained cells," in *Proceedings of the 12th Scandinavian Conference on Image Analysis (SCIA)*, Jun 2001.

[58] C. G. Loukas, G. D. Wilson, B. Vojnovic, and A. Linney, "An image analysis-based approach for automated counting of cancer cell nuclei in tissue sections." *Cytometry A*, vol. 55, no. 1, pp. 30–42, Sep 2003. [Online]. Available: http://dx.doi.org/10.1002/cyto.a.10060

[59] M. Lundin, J. Lundin, H. Helin, and J. Isola, "A digital atlas of breast histopathology: an application of web based virtual microscopy." *J Clin Pathol*, vol. 57, no. 12, pp. 1288–1291, Dec 2004. [Online]. Available: http://dx.doi.org/10.1136/jcp.2004.018739

[60] N. Malpica, C. O. de Solórzano, J. J. Vaquero, A. Santos, I. Vallcorba, J. M. García-Sagredo, and F. del Pozo, "Applying watershed algorithms to the segmentation of clustered nuclei." *Cytometry*, vol. 28, no. 4, pp. 289–297, Aug 1997.

[61] A. Marian, F. Charrière, T. Colomb, F. Montfort, J. Kühn, P. Marquet, and C. Depeursinge, "On the complex three-dimensional amplitude point spread function of lenses and microscope objectives: theoretical aspects, simulations and measurements by digital holography." *J Microsc*, vol. 225, no. Pt 2, pp. 156–169, Feb 2007. [Online]. Available: http://dx.doi.org/10.1111/j.1365-2818.2007.01727.x

[62] D. Martin, C. Fowlkes, D. Tal, and J. Malik, "A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics," in *Proc. 8th Int'l Conf. Computer Vision*, vol. 2, July 2001, pp. 416–423.

[63] M. E. Martone, J. Tran, W. W. Wong, J. Sargis, L. Fong, S. Larson, S. P. Lamont, A. Gupta, and M. H. Ellisman, "The cell centered database project: an update on building community resources for managing and sharing 3D imaging data." *J Struct Biol*, vol. 161, no. 3, pp. 220–231, Mar 2008. [Online]. Available: http://dx.doi.org/10.1016/j.jsb.2007.10.003

[64] J. G. McNally, T. Karpova, J. Cooper, and J. A. Conchello, "Three-dimensional imaging by deconvolution microscopy." *Methods*, vol. 19, no. 3, pp. 373–385, Nov 1999. [Online]. Available: http://dx.doi.org/10.1006/meth.1999.0873

[65] X. Michalet, F. F. Pinaud, L. A. Bentolila, J. M. Tsay, S. Doose, J. J. Li, G. Sundaresan, A. M. Wu, S. S. Gambhir, and S. Weiss, "Quantum dots for live cells, in vivo imaging, and diagnostics." *Science*, vol. 307, no. 5709, pp. 538–544, Jan 2005. [Online]. Available: http://dx.doi.org/10.1126/science.1104274

[66] T. W. Nattkemper, "Automatic segmentation of digital micrographs: a survey." *Stud Health Technol Inform*, vol. 107, no. Pt 2, pp. 847–851, 2004.

[67] S. Nayar and Y. Nakagawa, "Shape from focus," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 16, no. 8, pp. 824–831, 1994.

[68] B. Neumann, M. Held, U. Liebel, H. Erfle, P. Rogers, R. Pepperkok, and J. Ellenberg, "High-throughput RNAi screening by time-lapse imaging of live human cells." *Nat Methods*, vol. 3, no. 5, pp. 385–390, May 2006. [Online]. Available: http://dx.doi.org/10.1038/nmeth876

[69] A. Niemistö, "Quantitative image analysis methods for applications in biomedical microscopy," Ph.D. dissertation, Tampere University of Technology, 2006.

[70] A. Niemistö, T. Korpelainen, R. Saleem, O. Yli-Harja, J. Aitchison, and I. Shmulevich, "A k-means segmentation method for finding 2-D object areas based on 3-D image stacks obtained by confocal microscopy," in *Proc. 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society EMBS 2007*, 2007, pp. 5559–5562.

[71] A. Niemistö, T. Aho, H. Thesleff, M. Tiainen, K. Marjanen, M.-L. Linne, and O. P. Yli-Harja, "Estimation of population effects in synchronized budding yeast experiments," in *Proc. SPIE, Image Processing: Algorithms and Systems II*, E. R. Dougherty, J. T. Astola, and K. O. Egiazarian, Eds., vol. 5014, no. 1. SPIE, 2003, pp. 448–459. [Online]. Available: http://link.aip.org/link/?PSI/5014/448/1

[72] J. P. Nolan, S. Lauer, E. R. Prossnitz, and L. A. Sklar, "Flow cytometry: a versatile tool for all phases of drug discovery." *Drug Discov Today*, vol. 4, no. 4, pp. 173–180, Apr 1999. [Online]. Available: http://dx.doi.org/10.1016/S1359-6446(99)01320-3

[73] K. Ohba, J. C. P. Ortega, K. Tanie, M. Tsuji, and S. Yamada, "Microscopic vision system with all-in-focus and depth images," *Mach. Vision Appl.*, vol. 15, no. 2, pp. 55–62, 2003.

[74] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans. on Systems, Man, and Cybernetics*, vol. 9, no. 1, pp. 62–66, Jan 1979.

[75] J. B. Pawley, *Handbook of biological confocal microscopy*, 3rd ed. Springer-Verlag New York, Inc., 2006.

[76] R. Pepperkok and J. Ellenberg, "High-throughput fluorescence microscopy for systems biology." *Nat Rev Mol Cell Biol*, vol. 7, no. 9, pp. 690–696, Sep 2006. [Online]. Available: http://dx.doi.org/10.1038/nrm1979

[77] Z. E. Perlman, M. D. Slack, Y. Feng, T. J. Mitchison, L. F. Wu, and S. J. Altschuler, "Multidimensional drug profiling by automated microscopy." *Science*, vol. 306, no. 5699, pp. 1194–1198, Nov 2004. [Online]. Available: http://dx.doi.org/10.1126/science.1100709

[78] T. D. Pham, D. I. Crane, T. H. Tran, and T. H. Nguyen, "Extraction of fluorescent cell puncta by adaptive fuzzy segmentation." *Bioinformatics*, vol. 20, no. 14, pp. 2189–2196, Sep 2004. [Online]. Available: http://dx.doi.org/10.1093/bioinformatics/bth213

[79] P. J. Phillips, H. Moon, S. A. Rizvi, and P. J. Rauss, "The FERET evaluation methodology for face-recognition algorithms," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 10, pp. 1090–1104, 2000.

[80] B. T. Phong, "Illumination for computer generated pictures," *Commun. ACM*, vol. 18, no. 6, pp. 311–317, 1975.

[81] J. S. Ploem, N. Verwoerd, J. Bonnet, and G. Koper, "An automated microscope for quantitative cytology combining television image analysis and stage scanning microphotometry." *J Histochem Cytochem*, vol. 27, no. 1, pp. 136–143, Jan 1979.

[82] L. Prescott, J. Harley, and D. Klein, *Microbiology*, 4th ed. The McGraw-Hill Companies, Inc., 1999.

[83] K. Rodenacker and E. Bengtsson, "A feature set for cytometry on digitized microscopic images." *Anal Cell Pathol*, vol. 25, no. 1, pp. 1–36, 2003.

[84] J. C. Russ, *The image processing handbook*, 3rd ed. Boca Raton, FL: CRC Press, Inc., 1999.

[85] P. Ruusuvuori, A. Lehmussola, J. Selinummi, T. Rajala, H. Huttunen, and O. Yli-Harja, "Benchmark set of synthetic images for validating cell image analysis algorithms," in *Proc. of the 16th European Signal Processing Conference EUSIPCO-2008*, 2008.

[86] P. Sarder and A. Nehorai, "Deconvolution methods for 3-D fluorescence microscopy images," *IEEE Signal Processing Magazine*, vol. 23, no. 3, pp. 32–45, 2006.

[87] J.-R. Sarkanen, J. Nykky, J. Siikanen, J. Selinummi, T. Ylikomi, and T. O. Jalonen, "Cholesterol supports the retinoic acid-induced synaptic vesicle formation in differentiating human SH-SY5Y neuroblastoma cells." *J Neurochem*, vol. 102, no. 6, pp. 1941–1952, Sep 2007. [Online]. Available: http://dx.doi.org/10.1111/j.1471-4159.2007.04676.x

[88] M. Schena, D. Shalon, R. W. Davis, and P. O. Brown, "Quantitative monitoring of gene expression patterns with a complementary DNA microarray." *Science*, vol. 270, no. 5235, pp. 467–470, Oct 1995.

[89] D. A. Schiffmann, D. Dikovskaya, P. L. Appleton, I. P. Newton, D. A. Creager, C. Allan, I. S. Näthke, and I. G. Goldberg, "Open microscopy environment and findspots: integrating image informatics with quantitative multidimensional image analysis." *Biotechniques*, vol. 41, no. 2, pp. 199–208, Aug 2006.

[90] J. Selinummi, "Cellc software," 2008, http://www.cs.tut.fi/sgn/csb/cellc/index.html.

[91] ——, "Automated quantitative analysis of color-stained cell images," Master's thesis, Tampere University of Technology, 2004.

[92] M. Sezgin and B. Sankur, "Survey over image thresholding techniques," *J. Electron. Imaging*, vol. 13, no. 1, pp. 146–165, Jan 2004.

[93] L. A. Shepp and B. F. Logan, "Reconstructing interior head tissue from X-ray transmissions," *IEEE Transactions on Nuclear Science*, vol. 21, no. 1, pp. 228–236, Feb. 1974.

[94] P. Shilane, P. Min, M. Kazhdan, and T. Funkhouser, "The Princeton shape benchmark," in *Proc. Shape Modeling Applications*, 2004, pp. 167–178.

[95] L. A. Sklar, M. B. Carter, and B. S. Edwards, "Flow cytometry for drug discovery, receptor pharmacology and high-throughput screening." *Curr Opin Pharmacol*, vol. 7, no. 5, pp. 527–534, Oct 2007. [Online]. Available: http://dx.doi.org/10.1016/j.coph.2007.06.006

[96] P. Soille, "On morphological operators based on rank filters," *Pattern Recognition*, vol. 35, pp. 527–535, Feb 2002.

[97] ——, *Morphological Image Analysis: Principles and Applications*, 2nd ed. Secaucus, NJ, USA: Springer-Verlag New York, Inc., 2003.

[98] B. L. Sprague, R. L. Pego, D. A. Stavreva, and J. G. McNally, "Analysis of binding reactions by fluorescence recovery after photobleaching." *Biophys J*, vol. 86, no. 6, pp. 3473–3495, Jun 2004. [Online]. Available: http://dx.doi.org/10.1529/biophysj.103.026765

[99] J. Staal, M. D. Abramoff, M. Niemeijer, M. A. Viergever, and B. van Ginneken, "Ridge-based vessel segmentation in color images of the retina," *IEEE Transactions on Medical Imaging*, vol. 23, no. 4, pp. 501–509, April 2004.

[100] E. H. K. Stelzer, "Beyond the diffraction limit?" *Nature*, vol. 417, no. 6891, pp. 806–807, Jun 2002. [Online]. Available: http://dx.doi.org/10.1038/417806a

[101] D. J. Stephens and V. J. Allan, "Light microscopy techniques for live cell imaging." *Science*, vol. 300, no. 5616, pp. 82–86, Apr 2003. [Online]. Available: http://dx.doi.org/10.1126/science.1082160

[102] Y. Sun, S. Duthaler, and B. J. Nelson, "Autofocusing in computer microscopy: selecting the optimal focus algorithm." *Microsc Res Tech*, vol. 65, no. 3, pp. 139–149, Oct 2004. [Online]. Available: http://dx.doi.org/10.1002/jemt.20118

[103] J. R. Swedlow, S. E. Lewis, and I. G. Goldberg, "Modelling data across labs, genomes, space and time." *Nat Cell Biol*, vol. 8, no. 11, pp. 1190–1194, Nov 2006. [Online]. Available: http://dx.doi.org/10.1038/ncb1496

[104] A. Tárnok, "Slide-based cytometry for cytomics–a minireview." *Cytometry A*, vol. 69, no. 7, pp. 555–562, Jul 2006. [Online]. Available: http://dx.doi.org/10.1002/cyto.a.20317

[105] ——, "A focus on high-content cytometry." *Cytometry A*, vol. 73, no. 5, pp. 381–383, May 2008. [Online]. Available: http://dx.doi.org/10.1002/cyto.a.20571

[106] D. Tomazevic, B. Likar, and F. Pernus, "Comparative evaluation of retrospective shading correction methods." *J Microsc*, vol. 208, no. Pt 3, pp. 212–223, Dec 2002.

[107] University of California at Santa Barbara, Center for Bio-Image Informatics, "Biscue," 2006, http://www.bioimage.ucsb.edu/bisque.

[108] Y. Usson, A. Guignandon, N. Laroche, M. H. Lafage-Proust, and L. Vico, "Quantitation of cell-matrix adhesion using confocal image analysis of focal contact associated proteins and interference reflection microscopy." *Cytometry*, vol. 28, no. 4, pp. 298–304, Aug 1997.

[109] G. Valet, J. F. Leary, and A. Tárnok, "Cytomics–new technologies: towards a human cytome project." *Cytometry A*, vol. 59, no. 2, pp. 167–171, Jun 2004. [Online]. Available: http://dx.doi.org/10.1002/cyto.a.20047

[110] G. M. P. van Kempen, "Image restoration in fluorescence microscopy," Ph.D. dissertation, Technische Universiteit Delft, 1999.

[111] G. M. P. van Kempen, J. van Vliet, P. Verveer, and H. van Der Coort, "A quantitative comparison of image restoration methods for confocal microscopy," *Journal of Microscopy*, vol. 185, no. 3, pp. 354–365, 1997. [Online]. Available: http://www.blackwell-synergy.com/doi/abs/10.1046/j.1365-2818.1997.d01-629.x

[112] L. van Vliet, P. Verbeek, and I. Young, "Quantitative imaging: how to measure size features in digitized images," in *Proc. IEEE International Symposium on Biomedical Imaging: Nano to Macro*, 2004, pp. 1227–1230 Vol. 2.

[113] L. Vega-Alvarado, I. Elezgaray, A. Hémar, M. Menard, C. Ranger, and G. Corkidi, "A comparison of image deconvolution algorithms applied to the detection of endocytic vesicles in fluorescence images of neural proteins." *Conf Proc IEEE Eng Med Biol Soc*, vol. 2007, pp. 755–758, 2007. [Online]. Available: http://dx.doi.org/10.1109/IEMBS.2007.4352400

[114] L. Vincent and P. Soille, "Watersheds in digital spaces: an efficient algorithm based on immersion simulations," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, no. 6, pp. 583–598, 1991.

[115] C. Vonesch, F. Aguet, J.-L. Vonesch, and M. Unser, "The colored revolution of bioimaging," *IEEE Signal Processing Magazine*, vol. 23, no. 3, pp. 20–31, 2006.

[116] D. Vranic, "An improvement of rotation invariant 3D-shape based on functions on concentric spheres," in *Proc. International Conference on Image Processing ICIP 2003*, vol. 3, 2003, pp. III–757–60 vol.2.

[117] N. J. Wade and M. Swanston, *Visual Perception : An Introduction*. New Fetter Lane, London, UK: Routledge, 1991.

[118] C. Wählby, I.-M. Sintorn, F. Erlandsson, G. Borgefors, and E. Bengtsson, "Combining intensity, edge and shape information for 2D and 3D segmentation of cell nuclei in tissue sections." *J Microsc*, vol. 215, no. Pt 1, pp. 67–76, Jul 2004. [Online]. Available: http://dx.doi.org/10.1111/j.0022-2720.2004.01338.x

[119] C. Wählby, J. Lindblad, M. Vondrus, E. Bengtsson, and L. Björkesten, "Algorithms for cytoplasm segmentation of fluorescence labelled cells." *Anal Cell Pathol*, vol. 24, no. 2-3, pp. 101–111, 2002.

[120] W. Wallace, L. H. Schaefer, and J. R. Swedlow, "A workingperson's guide to deconvolution in light microscopy." *Biotechniques*, vol. 31, no. 5, pp. 1076–8, 1080, 1082 passim, Nov 2001.

[121] H. Wallrabe and A. Periasamy, "Imaging protein molecules using FRET and FLIM microscopy." *Curr Opin Biotechnol*, vol. 16, no. 1, pp. 19–27, Feb 2005. [Online]. Available: http://dx.doi.org/10.1016/j.copbio.2004.12.002

[122] D. Webb, M. A. Hamilton, G. J. Harkin, S. Lawrence, A. K. Camper, and Z. Lewandowski, "Assessing technician effects when extracting quantities from microscope images." *J Microbiol Methods*, vol. 53, no. 1, pp. 97–106, Apr 2003.

[123] N. Weng, Y.-H. Yang, and R. Pierson, "Three-dimensional surface reconstruction using optical flow for medical imaging," *IEEE Transactions on Medical Imaging*, vol. 16, no. 5, pp. 630–641, Oct. 1997.

[124] D. B. Wheeler, A. E. Carpenter, and D. M. Sabatini, "Cell microarrays and RNA interference chip away at gene function." *Nat Genet*, vol. 37 Suppl, pp. S25–S30, Jun 2005. [Online]. Available: http://dx.doi.org/10.1038/ng1560

[125] T. Würflinger, J. Stockhausen, D. Meyer-Ebrecht, and A. Böcking, "Robust automatic coregistration, segmentation, and classification of cell nuclei in multimodal cytopathological microscopic images." *Comput Med Imaging Graph*, vol. 28, no. 1-2, pp. 87–98, 2004.

[126] G. Xiong, X. Zhou, A. Degterev, L. Ji, and S. T. C. Wong, "Automated neurite labeling and analysis in fluorescence microscopy images." *Cytometry A*, vol. 69, no. 6, pp. 494–505, Jun 2006. [Online]. Available: http://dx.doi.org/10.1002/cyto.a.20296

[127] Q. Yang, Q. Yang, and B. Parvin, "Harmonic cut and regularized centroid transform for localization of subcellular structures," *IEEE Transactions on Biomedical Engineering*, vol. 50, no. 4, pp. 469–475, 2003.

[128] J. Zhang, R. E. Campbell, A. Y. Ting, and R. Y. Tsien, "Creating new fluorescent probes for cell biology." *Nat Rev Mol Cell Biol*, vol. 3, no. 12, pp. 906–918, Dec 2002. [Online]. Available: http://dx.doi.org/10.1038/nrm976

[129] Z. Zhang and R. Blum, "A categorization of multiscale-decomposition-based image fusion schemes with a performance study for a digital camera application," *Proceedings of the IEEE*, vol. 87, no. 8, pp. 1315–1326, 1999.

[130] X. Zhou, X. Cao, Z. Perlman, and S. T. C. Wong, "A computerized cellular imaging system for high content analysis in Monastrol suppressor screens." *J Biomed Inform*, vol. 39, no. 2, pp. 115–125, Apr 2006. [Online]. Available: http://dx.doi.org/10.1016/j.jbi.2005.05.008

[131] X. Zhou, K. Y. Liu, P. Bradley, N. Perrimon, and S. T. C. Wong, "Towards automated cellular image segmentation for RNAi genome-wide screening." *Int Conf Med Image Comput Comput Assist Interv*, vol. 8, no. Pt 1, pp. 885–892, 2005.

[132] X. Zhou and S. Wong, "Informatics challenges of high-throughput microscopy," *IEEE Signal Processing Magazine*, vol. 23, no. 3, pp. 63–72, 2006.

[133] X. Zhou, X. Zhou, and S. Wong, "High content cellular imaging for drug development," *IEEE Signal Processing Magazine*, vol. 23, no. 2, pp. 170–174, 2006.

[134] J. Ziauddin and D. M. Sabatini, "Microarrays of cells expressing defined cDNAs." *Nature*, vol. 411, no. 6833, pp. 107–110, May 2001. [Online]. Available: http://dx.doi.org/10.1038/35075114