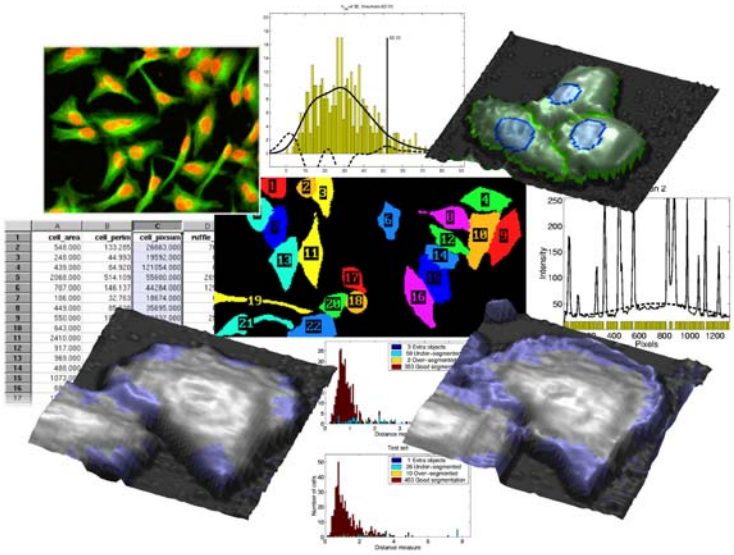
Development of Algorithms for Digital Image Cytometry

Joakim Lindblad



Comprehensive Summaries of Uppsala Dissertations
from the Faculty of Science and Technology 789

Development of Algorithms for Digital Image Cytometry

BY

JOAKIM LINDBLAD

ACTA UNIVERSITATIS UPSALIENSIS
UPPSALA 2002

Dissertation for the Degree of Doctor of Technology in Computerised Image Analysis presented at Uppsala University in 2003

Abstract

Lindblad J. 2003. Development of Algorithms for Digital Image Cytometry. Acta Universitatis Upsaliensis. *Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology* 789. 67 pp. Uppsala. ISBN 91-554-5497-6.

This thesis presents work in digital image cytometry applied to fluorescence microscope images of cultivated cells. Focus has been on the development and compilation of robust image analysis tools, enabling quantitative measurements of various properties of cells and cell structures. A significant part of the work has consisted of developing robust segmentation methods for fluorescently labelled cells. This, in combination with effort applied in the areas of feature extraction and statistical data analysis, has enabled the compilation of a complete chain of processing steps to produce a system capable of performing fully automatic segmentation and classification of fluorescently labelled cells according to their level of activation.

Two sequences of processing steps, both leading to automatic cytoplasm segmentation of fluorescence microscopy cell images are presented. In one of the sequences, an additional image of the nuclei of the cells is segmented. The nuclei are then used as seeds for the segmentation of the cytoplasm image. This solves the problem of oversegmentation of the cytoplasms in an efficient way. The other sequence uses merge and split algorithms on the cytoplasm image, in conjunction with statistical analysis of descriptive features. This analysis is used in a feedback system to improve the segmentation performance, and to give an overall quality measure of the segmentation.

A classification method that separates individual cells into three classes, depending on their level of activation, is described. The method is based on analysis of time series of images. Using both general purpose features and carefully designed problem specific features, in combination with a floating feature selection procedure, a Bayesian classifier is built. Evaluation showed that the performance of the fully automatic classification procedure was very close to the performance of skilled manual classification.

A novel method for performing estimation of intensity nonuniformites of microscope images is presented. Methods to solve many other problems related to image analysis of cell images are discussed and evaluated. All methods presented in this work are applicable to real-world situations. The two main projects of the thesis work have been performed in close cooperation with and according to demands of the biomedical industry.

Joakim Lindblad, Centre for Image Analysis, Uppsala University,
Lägerhyddsvägen 17, SE-752 37 Uppsala, Sweden

ISSN 1104-232X
ISBN 91-554-5497-6

Printed in Sweden by Fyris-Tryck AB, Uppsala 2002

Till Far och Mor

Acknowledgements

Since my father has taught me to always start with the most important, when presenting something, I will do so by expressing my sincere gratitude to all those, who in one way or another have helped me accomplish this work.

♡ Ingela, for finding me in the first place, for always being the greatest of support, for the vast amounts of help that you have given me.

♡ Ewert, my supervisor, for always being positive to my wild ideas but still (more or less successfully) trying to keep me on the track.

♡ Lina, the best of workmates, how much more interesting these projects have been, thanks to you.

♡ Gunilla, the boss, for always being there when needed, always ready to answer those vital questions.

♡ Lena, the real boss, making life at CBA feel like at a five star hotel.

♡ Bosse, for being a great soul-mate, to eat late with and to encounter in obscure C++ discussions.

♡ Lennart Björkesten, for finding me a very good project to work on.

♡ Olle, for keeping all computers of the house spick and span.

♡ Felix, always ready for a discussion about glurks of gingersnaps, and everything else for that matter.

♡ Lucia, Ida, Flinken, Rograrna, Stina, Mattiasar, Anna, Micke, Petra, Pettrarna, and all other past and present colleagues at CBA, for creating the best of working environments. Thanks!

♡ Heung-Kook Choi, HyunJu, Sun-Kyung, Byeong-Il, and the other extremely friendly faces on the other side of the world.

♡ The great guys in Cardiff, Gareth, Alla, Stuart, Simon, Dietrich, Nick and Liz. I hope that you'll find good use for the work we've shared.

♡ All my proof-readers, Ingela, Marit, Carolina, Nataša, Bosse, and Ewert, for lots and lots of helpful comments and suggestions.

♡ Marit, for being a wonderful dance partner and a lot more.

♡ Therése, for a well spent dinner.

♡ Patrick, for making me climb the walls.

♡ Nataša, for keeping me smiling during these last demanding weeks :-)

♡ Gabriel, Daniel, Johan, Stein and Henrik, for that joyous summer of '94. I still like to torture people around me with stories from that time.

♡ Algot, for being a part of those stories.

♡ Åsa, Jos, Ludde, Annika and other frinds of Brighton '97, making that spring one of the very best.

♡ Jonas, Anneli, Linda and all other mates of house one.

♡ Kerstin, for making me a better person, and being one of many reasons why this work came to be at CBA.

♡ Jocke B, best friends since forever ...

♡ My big big brother Stefan, the one and only, being the very best brother and friend one can have, in all possible ways.

♡ Mamma Lis and Pappa Pelle, for always believing in me, for all great encouragement, for bringing the wonders of science into my life, whether it being discussions of particle physics while eating "evighetsgrytor", or philosophizing about the nature of cosmos under starry nights aboard Islahåj. Without you I wouldn't be at all, much less a scientist.

Tusen tack!

Uppsala, December 3, 2002

/*Jocke*

Well, now when the important things are said, I guess we can go on with the less interesting part ...

List of enclosed papers

The thesis is based on the following publications, which will be referred to in the text by their Roman numerals:

I. J. Lindblad, C. Wählby, E. Bengtsson and A. Zaltsman. Image Analysis for Automatic Segmentation of Cytoplasms and Classification of Rac1 Activation. *Submitted for publication.*

II. C. Wählby, J. Lindblad, M. Vondrus, E. Bengtsson and L. Björkesten. Algorithms for Cytoplasm Segmentation of Fluorescence Labelled Cells. *Analytical Cellular Pathology*, 24(2,3):101–111, 2002.

III. J. Lindblad, C. Wählby, M. Vondrus, E. Bengtsson and L. Björkesten. Statistical Quality Control for Segmentation of Fluorescence Labelled Cells. In *Proceedings of the 5th Korea-Germany Joint Workshop on Advanced Medical Image Processing*, Seoul, Korea, May 2001. 11 pages.

IV. 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)*, pages 264–271, Bergen, Norway, June 2001.

V. J. Lindblad. Histogram Thresholding using Kernel Density Estimates. In *Proceedings of the Swedish Society for Automated Image Analysis (SSAB) Symposium on Image Analysis*, pages 41–44, Halmstad, Sweden, March 2000.

The author has significantly contributed to the work performed in relation to Papers I–V. Most computer implementations of the included algorithms have been performed by the author. The author has been deeply involved in algorithm discussions and method development which have taken place in connection to the work. Papers I–III have been produced in close cooperation with other researchers, chiefly C. Wählby.

Faculty opponent is Dr. Calum MacAulay, BC Cancer Research Centre, Vancouver, Canada.

Related work

In the process of performing the research leading to this thesis, the author has contributed also to the following publications, significantly so for papers i, iv-vi, and minor parts of papers ii and iii.

i. J. Lindblad and I. Nyström. Surface Area Estimation of Digitized 3D Objects using Local Computations. In *Proceedings of the 10th International Conference on Discrete Geometry for Computer Imagery (DGCI)* volume 2301 of *LNCS*, pages 267–278, Bordeaux, France, April 2002. Springer-Verlag.

ii. C. Wählby, F. Erlandsson, K. Nyberg, J. Lindblad, A. Zetterberg and E. Bengtsson. Multiple Tissue Antigen Analysis by Sequential Immunofluorescence Staining and Multi-dimensional Image Analysis. In *Proceedings of the 12th Scandinavian Conference on Image Analysis (SCIA)*, pages 25–32, Bergen, Norway, June 2001.

iii. C. Wählby, F. Erlandsson, J. Lindblad, A. Zetterberg and E. Bengtsson. Analysis of Cells using Image Data from Sequential Immunofluorescence Staining Experiments. In *Proceedings of the 5th Korea-Germany Joint Workshop on Advanced Medical Image Processing*, Seoul, Korea, May 2001. 10 pages.

iv. J. Lindblad. Perimeter and Area Estimates for Digitized Objects. In *Proceedings of the Swedish Society for Automated Image Analysis (SSAB) Symposium on Image Analysis*, pages 113–117, Norrköping, Sweden, March 2001.

v. C. Linnman-Wählby, J. Lindblad, M. Vondrus, T. Jarkrans, E. Bengtsson and L. Björkesten. Automatic Cytoplasm Segmentation of Fluorescence Labelled Cells. In *Proceedings of the Swedish Society for Automated Image Analysis (SSAB) Symposium on Image Analysis*, pages 29–32, Halmstad, Sweden, March 2000.

vi. J. Lindblad, Y. Sun and E. Bengtsson. Development of a system for real-time quality control of the superconducting LHC cable, In *Proceedings of the Swedish Society for Automated Image Analysis (SSAB) Symposium on Image Analysis*, pages 101–104, Uppsala, Sweden, March 1998.

Table of Contents

CHAPTER 1

Introduction

I keep six honest serving-men
(They taught me all I knew);
Their names are What and Why and When
And How and Where and Who.

— Rudyard Kipling *The Elephant's Child*

Image cytometry is the measurement of cell properties from images. Digital image analysis refers to the extraction of information from images with the aid of computers. We define the combination of these two concepts to be *digital image cytometry*.

This thesis is the result of work in digital image cytometry applied to fluorescence microscope images of cultivated cells. Focus has been on the development and compilation of robust image analysis tools, enabling quantitative measurements of various attributes and properties of cells and cell structures. Large parts of the work have been within two projects performed in cooperation with Amersham Biosciences, in Uppsala, Sweden, and Cardiff, Wales. Most of the work performed by the author has taken place at the Centre for Image Analysis, Uppsala University, Uppsala, Sweden.

1.1 Objective

The objective of the work leading to this thesis has been to perform image cytometry using digital image analysis. This includes applying existing image analysis techniques to the field of image cytometry, as well as, when required, the development of new combinations of processing steps to solve the problems encountered.

A significant part of the work has consisted of developing robust segmentation methods for fluorescently labelled cells. This, in combination with efforts applied in the areas of feature extraction and statistical data analysis, have enabled the compilation of a complete chain of processing steps to produce a system capable

of performing fully automatic segmentation and classification of fluorescently labelled cells according to their level of activation.

1.2 Motivation

Knowledge and understanding of the cell and the functionality of the cell is an essential part of the field of biotechnology. Recent improvements in the methods available for the mapping and description of the cell, the structures of the cell, and the functionality of these, is closely tied to the tremendous progress, chiefly in relation to medicine, experienced in that field. This progress is often referred to as the biotechnological revolution. Digital image cytometry is an important source of information to gain such knowledge about the cell. Applications of this knowledge range from pure research and understanding of the cell functionality, to cancer research and to drug discovery, the latter being one of the main driving forces behind the projects related to this thesis.

The study of genomics and proteomics has advanced our understanding of the field of biotechnology immensely over the past years. The transition from genome to proteome brought with it technological innovations at the diagnostic and functional level. The next important step is the linking of the proteome to the functionality of the cell, something that is related to the role of the cell as the final arbiter in the production of metabolic products. This is the field of cytomics. Describing the intricate pathways of specific cellular functions and their relations to different structures of the cell will be the next frontier in cellular engineering.

Genomics has become an important provider of drug discovery tools, aiding the identification of potential new drug targets. Proteomics has become indispensable in relating structure and function of protein targets to predict drug interactions. The next level of biological complexity is the cell. Efficient high throughput and high content cell assays will become more and more important in order to remove bottlenecks in the drug discovery process. The need for automated acquisition and analysis of multi-dimensional information from living (or dead) cells will then become crucial. An important source of such information is image cytometry. Fluorescence microscopy is a means of getting well specified, informative images of cells and cell structures, and is, as such, a very important modality for image cytometry. Digital image analysis in combination with fluorescence microscopy enables extraction of quantitative information about the cell through the detection and analysis of spatially, spectrally and temporally resolved signals.

CHAPTER 2

Background

Cytomics is the study of the functional relationships between the cell (Cytome) and the metabolic pathways (Proteomics-proteome) resulting from genetic control mechanisms (Genomics-genome). Some relate Cytomics to what is being termed functional genomics, i.e., linking the genome, via the proteome, to the functionality of the cell. A requirement for cytomics is the specialised technologies of cell analysis tools as present in the field of cytometry.

2.1 Cytometry

Cytometry refers to the characterisation and measurement (-metry) of cells (cyto-) and cellular constituents. These measurements may be of the physical properties (length, volume, etc.) of the cell, or of its biochemical properties (protein content, lipid content, etc.), or of a combination of these, e.g., distribution of various cellular constituents. The purpose of such measurements can be all aspects of analytical cytology, whether of biological, diagnostic, or therapeutic values. Two main techniques dominate the field of cytometry as sources of information about the cell. These are flow cytometry and image cytometry.

2.1.1 Flow cytometry

Flow cytometry is the technique of measuring optical properties of individual cells, or particles in general, in a flow stream rapidly passing one-at-a-time in front of a laser beam with detectors measuring fluorescence and light scatter, see Fig. 1. A flow cytometer can measure these parameters at rates of thousands of cells per minute. Being developed in the 1960s and 1970s, the technique has matured into a reliable, reproducible and quantitative method for studies of the phenotypes that compose a heterogeneous cell population. Although largely applied to the analysis of single cells, flow cytometry has some drawbacks. The mapping of functional activities is limited to cells in suspension, and therefore removed from the tissue structure, something that affects the behaviour of the cells and disables studies of inter-cellular phenomena. There is an uncertainty due

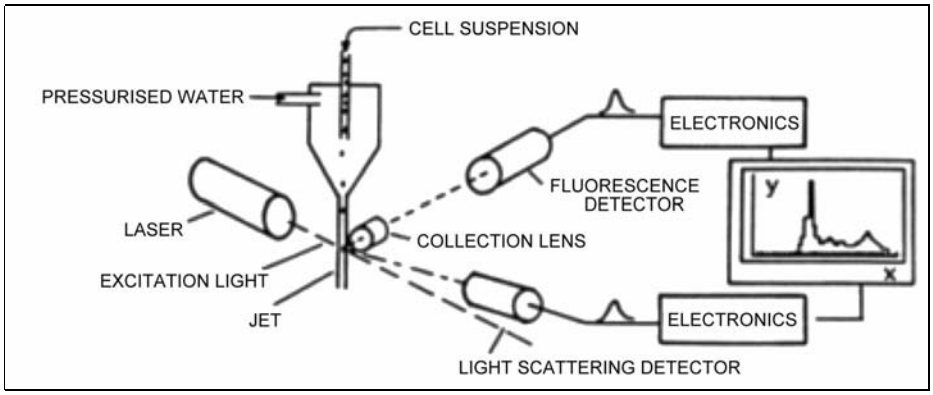


Fig. 1: Schematic representation of a flow cytometer with detectors for fluorescence and light scattering.

to artifacts such as cellular debris and clusters of cells in the cell suspension, and it is difficult to go back and have a closer look at signals which deviate from the normal after the analysis is completed, unless the full experiment is run again. It is also hard to make repeated measurements on the same cell, e.g., as to compare the response of a cell before and after the addition of insulin, as described in Paper I. It is neither possible, with flow cytometry, to define sub-cellular regions and measure shape and location of such, although some information of this type can be acquired from scatter analysis.

2.1.2 Image cytometry

Image cytometry is the measurement of cell properties from images. These can be images from a wide range of sensors, most of them being different types of digital microscopes. Chief among the advantages of this technique is spatial resolution, something that is required for studies involving features such as sub-cellular variations and populations of cells with heterogeneous responses. Most important is also the possibility to study temporal phenomena. Through the use of special stains, biological components of interest can be given specific colours (see Fig. 2) and this way, image cytometry does not only allow us to study the shape and morphometry of the cell itself, but also to study subparts of the cell and most importantly, to study transport of various substances inside the cell. It also enables the study of tissue and inter-cellular phenomena, something that is not possible with flow cytometry.

Qualitative evaluation of images of cells can be done visually. Manual interpretation of microscopic images is, however, a labour-intensive task, prone to error and operator fatigue. It is time-consuming, and consequently conclusions are drawn from measurements of, at best, a few hundred cells. To perform a quantitative evaluation, descriptive information, such as proportions of different regions or numbers of individuals in different classes, has to be extracted from

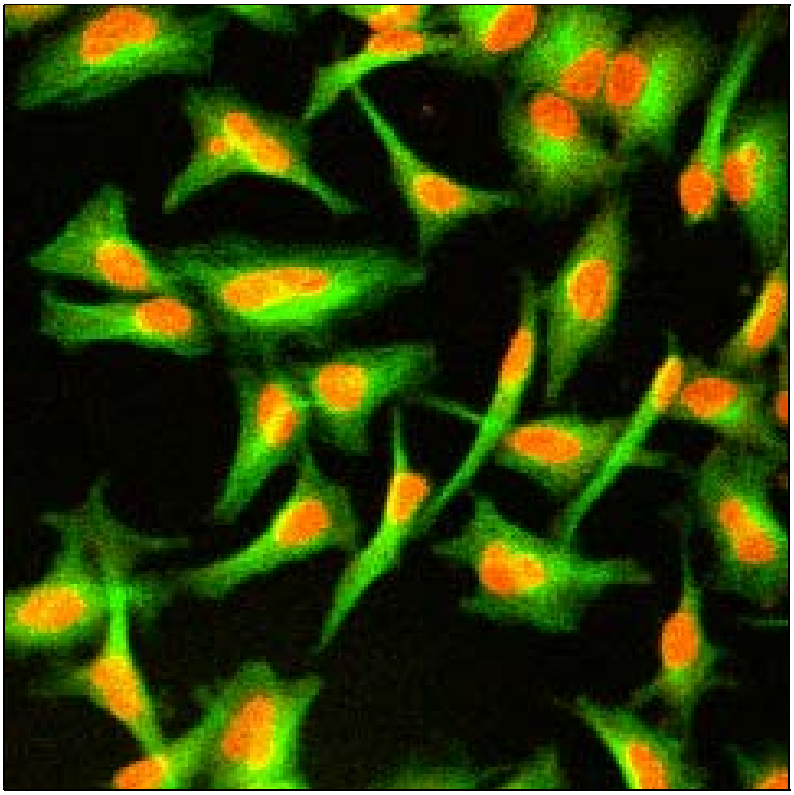


Fig. 2: Fluorescence microscope image of CHO-cells, with different stains for nuclei (red) and cytoplasms (green).

the images. Purely visual estimates of such information tend to be very difficult, and inter- and intra-observer variability is often high, cf. Paper I.

Using digital image analysis, i.e., computer assisted analysis of digitised images, cell populations can, with little or no manual intervention, be distinguished and enumerated, and specific properties of the individual cells can be measured quantitatively. Compared to manual interpretation this gives improvements in speed, reduction of cost, and reduction of subjectivity. To emphasise the use of computer algorithms to perform quantification and interpretation of cell properties from digitised images, we add the word digital to the term image cytometry, leading to the name Digital Image Cytometry, which is what this thesis is about.

2.1.3 Digital microscopy

Central to cytometry in bio-imaging is digital microscopy, i.e., the combination of a digital imaging device, e.g., a digital camera, and a microscope. This facilitates the direct production of digital images of biological states and processes.

These images may be acquired from optical emission and transmission microscopes, as well as from scanning and transmission electron microscopes. Image processing and analysis are tightly coupled to digital microscopy, in order to achieve quantisation, three-dimensional (3D) restoration, and object and pattern recognition. Digital microscopy adds high-resolution spatial and temporal dimensions to cytometry, something that is missing in flow measurements.

2.1.4 Confocal microscopy

In a conventional microscope set-up, not only is the plane of focus imaged, but a large part of the specimen above and below this point is also illuminated and imaged at the same time. This results in out-of-focus blur from these areas above and below the plane of interest, leading to a reduction in image contrast and a decrease in resolution.

Confocal microscopy is based on the principle that out-of-focus information can be rejected by the presence of pinholes in front of the detector and the light source. This implies that essentially only the part of the specimen that is in the focal plane of the microscope is imaged. Light collected from outside the focal plane is rejected by the pinhole. In addition, the beam of illuminating light diverges above and below the plane of focus, so that elements outside the focal plane receive less light. This further reduces the amount of out-of-focus information.

In this way an optical slice of an object may be imaged. This gives better contrast and enables higher resolution. In addition, if it is possible to scan the focus during the imaging, many slices of the object, at different focal depths, may be imaged. In this way high resolution, 3D images may be built up of a variety of specimens in a non-destructive fashion.

2.1.5 Fluorescence microscopy

A source of information about the cell that is gaining more and more in importance is the detection and analysis of fluorescence emitted when the cell has been exposed to fluorescence dyes attached to biologically active molecules and is exposed to excitation light. This is fluorescence microscopy, which is a technique that can be used to detect structures, molecules or proteins within the cell. Fluorescent molecules absorb light at one wavelength and emit light at another, longer wavelength. When fluorescent molecules absorb a photon of a specific energy (the absorption wavelength) for an electron in a given orbital, the electron rises to a higher energy level, the excited state. Electrons in this state are unstable and will return to the ground state, releasing energy in the form of light and heat. See Fig. 3. This emission of light is fluorescence. Because some energy is lost as heat, the emitted light contains less energy and therefore is of longer wavelength than the absorbed excitation light. In fluorescence microscopy, cells are stained with a dye and then illuminated with filtered light at the absorption wavelength; the light emitted from the dye is viewed through a filter that allows

only the emission wavelength to be seen, thus making the dye glow brightly against a dark background.

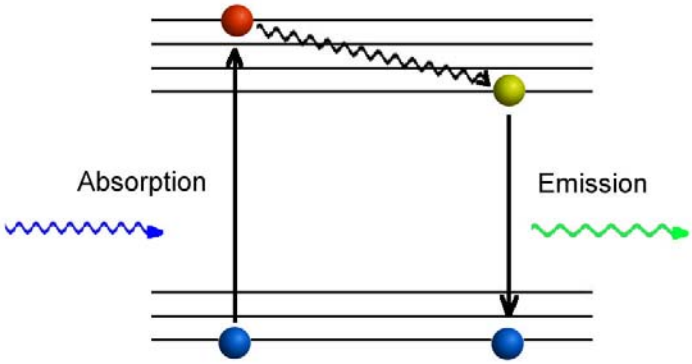


Fig. 3: Simplified Jablonski diagram depicting the fluorescence phenomenon.

The wavelength of light emitted is material dependent, and likewise the excitation light wavelength depends on the material. Fluorescence is always of a lower energy, and, hence, longer wavelength, than the exciting light. This separation in wavelength is known as the Stokes shift. The Stokes shift enables the exciting and emitted light to be separated by optical filters and thus the amount of fluorescence can be quantified.

Fluorescence microscopy allows you to attach fluorescent dye molecules to specific parts of your sample, so that only those parts are the ones seen in the microscope. You can also use more than one type of dye. By changing the excitation light, you can cause one type of dye to fluoresce, and then another, to distinguish two different parts of your sample. Fluorescent markers that bind specifically to a cellular component (e.g., DNA) may be added. Alternatively, stains that are excluded by living cells, but taken up by dead cells may be used. In this way many parameters can be measured simultaneously.

Staining methods

Several methods can be used to impart fluorescence to a cell. The simplest is autofluorescence, such as in chloroplasts, when cell parts fluoresce without any treatment. Next, different fluorochromes are available, such as 4',6-Diamidino-2-phenylindole (DAPI), which can bind directly to cell parts and organelles. It is also possible to attached fluorochromes to DNA, and use this as a probe that will hybridise with chromosomal DNA. This technique, called fluorescence in situ hybridisation (FISH), provides a sensitive method for the identification of, e.g., chromosomal abnormalities.

Through the process of transgenic engineering, cells can be created which express fluorescently labelled protein in vivo. This is one of the most powerful techniques of fluorescence microscopy, since it allows the cells to be examined while still alive with a minimum of interference.

Fig. 4: Aequorea victoria (illuminated with white light) and a rendering of the structure of GFP.

Green fluorescent protein (GFP) is a remarkable protein originally isolated from Aequorea victoria - a luminescent jellyfish (Fig. 4, left). It is a fluorescent protein that absorbs blue light and re-emits it as green fluorescence (Fig. 4, right). The gene coding for GFP has been isolated and has become a very useful tool in cellular and molecular biology research [9, 41]. The DNA sequence for GFP can be inserted into the DNA of an organism in such a way that GFP is produced whenever the gene of interest is expressed. The ability to detect very small amounts of a compound by fluorescence microscopy enables a very sensitive assay for gene expression and the localisation of gene products. One may thus have an in vivo fluorescent protein which may be followed in a living system (see, e.g., Paper I). Numerous modified varieties have been produced for optimised expression in many organisms. In addition, colour variants have been isolated which provide the opportunity for dual-labelling studies.

Immunocytochemical staining is another useful method for visualisation of structures in the cell. Here, fluorochromes are attached to antibodies, as depicted in Fig. 5. The structure that the antibody reacts with will then fluoresce when exposed to excitation light. This gives us a way to measure the exact location of structures reacting with specific antibody. This type of staining can be used to identify a wide variety of molecular epitopes (a molecular region on the surface

of an antigen capable of eliciting an immune response and of combining with the specific antibody produced by such a response), representing an equally wide range of cell functions. Attaching the fluorochrome label directly to the primary antibody provides little signal amplification and is therefor rarely used. By instead labelling a second antibody that is directed against the primary antibody, we get a more versatile method, since a variety of primary antibodies from the same species can be used with the same conjugated secondary antibody. It will also give a stronger signal, as several secondary antibodies are likely to react with one primary antibody, thus amplifying the signal. Three-step methods also exist, giving further increase in staining intensity.

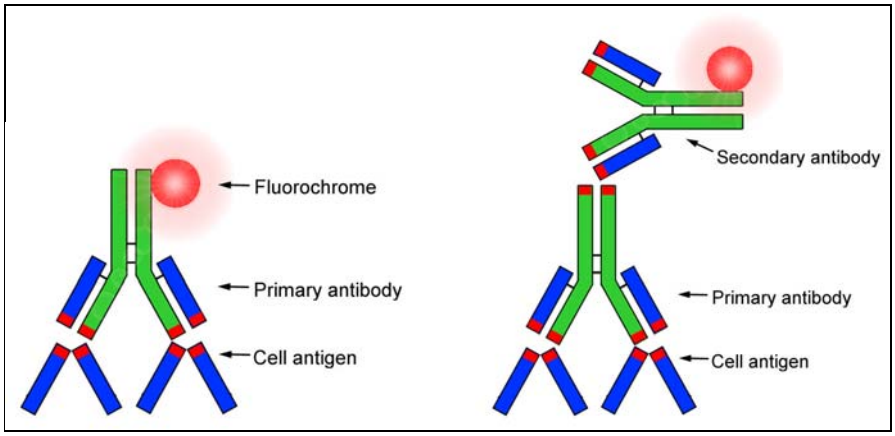


Fig. 5: Direct (left) and two-step indirect (right) method for antibody-fluorochrome immunocytochemical staining.

Fluorescent labels, in combination with narrow band pass filters for limitation of fluorochrome excitation and emission, enables quantitative imaging of several antigens in parallel in a single multiply stained specimen. Double and triple stains have been used to observe as many as three different antigens simultaneously [16]. However, many problems occur when a large number of antigens are stained at the same time. The different primary and secondary antibodies can interact and "bleed-through" of contrasting emission colours as well as autofluorescence of the stained material can disturb the obtained images. An alternative approach is sequential staining [ii,iii], which could greatly increase the number of antigens possible to study simultaneous.

2.2 Digital image analysis

As you read this text, an image of this page is formed upon your retina. Your eye, and other components of the visual system, form a biological device for capturing

images. Your brain performs the task of image analysis, when it interprets and analyses this image to extract useful information from it, most notably in this case, the words of the text.

Although the visual system is superbly adapted to making qualitative judgements about features present in the image, it is not well suited for making quantitative judgements. For example, your visual system may tell you that an object in your field of view is very big, or that it is very bright, but it cannot tell you exactly how big or how bright it is. Computers on the other hand are very good at making quantitative judgements, something that is often required in the scientific world. Unfortunately, they are not that good at making qualitative judgements.

Image analysis refers to the extraction of information from images. Digital image analysis is the extraction of quantitative information (i.e., numerical data) from digital images with the aid of computers. In this thesis we will, however, quite often omit the word digital, although referring to computer assisted image analysis. Hopefully, it will be clear from the context what is meant.

2.2.1 Digital images

A digital image is a computer representation of an image. Digital images are not continuous, but are made up of large numbers of elements, each of which can take a value from a limited range, which describes its appearance. This value is often a grey level or a colour, but could also be a height or a pressure, for example. A digital image represents an image as an array of numbers, assigning a value to each position in the image. The array may be 2D, e.g., for a representation of an ordinary black and white photograph, but can also be 3D for volume images (imagine a stack of images). If we have a time series of images, we can interpret this as one, higher dimensional image, where time is an additional dimension. Also, we can have images representing light at different wavelengths (colours). Often a colour image is made up of three images*, one red, one green, and one blue image. However, there is nothing restricting us to only three wavelength bands, and if we have a good sampling of this wavelength space, we can think of this as yet another dimension, leading us to 4D or 5D images. The use of such high dimensional images is, however, not that common yet, as the very large amount of information contained soon becomes unmanageable (perhaps in a not too distant future though). Most images used in this thesis have been 2D images, but also 3D volume images (from confocal microscopes), and time series of images have been used.

Each element of a 2D array representing a digital image is called pixel, a short form for picture element. The corresponding name for a 3D image element is voxel, for volume picture element. The number of pixels in an image determines the limit of resolution of the image, and the range of values that each pixel can take determines the limits for contrast and brightness. The elements of the image

* This corresponds to the fact that the retina of our eyes have three type of cones for sensing colour.

are usually arranged in a rectangular grid, the images used in this thesis will all be of this type, but other grid configurations exist (hexagonal, body-centred cubic, etc.), and are sometimes preferable.

Manipulation of images is often referred to as image processing, where the input is an image and the output is also an image. Image analysis on the other hand deals with the computation of image properties; the blood cells in a cell image, their image locations, object areas, etc. Thus, the input to a digital image analyser consists of an image and the output is comprised of a set of numbers representing certain image properties.

2.2.2 Image analysis methodology

Although image analysis problems range over a very large scope of different applications, the problems encountered stay fairly much the same, independent of the application and the specific task that is to be performed. This in turn leads to that the approach taken when solving one image analysis problem much resembles the approach taken when solving another, although the specific algorithms used may be very different.

The methodology we have used to accomplish the tasks of the projects of this thesis can be summarised in the following steps.

Sample preparation What is done before the images are acquired. In relation to microscopy this could be cell culture preparation, staining, etc.

Image acquisition The depicting of the objects of interest, creating the digital images we are to work with. The higher the quality of the images, the better the results of the analysis.

Image pre-processing Processing the raw image so that it is best suited for further analysis. This could, e.g., be image registration of 2D slices to a 3D volume image, or shading correction to compensate for imperfections in the acquisition step.

Image segmentation One of the most important, but also one of the most difficult tasks of image analysis is the one of segmentation, i.e., to find and outline the individual objects of the image.

Feature extraction To extract descriptive quantitative features from the images. These feature measures will then serve as the foundation for the data analysis.

Data analysis Transforming the direct measures of the image, the feature values, into something more comprehensible for the end user. This often involves dimensionality reduction and/or classification.

Evaluation Last, but by no means least important, for any scientific study there has to be an evaluation of the result. How good the method is, how well it performs in comparison to alternative methods.

The names of the different steps vary throughout the image analysis community, (someone mostly involved with image acquisition will probably not call, what here is referred to as image pre-processing, by that name, but rather image post-processing) the contents will, however, stay much the same. The algorithms that we have used in the different processing steps are described in Chapt. 4.

CHAPTER 3

Projects

It's a dirty work, but someone's got to do it.
—

This thesis deals with problems related to cell image analysis, i.e., using digital image analysis to extract information about cells from images. The work constitutes a contribution to the field of digital image cytometry. The work presented consists mostly of combinations of existing methods from different areas of the image analysis world applied to the field of image cytometry, but also, of improvements of existing algorithms, development of new algorithms, and novel combinations of algorithms. The algorithm compilations, together with the evaluations of robustness and performance on real-world applications, that have been carried out, would hopefully be valuable for the scientific community.

During the time of the doctoral studies, effort has also been put down into more theoretical works in the areas of digital geometry and more specifically on estimators of length and surface area of objects. The inclusion of these works would make this thesis a bit too heterogeneous and are therefore only briefly mentioned in this text.

Large parts of the work leading to this thesis have been performed within two projects, aimed at the development of algorithms for segmentation and quantification of properties of fluorescently labelled cells. The first project, funded by Amersham Pharmacia Biotech, Uppsala (now Amersham Biosciences), aimed primarily at developing methods for the fully automatic segmentation of cytoplasms of fluorescently labelled cells. The second project, funded by Amersham Biosciences, Cardiff, was set up both to study the feasibility of real-time analysis of fluorescence images, and to quantify the activation of Rac1-GFP, induced by the addition of insulin-like growth factor 1 (IGF-1) in living cells.

Both these projects have been performed together with Carolina Wählby (*née* Linnman) and my supervisor, Ewert Bengtsson, the first one also together with Mikael Vondrus. Close cooperation with Amersham Biosciences has been of great value. The author has not been involved so much in the biology and microscopy leading to the images, but more so at the steps taken once the images were acquired. Following are short descriptions of the projects the author have

been involved with during the work of this thesis, starting with the ones most significant for this text.

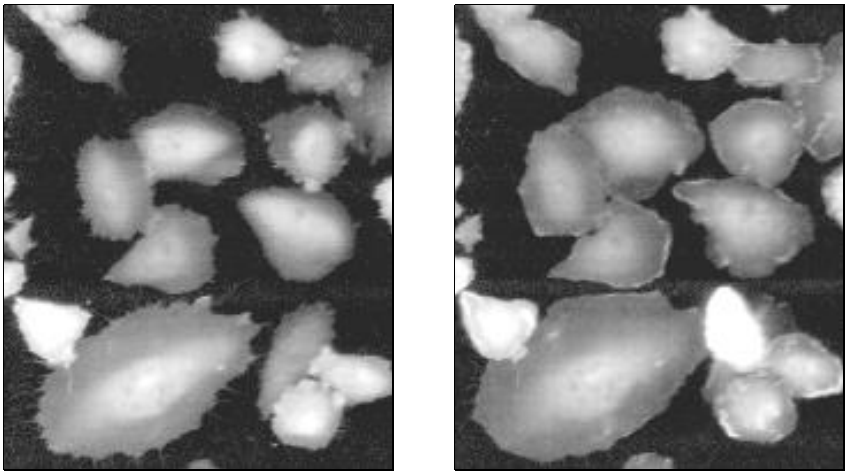


Fig. 6: CHO-hIR cells expressing GFP-Rac1 fusion protein, imaged on IN Cell Analyzer. Cytoplasms, before (left) and 4.3 min after (right) incubation with IGF-1. Ruffles appear as bright formations along the edges of the cells.

Algorithms for automatic segmentation of cytoplasms and classification of Rac1 activation

Rac1 is a small GTP-binding molecule that is involved in a wide range of cellular processes, such as actin reorganisation, cell cycle progression, gene transcription, cell adhesion and migration, and therefore of great interest to study. The aim of this project was to develop algorithms enabling fully automatic, real-time segmentation and analysis of fluorescence images of cells so as to quantitatively estimate the IGF-1 induced translocation of GFP-Rac1 to the cellular membrane for individual cells. See, Fig. 6. Due to the ultimate goal, of analysing one image containing roughly 200 cells in less than two seconds, effort was taken not to use algorithms of high time complexity. Digital image analysis methods for segmentation of cell nuclei and cytoplasms, feature extraction, and classification of cells according to their activation, i.e., GFP-Rac1 translocation and ruffle formation at stimuli, have been presented. Selected general and problem specific features are extracted from the segmented cells and non-informative features are removed from the feature set, using a floating feature selection procedure. Based on training data, from visual annotation of a set of images, a classifier is created. The results of the classification made by the computer have been compared with results from visual inspection of the time sequences. The automatic classification

differed from the visual classification at about the same level as visual classification performed by two different skilled professionals differed from each other. The results of this project have been presented in Paper I and at a seminar at the Amersham Biosciences site (Maynard Centre) in Cardiff, Wales, in October 2002.

Algorithms for segmentation of fluorescently labelled cells

The interaction with and effect of potential drugs on living cells can be observed by fluorescence microscopy. Automated methods for feature extraction from fluorescence microscopy images of cells can be used as a tool in the drug discovery process. The cell nucleus has a well-defined shape and is relatively easy to detect. The cytoplasm is, however, more complex. The goal of this project was to develop a fully automatic method for cytoplasm segmentation. The suggested algorithm, inspired by literature and previous experience, consists of an image pre-processing step, a general segmentation and merging step followed by a quality measure and a splitting step. By training the algorithm on one image, it is made fully automatic for subsequent images created under similar conditions. This method was presented at an internal Amersham Pharmacia Biotech R&D conference in Uppsala in late 1999. During 2000 the algorithms were improved through a more elaborated shape analysis and a more consistent feature extraction and quality evaluation step. The results were documented in Paper II, with a more detailed description of the quality measure feedback system in Paper III.

Estimation of intensity nonuniformities in 2D and 3D microscope images

Intensity nonuniformity (INU), also commonly referred to as shading artifacts, is a common problem in digital image analysis. Although the INU has limited impact on visual interpretation, it may have a negative effect on image segmentation, as well as on the interpretation of image intensity values. There exists a range of methods to perform correction of INUs. This project is aimed both at evaluating which methods perform best and under what conditions, and toward finding possible improvements to the existing methods. Although trying to stay general, the project is biased toward the applicability to microscope images, mainly of fluorescence stained cells. Under the scope of this project a method to reduce the parameter space of a B-spline based iterative shading estimation algorithm has been suggested and evaluated. The results of this evaluation were presented at the 12th SCIA conference in Bergen, Norway (Paper IV). The developed method has proved to be robust and has been used with good results in Papers I–III.

Methods to perform automatic selection of optimal or near optimal threshold values

The simplest and by far most popular method of separating data into different classes is by thresholding. Although not very elaborate and often too abrupt, it is

robust and performs well in many different situations. The problem of selecting thresholds is very general and applicable in a vast range of problems. Still, there exist no satisfactory general solution to the threshold selection problem. This project addresses the task of finding robust all-purpose methods for selecting good thresholds in a general distribution. Inspired by statistics, the use of kernel density estimates (KDE) to threshold distributions at locations of high second derivative has been investigated and shown to be useful. The difficult problem still to be solved relates to optimal selection of the width of the kernel function for the estimate. The method was originally developed to separate cells of different cyclin contents [16]. The further refined KDE approach was presented at the SSAB conference in Halmstad, March 2000 (Paper V). The thresholding technique developed in this project has proved to be most useful in the INU estimation method described in Paper IV, as well as for direct data analysis, as performed in papers ii and iii, and [68].

Accurate and precise size estimators for digitised 2D and 3D objects using local computations

The part of my work that has been connected to this project, is deliberately excluded from this thesis, in an effort to make it more homogeneous and comprehensible for the reader.

Information is irrevocably lost in the process of digitising a continuous object of the real world to fit the digital world of the computer. Therefore, feature measurements of digitised objects can be no more than estimates. Good estimators are those that approach the corresponding feature value of the continuous original object. The possibility to use only local computations is a desirable property in digital image analysis, both to keep the complexity level at a minimum, and to enable parallelism in various ways. This project aims at finding good local estimators for size related measures of digitised objects, i.e., perimeter and area of 2D objects, and surface area and volume of 3D objects. Statistical validation of the estimators have been performed on large numbers of computer generated digitised objects. The breakdown behaviour at very low resolution, as well as the asymptotic behaviour at high resolution have been studied. The 2D estimators were presented at the SSAB conference in Norrköping, March 2001 (paper iv). A surface area estimator, inspired by the Marching Cubes algorithm [34, 71], with improved precision and accuracy obtained by optimising the area contribution locally, was presented at the DGCI 2002 conference in Bordeaux (paper i).

Inspecting the LHC superconducting cable

In the beginning of my doctoral studies, a considerable amount of work was also put down into this project, which was an extension of my Masters Thesis work [32].

In the development of the next generation elementary particle research accelerator at CERN, Geneva, the Large Hadron Collider (LHC), 6,000 km of superconducting cable has to be manufactured. This cable has to be "perfect" without any crossing strands or other irregularities. In order to ensure this, the

cable is to be inspected through digital image analysis techniques. This project consisted of the assembly of an image analysis system capable of performing real-time analysis of the cable as it passes in front of a line scanning camera, when produced at the manufacturing plant. This work has included the writing of special purpose image analysis routines, to find imperfections in the cable, as well as routines to interface with the camera and a devices to measure the speed of the cable. An operator interface was designed to enable user interaction and supervision of the process. Due to the stringent time requirements of processing about 2000 scan-lines per second, the image analysis was done in a two-step procedure, where the first part of the analysis was done on one computer and images of suspicious parts of the cable where sent over to another computer, where further analysis took place under more relaxed time requirements. The operator interface is also run on this second computer. In the end of 1998 this system was tested at CERN. The project is described in paper vi and summarised in the technical report [33].

CHAPTER 4

Material and Methods

Do, or do not. There is no try.
— Yoda *The Empire Strikes Back*

This chapter will give a brief summary of the methods used throughout the work leading to this thesis.

4.1 Sample preparation

Before images can be acquired, one must, of course, decide what to image. Careful planning and selection of imaging modalities, staining methods, etc., can significantly simplify the following image analysis part. If we know what we wish to analyse, putting down effort in finding the best way to capture vital information, without introducing a redundant or disturbing background, is often beneficial.

The material that we have worked on has mainly been fluorescence images of Chinese hamster ovary (CHO) cells. That is, cultured cells grown on glass, as opposed to tissue samples. For the first project, which was aimed at cytoplasm segmentation, we tested the algorithms on two sets of images showing CHO-cells cultivated for 24 hours and stained with calcein dietoxy methyl ester, an indicator of living cells.

For the second project, CHO cells, stably expressing the human insulin receptor (CHO-hIR cells), stably transfected with GFP-Rac1 construct were used. The cells were seeded in a 96-well plate and after 24 hours of incubation at 37°C, the agonist; insulin-like growth factor 1 (IGF-1), was added, activating the relocation of GFT-Rac1 to the cell membrane. To simplify the segmentation, (see Sect. 4.4), Nuclei dye Hoechst^{TM}33258, was added to each well 30 min prior to imaging.

4.2 Image acquisition

The first sets of images were acquired by fluorescence microscopy and a CCD-camera equipped with a narrow green filter. A 10× objective and a 20× objective

Fig. 7: The IN Cell Analyzer (Amersham Biosciences), the microscope that the Rac1 images where acquired on.

were used for the first and second set of images respectively. One image at each magnification was used as a training image for parameter optimisation. These parameters were then used in the segmentation of the similar test images. All objects cut by the image border were removed from analysis to simplify matters.

For the second project, time series of images of the cells were taken before and after the addition of IGF-1 on an IN Cell Analyzer (Amersham Biosciences, Cardiff, Wales), automated confocal high-speed fluorescent system (Fig. 7). Sequential excitation at 488 and 364 nm was used and green and blue emissions were collected on two CCD cameras. Images of each well were taken at regular intervals, over a time period of 7 minutes after cell stimulation with IGF-1. Each image of 1280×1280 pixels represents a 0.75 mm$\times$0.75 mm field of view.

4.3 Image pre-processing

What we collect under the heading image pre-processing are the steps taken to make the acquired images look the best for the further image analysis to

be performed, i.e., to reduce the effect of undesired imperfections of the image acquisition. Depending on the quality of the images acquired, the segmentation method used, and the features extracted from the images, there are more or less requirements for pre-processing of the images.

Pre-processing can, e.g., be smoothing of the image to reduce the effect of noise in the image, or deconvolution filtering to reduce the effect of smoothing. Filtering techniques not primarily intended to make the image look better does not belong here. For example, when using scale space based methods, smoothing is often a part of the algorithm, and is then no longer a pre-processing step. Registration of time-series of images or 2D slices of a 3D volume image is another example of pre-processing steps.

The major pre-processing that has been applied in this thesis, is for the correction of intensity nonuniformities (INU) present in the images. If we wish to compare intensity measures from different parts of the image, e.g., by using a global threshold, it is very important that no large scale variations in intensity are present in the image. Such intensity variations are, however, often present in microscope images, as it is quite difficult to achieve perfect uniformity in illumination of the specimen.

The impact of INU on visual interpretation of images is limited. It does, however, affect all image segmentation methods that, in some way, rely on absolute intensity values. Moreover, INU strongly affects the possibility to use intensity values of the segmented images as quantitative measures in further data processing.

4.3.1 INU estimation using additional images

Estimation of the INU field (often called bias field) can either be based on specialised acquisition protocols, such as imaging of homogeneous phantoms or black-level images in relation to the ordinary imaging step, or be based on analysis of the image data itself. Whether it is feasible to acquire a good background image directly, in conjunction with the imaging procedure, or not, is very much dependent on the imaging environment. Data-driven approaches are usually simpler from a practical viewpoint, but may introduce artifacts if the algorithm used cannot reliably solve the estimation problem.

Given that the uneven background is due to uneven illumination and the intensity of the fluorescence in each cell is a linear function of the light illuminating the cell, the usually accepted illumination correction is:

$$I_{cor} = (I_{raw} - I_{dark})/(I_{blank} - I_{dark}), \tag{1}$$

where I_{dark} is an image without the illumination and I_{blank} is an image from a blank part of the slide with illumination.

Using a global threshold for segmenting the image can be very sensitive to variations in intensity over the image, especially if the contrast between foreground and background is low, as is often the case, e.g., for cytoplasm images. Experience has shown that in such cases, even if shading images are available,

the application of a data-driven algorithm can be beneficial, to fine tune the background level after the initial illumination correction, and to compensate for imperfections in the shading images, see, e.g., Paper I.

4.3.2 Data-driven INU estimation

A number of authors have proposed image filtering (often homomorphic filtering [21]) as a means to estimate the INU field [8]. The disadvantage of the spatial filtering methods is that they generally rely on the assumption that the frequency spectrum of the INU field and that of the true image are separable, which is typically not the case. As such, spatial filtering tends to introduce severe, undesirable filtering artifacts.

Rolling ball algorithms may perform better than just a flat minimum filter, but they do have the same inherent properties and corresponding problems as a minimum filter (dark background case). If the filter is too small, it will include the foreground objects. If the filter is too large, it will smooth the background. Also, using the extreme values, as the rolling ball and minimum filter approach do, is noise sensitive. A better approach is to use a small percentile, e.g., the 5% percentile. Unfortunately, the choice of an optimal percentile is dependent on the amount of noise present in the image [Paper IV].

Another approach is to try to estimate the background field as a smoothly varying function [6]. Many different functions have been proposed in the literature, but the most popular is probably the class of uniform cubic B-spline functions (see, e.g., [30]). Cubic B-spline functions have many nice properties, they guarantee $\mathcal{C}^2$ continuity, i.e., they are always smooth, and different levels of flexibility can be allowed by varying the number of control points used for the function. The task of fitting a B-spline function to data can be formulated as an ordinary least squares problem, and is, thus, analytically solvable in an efficient way, which is, of course, also a very nice property.

The approach taken by the data-driven shading correction methodology described in Paper IV, is to iteratively make improved B-spline approximations of the bias field present in the image. We have used this method in Papers I and II with good results.

The algorithm works by iteratively improving the estimates of the background of the image. The background is assumed to be smooth and slowly varying, and a cubic B-spline surface should therefore be suitable as a model for the background shading. A surface patch S of the background, is modelled by a tensor product of spline functions. That is, a surface point $S(u, v)$ will be written as

$$S(u, v) = \sum_{kl} B_k(u) B_l(v) x_{kl}, \tag{2}$$

where B_k are the B-spline blending polynomials and the x_{kl} are the control points of the surface. The number of control points for the B-spline surface defines how flexible it is. If the background varies a lot, many control points are needed, but in most cases only a few are enough. 5×5 evenly spread control

points has demonstrated to work well in all of the tested images of Papers I and II.

To get a first estimate of the background, the spline surface is initially fitted to the whole image. The distance between the spline surface and the background is minimised by least squares regression. This first estimate will give a too bright image of the background, as it also includes the brighter cells into the background. This background is then subtracted from the original image to get a first estimate of a background compensated image. This image is then thresholded, so as to separate the foreground from the background, and the foreground is masked away. The thresholding can, e.g., be at a fixed level [19], but better is to use a constant times the estimated standard deviation of the background (see Paper IV), or, to totally avoid the need for constants, the point of maximal change of trend in a kernel density estimate (KDE) of the intensity distribution (see Paper V and also Sect. 4.4).

The second iteration also starts with the original image, but this time the spline surface is only fitted to the pixels that have not already been masked away as foreground pixels. Therefore this second estimate will be a little bit better than the first estimate of the background. Once again this new background is subtracted from the original image and more foreground pixels are found and masked away. This iterative procedure continues until the maximal change in pixel value between two successively calculated backgrounds is less than half the original quantisation step of the image. Convergence is fast and the stop criterion is usually reached after 4–10 iterations. Fig. 8 shows an image cross section after 1, 2, and 4 iterations of the algorithm.

Visual inspection shows that this algorithm performs well on all tested images. See, Fig 9 The evaluation performed in Paper IV indicates that the performance of the algorithm also stays fairly robust under varying imaging conditions. When comparing homomorphic filtering with the above described iterative B-spline fitting algorithm, for estimation of shading artifacts in fluorescence images, the B-spline estimate outperformed the optimally tuned spatial filtering in all but one case [Paper IV]. Using a kernel density estimate technique to find the threshold needed in the B-spline estimation algorithm seems to be a fairly good choice. It is not always optimal, but in general not far behind the optimal threshold choice and sometimes even outperforming it (this is possible since the automatic thresholding scheme changes the threshold dynamically during the iterative process).

In Paper I, the above algorithm is not only used to reduce the variations of the background, but the internally calculated foreground/background threshold is also used for the segmentation of the cytoplasms (see Sect. 4.4).

4.4 Image segmentation

One of the most important, but also one of the most difficult tasks of image analysis is the one of segmentation. That is, to find and outline the individual objects of the image.

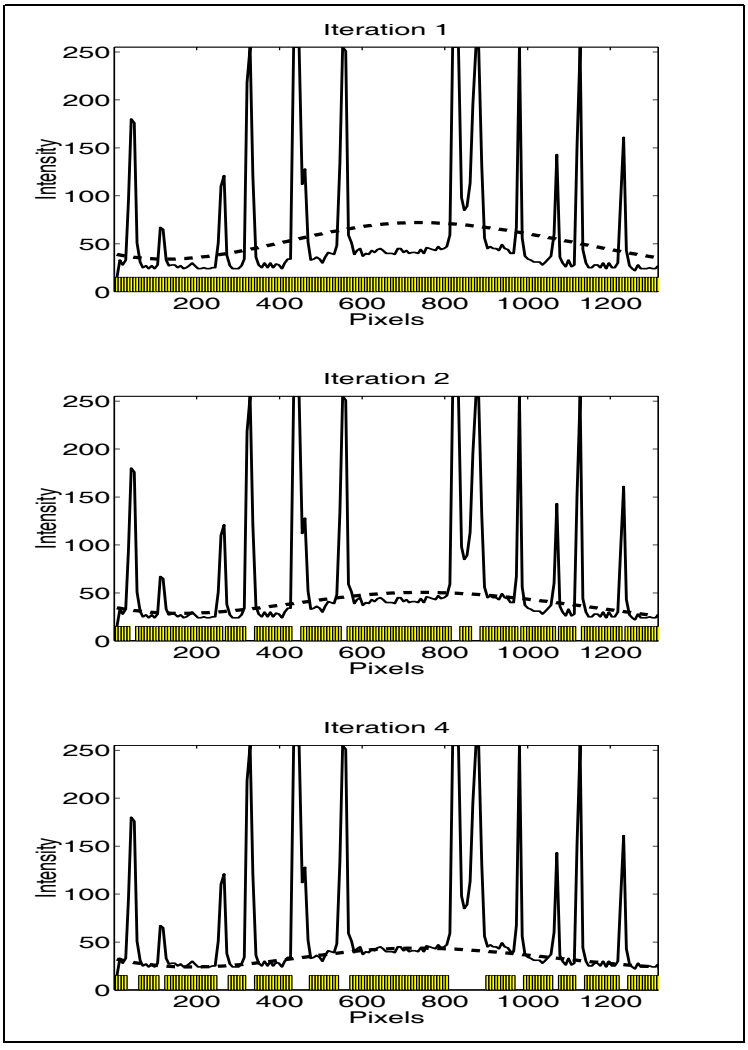


Fig. 8: An intensity profile across one image with fitted background (dashed) after the first, second, and fourth (final) iteration of the background approximation algorithm. At the bottom of each graph is indicated which pixels from the original image that have been used for fitting the spline surface to the background.

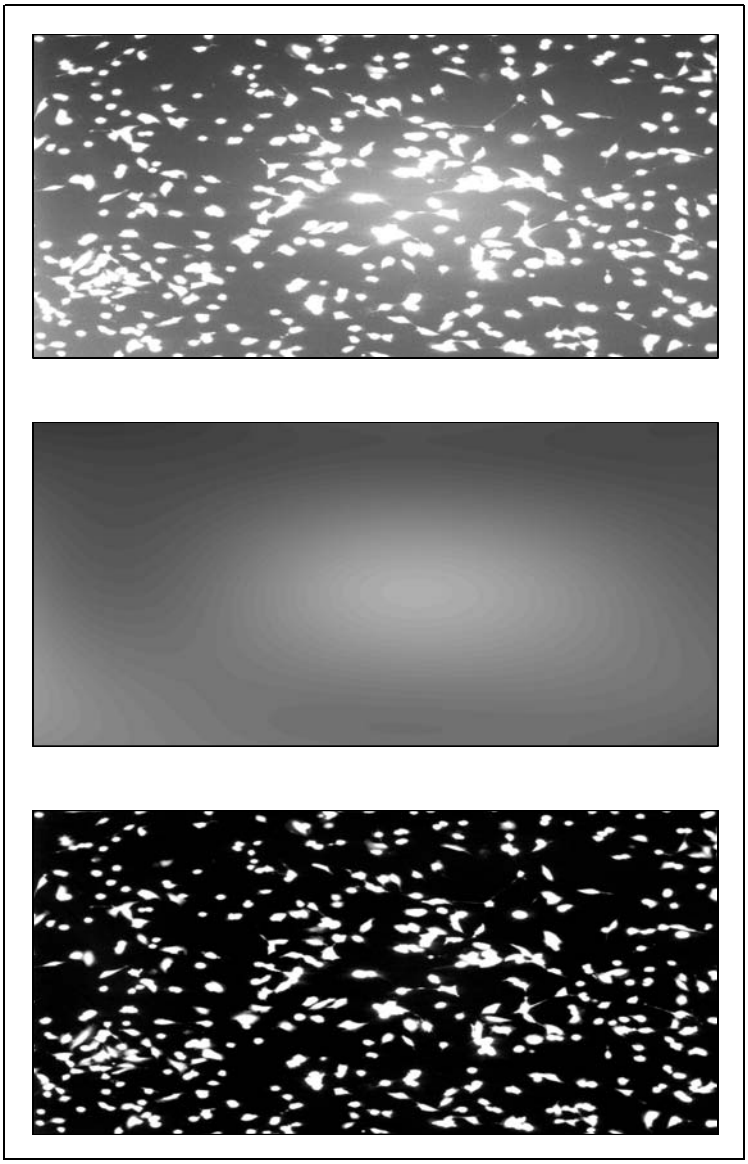


Fig. 9: Example microscope image with fitted background. Original image before background correction (top), the fitted spline surface of the background (middle) and the image after background subtraction (bottom). The intensity scale (same for all images) has been set to enhance the contrast of the darker parts of the images.

The segmentation problem has been present since the beginning of image analysis, and still no good general purpose solution has been found. In this section we will just briefly touch upon some of the methods that we consider relevant for image cytometry, and describe in a little more detail the ones that we have used in the work leading to this thesis.

Most segmentation methods are either region based or edge based. Edge based methods consist of finding and tracking edges and borders in the image, and in some way connecting them to create enclosed regions. Region based methods, on the other hand, usually find connected regions of foreground and split these up into individual objects. Of course, there also exist combinations of the two. Edge based methods suffer from the fate of being inherently built on contrast, i.e., differences or derivatives, in the image. This makes them more noise sensitive by definition. On the other hand, edge based methods are better to cope with large scale variations in the image, such as shading or illumination variations. When using a purely edge based method there is the non-trivial problem of connecting the edges to form connected objects. Therefore, edge based methods are often combined with region based methods, to decide what is object and what is background. Instead of connecting the edges when their correct positions are found in the image, one can start with connected edges and then try to find their correct position, as is done, e.g., when using snakes or active shape models.

The method that is most suitable for a specific situation is highly image dependent. If the edges of the objects are of high contrast, an edge based method may produce good results. However, if the borders are fuzzy, but the interior of the objects looks significantly different from the background, a region based method may be preferable. One should also notice that it is possible to apply a region based method to edge image data by first applying, e.g., a distance transform to the edge image, making it more suitable for a region based segmentation.

Depending on the imaging conditions and the staining methods applied, different methods may be more or less well suited for cell analysis. One of the more difficult problems encountered in the work of this thesis has been the one of cytoplasm segmentation. As preliminary studies showed that with the stains used, there was not enough edge contrast for reliable edge based segmentation, work was on an early stage led in the direction of region based segmentation; something that proved to be useful.

The time complexity of the methods may also be of importance. In this thesis we aim at segmenting hundreds of objects in less than a second, somewhat restricting the number of applicable methods. This is the main reason why we will not discuss things like iteratively refined active shape models, which tend to be orders of magnitude slower than the more direct methods that we have applied.

4.4.1 Segmentation into foreground and background

One important part of segmentation is to separate the image into foreground and background, i.e., to binarize the image. For the edge based methods, we

trace a border between the two, for the region based ones, we try to find a property that separates the two in a well-behaved manner. Note also that in the field of fuzzy set theory [72], the foreground/background binarisation is fuzzified into a sometimes better posed problem of assigning membership values. Fuzzy approaches are not covered in this thesis.

If the image is, in some sense, homogeneous over the field of view, foreground/background segmentation can be achieved with a thresholding operation. In the simplest case, saying, e.g., that everything in the image that is brighter than the threshold is object and the rest is background. When using a global threshold for the whole image the shading correction is crucial. Alternatively, one can use local thresholds, reducing the need for good shading correction. However, the use of local thresholds reduce the amount of data available to build statistics on for the thresholding. This makes local approaches more sensitive to noise. If good shading correction can be achieved, a global threshold may be preferable.

Nothing restricts the thresholding to be done on the intensity image. Anything that separates foreground and background can be used; colour, texture, shape (e.g., convexity), etc. In addition, multivariate approaches such as maximum likelihood classification of pixels, etc., can be used to combine different properties.

Despite (or perhaps, due to) its simplicity, thresholding can be a very powerful method to separate foreground from background. However, one question remains; what is the best threshold to pick? Just as for the segmentation problem as a whole, there are hundreds of methods also for thresholding, none being the perfect one. Histogram based methods are popular, thanks to simplicity and being easy to comprehend, but other methods incorporating the spatial information of the image are also quite common.

One of the most popular ways to threshold an image is to find a minimum in the histogram. This corresponds to the most stable point of the thresholding, so that moving the threshold up or down will affect the fewest pixels at that point. Two problems arise with this method, i) what to do if there are many minima, ii) what if the distribution is unimodal (no significant minimum). The first problem can quite easily be solved by, in some way, finding the most significant minimum, e.g., by iterative smoothing until only one minimum remains (care not to move the minimum must be taken, though), or to directly pick the deepest minimum in some sense. The second problem is a bit harder to solve. Different transformations can be applied to the data, e.g., transforming the intensity scale in such a way that the histogram is no longer unimodal, but this also implies that we will no longer threshold at the most stable point in the image.

Alternative methods that make some model assumptions about the grey level distributions of the image have been suggested. A commonly mentioned one is the method of Otsu [39] or the equivalent iterative version proposed by Ridler and Calvard [53]. These methods split the histogram into two parts, so that the threshold is located in the middle between the means of the two classes. This works well if the distribution is made up of two classes of equal variance.

However, this is rarely the case. Better performance [20] has been observed using the method suggested by, Kittler and Illingworth [26], that minimises the number of misclassifications between two Gaussian distributions.

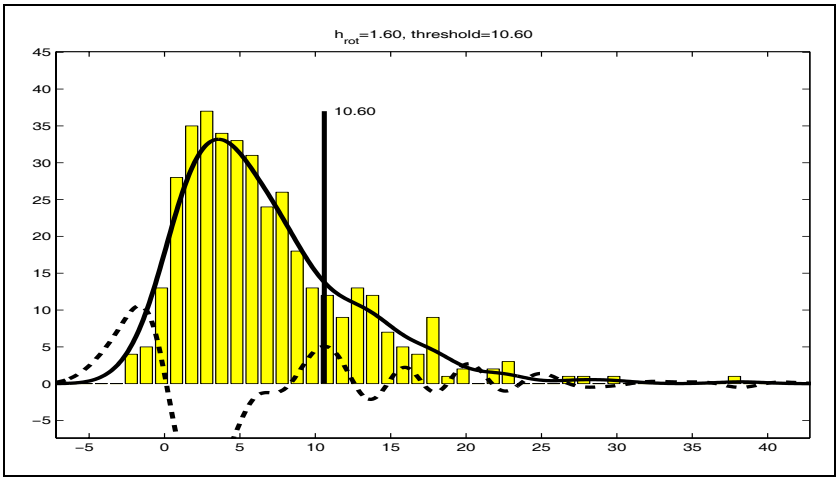


Fig. 10: Example of a histogram thresholded with the KDE method described in Paper V. Background histogram is number of cells plotted against average internal fluorescence intensity. Solid line is the KDE $\hat{f}$ and the dashed line is the second derivative $\hat{f}''$ (scaled). Suggested threshold is, the largest peak in the second derivative that is to the right of the main peak in the KDE.

A methods which has proved to work well in various situation is the one described in Paper V, which puts the threshold at the point of maximal change of trend in the distribution, as modelled by a kernel density estimate (KDE). See Fig. 10. It seems to be specifically well suited when applied to highly overlapping distributions, where many of the other methods fail. Unfortunately, it does not work that well on well separated distributions of very different magnitude, where it tends to assign too much importance to the larger one.

A reasonable way to perform thresholding is to maximise the difference between foreground and background. That is, find a threshold that maximises the difference of the mean of the foreground and the mean of the background. This can be calculated from the histogram. On the fluorescence images studied in this work, it proved better to maximise the ratio between the two. Even better results were achieved by only looking at the pixels of the border of the two sets. That is, maximising the ratio between the average pixel value of the bright side of the threshold border and the average at the dark side of the border, for each possible threshold and then pick the threshold so as to maximise this contrast. As this method uses spatial information of the image, it can no longer be done only by looking at the histogram. One pass through the image is required,

but all thresholds can be calculated at once. This is the algorithm we chose for segmenting the nuclei in Paper I.

There is nothing that says that we should use the intensity distribution as the basis for our thresholding. We might, e.g., calculate histograms based on feature values of the thresholded objects. These method, called global feature histograms by Ranefall [46] is, e.g., used for foreground/background thresholding in [48–50]. The brute force version of this method is, to apply all possible thresholds and then, for each threshold calculate the desired feature value, to get a curve of the feature measure versus applied threshold value. Using that approach, this of course becomes a very slow method, but with intelligent treatment of the data, many features can actually be calculated for all possible thresholds at once, with only one pass through the image, without the need for the explicit thresholding of the image at each grey level. Ranefall uses this method to get histograms of total perimeter of the thresholded image, total area and total integrated intensity of the foreground (this can be calculated directly from the histogram), and number of connected components in the image. Various combinations of these can also be formed, such as, average component size etc. Any of the histogram based thresholding methods may then be applied to the feature histograms to find a good threshold.

Anyone of the methods mentioned above should be applicable to iteratively make refined estimates of the image background, such as described in Paper IV. We have only tested to use the KDE second derivative thresholding, which showed good results.

4.4.2 Object separation

When one knows what is foreground and what is background, what remains is to split the foreground into individual objects. This is one of the most problematic parts of segmentation, especially for image cytometry, and most notably so for the segmentation of cytoplasms, where the objects can be close together and with no distinct border in between.

Watershed segmentation

A method which has proved to be very useful in many areas of image analysis is the so called watershed algorithm, originally suggested by Digabel and Lantuéjoul [13, 31] and then refined and used in very many situation, see ,e.g., [55] for an overview. The watershed algorithm splits the image into regions similar to the drainage regions of a landscape. If the intensity of the image is interpreted as elevation in a landscape, the watershed algorithm can then be used to find mountains, lakes and catchment basins in this landscape. This can serve as a ground for segmentation algorithms, but it can also be a useful tool for describing the internal structure of objects.

The watershed segmentation can be thought of in the following way. Given that the intensity image is interpreted as a 2D surface in a 3D space, where the intensity encodes the height coordinate. We take this surface and pierce a small

hole at every local minimum of the surface. We then slowly immerse the surface into a container filled with water. The water will start to rise from the small holes at the minima of the image. As the water level rises, water from adjacent minima will eventually meet. At the border where the water meets, we build a dam, to prevent the mixing of the water from the different minima. We immerse the surface further until we have a pattern of dams covering the image. These are the watersheds, surrounding the catchment basins of the respective local minima. This separation of the image into catchment basins is what we will refer to as watershed segmentation. The border between adjacent catchment basins will be at the ridges in the intensity landscape. Watershed segmentation can be implemented with sorted pixel lists so that essentially only one pass through the image is required [66]. This implies that the segmentation can be done very fast. The algorithm is straightforward and fairly easy to extend to n-dimensional images.

In the paper by Vincent and Soille [66] the water level rises from the local minima of the image. When segmenting fluorescence images. It is in most cases more natural to turn the algorithm upside down, so that we start at the local maxima in the image and separate objects at the dark valleys in the intensity landscape instead. We are not interested in thick watershed lines, leading to pixels that are not part of any catchment basin. In our implementation of the watershed algorithm, we keep track of the pixels that are ambiguous, i.e., located at an equal distance between two catchment basins, and let the water flow around those pixels, leading to much thinner watershed lines than in the original version of the algorithm. Still a smaller number of ambiguous pixels will occur, we have simply assigned those to the adjacent region with the brightest peak, so that we get a complete and well defined tessellation of the image.

Watershed segmentation is a fast and robust tool for performing segmentation, and defining regions of interest. When applied directly to an intensity image, or any other not perfectly noise free image, the algorithm often results in severe over-segmentation, i.e., the image is segmented into too small parts, splitting the objects of the image. This implies that additional pre- or post-processing is, in most cases, required. If it is possible to find an unique identifier inside each object, a seeded version of watershed segmentation can be applied, where we only pierce holes in the landscape at the seed points. This removes the problem of over-segmentation.

In addition to segmentation, one can also derive many textural features using watershed segmentation and looking at the watershed regions. Upper and lower rice field and half height partitions are some examples [54]. In addition, internal watershed regions can be useful to define regions of interest for feature extraction, see Sect. 4.5.1.

It can also be noted that seeded watershed segmentation is closely related to segmentation by fuzzy connectedness [56, 63]. The fact that so similar methods have arisen from two different views just verifies that the usefulness of the approach is great.

The watershed algorithm can be applied to any type of grey level image, of any dimension. Directly on the intensity image, on distance transformations of images, on images created by filter matching, on curvature images, on membership images, etc. Of course one can also make algebraic combinations of images. This enables us to design many different algorithms bases on the watershed algorithm, making it a powerful tool of digital image analysis.

4.4.3 Merge and split

In the case of the seeded watershed segmentation, we solved the problem of over- and under-segmentation by having one marker inside each object. As it is not always possible to find such a marker, there is a demand for other methods to solve the problem. In Paper II, there were no seeds available for the segmentation of cytoplasms. Starting from a watershed segmentation of the image, rule based merging and splitting of these regions improved the situation. Trying to remove as much of the heuristics from the algorithms as possible, we developed a method where each object was classified as cell-like or not. For the, in that case, more difficult problem of splitting, we calculated the cell-resemblance of the resulting objects, for each potential split. Finally, the split that lead to the most cell-like objects was chosen. An additional requirement was that the result had to be more cell-like than what we started with, or no split at all was performed. However, it should be noted that more material to train the cell-resemblance classifier on is necessary, to be able to draw conclusions of the true usefulness of that method.

4.4.4 Segmentation of nuclei

Several fully automatic methods for segmentation of cell nuclei in both 2D and 3D images have been described [1, 3, 12, 35, 64]. The nuclei can be stained with good contrast using fluorescent nuclear markers. The fairly round shape of the nucleus is of great benefit when segmenting it from the background and separating clusters of nuclei. After intensity thresholding, clustered nuclei can be separated based on the indentations on the contours of the cluster [1]. Another approach is to create a distance image [5] from the thresholded image of the clustered nuclei. Watershed segmentation [4, 66] can then be applied directly to the distance image [47], or the maxima in the distance transform can be used as markers for subsequent watershed segmentation of the original image of the cluster [35]. The watershed segmentation can also be applied directly to the intensity image of the nuclei, but this often results in over-segmentation. Over-segmentation can, however, be reduced by rule-based merging of the fragmented objects [64].

The approach taken by us for the nuclei segmentation is the following: The objects are separated from the background by thresholding. An image specific threshold was set so that the contrast between object pixels and background pixels at the object border was maximised, see Sect. 4.4.1. After thresholding, a Euclidean distance transform was calculated on the objects [7]. The distance

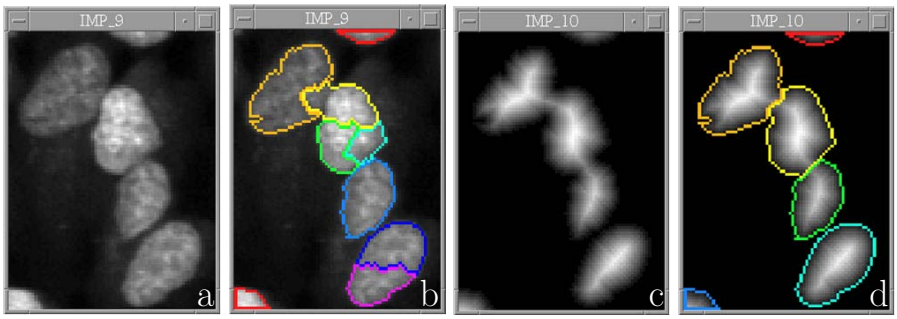


Fig. 11: Examples of nuclei segmentation. Original image (a). Directly applied watershed segmentation (b) leading to over-segmentation. By instead applying the watershed segmentation to a distance transform image (c), we get a much better result (d).

transform assigns, to each pixel belonging to an object, a value that represents the distance to the image background. A watershed segmentation was thereafter applied on the result of the distance transformation [47, 66]. Thus, touching circular nuclei were separated. See Fig. 11.

4.4.5 Segmentation of cytoplasms

Segmenting the cytoplasm is a much more challenging task than the segmentation of the nucleus. First of all, the contrast achieved from cytoplasm stainings are usually not as high as the contrast of the nuclear stainings. Second, the cytoplasm is much more irregular in shape, and in many cases far from convex. This reduces the usefulness of the distance transform, which was so fruitful when separating touching nuclei. On the good side can be said that it is often possible to get a fairly smooth staining of the cytoplasm. This allows us to apply the watershed algorithm directly on the intensity image. However, in the common case, the image is still severely over-segmented by the watershed algorithm.

The problem of over-segmentation can be approached with a number of different methods. One method that solves the over-segmentation problem in an elegant way, is the seeding method. If it is possible to get one single identifier inside each cytoplasm, it is a much simpler task to separate the cytoplasms. This is the approach taken by us in Paper I, where we had a parallel staining of the nuclei of the cells. We perform a segmentation of the nuclei as described above and use this segmentation as seeds for a following segmentation of the corresponding cytoplasms. This relies on the assumption that one cell has one nucleus, which may not always be the case. Performance is very good, but the need for the nuclear stain somewhat reduces the usefulness of the method.

Watershed segmentation using seeds has previously been described, e.g., in [65], but here, the markers are created from the image that is to be segmented and not from a parallel image. Cytoplasm segmentation by detecting

the boundaries of the cytoplasms of living cells, imaged using modulation contrast and differential interference contrast microscopy, has been described in [59], where the cell boundaries were found by image pre-processing and thresholding, followed by morphological operations removing noise and filling holes in the objects. Segmentation of cytoplasms of living cells is also described in [70], where the cells were imaged by brightfield optics. Segmentation of fluorescence images of cytoplasms has previously been described in [12], but in this case, membrane markers were used to show the borders of the cytoplasm, and seeds were drawn manually within each cytoplasm, making the method semi-automatic.

With or without seeds, there is still a need for a foreground/background segmentation. As the contrast of the cytoplasm stain is lower, this is a more delicate problem than the corresponding situation for the nuclear images. Of the thresholding methods mentioned in Sect. 4.4.1, the one that showed to work best for the cytoplasm images of Paper I was to use the threshold that the shading correction step converged to (see Sect. 4.3).

A morphological opening was applied to the watershed regions to remove jagged edges and thin structures. A result of the seeded watershed segmentation of the cytoplasms can be seen in Fig. 12(c), where the nuclei are outlined in blue and the cytoplasms are outlined in green. Note that all cytoplasms are nicely separated by the algorithm.

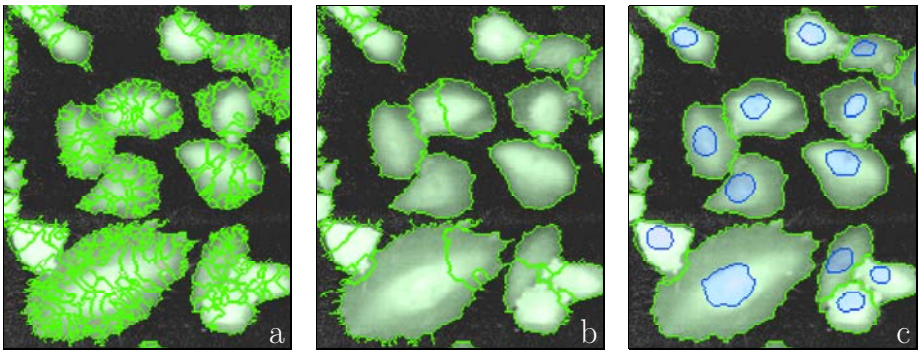


Fig. 12: Examples of cytoplasm segmentation results. A directly applied watershed segmentation leads to sever over-segmentation (a). Merging shallow valleys in the image improves the situation somewhat (b). When using nuclei (blue overlay) as seeds (c), the problem of over-segmentation disappears.

If seeds are not available, hysteresis can be included in the watershed algorithm (see Paper II) to reduce the over-segmentation. Similar results can be achieved by either applying a pre-filtering to reduce the number of local maxima for the watershed algorithm to pick up, e.g., by using morphological reconstruction [65], or by a post-processing step where regions from the watershed

segmentation are merged if the depth of the separating valley between them is smaller than a certain level h, see Fig 12(b).

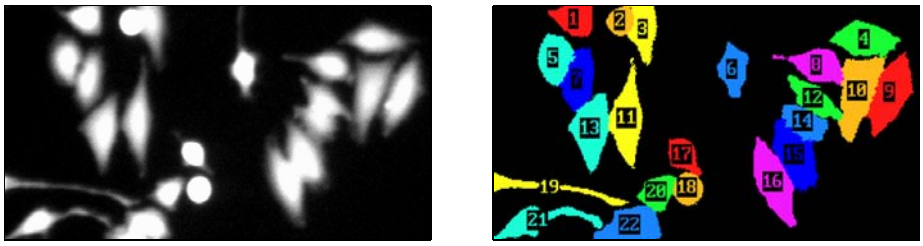


Fig. 13: Examples of cytoplasm segmentation results using the merge and split approach described in Papers II and III.

In general, this is not enough to get a good cytoplasm segmentation, and some more advanced methods have to be applied. In Papers II and III, we describe such a situation, where merge and split algorithms were used to improve the segmentation, see Sect. 4.4.3. An example result can be seen in Fig. 13.

4.5 Feature extraction

The next step in the analysis process is to extract descriptive feature measures from the segmented cells. There exists thousands of different, more or less general, features described in the literature, see, e.g., Rodenacker and Bengtsson [54] for a large list of general features for cell image analysis. However, as is also pointed out in that paper, the features chosen, if to be of any use in the further analysis, have to reflect the property of interest. When working on a real-world application, it is rarely enough to only use general purpose features. To achieve good results from the data analysis, it is almost always fruitful to measure additional features specifically designed to capture the property of interest.

4.5.1 Regions of interest

Problem specific features is often a matter of calculating general feature measures on specific regions of interest (ROIs). For example, if we can find the region of increased activity of a cell, we can measure location, area, and intensity of this region. The task of finding good features is therefore much a task of defining good ROIs, leading us back to the segmentation problem.

The watershed algorithm proved to be most useful when segmenting both the nucleus and the cytoplasm. It therefore seems quite natural to see if the same algorithm can also be used to define ROIs inside the cell. This is the course of actions we have taken in Paper I, and it turned out to be a useful approach, see

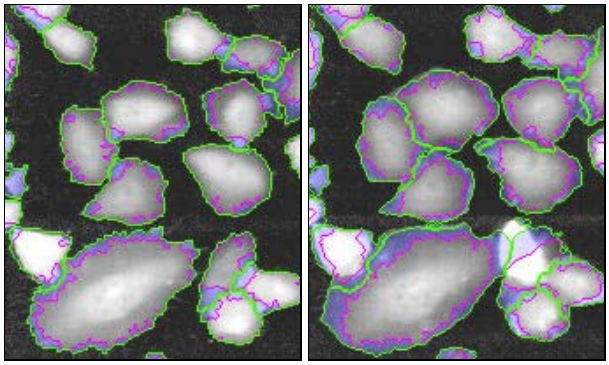


Fig. 14: Example of ROIs (here, 'ruffle regions', as defined in Paper I), marked in blue, before (left) and after (right) stimulation with IGF-1. In the lower right of the right image, debris has complicated the situation.

Fig 14. There we used both seeded and non-seeded watershed segmentation to find the locations of increased activity at the border of the cells, the ruffles.

When good ROIs have been found, what remains is to calculate useful quantitative measures from these regions, to pass on to the data analysis step.

4.5.2 Selected general features

Features describing shape and distribution of intensity have been of importance for the analysis performed in relation to this thesis, whereas texture features have not had any prominent roles. This choice of features is closely related to the projects that we have worked on, and says little about their usefulness in other situations. In Paper I, where sub-cellular measures were to be done, our approach was to define sub-cellular ROIs and to perform measures on them, rather than to try to describe the local situation using texture analysis.

Just to mention some of the features used, starting from the less complex ones, we will begin with the ubiquitous measures of area and integrated intensity of the objects. Being simple and straightforward to calculate, they appear in most image analysis systems. Both give measures of size, and are, e.g., most useful for early rejection criteria for the segmentation algorithm not to include small noise objects. In Paper II, we notice that for this purpose, in connection with cytoplasm segmentation, the integrated intensity was a better feature to use than the area, as it stayed more stable during the cell cycle. However, it should be noted that the integrated intensity, as all intensity based measures, is sensitive to the staining level of the specimen, thus reducing the stability of the measure.

Another feature of interest is the perimeter of the objects. Being dependent on the border of the segmented object, this feature is more unstable than the

previously mentioned. It can still give some useful information, e.g., when used together with area to produce a compactness measure.

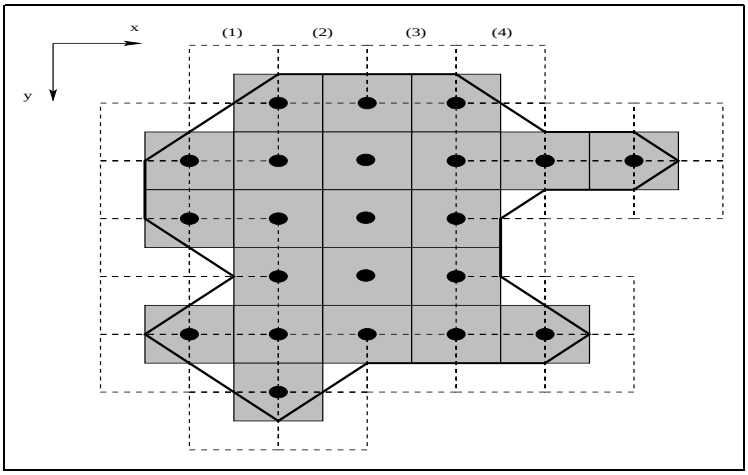


Fig. 15: The area of an object can be measured by counting the number of pixels inside the object. We can measure the perimeter of the object, by following the border toward the background. Using appropriate weights for straight and diagonal steps, and unbiased estimate can be achieved.

Measuring the perimeter can be done in many ways. We have used the sum of steps taken when walking around the edge pixels of the object (Fig. 15), where horizontal and vertical steps are weighted by $a = 0.948$ and diagonal steps are weighted by $b = 1.343$, see paper iv and [14, 28].

If we use 3D data, the corresponding measure is surface area, which can be measured, e.g., using the approach described in paper i. Surface area may also be of interest on 2D images, if we interpret the grey levels of the image as height in a landscape, and measure the surface area of the created 3D object. However, the scaling of the intensity dimension relative to the spatial dimensions is not well defined.

Also closely related to the border of the object set of pixels, are various representation of the shape of the border of the region. The most popular are possibly the Fourier descriptors, quantifying the frequency properties of the border shape.

A set of shape related measures, that we have found useful, are the grey scale and silhouette moments of the intensity distribution and the object pixels respectively. There exists a rich plethora of different moments, from the raw moments, which are dependent on scale and rotation, to the normalised moments of Hu [24, 67], various complex moments [57] and many more [42]. To get a feeling of what the different moments really describe, we have found it fruitful to look at moments derived from multivariate statistics, e.g., multivariate kurtosis [36].

It can be noted that the area of the objects as well as the integrated intensity can also be expressed as moments, i.e., silhouette moment m_{00}, and grey scale moment m_{00}, respectively.

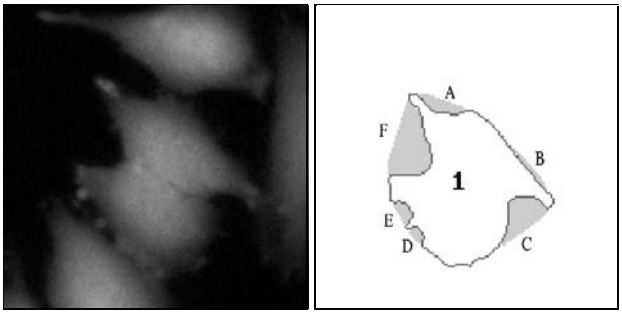


Fig. 16: Cutout of a part of a cytoplasm image (left). The segmentation result where two cells have incorrectly been segmented as a single object (right). The convex deficiency of the object is shown as the shadowed areas around the borders of the object.

A ROI providing us with a whole new set of shape measures, is the convex hull of the object. This could be the 2D convex hull of the pixels of the object, or a 3D convex hull calculated on the subgraph of the object. This allows, among other things, the use of area and volume of the corresponding convex object. Comparing these measures with those of the original object, i.e., creating convex deficiency like measures, can be useful, e.g., when deciding whether to split objects into two. See Fig. 16 and Papers II and III.

In addition to these, one can then add features from the very large set of texture measures. As mentioned earlier, we have not considered this to be useful in relation to this work. The problem specific feature regions that we define in Paper I, are, however, closely related to the rice field [54] type of textural features.

4.6 Data analysis

When useful descriptive features have been measured, some sort of data analysis is to follow. Depending on the ultimate goal of the study, this could be a reformulation of the measured values, e.g., a dimensionality reduction, or a classification procedure to divide the objects into different classes.

There exist far too many different methods to perform data analysis to even try to mention them all in this text. We will just present some that have been relevant for the work of this thesis.

4.6.1 Dimensionality reduction

When the desired output of the method is to assign a scalar value to each object in the study, some kind of dimensionality reduction is in most cases needed. In the very rare occasion that the property of interest is directly measurable from the image, we are done. Interesting measures are, however, seldom related to the apparent properties of the image, such as area and perimeter, but more often to much subtler properties, such as degree of activation, or the deviation from a normal situation.

It could be that the desired property to measure is well correlated to one of the more direct features of the image, but in the great majority of situations, the desired property is more vaguely related to a larger number of image features, and some kind of multivariate approach is needed. That is, we need to go from a higher dimensional space, consisting of the measured feature values, to a lower dimensional space, describing the property or properties of interest.

This dimensionality reduction could, e.g., consist of a distance measure, to say if an object is far away from a known population. It could also be a regression procedure, to relate different variables. Canonical correlation is a useful method if more variables are to be described. If the problem is non-linear, more non-linear function approximators than regression would be useful, e.g., artificial neural networks.

Unsupervised methods, such as principal component analysis (PCA), independent component analysis (ICA), and maximum autocorrelation factors (MAF) could also be useful. However, if it is possible to find some kind of ground truth, i.e., something to aim at, the supervised methods will, with very few exceptions, outperform the unsupervised ones. This is an important thing to notice, as unsupervised methods are often used in situations where a supervised method would be much better. PCA is, e.g., all too often used for dimensionality reduction before performing a classification, where it is assumed that the interesting properties will show in some of the first principal components. There is, however, rarely support for this assumption, since what is present in the first principal components are the directions of most variation in the data set. A direction of large variation could well mean that this is a direction of large noisiness. It is quite possible that the last principal component is actually the best variable for classification.

4.6.2 Classification

This leads us to the task of classification, which really is just a discretisation of the previous situation. However, the methods available to perform classification can be quite different from the methods to perform dimensionality reduction.

Common methods for classification range, from the simplest one of using a single threshold on a single variable, to complex non-linear multivariate learning systems. In the univariate threshold case, an object is classified into one of two possible groups depending on, whether the value of the variable for that object is smaller or larger than the threshold value. The single variable could be a direct

image feature, or, more probably, a result of dimensionality reduction. If more classes are required, more thresholds are introduced. The naive extension of this to more input variables is a box classifier, where the feature space is separated into rectangular subsets representing the different classes. However, this is still an intrinsically univariate approach. The classification regions will always be aligned with the axes of the feature space, and we use only one dimension (i.e., feature) at a time. This leads to that co-variations between the measured features are not fully utilised. Better statistical properties can be achieved by assuming that the different classes are not boxes in the feature space, but instead multivariate normal distributions. This leads to the family of Bayesian discriminant classifiers.

Derived from Bayes decision theory, see, e.g. [15], Bayesian classifiers are designed, so as to minimise the error rate, i.e., minimising the number of misclassifications the classifier will make. There are many ways to represent pattern classifiers. One way is to use a set of discriminant functions

$$g_i(\boldsymbol{x}),\ i = 1, 2, \ldots, c, \tag{3}$$

where $\boldsymbol{x}$ is a vector of feature values from the object to classify and c is the number of classes. The classifier assigns a feature vector $\boldsymbol{x}$ to class ω_i if

$$g_i(\boldsymbol{x}) > g_j(\boldsymbol{x}) \quad \forall \quad j \neq i. \tag{4}$$

From Bayes theorem, one can derive that general minimum-error-rate classification can be achieved by use of the discriminant functions

$$g_i(\boldsymbol{x}) = \log p(\, \boldsymbol{x} \,|\, \omega_i \,) + \log P(\omega_i), \tag{5}$$

where $P(\omega_i)$ is the a priori probability of class i.

Assuming that the classes are multivariate normal distributed, and evaluating this expression for a d-dimensional case, we arrive at the following set of discriminants

$$g_i(\boldsymbol{x}) = -\frac{1}{2}(\boldsymbol{x} - \mu_i)^t \Sigma_i^{-1}(\boldsymbol{x} - \mu_i) - \frac{d}{2}\log 2\pi - \frac{1}{2}\log|\Sigma_i| + \log P(\omega_i), \tag{6}$$

where μ_i and Σ_i are the mean value and covariance matrix of the distribution of class i, respectively.

Assuming that the different classes c all have the same covariance behaviour, $\Sigma_i = \Sigma$, using equal a priori probabilities $P(\omega_i)$ for the different classes, and substituting the mean and covariance matrix for the sample mean m_i and the sample covariance matrix S_i, the minimum error rate classifier discriminants then becomes

$$g_i(\boldsymbol{x}) = -\frac{1}{2}(\boldsymbol{x} - m_i)^t S_i^{-1}(\boldsymbol{x} - m_i) \tag{7}$$

which gives linear decision boundaries between the different classes. This is the classifier of linear discriminant analysis (LDA). The LDA classifier of equation (7) can also be described as a minimum-Mahalanobis-distance classifier. Classifying the sample $\boldsymbol{x}$ as belonging to the class ω_i with the smallest Mahalanobis distance to $\boldsymbol{x}$.

In the general multivariate case, the covariance matrices are different for each category. This leads to the following discriminant functions

$$g_i(\boldsymbol{x}) = -\frac{1}{2}(\boldsymbol{x} - m_i)^t S_i^{-1}(\boldsymbol{x} - m_i) - \frac{1}{2}\log|S_i| + \log P(\omega_i). \tag{8}$$

Here, the discriminant functions are inherently quadratic, giving decision boundaries which are hyperquadrics. This is the quadratic discriminant analysis (QDA) classifier. Both LDA and QDA are well known and well tested methods for performing automatic classification.

Alternative classification methods include the k-nearest neighbour (k-NN) classifiers, support vector machines (SVM), learning vector quantisation (LVQ), artificial neural networks (ANN) and many more. Of course there is also the unsupervised version of classification, i.e., various clustering and grouping algorithms, but again, if we know what we are looking for, and can find some kind of ground truth, the supervised version will, in most cases, outperform the unsupervised one.

The free lunch theorem for optimisation [69] indicates that no method may outperform all the other methods on all problems. In real world testing, one classifier rarely is significantly better than another one. If the classes are not well separated by the measured features, no classifier will give a good result. Our experience has shown that, for a certain situation, one can often improve the performance ever so slightly by using, e.g., neural networks instead of the Bayesian classifiers. However, the decrease in predictability, and increase in training time, is rarely justified by this gain in classification performance. It is often more fruitful to try to find feature measures that better describes the data instead.

4.6.3 Ground truth

Supervised methods require some kind of training set. This should be a representative sample of the data to analyse. Parameters of the data analysis model are derived from the combination of the feature values of this training set, and the known ground truth for the same set.

Having a good goal to aim at is of great importance when using supervised training. The performance of the automatic procedure can never be better than the ground truth we have used for the training. Unfortunately, when working with image cytometry, it is seldom possible to acquire quantitative measures of the interesting properties, which could be measures like amount of activation or degree of strangeness. Often, the only way to get a goal to aim at, is to have individuals manually try to judge the desired property. This is often both tedious and difficult. Therefore, there will always be a trade-off between reliability and cost. In most cases, it is valuable to have a few people assessing the same material, so that one can at least get an estimate on how stable the ground truth is. The more the individual judges concur, the better and the fewer parallel assessments are needed.

In Paper I, we used LDA and QDA classifiers, where the estimates of the class mean values and covariance matrices were derived from a set of training images. The ground truth in that project was a weighted summary of three manual classifications of the individual cells.

4.6.4 Feature selection

> *Perfection has been reached, not when there is nothing left to add, but when there is nothing left to take away.*
>
> —

Having a classifier or a regression model, and a set of features, it is tempting to directly use the complete set of features in the data analysis, as every included feature should, in the ideal case, contribute to reducing the probability of error. Unfortunately, it has frequently been observed that, beyond a certain point, the inclusion of additional features leads to worse rather than better performance [15]. This can be seen as a type of over-training, where the classifier "learns" the sample set well, but does not manage to generalise to new samples. The source of this problem lies in the fact that the number of design samples is finite. Including more features in the classifier implies that there are more unknown parameters to estimate from the same set of samples. An inaccurate estimation of parameters increases the number of classification errors. If this increase is larger than the decrease in classification errors produced by the addition of the new feature, the net effect is a decrease in the overall performance of the classifier. Therefore, in the finite-sample-size case, there exist a "peaking" phenomenon [51], where the performance of the classifier first increases with the inclusion of additional features, then attains a maximum, and then begins to decrease.

For a given situation, there exists an optimal set of features that maximises the performance of the classifier. In fact, for many problems, finding the proper set of features is much more important than selecting the best algorithm for doing the classification. In addition, if we can reduce the number of features used for the classification, we no longer need to measure these features, leading to an overall speed improvement of the algorithm.

To find the optimal set of features to use, several feature selection methods have been suggested in literature [11]. Feature selection algorithms can be characterised by a *search rule*, a *selection criterion*, and a *stopping rule* [45].

The optimal search rule is to perform an exhaustive search over all possible combinations of the features available. As the number of possible combinations increases exponentially with the number of features, it is not feasible to do this for more than a very limited number of features. Instead, the standard procedure is to apply some sub-optimal search rule, to traverse the space of feature combinations. The search rule that we have applied in Papers I, II, and III, is the sequential floating backward selection procedure (SFBS) [43, 44], which has proved to be a good trade-off between speed and performance [17, 25, 27].

As we in all cases had a small enough number of features to use the backward version, no problem of stopping too early, as mentioned in [25] is present.

The outline of the algorithm is as follows. First, all features thought to be valuable for classification are included in the classifier. This will most certainly result in over-training, as it gives too many degrees of freedom for the classifier. One feature at a time is therefore excluded temporarily, and the performance of the classifier is tested. The feature that contributed the least to the classification performance is then removed. This is done over and over again until there is only one feature left. To be sure not to accidentally remove the best feature in the beginning, before removing another feature, it is always checked if the inclusion of one of the previously removed features will give a performance strictly higher than what we had before with the same number of features. That is, we remove and put back features alternating, but when we put features back, we always make sure that the performance goes up. Therefore, the process will not go on forever, and finally we will have only one feature left.

The selection criterion decides which features to include in the classifier. The natural selection criterion to use in the feature selection procedure is the increase in performance of the classifier [23]. Alternative selection criteria include the use of F-statistics or Wilks' Λ, (which are not optimal for the task [23, 52]), or the more advanced Bhattacharyya distance measure [22]. The main motivation to apply any of these alternative criteria is, if the direct measure of the classification performance is too time consuming. This has not been the case in the work related to this thesis. Hence, the selection criterion of our choice has been the increase in classification performance.

The stopping rule for the feature selection procedure tells us how many features to use. Most feature selection procedures use a test statistic applied to the difference in score on adjacent levels in the feature selection procedure. There exist many versions of such stopping rules, expressing different properties [10]. When using the classifier performance as a selection criterion, the obvious rule is to pick the set of features leading to the best performing classifier, according to our performance measure. This is what we have done in the work of this thesis. During the feature selection, we keep track of which features are included in the best performing classifier, for every possible number of features. We can then backtrack this list and pick the feature set that performs best on the test set.

Performance measures

The direct use of the error rate as a performance measure may not be the best on skewed distributions. If, e.g., 90% of the samples belong to class 1 and 10% belong to class 2, having a classifier that classifies everything as belonging to class 1 will only get an error rate of 10% for this data set. This looks like a fairly good performance of the classifier, when just looking at the error rate, still the classifier is useless.

Alternative performance measures, which take into account the proportions of the different classes in the sample, exist. Cohen's Kappa and Cohen's Weighted

Kappa [18] provide a measure of the degree to which two judges concur in their respective sortings of N items into k mutually exclusive categories.

Cohen's Weighted Kappa can be summarised in the following: Assume that N objects are distributed into one of k classes by one classifier and, independently, to one of the same k classes by a second classifier. This results in a classification matrix with k^2 cells. Let p_{ij} be the proportion of objects placed in the i,jth cell. Let $p_{i\cdot}$ be the proportion of objects in the ith row and let $p_{\cdot j}$ be the proportion of objects in the jth column. Then

$$\kappa_w = \frac{\sum_{i=1}^{k}\sum_{j=1}^{k} w_{ij}p_{ij} - \sum_{i=1}^{k}\sum_{j=1}^{k} w_{ij}p_{i\cdot}p_{\cdot j}}{1 - \sum_{i=1}^{k}\sum_{j=1}^{k} w_{ij}p_{i\cdot}p_{\cdot j}} \tag{9}$$

This gives us a useful performance measure, both when comparing the classification made by the computer with the "ground truth", and also when comparing the classifications made by two individuals [61]. In Paper I, Cohen's Weighted Kappa κ_w was used, both for the feature selection procedure, and to evaluate the performance of the overall result.

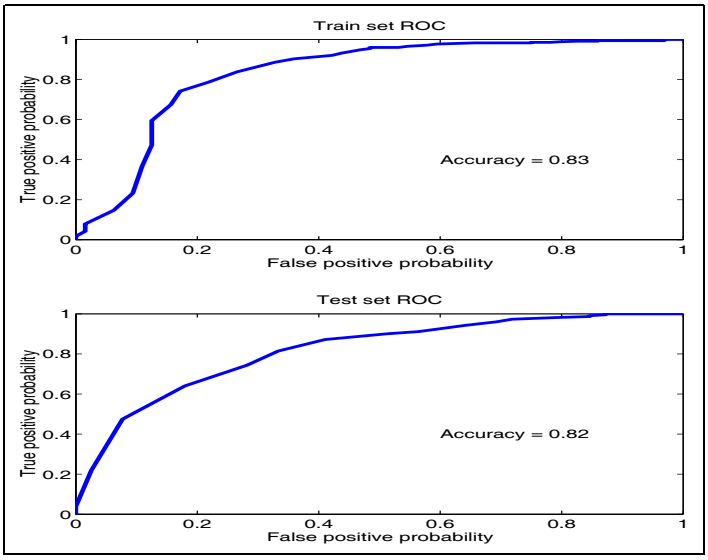


Fig. 17: ROC curves for a training image and a test image, using the best performing classifier of Paper II.

If the desired output of the data analysis is a descriptive scalar measure, e.g., a distance measure telling how far from the normal population of cells the current individual is (see, e.g., Paper III), we can often turn this into a classification problem, e.g., by selecting a threshold for rejection of cells. It is,

however, not always desirable to do so, as the selection of the specific threshold value to use is not well defined and is highly dependent on the application. We would like to be able to judge the discriminating power of the measure without having to select a specific threshold. One commonly used way to do this, and to get a unified measure of the discriminating power of such a scalar measure, is to draw a so called receiver operating characteristic (ROC) curve [60, 62]. Here we do not select any certain threshold, but rather use all possible thresholds. An ROC curve is a plot, where for each possible threshold, the percentage of true positives is plotted against the percentage of false positives. If the discriminating power is large, the ratio of true positives to false positives should be large, and a descriptive scalar measure is correspondingly characterised by a large area under the ROC curve. This area under the ROC curve is called the accuracy (A) of the classifier. A bad classifier has an accuracy A near 0.5, which corresponds to random selection, whereas a perfect classifier has $A = 1$, i.e., no false positives. In Papers II and III, we have used ROC accuracy as the performance measure, which gave better results than the direct use of error rate. See Fig. 17.

Curse of dimensionality

The term curse of dimensionality [2] generally refers to the difficulties involved in fitting models, estimating parameters, or optimising a function in many dimensions. As the dimensionality of the input data space (i.e., the number of features) increases, it becomes exponentially more difficult to find global optima for the parameter space. The number of test samples needed to fill this space also grows at exponential rate. This means that, even when applying feature selection procedures, more features are not always better [58]. The original feature space, before feature selection, has to be filled, at least to such an extent, so that the feature selection procedure has enough data to work with. Hence, it is more or less a necessity to pre-screen and pre-select, from among a large set of input variables, those that are likely to be of use for predicting the outputs of interest.

4.7 Evaluation

96.37% of all statistics are made up.
—

To evaluate the result in an unbiased way, is of utmost importance [37, 58]. In both of the cell image studies performed, tedious manual annotation was required to get a ground truth to compare with. This in turn lead to that the total number of cells to test the algorithms on has been very limited. This increases the need for careful use of the available data.

Several studies has shown that using the resubstitution error, i.e., testing the methods on the same data that it is trained on, is severely biased. Partitioning the data into two separate sets, training on one and using the other for validation (the so called hold out method), is a better alternative. Unfortunately, one does not utilise all the available data this way. A solution to this is to use cross

validation, that is, to remove n samples from the data, train the algorithm on the rest, and validate on the removed samples. Then remove n other samples and do the same thing again. If we choose n to be one, we have the popular leave-one-out method [29].

Instead of removing one cell at a time from the data set, we have partitioned the data used into the separate images acquired. This corresponds to k-fold cross validation, where k is the number of partitions. Here, one partition at a time is removed and the algorithm is trained on the remaining $k - 1$. The classifier is then set to classify the unseen removed data set.

The reason for this partitioning is, first that it is a natural splitting of the data, but also that we did not want to mix cells from different images. Such a mixing, that, e.g., the leave-one-out method would result in, would most probably introduce a bias, as the cells from the same image can hardly be considered to be uncorrelated. That is, we see one image as one sample, rather than seeing one cell as one sample.

Due to the fact that we used the performance of the classifier, as selection criterion in the features selection procedure, a rather complicated scheme was needed. What we did in Paper I, was to have two cross-validation procedures at different levels. Starting with four images, we removed one, as a validation set, from the remaining three images, we then used 3-fold cross validation for the feature selection procedure. The classifier was then trained, using the three images, and the performance was tested on the fourth image. The resulting classifications of this image were compared with the ground truth using Cohen's weighted Kappa. This was then repeated four times in a 4-fold cross validation procedure.

The overall classification results were compared with those achieved by visual inspection. Variation between visual classifications performed by different persons, as well as the variation between classifications performed by the same person at two different time points, i.e., inter- and intra observer variability, were also observed and compared with the results from the automatic classification.

4.8 Implementation

4.8.1 Software

All algorithms tested in this thesis have been implemented, by the author, solely or in close cooperation with C. Wählby. Most implementations have been performed in the Matlab (The MathWorks, Inc., Natick, MA) environment. Most algorithms have also been implemented in the IMP environment [38] using C or C++. IMP is a general purpose image analysis software for up to 5D images, developed at the Centre for Image Analysis. Besides the methods described in this text, many other tools for manipulation and visualisation of images have been implemented during the project, including, e.g., an interactive 3D visualisation engine.

4.8.2 Real-time aspects

We have focused on the use of fast algorithms, especially in Paper I, where specific real-time requirements were a part of the project goal. To enable large scale screening studies, the processing of cell images has to be done in a few seconds, or preferably, less time. Therefore, only limited consideration where given to algorithms which are intrinsically slow, e.g., diffusion models (which, however, often can be parallelised) and iteratively refined active shape models. If the time restrictions allows for it, additional refinement of the segmentation results can most probably be achieved by post-processing the segmentation with, e.g., a snake algorithm [40]. This has not been tested in the scope of this thesis.

Very limited efforts on optimising the code have been taken. Some preliminary performance figures can still be informative. With the current implementation in C++, the nuclear segmentation, as described in Paper I, when performed on an image of size 640×640 pixels, containing 153 cells, takes 0.27 seconds on a 1.7 GHz Pentium. The seeded cytoplasm segmentation of the same image takes 0.68 seconds. Timing studies of the complete chain of processing steps that is suggested in Paper I, i.e., including feature extraction and classification, has not yet been performed, as some of the steps still only exist as Matlab code.

CHAPTER 5

Results

Results! Why, man, I have gotten a lot of results.
I know several thousand things that won't work.
— Thomas Alva Edison

During the work of this thesis, two sequences of processing steps leading to automatic cytoplasm segmentation of fluorescence microscopy cell images, have been developed. The approach described in Paper I uses a parallel image of the nuclei of the cells, which is segmented and the nuclei are used as seeds for the segmentation of the cytoplasm image. This solves the problem of over-segmentation of the cytoplasms in an efficient way. The approach described in Papers II and III does not rely on a parallel nuclear staining, but instead uses merge and split algorithms, in conjunction with statistical analysis of descriptive features. This analysis is used in a feedback system to improve the segmentation performance, as well as to give an overall quality measure of the segmentation.

The performance of the seeded cytoplasm segmentation has not been quantitatively evaluated, but experience indicates that it performs considerably better than the unseeded version. The latter one reached a correct segmentation rate of 89–97% for the fully automatic segmentation of cytoplasm of fluorescently labelled cells [Paper II].

In Paper I, a classification step was also included, to separate the individual cells into three classes depending on their level of activation. The classification was based upon time series of images, where the first sets of images were acquired before addition of the agonist (IGF-1). Additional images were then taken at regular intervals, following the translocation of Rac1 taking place as a response to the IGF-1 treatment. By using both general purpose features and carefully designed problem specific features, in combination with a floating feature selection procedure, a Bayesian classifier was built. Evaluation showed that the performance of the fully automatic classification procedure was very close to the performance of skilled manual classification. It should be noted that the specific cell response studied is quite difficult to quantify, whether done manually or with the aid of a computer.

The problem specific features were built upon the notion of ROIs inside the studied cell. By using various combinations of image analysis methods, chiefly watershed segmentation, sub-regions of increased activity inside the cell were defined and measured. By examination of the feature selection results, it was concluded that the crafted problem specific features were of the most importance for the classification. LDA and QDA were tested as classifiers for the problem. Both performed well, but the more flexible QDA was the one that best managed to capture the variations expressed.

Many problems were encountered on the way, while designing the systems described in Papers I and II, some requiring development of new methods. Most notably, the requirement for robust correction of intensity variations was observed. This requirement lead to the design of the method for estimation of intensity nonuniformities that is described in Paper IV. The method, which is based on an iteratively refined B-spline estimate of the background shading, has been quantitatively evaluated in Paper IV, and it has shown robustness and versatility in a range of different situations.

The need for various thresholding methods has surfaced in many situations during the work. In Paper V, a method for performing thresholding of a distribution, at the peak of the second derivative, based on a kernel density estimate, is presented. Although first designed for the unsupervised separation of a cell population into two classes, depending on cyclin content [16], it has proved to be useful in very many, quite different, situations; as a part of the INU estimation method of Paper IV, for foreground/background segmentation of Paper I, for direct data analysis, as performed in papers ii and iii, just to mention a few.

Accurate and precise feature estimators were also found to be required, which lead to the studies of performance of perimeter estimators in paper iv. This in turn lead to the development of the novel methods for surface area estimation described in paper i.

The importance of careful data analysis, including the indispensable use of feature selection, has been very clear. Nothing really new has come out of this part of the work, but the results can still be seen as yet another indicator of the significance of this step. The floating feature selection procedure was found to perform well in the encountered situations, and the requirements for well discriminating carefully designed features can be deduced from the results of Paper I.

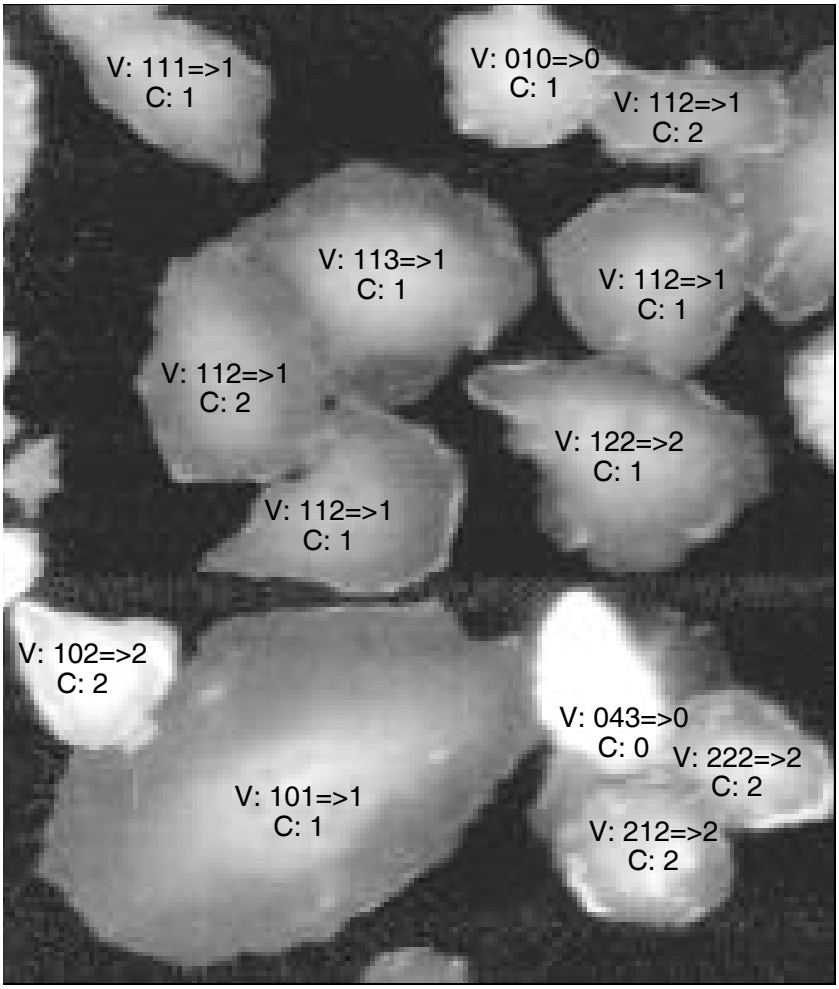


Fig. 18: Example of classification results from Paper I. V: XXX corresponds to the visual classes by three different manual judgements. The number after the arrow is the weighted summary result of the three visual classifications, which is considered to be the "ground truth" in that study. C: X corresponds to the automatic classification result when using the suggested method.

CHAPTER 6

Concluding remarks

Life's like a bottle of ketchup; nothing, nothing, nothing,
and then, suddenly, everything!
—

6.1 Discussion

The need for fast and accurate segmentation of cells is ubiquitous in the field of digital image cytometry. The segmentation methods described in Papers I and II show promising robustness and versatility. The seeded watershed segmentation of cytoplasms that is suggested is elegant in its simplicity. The non-seeded version is less reliable, but has a larger range of applications. The number of possible applications for the segmentation methods developed within the scope of this thesis, should be considerable.

Applications of the methods developed within the more problem specific study of Rac1 activation and translocation, described in Paper I, would, when combined with high-speed high-resolution imaging platforms, allow efficient screening studies, revealing how drug candidates affect the dynamics of signalling in living cells. Although being focus on a particular problem, i.e., the classification of Rac1 activation and translocation, the overall methodology is very general and should be useful in a great variety of situations. While the problem related features, based on ROIs inside the cells, may be of limited value in other studies, the use of watershed segmentation for finding and defining sub-cellular regions, should be of importance also in other situations.

During the work, many sub-problems have been encountered, and the solutions described should be applicable in many, more or less similar, image analysis projects. These solutions range from the very general ones, such as finding thresholds of a general distribution, via more specific ones, such as estimating intensity variations over an image, to the very specialised ones, such as the imaging device specific task of reducing artifacts of the acquired images.

All the methods developed in this work are applicable to real-world situations. The two main projects have been performed in close cooperation with and

according to demands of the biomedical industry (here, represented by Amersham Biosciences).

In all, the methods described, as well as the general methodology applied to solve the problems encountered, should be useful for a wide range of image analysis problems, not necessarily related to the fields of digital microscopy and image cytometry.

6.2 Future

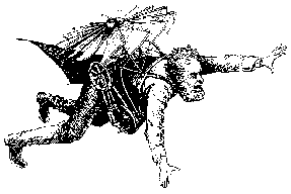
The one who lives, will see...
—

This thesis does by no means present any final solutions. The general problem of segmentation still requires lots of work, and will, probably, never reach the level of being solved. Better methods for finding the sub-distributions of a mixed data set will always be desired. The need for more robust feature estimators is still present. We could continue this list forever, but the point is already clear.

A nice property of science and knowledge is that it is like a growing ball; the more your knowledge expands, the larger is the frontier toward the unknown. Admittedly, it does feels like the area of this frontier grows faster than the surface area of an expanding ball in our ordinary 3D space. Consequently, during this work, there have come up numerous side tracks, most of them still remaining to be explored. Also, as our knowledge about existing methods has increased during the work, similarities and connections between the different methods become apparent and call for attention. Paths that would be most interesting to follow, and to mix with the work done, include fuzzy techniques, scale space methods, level set theory, active shape models, and Markov fields; just to mention a few. To dig deeper into the problematics of the morphology of the digital space, as well as trying to find and explore connections between fuzzy approaches, grey-scale morphology, and statistical descriptions, would be most interesting.

Combining all these possibilities, with the opportunity of applying the methods developed, to the field of biomedicine, with its possibilities of being useful both for humankind as a whole, but also for individuals, all over the world
— this cannot be defined as anything less than a dream situation.

Hello World!
Here I come!



References

The secret to creativity is knowing how to hide your sources.
— Albert Einstei

[1] J. A. M. Beliën, H. A. H. M. van Ginkel, P. Tekola, L. S. Ploeger, N. M. Poulin, J. P. A. Baak, and P. J. van Diest. Confocal DNA cytometry: A contour-based segmentation algorithm for automated three-dimensional image segmentation. *Cytometry*, 49:12–21, 2002.

[2] R. E. Bellman. *Adaptive Control Processes: A Guided Tour*. Princeton Univ. Press, 1961.

[3] S. Beucher. The watershed transformation applied to image segmentation. *Scanning Microscopy*, 6:299–314, 1992.

[4] S. Beucher. The watershed transformation applied to image segmentation. *Scanning Microscopy*, 6:299–314, 1992.

[5] G. Borgefors. Distance transformations in digital images. *Computer Vision, Graphics and Image Processing*, 34:344–371, 1986.

[6] C. Brechbühler, G. Gerig, and G. Székely. Compensation of spatial inhomogeneity in MRI based on a parametric bias estimate. In *Proceedings of the 4th International Conference on Visualization in Biomedical Computing (VBC)*, pages 141–146, Hamburg, Germany, Sept. 1996. Springer-Verlag.

[7] H. Breu, J. Gil, D. Kirkpatrick, and M. Werman. Linear time Euclidean distance transform algorithms. *IEEE Trans. on Pattern Anal. and Machine Intelligence*, 17(5):529–533, 1995.

[8] B. H. Brinkmann, A. Manduca, and R. A. Robb. Optimized homomorphic unsharp masking for MR grayscale inhomogeneity correction. *IEEE Trans. on Medical Imaging*, 17:161–171, April 1998.

[9] M. Chalfie, Y. Tu, G. Euskirchen, W. W. Ward, and D. C. Prasher. Green fluorescent protein as a marker for gene expression. *Science*, 263:725–888, 1994.

[10] M. C. Costanza and A. A. Afifi. Comparison of stopping rules in forward stepwise discriminant analysis. *Journal of the American Statistical Association*, 74:777–785, 1979.

[11] M. Dash and H. Liu. Feature selection for classification. *Intelligent Data Analysis*, 1(3):131–156, 1997.

[12] C. de Solorzano, R. Malladi, S. Lelievre, and S. Lockett. Segmentation of nuclei and cells using membrane related protein markers. *Journal of Microscopy*, 201(3):404–415, 2001.

[13] H. Digabel and C. Lantuéjoul. Iterative algorithms. In J.-L. Chermant, editor, *Eropean Symp. Quantitative Analysis of Microstructures in Materials Sciences, Biology and Medicine, Caen, France, Oct. 1977*, pages 85–99. Riederer-Verlag, 1978.

[14] L. Dorst and A. W. M. Smeulders. Length estimators for digitized contours. *Computer Vision, Graphics and Image Processing*, 40:311–333, 1987.

[15] R. O. Duda and P. E. Hart. *Pattern Classification and Scene Analysis*. Wiley, New York, 1973.
[16] F. Erlandsson, C. Linnman-Wählby, S. Ekholm, E. Bengtsson, and A. Zetterberg. A detailed analysis of cyclin a accumulation at the G_1/S border in normal and transformed cells. *Experimental Cell Research*, 256:86–95, 2000.
[17] F. Ferri, P. Pudil, M. Hatef, and J. Kittler. Comparative study of techniques for large-scale feature selection. In E. S. Gelsema and L. N. Kanal, editors, *Pattern Recognition in Practice IV: Multiple Paradigms, Comparative Studies and Hybrid Systems*, pages 403–413. Elsevier, 1994.
[18] J. L. Fleiss, J. Cohen, and B. S. Everitt. Large sample standard errors of kappa and weighted kappa. *Psychological Bulletin*, 72:323–327, 1969.
[19] S. Gilles, M. Brady, J. Declerck, J. Thirion, and N. Ayache. Bias field correction of breast MR images. In *Proceedings of the 4th International Conference on Visualization in Biomedical Computing (VBC)*, pages 153–158, Hamburg, Germany, Sept. 1996. Springer-Verlag.
[20] C. A. Glasbey. An analysis of histogram-based thresholding algorithms. *Computer Vision, Graphics and Image Processing*, 55(6):532–537, Nov. 1993.
[21] R. C. Gonzales and R. E. Woods. *Digital Image Processing*. Addison-Wesley, Reading, 1993.
[22] X. Guorong, C. Peiqi, and W. Minhui. Bhattacharyya distance feature selection. In *Proceedings of the 13th International Conference on Pattern Recognition*, volume 2, pages 195–199. IEEE, Aug. 1996.
[23] J. D. F. Habbema and J. Hermans. Selection of variables in discriminant analysis by F-statistic and error rate. *Technometrics*, 19(4):487–493, 1977.
[24] M. Hu. Visual pattern recognition by moment invariants. *IRE Trans. on Information Theory*, 8(2):179–187, 1962.
[25] A. K. Jain and D. Zongker. Feature selection: Evaluation, application, and small sample performance. *IEEE Trans. on Pattern Anal. and Machine Intelligence*, 19(2):153–158, 1997.
[26] J. Kittler and J. Illingworth. Minimum error thresholding. *Pattern Recognition*, 19(1):41–47, 1986.
[27] M. Kudo, P. Somol, P. Pudil, M. Shimbo, and J. Sklansky. Comparison of classifier-specific feature selection algorithms. In F. Ferri, J. Iñesta, A. Amin, and P. Pudil, editors, *Advances in Pattern Recognition, Joint IAPR International Workshops SSPR 2000 and SPR 2000*, volume 1876 of *LNCS*, pages 677–686, Alicante, Spain, August/September 2000. Springer-Verlag.
[28] Z. Kulpa. Area and perimeter measurement of blobs in discrete binary pictures. *Computer Graphics and Image Processing*, 6:434–454, 1977.
[29] P. Lachenbruch and M. Mickey. Estimation of error rates in discriminant analysis. *Technometrics*, 10(1):1–11, 1968.
[30] P. Lancaster and K. Šalkauskas. *Curve and Surface Fitting, an Introduction*. Academic Press, London, 1986.
[31] C. Lantuéjoul. *La squelettisation et son application aux mesures topologiques des mosaïques polycristallines*. Phd thesis, Ecole des Mines de Paris, 1978.
[32] J. Lindblad. Image analysis for real-time quality control. Master's thesis, Centre for Image Analysis, Uppsala University, Sweden, 1997.
[33] J. Lindblad. Technical documentation of the CERN LHC cable inspection system. Technical report, Centre for Image Analysis, Sept. 1999.
[34] W. E. Lorensen and H. E. Cline. Marching Cubes: A high resolution 3D surface construction algorithm. In *Proceedings of the 14th ACM SIGGRAPH on Computer Graphics*, volume 21(4), pages 163–169, July 1987.

[35] N. Malpica, C. O. de Solorzano, J. J. Vaquero, A. Santos, I. Vallcorba, J. M. Garcia-Sagredo, and F. del Pozo. Applying watershed algorithms to the segmentation of clustered nuclei. *Cytometry*, 28(4):289–297, 1997.

[36] K. V. Mardia. Measures of multivariate skewness and kurtosis with applications. *Biometrika*, 57(3):519–530, 1970.

[37] G. D. Murray. A cautionary note on selection of variables in discriminant analysis. *Applied Statistics*, 26(3):246–250, 1977.

[38] B. Nordin. IPAD, version 2.0 & IMP - an IPAD application. Internal Report 6, Centre for Image Analysis, Uppsala, Sweden, 1997. Available from the author.

[39] N. Otsu. A threshold selection method from gray-level histograms. *IEEE Tran. on System Man and Cybernetics*, 9(1):62–69, 1979.

[40] J. Park and J. M. Keller. Snakes on the watershed. *IEEE Trans. on Pattern Anal. and Machine Intelligence*, 23(10):1201–1205, Oct. 2001.

[41] D. C. Prasher, W. K. Eckenrode, W. W. Ward, F. G. Prendergast, and M. J. Cormier. Primary structure of the aequorea victoria green fluorescent protein. *Gene*, 111:229–233, 1992.

[42] R. J. Prokop and A. P. Reeves. A survey of moment-based techniques for unoccluded object representation and recognition. *CVGIP: Graphical Models and Image Processsing*, 54(5):438–460, 1992.

[43] P. Pudil, F. J. Ferri, J. Novovičová, and J. Kittler. Floating search methods for feature selection with nonmonotonic criterion functions. In *Proceedings of the 12th International Conference on Pattern Recognition*, pages 279–283, Oct. 1994.

[44] P. Pudil, J. Novovičová, and J. Kittler. Floating search methods in feature-selection. *Pattern Recognition Letters*, 15(11):1119–1125, November 1994.

[45] C. E. Queiros and E. S. Gelsema. A note on some feature selection criteria. *Pattern Recognition Letters*, 10(3):155–158, 1989.

[46] P. Ranefall. *Towards Automatic Quantification of Immunohistochemistry Using Colour Image Analysis*. PhD dissertation No. 347, Uppsala University, Uppsala, Sweden, 1998.

[47] P. Ranefall, B. N. L. Egevad, and E. Bengtsson. A new method for segmentation of color images applied to immunohistochemically stained cell nuclei. *Analytical Cellular Pathology*, 15:145–156, 1997.

[48] P. Ranefall, K. Wester, A.-C. Andersson, C. Busch, and E. Bengtsson. Automatic quantification of immunohistochemically stained cell nuclei based on standard reference cells. *Analytical Cellular Pathology*, 17(2):111–123, 1998.

[49] P. Ranefall, K. Wester, and E. Bengtsson. Automatic quantification of immunohistochemically stained cell nuclei using unsupervised image analysis. *Analytical Cellular Pathology*, 16(1):29–43, 1998.

[50] P. Ranefall, K. Wester, C. Busch, P.-U. Malmström, and E. Bengtsson. Automatic quantification of microvessels using unsupervised image analysis. *Analytical Cellular Pathology*, 17(2):83–92, 1998.

[51] S. J. Raudys and A. K. Jain. Small sample size effects in statistical pattern recognition: Recommendations for practitioners. *IEEE Trans. on Pattern Anal. and Machine Intelligence*, 13(3):252–264, 1991.

[52] A. C. Rencher and S. F. Larson. Bias in Wilks's Λ in stepwise discriminant analysis. *Technometrics*, 22(3):349–356, 1980.

[53] T. Ridler and S. Calvard. Picture thresholding using an iterative selection method. *SMC*, 8(8):630–632, 1978.

[54] K. Rodenacker and E. Bengtsson. A feature set for cytometry on digitized microscopic images. *Accepted for Analytical Cellular Pathology.*

[55] J. B. T. M. Roerdink and A. Meijster. The watershed transform: Definitions, algorithms and parallelization strategies. *Fundamenta Informaticae*, 41(1-2):187–228, 2000.

[56] A. Rosenfeld. Fuzzy digital topology. *Information and Control*, 40:76–87, 1979.

[57] I. Rothe, H. Süsse, and K. Voss. The method of normalization to determine invariants. *IEEE Trans. on Pattern Anal. and Machine Intelligence*, 18(4):366–376, 1996.

[58] H. Schulerud and F. Albregtsen. Many are called, but few are chosen.[MAT 22:14] - feature selection and error estimation in high dimensional spaces. *Submitted to Computer Methods and Programs in Biomedicine*.

[59] I. Simon, C. R. Pound, A. W. Partin, J. Q. Clemens, and W. A. Christens-Barry. Automated image analysis system for detecting boundaries of live prostate cancer cells. *Cytometry*, 31:287–294, 1998.

[60] M. Sonka and J. M. Fitzpatrick. *Handbook of Medical Imaging*, volume 2. SPIE Press, Washington, 2000.

[61] B. Stenkvist, E. Bengtsson, O. Eriksson, T. Jarkrans, B. Nordin, and S. Westman-Naeser. Histopathological systems of breast cancer classification: reproducibility and clinical significance. *Journal of Clinical Pathology*, 36:392–398, 1983.

[62] J. A. Swets, R. M. Dawes, and J. Monahan. Better decisions through science. *Scientific American*, 283(4):70–75, 2000.

[63] J. Udupa and S. Samarasekera. Fuzzy connectedness and object definition: theory, algorithms, and application in image segmentation. *Graphical models and image processing*, 58(3):246–261, 1996.

[64] P. S. Umesh Adiga and B. B. Chaudhuri. An efficient method based on watershed and rule-based merging for segmentation of 3-D histo-pathological images. *Pattern Recognition*, 34:1449–1458, 2001.

[65] L. Vincent. Morphological grayscale reconstruction in image analysis: Applications and efficient algorithms. *IEEE Trans. on Image Processing*, 2(2):176–201, 1993.

[66] L. Vincent and P. Soille. Watersheds in digital spaces: An efficient algorithm based on immersion simulations. *IEEE Trans. on Pattern Anal. and Machine Intelligence*, 13(6):583–597, 1991.

[67] K. Voss, H. Süsse, and C. Bräuer-Burchardt. Affine point pattern matching. In *Pattern Recognition, Proceedings of the 23rd DAGM Symposium*, volume 2191 of *LNCS*, pages 155–162, Munich, Germany, Sept. 2001. Springer-Verlag.

[68] C. Wählby, F. Erlandsson, E. Bengtsson, and A. Zetterberg. Sequential immunofluorescence staining and image analysis for detection of large numbers of antigens in individual cell nuclei. *Cytometry*, 47(1):32–41, 2002.

[69] D. H. Wolpert and W. G. Macready. No free lunch theorems for optimization. *IEEE Trans. on Evolutionary Computation*, 1(1):67–82, April 1997.

[70] K. Wu, D. Gauthier, and M. D. Levine. Live cell image segmentation. *IEEE Trans. on Biomedical Engineering*, 42(1):1–12, 1995.

[71] G. Wyvill, C. McPheeters, and B. Wyvill. Data structures for soft objects. *The Visual Computer*, 2(7):227–234, 1986.

[72] L. Zadeh. Fuzzy sets. *Information and Control*, 8(3):338–353, 1965.

Brief summary of the enclosed papers

Paper I: "Image Analysis for Automatic Segmentation of Cytoplasms and Classification of Rac1 Activation"

This paper describes a fully automatic image analysis method for segmentation of cell nuclei and cytoplasms, feature extraction, and classification of cells according to their activation, i.e., GFP-Rac1 translocation and ruffle formation at stimuli. Using a seeded watershed algorithm, cytoplasms are segmented. Selected general and problem specific features are extracted from the segmented cells and non-informative features are removed from the feature set using a floating feature selection procedure. Based on training data produced by visual annotation of a set of images, a classifier is created. The results of the classification made by the computer were compared with results from visual inspection of the time sequences. The automatic classification differed from the visual classification at about the same level as visual classification performed by two different skilled professionals differed from each other.

Paper II: "Algorithms for Cytoplasm Segmentation of Fluorescence Labelled Cells"

A new combination of image analysis algorithms for segmentation of cells imaged by fluorescence microscopy is presented. The algorithm consists of an image pre-processing step, a general segmentation and merging step followed by a splitting step with feedback from a segmentation quality measurement. The segmentation is based on a dual threshold hysteresis version of watershed segmentation. The quality measurement consists of a statistical analysis of a number of shape descriptive features. Objects that have features that differ to that of correctly segmented single cells can be further processed by a splitting step. After the segmentation is completed, the quality of the final segmentation is evaluated. By training the algorithm on a representative set of training images, the algorithm is made fully automatic for subsequent images created under similar conditions. Automatic cytoplasm segmentation was tested on CHO-cells stained with calcein. The fully automatic method showed between 89% and 97% correct segmentation as compared to manual segmentation.

Paper III: "Statistical Quality Control for Segmentation of Fluorescence Labelled Cells"

This paper goes a bit deeper into the statistical quality control applied in Paper II. In this study, we have assembled a sequence of processing steps that lead to an automatic cytoplasm segmentation of fluorescence microscopy cell images. By statistical analysis of measured object features and a feedback system for separation of clustered objects, the final result is improved. The algorithm consists of an image pre-processing step, a general segmentation and merging step followed by a splitting step. Along the chain of processing the quality of the segmentation is estimated, so that the proper next step can be applied. The quality measure is both judging if the cells should be split or not, as well as giving an assessment of the overall result at the end. The quality measurement consists of a statistical analysis of a selected set of descriptive features. A fully automatic method to select a set of features to include in the quality measure is presented. The feature selection is devised to avoid over training, but still facilitate a powerful and flexible classifier.

Paper IV: "A Comparison of Methods for Estimation of Intensity Nonuniformities in 2D and 3D Microscope Images of Fluorescence Stained Cells"

This paper describes the shading correction method applied in the projects described in Papers I and II. A comparison of the accuracy and robustness of different data driven methods for intensity nonuniformity field estimation (background shading estimation) on simulated and real images of fluorescence stained cells is presented. A novel attempt to reduce the parameter space of a B-spline based algorithm for shading estimation, using automatic thresholding with a kernel density estimator, is tested and compared with the optimal parameter value, found by exhaustive testing on simulated data.

Paper V: "Histogram Thresholding using Kernel Density Estimates"

This paper describes the thresholding methodology applied in Paper IV. A nonparametric method for thresholding of data which makes very weak assumptions about the underlying distribution is presented. The output consists of one or more recommended threshold points, along with a scalar measurement of significance which enables easy comparison of different thresholds. The suggested thresholds are given as real-valued scalars, and are not restricted to integer values or anything similar to bin-size of an ordinary histogram. The method has been shown to work well in different situations, including cases where it is hard to find a good threshold by visual inspection.