UPPSALA
UNIVERSITET

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology 856

Methods for 2D and 3D Quantitative Microscopy of Biological Samples

AMIN ALLALOU

ACTA
UNIVERSITATIS
UPSALIENSIS
UPPSALA
2011

ISSN 1651-6214 0346-5462
ISBN 978-91-554-8167-4
urn:nbn:se:uu:diva-159196

Dissertation presented at Uppsala University to be publicly examined in room 2446, Polacksbacken, Lägerhyddsvägen 2, Uppsala. Friday, November 11, 2011 at 10:15 for the degree of Doctor of Philosophy. The examination will be conducted in English.

Abstract
Allalou, A. 2011. Methods for 2D and 3D Quantitative Microscopy of Biological Samples. Acta Universitatis Upsaliensis. *Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology* 856. 72 pp. Uppsala. ISBN 978-91-554-8167-4.

New microscopy techniques are continuously developed, resulting in more rapid acquisition of large amounts of data. Manual analysis of such data is extremely time-consuming and many features are difficult to quantify without the aid of a computer. But with automated image analysis biologists can extract quantitative measurements and increases throughput significantly, which becomes particularly important in high-throughput screening (HTS). This thesis addresses automation of traditional analysis of cell data as well as automation of both image capture and analysis in zebrafish high-throughput screening.

It is common in microscopy images to stain the nuclei in the cells, and to label the DNA and proteins in different ways. Padlock-probing and proximity ligation are highly specific detection methods that produce point-like signals within the cells. Accurate signal detection and segmentation is often a key step in analysis of these types of images. Cells in a sample will always show some degree of variation in DNA and protein expression and to quantify these variations each cell has to be analyzed individually. This thesis presents development and evaluation of single cell analysis on a range of different types of image data. In addition, we present a novel method for signal detection in three dimensions.

HTS systems often use a combination of microscopy and image analysis to analyze cell-based samples. However, many diseases and biological pathways can be better studied in whole animals, particularly those that involve organ systems and multi-cellular interactions. The zebrafish is a widely-used vertebrate model of human organ function and development. Our collaborators have developed a high-throughput platform for cellular-resolution *in vivo* chemical and genetic screens on zebrafish larvae. This thesis presents improvements to the system, including accurate positioning of the fish which incorporates methods for detecting regions of interest, making the system fully automatic. Furthermore, the thesis describes a novel high-throughput tomography system for screening live zebrafish in both fluorescence and bright field microscopy. This 3D imaging approach combined with automatic quantification of morphological changes enables previously intractable high-throughput screening of vertebrate model organisms.

Keywords: Image analysis, cytmetry, model organism, zebrafish, screening

Amin Allalou, Uppsala University, Centre for Image Analysis, Lägerhyddsv. 3, SE-752 37 Uppsala, Sweden.

ISSN 1651-6214 0346-5462
ISBN 978-91-554-8167-4
urn:nbn:se:uu:diva-159196 (http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-159196)

My family

List of Papers

This thesis is based on the following papers, which are referred to in the text by their Roman numerals.

I A. Allalou, F. M. van de Rijke, R. J. Tafrechi, A. K. Raap, and C. Wählby. Image Based Measurements of single cell mtDNA mutation load. In *Proceedings of the 15th Scandinavian Conference on Image Analysis (SCIA)*, Aalborg, Denmark. Published in Lecture Notes in Computer Science (LNCS) 4522, pp. 631-640, 2007.

II R. J. Tafrechi, F. M. van de Rijke, A. Allalou, C. Larsson, W. C. R. Sloos, M. van de Sande, C. Wählby, G. M. C. Janssen, and A. K. Raap. Single-cell A3243G mitochondrial DNA mutation load assays for segregation analysis. *Journal of Histochemistry and Cytochemistry*, 55: 1159-1166, 2007.

III A. Allalou, and C. Wählby. BlobFinder; a tool for fluorescence microscopy image cytometry. *Computer Methods and Programs in Biomedicine*, 94(1):58-65, 2009.

IV A. Pinidiyaarachchi, A. Zieba, A. Allalou, K. Pardali, and C. Wählby. A detailed analysis of 3D subcellular signal localization. *Cytometry Part A*, 75(4):319-328, 2009.

V A. Allalou, A. Pinidiyaarachchi, and C. Wählby. Robust signal detection in 3D fluorescence microscopy. *Cytometry Part A*, 77(1):86-96, 2010.

VI C. M. Clausson, A. Allalou, I. Weibrecht, S. Mahmoudi, M. Farnebo, U. Landegren, C. Wählby and O. Söderberg. Increasing the dynamic range of *in situ* PLA. *Accepted for publication in Nature Methods*, 2011.

VII C. Pardo-Martin*, T. Y Chang*, A. Allalou*, C. Wählby and M. F. Yanik. High-throughput cellular-resolution *in vivo* vertebrate screening. *In Proceedings of the the 15th International Conference on Miniaturized Systems for Chemistry and Life Sciences*, Seattle, USA,

October, 2011.

VIII T. Y. Chang*, C. Pardo-Martin*, A. Allalou, C. Wählby, and M.F. Yanik. Fully automated cellular-resolution vertebrate screening platform with parallel animal processing. *Submitted for journal publication*, August 2011.

IX C. Pardo-Martin*, A. Allalou*, P. Eimon, C. Wählby, and M.F. Yanik. High-throughput *in vivo* optical projection tomography of small vertebrates. *Manuscript*, 2011

Reprints were made with permission from the publishers.

The method development and writing for Paper I was performed almost entirely by the author. The author was only involved in the image analysis section of Paper II. Paper I and Paper II have a lot of content overlap, but the author chose to include both papers since they have unique parts important to this thesis. In Paper III the method development and writing was performed mainly by the author. In Paper IV the main work was performed by Amalka Pinidiyaarachchi, the author contributed with the signal detection part. For Paper V method development and writing was done mainly by the author with contributions from Amalka Pinidiyaarachchi. The author contributed with the method development and writing for the section containing image analysis in Paper VI. Paper VII and VIII were performed by the author in close corporation with Carlos Pardo and Tsung-Yao Chang. The main focus for the author was the part containing image processing. Paper VII and VIII also have a lot of content overlap. However, the author feels that both papers are contributing to the thesis and therefore chose to include both papers. Paper IX was done in close corporation between the author and Carlos Pardo.

For color versions of Papers II-VIII visit the corresponding journal or publisher's website.

* means that the authors contributed equally to the work.

Related work

In addition to the papers included in this thesis, the author has also written or contributed to the following publications:

A. Allalou, F. M. van de Rijke, R. J. Tafrechi, A. K. Raap, and C. Wählby. Segmentation of cytoplasms of cultured cells. *In Proceedings of the Swedish Society for Automated Image Analysis (SSBA) Symposium on Image Analysis*, Linköping, Sweden, March 2007.

A. Allalou, and C. Wählby. Image based measurements of single cell mtDNA mutation load. *In Proceedings of Medicinteknikdagarna*, Örebro, Sweden, October 2007.

A. Allalou, and C. Wählby. Signal detection in 3D by stable wave signal verification. *In Proceedings of the Swedish Society for Automated Image Analysis (SSBA) Symposium on Image Analysis*, Halmstad, Sweden, March 2009.

A. Allalou, V. Curic, C.P.-Martin, M. F. Yanik, and C. Wählby. Approaches for increasing throughput and information content of image-based zebrafish screens. *In Proceedings of the Swedish Society for Automated Image Analysis (SSBA) Symposium on Image Analysis*, Linköping, Sweden, March 2011.

Contents

1. Introduction

1.1 Objective and motivation

In recent years, great improvements have been made in the field of microscopy, both microscope hardware and staining techniques. This has allowed scientists to generate vast amounts of high content data in a short period of time. Manual analysis of the data is time consuming and in some cases even impossible. Together with the rapid growth in computer power, digital image analysis has become a part of everyday work in biology and medicine.

Microscopy provides a excellent tool for studying gene and protein expressions in cells. Examining a single cell or organism often provides little significant knowledge due to sample heterogeneity. Even a homogeneous population will show a certain degree of variability. Increasing the number of examined cells/organisms will provide better statistics and an opportunity to study the variation in the population. However, the high number of examined samples will result in large amounts of data that are extremely time consuming to manually quantify. Furthermore, some characteristics are almost impossible to manually quantify in a consistent and unbiased manner. Image analysis provides a tool for fast measurements on huge amounts of data. In addition, the unbiased nature of an analysis performed by a computer provides the opportunity to perform an analysis that is neutral towards the outcome and fully reproducible.

Acquisition of cell images with point like signals as markers for specific DNA sequences or proteins are common in biomedical research. Cells in a sample will always show some degree of variation in their characteristics. To catch these variations the analysis must be performed on single cells, since individual characteristics will be lost in an average of the image. Methods and software that can accurately detect and quantify the number of signals for individual cells will provide extremely useful tools in the biomedical research field.

High-Throughput Screening (HTS) is a technique for searching large libraries of chemical or genetic perturbants, to find new treatments for a disease or to better understand disease pathways. HTS of cell-based assays has been widely used [10, 31]. However, studying the effects of a disease on isolated cells will not always reveal information on the effects on the whole animal. As a result, HTS on whole animal is becoming more popular and new techniques are emerging rapidly [14]. The zebrafish is one of the animal models

that is becoming increasingly popular for use in HTS [78]. These new systems require new image analysis methods that can work in a fully automated manner.

1.2 Thesis outline

This thesis is focused on the development of image analysis methods for detecting and quantifying signals and structures in 2D and 3D image data from cells and zebrafish.

Chapter 2 consists of a brief introduction to the different microscopy and staining techniques used throughout the thesis for image acquisition. In addition, the basic terminology of digital image analysis is introduced. Furthermore, the main image analysis methods in this thesis are introduced and briefly explained.

Chapter 3 contains a description and discussion of the papers included in this thesis. This chapter is divided in two parts; methods and applications for cell images, and methods and applications for zebrafish images.

Chapter 4 contains concluding remarks and future work.

2. Background

This chapter will provide some background information in order to get a better understanding of the methods and applications discussed in chapter 3. First, the two different types of biological models used in the thesis are described briefly. Secondly, the different image acquisition methods are presented and described in brief. The last part consists of an introduction to some basic concepts in digital image analysis and a more in depth description of some of the methods commonly used throughout the thesis.

2.1 Models

The human body is a system built up of many small complex systems. Depending on the biological question we can observe this system on different scales. Starting from subatomic particles that build up the atoms. The atoms in turn are the components that build up molecules. Molecules together form the next level, which contains macromolecules. The macromolecules, DNA and proteins, are the basis of cells. Numerous cells then form tissues that together with other tissues form organs. Finally, all the organs in the end build up the whole organism [43].

By studying the subatomic level, interaction of protons and electrons, we will gain knowledge of how different molecules are built up. However, if we want to know how organs in the organism function together, studying the subatomic level will provide us little useful information. In a system, like the human body, the combination of lower level systems provides a higher order system with new and unique properties that could not have been predicted, *a priori*, from the laws of the lower level [3]. This feature leads to a loss of linearity when moving from one level to another, and as a consequence we need to study several levels or choose the level that is best suited for the question asked.

A model system can be thought of as a simplification of a complex system. In biology it is common to perform experiments on a model system with the expectation that the discoveries can be translated into more complex systems. The model system can be, e.g., cell cultures [22], unicellular organisms [44] or mutli-cellular organisms [48]. The choice of model system will depend on the question asked. If we want to identify a protein change as a result of a certain mutation, an isolated cell culture is an appropriate model. However, if

we aim to examine the phenotype change in an animal as a result of the same mutation we need to use the whole animal as a model.

2.1.1 The cell

The cell is the functional and structural unit in all living organisms and is often called the building block of life. The human organism consists of $\sim 10^{14}$ cells. The average diameter of cells in a multicellular organism is between 10 and 30μm [68]. The DNA is contained inside the cell and is responsible for all genetic instructions in the development and function of the cells. Through a complex system of chemical signals the cells divide and differentiate into building blocks of different parts of an organism.

Living organisms are built up by cells, therefore, a defect in any or multiple cells will affect the entire organism. For example, in tumors, the cell cycle is disrupted and this is causing the individual cells to grow uncontrollably. This leads to failure in regulating the tissue growth in an individual. The lack of control in growing tissue may cause the tumor to invade and destroy adjacent tissues, causing malignancy. By studying how and why the cells start to lose control of their cell cycle, it is possible to get a better understanding of the causes of cancer tumors. For all types of diseases, investigating on a cellular level will provide information of the cellular mechanisms of the disease.

2.1.2 The zebrafish

The zebrafish (*Danio reiro*) is a very good model organism for vertebrate development. The zebrafish embryos develop externally and therefore all stages of the development can be easily viewed and manipulated. The organization of the embryo is simple and the body is transparent, making it easy to study by microscopy. In addition, the embryonic development is very rapid. After 5-6 days post-fertilization (dpf) all major organs are present in the larvae. Furthermore, after 3-4 months the zebrafish is able to generate new offspring. A female zebrafish produces hundreds of eggs each week. The zebrafish genome has been fully sequenced, and many transgenic lines with different mutations are available. All these features make the zebrafish an ideal model organism for studying organ development and pathways related to human disease [6, 59].

2.2 Microscopy

Optical microscopy is a technique used to magnify small objects, making it essential for the study of cells or small organisms. Visible light and a system of lenses, is used to magnify the sample. All the images used in this thesis are

acquired with some type of microscopy. A brief introduction to the different types of microscopes is presented in this section.

2.2.1 Bright field microscopy

In bright field microscopy white light is transmitted through a sample and the differences in absorptions are visualized. Only specimen that have properties that affect the amount of light that passes through can be visualized with this type of microscope. The setup consists of a light source, a condenser lens that concentrates the light onto the specimen, an objective lens that collects the light and magnifies the sample, and a detector consisting of oculars or a camera [9]. Many unstained cells are transparent and therefore have little contrast in bright field microscopy. Staining adds color to a sample and enhances contrast. Unfortunately, many dyes are toxic and can only be used with fixed (dead) cells, thus limiting the use of bright field microscopy.

2.2.2 Fluorescence microscopy

In fluorescence microscopy a sample is irradiated with a specific band of wavelengths. These wavelengths are absorbed by fluorophores and light of longer wavelengths are emitted. Through optical filters only specific wavelengths of the emitted light reach the detector (see Fig. 2.1). In fluorescence microscopy, the use of a fluorophore capable of emitting light in the detectable visible range, defined by the filters, is required to visualize the sample [58].

In contrast to bright field microscopy, where the sample is observed together with the incident light, fluorescence microscopy makes use of the difference in excitation and emission wavelengths to block the incident light. This results in an image with high contrast between sample and background.

2.2.3 Confocal microscopy

In a wide field fluorescence microscope (Sec. 2.2.2) the entire sample is illuminated at the same time, all the emission from the specimen is collected including the unfocused background light. A confocal microscope eliminates the out of focus light by using point illumination and a pinhole in an optically conjugate plane in front of the detector. In order to create an image the focused spot of light must be scanned across the specimen. The use of a focused spot of light enables the control of imaging different depth of the specimen. By imaging several different focal depths a 3D image can be acquired [53]. Confocal microscopy poses several advantages over conventional wide field optical microscopy: reduction of information outside the focal plane, depth of field control and the ability to collect serial optical sections.

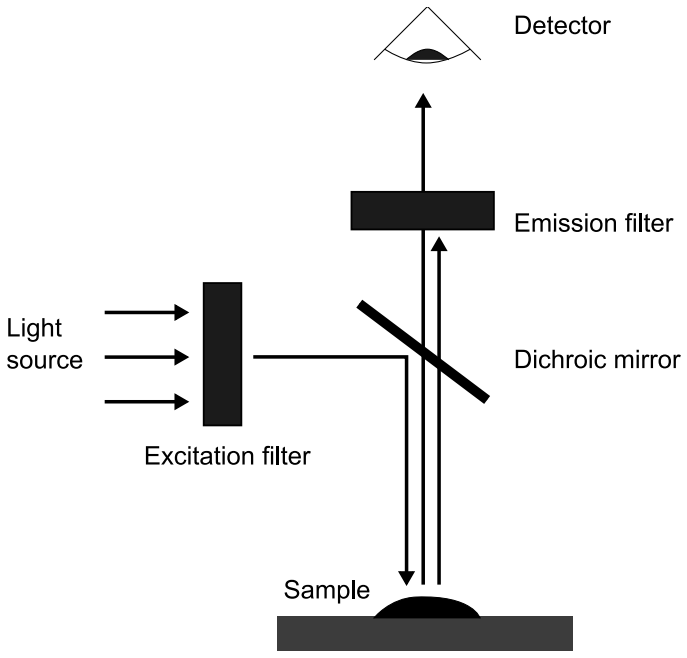


Figure 2.1: Simplified diagram of a fluorescent microscope setup.

2.2.4 Point spread function

In microscopy, a point source or point object will be seen as a blurred spot in the acquired image. The point spread function (PSF) describes the relationship between the point object and the blurred response produced in the microscope. An image from a microscope consists of a sum of all PSF from the point objects in the scene. A wider PSF will decrease the resolution in the acquired image.

The PSF differs between different microscopes and microscopy techniques. More blurring is usually seen in the z-direction (axial) than in the $x-y$ direction (lateral). Confocal microscopes decrease the size of the PSF in all directions, i.e., improving the resolution in all directions. Even though the axial resolution is improved, it is still lower than the lateral [7]. If the PSF of the microscope is known, deconvolution methods can be used to reduce the blurring effect caused by the PSF [74].

2.2.5 Optical projection tomography

Optical Projection Tomography (OPT) was invented by Dr James Sharpe at the Medical Research Council, Human Genetics Unit, in Edinburgh 2001 [63]. OPT combines conventional light microscopy with tomography to acquire 3D

images of samples that are too large (>1mm) to be imaged with confocal microscopy. In Computerized Tomography (CT), images are acquired from different angles of a sample and combined to produce a 3D image [32]. CT has been used with X-ray images for a long time, but in OPT a conventional light microscope is used. By imaging the whole sample with a large depth of field, keeping the whole sample in focus, specimen between 1 and 15 mm can be reconstructed in 3D. This technique can also be combined with staining techniques to label internal structures and complex gene activity.

2.3 Labeling techniques

Cells and cellular structures are often colorless or transparent making them hard to see in a microscope. Labeling techniques can be used to enhance, e.g., specific detection events, parts of a cell, cell populations or tissues.

2.3.1 Fluorescent labels

In order for a sample to be visible in fluorescence microscopy it must be fluorescent. This can be done by using a fluorescent stain or through the expression of a fluorescent protein. Sometimes the intrinsic fluorescence of a sample can also be used, called autofluorescence.

Some fluorescent stains consist of molecules that are intrinsically fluorescent and can bind to biological molecules of interest, e.g., DAPI binds to DNA and labels the nucleus of a cell [33]. Immunofluorescence is an antibody-based labeling technique. A highly specific antibody, conjugated to a fluorophore, binds to its antigen and labels specific proteins or molecules within the cell. An alternative approach is to use a secondary antibody that is conjugated to a fluorophore and binds specifically to the unlabeled primary antibody, raised in another species [34].

The green fluorescent protein (GFP) produced by the jellyfish *Aequorea victoria* is commonly used in labeling. GFP emits bright green light when exposed to blue light. The gene for the production of GFP has been isolated and chimeric genes, artificial genes consisting of fragments of unrelated genes or other DNA segments, can be constructed of the GPF gene and a gene of interest. This makes it possible to have an *in vivo* fluorescent protein that can be followed in a living system [69].

In DNA and RNA analysis, a small number of target sequences must be accurately detected among a large background of irrelevant nucleic acids. Padlock probes together with rolling-circle amplification (RCA) exhibit very high specificity and are a good tool for this type of task [5]. Padlock probes are oligonucleotides that become circularized when an appropriate target DNA or RNA sequence is present, see Fig. 2.2(a-c). The reaction is highly specific since it requires a perfect match at both ends of the probe to join the ends and

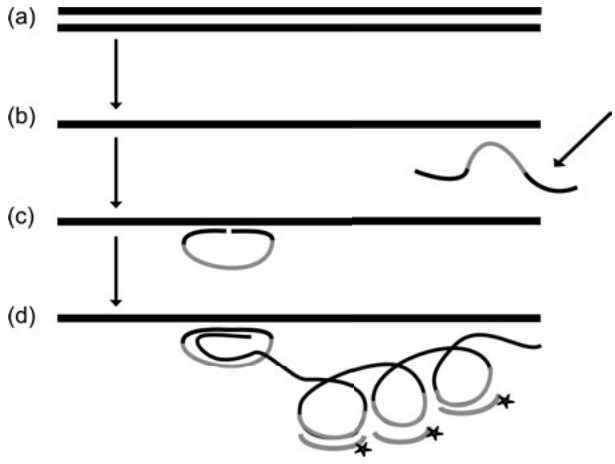


Figure 2.2: Overview of padlock probes with rolling-circle amplification. (a) Target DNA. (b) The target DNA is made single stranded and padlock probes are added. (c) The target-matched padlock probes are circularized after perfect match. (d) Amplification of the signal through rolling-circle amplification, and subsequent detection by fluorescently labeled oligonucleotides

after ligation produce a circularized probe around the target sequence. RCA is a process of synthesizing multiple copies of circular DNA or RNA by replication. When the padlock probe has ligated and formed a circular DNA or RNA strand, RCA will create a long strand of multiple copies of the first circle. These copies are detected by fluorescently labeled oligonucleotides creating a strong fluorescent signal, see Fig. 2.2d. Another advantage is that when multiple probes are added simultaneously, unlike Polymerase Chain Reaction (PCR), cross-reactions are unlikely to arise and the risk of false products are low [5].

In situ Proximity Ligation Assay (PLA) can be used to visualize proteins, protein-protein interactions, and post-translational modifications in cells and tissues. The method was developed by Professor Ulf Landegren *et al*., and commercialized by Olink Biosciences (Uppsala Science Park, Sweden) [66]. PLA is a method that uses two primary antibodies that recognize the target antigen or antigens of interest. Secondary antibodies, PLA probes, with a unique short DNA strand attached to it, bind to the primary antibodies. When the PLA probes are in close proximity they will, together with added connector oligonucleotides, form circular DNA. After ligation, similar to the padlock probes, the circular DNA strand will be amplified with RCA and produce a strong signal, when detected. Since the *in situ* PLA technology requires positive identification of two different proteins or epitopes on the same protein, specificity is enhanced compared to assays that depend only on single binding

recognition [39]. When two primary antibodies are used in *in situ* PLA, they must be raised in different species.

2.3.2 Bright field labels

Many different light absorbing and light scattering stains exist for bright field microscopy. Here we focus on two stains used in this thesis work.

Alcian blues are a family of polyvalent basic dyes. Alcian blue 8GX is the most commonly used member of the family. The dye can be used to visualize glycosaminoglycans with a light or electron microscope [28]. The stain is commonly used in histochemistry and cytochemistry. Furthermore, alcian blue can be used as a bone marker in zebrafish [17].

Similar to PLA in fluorescence microscopy, Zeiba. et. al developed a method for the use of PLA with bright field microscopy. Horseradish peroxidase (HRP) is conjugated to oligonucleotides similar to fluorescent molecules in the fluorescence-based readout. HRP/NovaRED uses enzymatic conversion of NovaRED substrate by HRP to a colored product visualizing the proteins *in situ*. The bright field PLA shows equivalent results to the fluorescent method. The staining is compatible with conventional histologic staining [77].

2.4 Digital image analysis

A digital image is an image represented in a computer or any other digital device. It is a discrete representation of the continuous scene that was imaged. Digital image analysis extracts information from a digital image through the aid of a computer. Image processing is closely related to image analysis but here the output is a processed image, enhanced in some way, instead of information.

This section describes some of the general concepts in image processing and image analysis. Some concepts and methods that are used throughout the thesis are described in more detail.

2.4.1 Basic concepts

Every image analysis task is unique but there is a general scheme with some fundamental steps that are common for almost all image analysis problems:

Image acquisition	Image acquisition is the process of digitizing a scene into a digital image. This is a crucial step since the acquired image reflects the scene and limits the information that subsequently can be extracted from it. It is important to acquire the image in such a way that it is optimally suited for the information that is intended to be extracted.
Pre-processing	Pre-processing, as the name suggests, is processing done to the image prior to analysis. This step alters the image to make it more suitable for analysis through, e.g., reduction of noise, normalizing intensity non-uniformities, enhancing edges, aligning images, etc.
Segmentation	Segmentation refers to the process of partitioning the image into multiple segments. Segmentation is often used to separate objects from the background or to identify different objects in the image. Segmentation is further described in section 2.4.2.
Feature extraction	Feature extraction consists of the extraction of interesting features from the objects of interest in the image. The set of relevant features has to be selected based on the information that is of interest in the analysis. Features can be, e.g., color, shape, size, texture, etc.
Classification	Classification is the step of separating the segmented objects into different categories. For example, a cell culture might contain large and small cells. After segmentation, the size feature can be extracted from each cell. The classification then divides the segmented cells into groups of large and small cells based on this size feature.
Data analysis	Data analysis consists of gathering information from the previous steps and representing them as an information output, e.g., number of large cells vs. number of small cells.
Evaluation	Evaluation is necessary during development. This step consists of evaluating the method and results through statistical tools to validate the analysis. This can be done by comparing the method with ground truth data or a manual/visual expert analysis.

2.4.2 Segmentation

The aim of segmentation is to divide the image into different regions that are homogeneous with respect to certain criteria. The segments are often referred to as objects and can be labeled with a unique number identifying the different segments in the image.

2.4.2.1 Thresholding

Thresholding is the simplest method of image segmentation. It is a pixel-based segmentation method, where each pixel is segmented based only on the individual pixel value. The output from the thresholding is a binary image $g(x,y)$ (an image consisting of 0 and 1). A threshold value T is defined and pixels with values above T are considered to be part of the object in the image $f(x,y)$:

$$g(x,y) = \begin{cases} 1 & \text{if } f(x,y) \geq T \\ 0 & \text{if } f(x,y) < T \end{cases}. \tag{2.1}$$

The value T can be set manually or selected automatically from certain criteria. As a guide for choosing an appropriate threshold value, an image histogram is often used. The image histogram $p(f)$ of the image $f(x,y)$ is the probability density function which gives the frequency of the different pixel values in $f(x,y)$.

Often, using a manual threshold is not the most desirable way of thresholding an image. Manual input is undesired, as it is time consuming and may vary depending on the user, and should be minimized. Many different automatic thresholding techniques have previously been developed [61]. One commonly used method was developed by Otsu in 1979 [52]. From the histogram, Otsu's algorithm selects a threshold that minimizes a weighted intra-class variance of the background and foreground, see Fig. 2.3. The method relies on the assumption that all the pixels in the image belong to either the background or the foreground (i.e. objects). Otsu's method of thresholding is used for separation of foreground and background in papers I, II, III and IX.

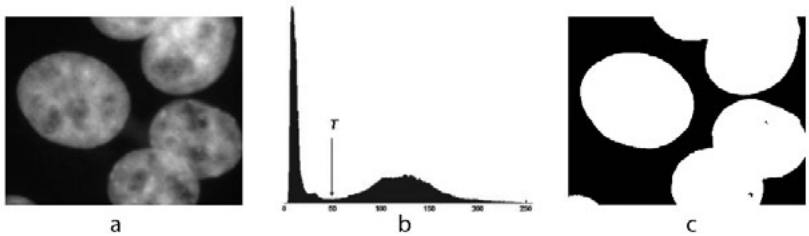


Figure 2.3: (a) Original image $f(x,y)$. (b) Histogram $p(f)$ with T representing the selected threshold for image segmentation. (c) Binary image $g(x,y)$ after thresholding with value T.

2.4.2.2 Watershed

The watershed algorithm was originally presented by Beucher and Lentuéjoul (1979) [37] and later refined in a more efficient implementation by Vincent and Soille (1991) [72]. The watershed segmentation is a region-based segmentation method and has been extensively used in many areas of image analysis, e.g., cell segmentation [11, 45, 73].

The watershed segmentation can be understood by seeing the image as a landscape. The gray-level intensity represents the elevation in this landscape. The landscape image can be, e.g., the original gray-level image, a distance transformed image (an image containing a distance value to the nearest object pixel or nearest background pixel) [47], or a gradient magnitude image (edge image), see Fig. 2.4b. The algorithm can be illustrated by letting water enter through the local minima and start to rise. A lake around a local minimum is created and referred to as catchment basin. When the water fronts from different catchment basins meet they form a dam or watershed that separates the catchment basins. All that is left after the watershed segmentation are watershed lines separating the objects, see Fig. 2.4. If objects of interest are bright rather than dark, the image is inverted before applying watershed segmentation.

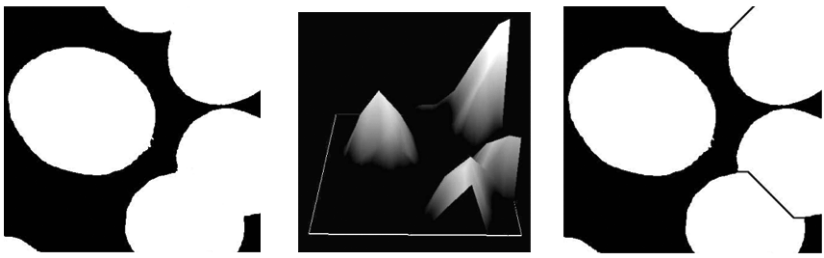


Figure 2.4: (a) Binary image before watershed segmentation. (b) Landscape like representation of the distance transformed image. (c) Image after watershed segmentation.

Watershed segmentation is commonly used in cell-segmentation. In papers I, II, III, IV and VI, watershed is used to separate clustered nuclei and/or to segment the cytoplasm.

2.4.2.3 Level sets

Level set methods present a powerful tool for image segmentation [12, 13, 41, 51]. In level set methods, a level set function is defined as $\phi(i,j,t)$, where (i,j) are coordinates of the image and t is time. At any given time the level set function defines a contour at $\phi = 0$ and the segmentation regions are defined by $\phi < 0$ and $\phi \geq 0$. The contour of the function evolves according to some

partial differential equation and finally reaches a steady state, corresponding to the final segmentation.

The Chan-Vese level set algorithm uses variational calculus to evolve the level set function ϕ [13]. This method differs from most other level set methods in the sense that here the curve evolution is driven by region-based energy, instead of edge-based energy. The method works by seeking a function that minimizes a functional. The goal of the segmentation is to minimize the functional for a given image, and the segmentation will be defined by the level set function ϕ. A general form of the functional is

$$F(c_1,c_2,\phi) = \mu \int_\Omega |\nabla H(\phi(x,y))| dxdy + \nu \int_\Omega H(\phi(x,y)) dxdy + \lambda_1 \int_\Omega |I(x,y)-c_1|^2 H(\phi(x,y)) dxdy + \lambda_2 \int_\Omega |I(x,y)-c_2|^2 (1-H(\phi(x,y))) dxdy, \tag{2.2}$$

where μ, ν, λ_1 and λ_2 are parameters that are set by the user to fit a certain type of image. I is the image and c_1 and c_2 are the averages of the image in the regions inside and outside of the contour, respectively. H is the Heaviside function.

$$H(\phi) = \begin{cases} 1 & \text{if } \phi \geq 0 \\ 0 & \text{if } \phi < 0 \end{cases} \tag{2.3}$$

The first term in (2.2) represents the length penalty term. If regions are expected to have smooth borders then this term should be heavily weighted with parameter μ. Similarly, the second term in (2.2) is a penalty term for the total area of the foreground, regulated by parameter ν. The third term is a variance measure of the intensities in the foreground and the fourth term is the variance of the intensities in the background. The final segmentation is acquired by minimizing the sum of all these terms. This should lead to a segmentation that has a background and foreground region that are as uniform as possible. The Chan-Vese level set method is used in paper VI to separate the cells from the background.

2.4.2.4 Point detectors

Fluorescent biomarkers make biomolecules visible as point-like signals (PLS) in the captured microscope image data. Large experiments of this type produce huge amounts of data where manual detection of the signals is a time consuming task. Several automatic methods to detect PLS have previously been presented [65]. A PLS in this context can generally be defined as a small object, relatively higher in intensity than the image background. For a point detector to work well the method should be robust to different intensities and sizes of

the signals. In addition, the method should be able to distinguish individual signals that are clustered together.

A few of common point detector methods, with which we compare our novel method in Paper V, are described here in more detail.

TopHat transformation. The morphological TopHat transformation (TH) makes use of the difference between a morphologically opened image and the original image [49]. The TopHat transformation to find signals of high intensity is defined by

$$I_{TopHat} = I - max_B(min_B(I)), \tag{2.4}$$

where I is the input image and max_B is the maximum in a neighborhood defined by structuring element B, and min_B is the minimum. The size and shape of the structuring element is crucial for the performance of the method. The signal to be detected should fit inside the structuring element for the method to enhance the signal. The structuring element must take the PSF of the microscope into consideration. Often the signal is more smeared along the axial (z) direction than the lateral ($x-y$), and the structuring element should therefore be larger in the axial direction. The output from the algorithm is an image containing all the local maxima of the input image that fit inside the structuring element.

Difference of Gaussian. The Difference of Gaussian (DoG) is an edge and signal detection filter that works by blurring the image with a wide and narrow Gaussian kernel [60]. The result after applying the wider Gaussian is subtracted from the result after applying the narrow Gaussian and an image with enhanced signals is created. Setting the different sizes of the two Gaussians is crucial to get a good detection of signals of a certain size. A 2D DoG proposed in [35] has the following form:

$$DoG_{\sigma,\gamma}(x,y) = \frac{1}{2\pi\sigma^2\gamma^2}\exp^{-\frac{x^2+y^2}{2\gamma^2\sigma^2}} - \frac{1}{2\pi\sigma^2}\exp^{-\frac{x^2+y^2}{2\sigma^2}}, \tag{2.5}$$

where σ is the standard deviation of the wide Gaussian and γ ($0<\gamma<1$) is the ratio of the standard deviation of the wide and narrow Gaussian. The zero crossing of the DoG is important when choosing the correct value of the two Gaussians. Setting the equation $DoG_{\sigma,\gamma}=0$ a relationship between radius r, distance to zero crossing, and σ for any value of γ ($0<\gamma<1$) can be derived.

$$r = 2\gamma\sigma\sqrt{\frac{-\ln\gamma}{1-\gamma^2}}, \tag{2.6}$$

$$\sigma = \frac{r}{2\gamma}\sqrt{\frac{1-\gamma^2}{-\ln\gamma}}. \tag{2.7}$$

From (eq. 2.7) it is easy to get the σ from a desired radius r of the zero crossing. The γ parameter can be used to regulate the valleys of the DoG. A high value of γ will result in narrow valleys and will make the filter smaller and consequently better at separating signals that are close in space. On the other hand a low value of γ will produce wider valleys that provide more smoothing.

In [35] the DoG is defined in 2D only. In paper V we use the same derivation as above to yield a relationship between σ and the desired distance to the zero crossing r in 3D.

Multiscale Product. The Multiscale Product (MP) was proposed by Olivo-Marin [50] and detects signals with a Multiscale Product from the á trous wavelet decomposition of the original image. The original method was proposed in 2D but was extended to 3D in [71]. The method is designed to detect signals that resemble a 3D Gaussian intensity profile of a given lateral size. The 3D confocal PSF can be approximated by a 3D Gaussian [76]. Since the PSF is often anisotropic, an anisotropic 3D Gaussian is required. In [71] the axial to lateral ratio has been set to 3, a typical ratio in confocal microscopy. This gives

$$G(\sigma_i) = \frac{1}{3\sigma_i^3(2\pi)^{3/2}} \exp\left(-\frac{1}{2}\left(\frac{x^2}{\sigma_i^2} + \frac{y^2}{\sigma_i^2} + \frac{z^2}{9\sigma_i^2}\right)\right). \tag{2.8}$$

A Gaussian scale space g_i is produced, by convolving the image with Gaussian filters of different widths. A scale base, b, is used to define the different widths of the the Gaussian,

$$\sigma_i = b\sqrt{2}^i. \tag{2.9}$$

The last step consists of multiplying the differences in the Gaussian scale space, resulting in an image, I_{MP}, with enhanced contrast where signals are present,

$$I_{MP} = (g - g_0)(g_0 - g_1)(g_1 - g_2). \tag{2.10}$$

2.4.3 Tomographic reconstruction

When imaging a sample in 2D, underlying structures will be obscured by the overlaying structures. Tomography is a non-invasive 3D imaging technique that allows visualization of internal (underlying) structures by acquiring projection images from several different angles. There are many different tomography techniques available, e.g., X-ray Computed Tomography (CT), Single-Photon Emission Computed Tomography (SPECT) and Positron Emission Tomography (PET) [16]. Reconstructing the 3D image can be done by backprojection (Sec. 2.4.3.1) or with iterative methods (Sec. 2.4.3.2).

2.4.3.1 Filtered Backprojection

Backprojection is a direct implementation of the inverse radon transform [56]. Each single projection corresponds to an absorption pattern for a specific direction through the sample. The projection data is projected back across the image in the same direction it was acquired. This is done for all angles used in the acquisition, Fig. 2.5.

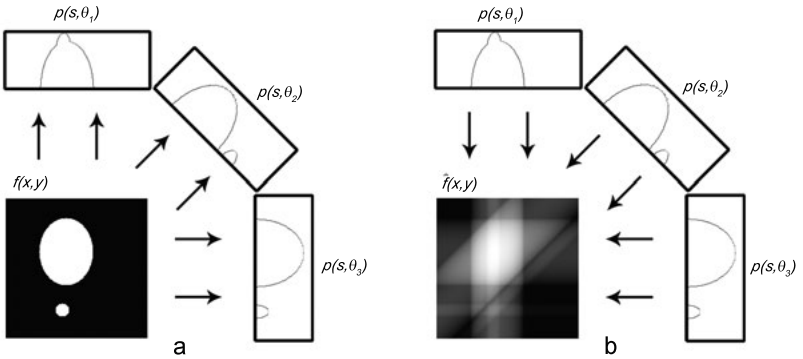


Figure 2.5: (a) Projections $p(s,\theta)$ from image $f(x,y)$. (b) Backprojection of $p(s,\theta)$ to reconstructed image $\hat{f}(x,y)$

.

We define $p(s,\theta)$ as the line integral of the image $f(x,y)$, along a line with distance s from the origin and an angle of θ off the x axis. Backprojection assigns an equal weight to all the pixels contributing to a particular projection. This process is repeated for all projections at all angles producing the reconstructed image $\hat{f}(x,y)$ [8].

$$\hat{f}(x,y) = \int\int p(s,\theta)\delta(x\cos\theta + y\sin\theta - s)dsd\theta, \tag{2.11}$$

where δ (Dirac delta function) is nonzero along the line $x\cos\theta + y\sin\theta = s$. This method produces a blurred image with low contrast as can be seen in Fig. 2.7b. In order to reduce the artifacts associated with the backprojection, a method called Filtered Backprojection (FBP) [8] is commonly used. The FBP consists of applying a filter to each projection before performing the inverse radon transform. One commonly used filters is the Ramachandran-Lakshminarayanan (Ram-Lak) filter [57]. This filter suppresses low frequencies and amplifies high frequencies. Some other commonly used filters are Shepp-logan, low-pass cosine and generalized Hamming [15]. Since convolution is computational intensive, in practice FBP is performed in the frequency domain. The frequency responses for the different filters can be seen in Fig. 2.6. The result of the FBP, with the Ram-Lak filter, can be seen in figure Fig. 2.7c.

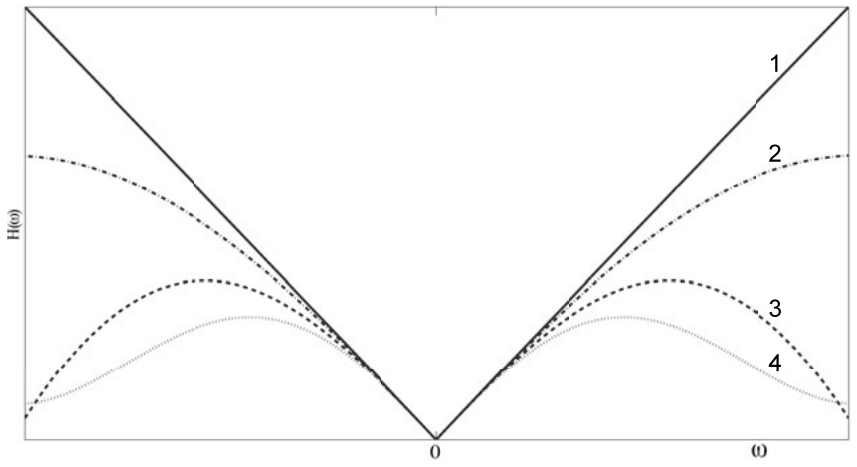


Figure 2.6: Some common filter functions used in Filtered Backprojection. 1=Ram-Lak, 2=Shepp-logan, 3=Low-pass Cosine, and 4=generalized Hamming

Figure 2.7: (a) Original Shepp-Logan phantom (b) Reconstruction with backprojection (c) Reconstruction with Filtered Backprojection

2.4.3.2 Iterative reconstruction

Iterative reconstruction methods are commonly used in tomographic reconstruction today. These methods are more computationally expensive but produce better reconstructions than the FBP in some cases, e.g., if there are a small number of projections, or if the projections are not evenly distributed over 180° or 360°. Furthermore, the iterative methods make it easier to compensate for ray bending from refraction and attenuation along the ray paths [32].

The basic principle of the iterative methods consists of constructing a system of linear equations according to the imaging geometry and physics. The system of equations is often overdetermined. In most cases, it is not possible to exactly solve the equation system. Therefore, iterative algorithms that converges to the correct solution are used. There are numerous methods avail-

able to solve these types of system for tomographic reconstruction, e.g., ART, SIRT, SART, MLEM, and OSEM [4, 23, 29, 30, 36].

One commonly used iterative method for tomographic reconstruction is the Maximum Likelihood Expectation Maximization (MLEM) algorithm [24]. The MLEM algorithm computes

$$x_i^{n+1} = \frac{x_i^n}{\sum_p A_{pi}} \sum_p \left[\frac{A_{pi} y_p}{\sum_{i'} A_{pi'} x_{i'}^n} \right], \tag{2.12}$$

where x_i^n is the estimated ith pixel of the reconstructed image in the nth iteration, and A_{pi} is the element in the system matrix describing the relationship between pixel i in the image and the projection p. The variable y_p is the measured projection pixel at position p.

The system matrix contains the pixel coverage for each projection that reaches the detector. Depending on the system set-up used to acquire the images, the system matrix can be constructed to optimally represent the path for each ray passing through the sample.

A method called Ordered Subset can be used to optimize iterative methods. The projection data is grouped into an ordered sequence of subsets. The iterative reconstruction is then performed on the subsets. Using Ordered Subsets Expectation Maximization (OSEM) provides a significant increase in speed compared to MLEM [30]. Furthermore, the quality of the reconstruction from OSEM is comparable with reconstruction from MLEM. In paper IX, OSEM is used in the tomographic reconstruction of zebrafish.

3. Methods and applications

In this section the methods and applications of the included papers are presented briefly. In addition, some further development of previously published results are presented. The work can be divided into two groups; the first part focuses on image analysis on cells, and the second part on image analysis on zebrafish.

3.1 Digital image cytometry

Digital image cytometry is the field in image analysis that deals with automated measurements and extraction of quantitative data from images of cells.

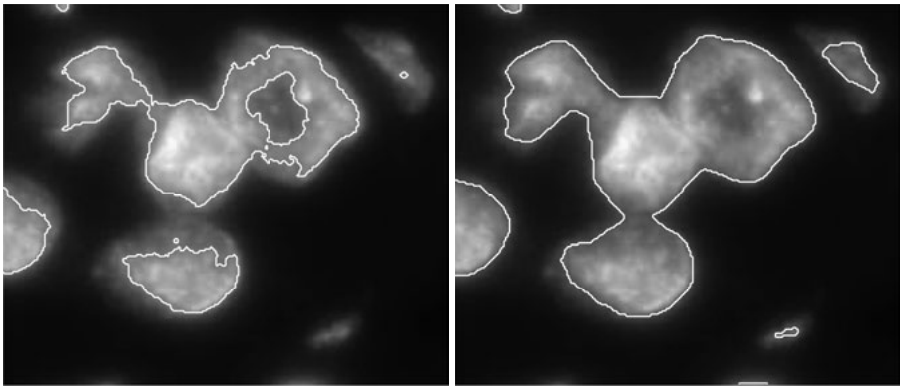
3.1.1 Cell segmentation

Cell cultures, as well as cells in tissue, always show a certain degree of variability, and measurements based on cell averages will miss important information contained in a heterogeneous population. Single cell analysis performs analysis on each individual cell instead of an average over the image. The cell segmentation is therefore a crucial step in single cell analysis. Depending on the sample type and the staining the difficulty in segmentation may vary, e.g., tissue samples generally contain more clustered cells as compared to cultured cells making the segmentation increasingly difficult.

3.1.1.1 Nucleus segmentation

In fluorescence microscopy, DAPI is commonly used to stain cell nuclei, see Fig. 3.1. If the nuclei are similar in intensity level, a threshold is often good enough to separate the cells from the background. Otsu's method of thresholding works well due to the assumption that only background and foreground is present in the image. Sometimes, dark areas inside the nuclei appear as holes after the threshold. These holes can be filled using a flood fill algorithm. The algorithm floods the background from all background image border pixels. All pixels reached by the flood fill will be set as background while all pixels not reached by the flood fill will be set as objects. This method of separating nuclei from background is used in Paper I, II and III.

Sometimes Otsu's thresholding method does not provide a suitable initial separation of foreground and background. This is the case in paper VI where



(a) (b)

Figure 3.1: (a) Contour from segmentation by Otsu's threshold. (b) Contour from segmentation by Chan-Vese level set.

we used the Chan-Vese level set method to detect the objects in the image [13]. The Chan-Vese method and Otsu's method are based on the same idea, but Chan-Vese is trying to solve a local minimization problem rather than, as for Otsu, a global one. To get the level set method to work optimally for nuclei images we have to adjust the parameters of the level set. The weight for the area term, ν, is set to zero since we don't want to have a restriction on the total area of the nuclei. We set some weight on the first term, μ, to achieve smoother borders of the segmentation. There is some variation in the intensities of the nuclei while the background is more uniform. Therefore we use a lower weight on the foreground than the background, $\lambda_1 < \lambda_2$. This allows more variation of the intensities inside the cell than outside. In Fig. 3.1, a comparison of the initial separation of foreground and background with Otsu's and Chan-Vese level set can be seen.

The binary image from the initial separation of the nuclei and background is transformed to a landscape-like image using a distance transform [47]. Seeds, entry points for the water in watershed segmentation, representing the different nuclei, are needed in order to separate clustered nuclei into different objects. The seeds are the local maxima of the distance transform. Due to imperfect circularity of the nuclei, distance transformation may lead to multiple seeding points for the same nucleus. This will result in over-segmentation. The h-maxima transform is able to suppress maxima whose height is smaller than a given threshold [67]. By suppressing all small maxima several adjacent local maxima are merged into one regional maximum, i.e., one seed point for each nucleus is achieved. This method of separating clustered nuclei with watersheds is used in Papers I, II, III and VI.

3.1.1.2 Delineation of cytoplasm

In single cell analysis it is necessary to assign a detected fluorescent probe to a specific cell. Biomolecules are not always contained inside the nucleus, instead located in the cytoplasm. If no cytoplasmic or membrane stain is present, as is often the case, the delineation of the entire cell must be made with some assumptions. One commonly used assumption is that a signal belongs to its closest cell nuclus [11]. In paper I we examine how well this segmentation (*NCS*) performs in comparison to manual segmentation and a segmentation constrained by a cytoplasmic stain (*CS*). The main goal of the segmentation is to, in the end, accurately count the number of signals per cell rather than achieving a perfect delineation of the cytoplasm.

The evaluation of the segmentation methods consists of three parts. The first part is the segmentation evaluation of the manual segmentation and the two automatic segmentation methods (*NCS* and *CS*). In the second part we evaluate the single cell analysis based on the segmentation *NCS* on cells with known distribution of mutated mtDNA (mitochondrial DNA). In the third part, the same single cell analysis method is compared to a biomedical method, Polymer Chain Reaction-Restriction Fragment Length Polymorphism (PCR-RFLP).

To evaluate the three different cytoplasmic segmentation methods a framework for evaluating segmented images presented by Udupa et. al. [70] was used. A total of 56 cells were used in the evaluation. Accuracy (agreement with truth), precision (reproducibility) and efficiency (time) are compared for the different segmentation methods. Accuracy is based on three different quantities: False Negative Area Fraction (FNAF), False Positive Area Fraction (FPAF), and True Positive Area Fraction (TPAF). We define S the result from automated segmentation, and compare it to S_t, the true segmentation. There is no true segmentation available, therefore we use one of the manually segmented results as S_t. FNAF is the fraction of S_t that is not included by S. FPAF is the area that is falsely identified by S as a fraction S_t. In our case, the parts of S that overlap with the image background, as defined in S_t, are not a part of the FPAF because the background does not give rise to any signals and hence will not affect the calculations of signals per cell. TPAF describes the total amount of cytoplasm defined by S that coincides with S_t as a fraction of S_t. The precision is a measure of reproducibility and naturally the automated methods in this comparison will always reproduce the same results. The efficiency measure consists of the time spent by computer or human to achieve the segmentation.

Regarding the accuracy, there are small advantages seen in the manual delineation for the TPAF and FNAF. The biggest difference can be seen in the FPAF, where both automatic methods have significantly higher values. On the other hand, the precision is lower for the manual delineation, due to a high degree of inter and intra observer variability, see Table. 3.1.

Table 3.1: *Comparison of segmentation methods. O_{1a} is the first manual segmentation performed by observer 1. O_{1b} is the second manual segmentation performed by observer 1. O_2 is the manual segmentation performed by observer 2. NCS is segmentation without cytoplasmic stain and CS is the segmentation with cytoplasmic stain.*

Method	Accuracy				Precision	Efficiency
	vs.	TPAF	FNAF	FPAF	(%)	Cells/min
NCS	O_{1a}	0.87±0.03	0.14±0.03	0.12±0.04	100	30
CS	O_{1a}	0.85±0.03	0.16±0.03	0.11±0.03	100	30
O_2	O_{1a}	0.84±0.02	0.16±0.02	0.02±0.01	79	1
O_{1b}	O_{1a}	0.90±0.02	0.10±0.02	0.03±0.01	84	1

The single cell analysis used in this evaluation is based on the cytoplasm segmentation without a cytoplasmic stain *NCS*. Mutation load is the proportion of mutated mtDNA compared to wild type mtDNA per cell. Two different cell cultures where used to evaluate the single cell analysis; a co-culture of cells with either no or all mtDNA mutated and a culture containing cells with approximately equal number of mutated and wild type mtDNA. The results from the single cell analysis on the co-culture shows distinct distributions at the extremes, indicating that we mainly have cells with either no or all mtDNA mutated. For the second cell culture we observe a peak close to 50%, which is expected.

The quantification of mtDNA mutation with the image-based method is compared to quantification performed by PCR-RFLP. In Paper I, we show that the results from the two different quantification methods are comparable. Furthermore, in Paper II the image-based method is compared to another biomedical method, PCR-RFMT, a PCR-RFLP mutation load assay based on melting curve analysis. The result shows good agreement between the image-based method and PCR-RFMT.

The measurements of mutation load from the image-based single cell analysis without a cytoplasmic stain shows good agreement with the *a priori* known mutation loads for the two different cultures. In addition it shows good agreement with the measurements performed by PCR-RFLP and PCR-RFMT. The presented automatic image-based single cell quantification provides a good segmentation method that agrees with the predictions.

The described method for single cell analysis is used in Papers I, II, III and VI.

3.1.1.3 3D cell segmentation

Cell segmentation in 2D is often enough, but some applications, such as exact localization of signals, require a segmentation of the cell in 3D. In Paper IV a semi automatic method is used to delineate the cell nuclei in 3D. The method

was applied to biological data studying the localization of Smad2-Smad4 protein complexes in relation to the nuclear membrane over time.

The algorithm starts by allowing the user to define background and object seed points through a graphical user interface. The seeds are marked on a maximum intensity projection of all z-slices in the channel containing the nuclei. No extensive requirements on precision are needed in the marking process. The volume image is smoothed by using anisotropic diffusion [55], that smooths while preserving the important edge information. The edge information is crucial since it will be used as a merging and splitting criterion. A 3D Sobel filter that approximates the gradient is then applied on the smoothed image. The Sobel filter is applied in all three dimensions and combined to an edge enhanced image. A seeded watershed algorithm is applied to the edge image. The seeds will grow and eventually meet at the borders, which are intensity maxima of the edge image. The texture inside the nuclei often produce maxima in the edge image resulting in more than one region for each object, i.e., oversegmentation. A merging criterion that preserves the strongest edges is applied to the oversegmented image [73].

The signal detection is performed with the method 3DSWD, described in section 3.1.2.2. The subnuclear positions of the signals are of interest and therefore subnuclear regions must be defined. A distance transform from the borders of the nuclei is applied. The distance at the border will be zero, and the inside if the nucleus will have negative distance values while the outside will have positive distance values. From the distance transform there will be shells of different distance values in and around the nucleus. The number of signals detected in each shell is measured. Since the volumes of all the shells will differ, a normalization is done with the shell volume, resulting in a concentration measurement. For a given time point, a measure of signal concentration in each shell for each individual nucleus is obtained.

Paper IV provides a method for detecting and localizing of Smad-complexes inside the cell over time. The results show that the Smad complexes are formed at a very early time point after stimulation. However, more data is needed to draw conclusions regarding the their spatial location inside the cell over time.

3.1.2 Point-like signal detection

Fluorescent markers are often used to identify subcellular structures such as protein complexes, chromosomes, genes and mutations in genes. There are many different methods utilized to detect these subcellular structures, but the images produced share many similarities. The fluorescent biomarkers make detection events visible as point-like signals (PLS) in the captured image data. Quantification of these PLSs is an essential part of analyzing these types of images. There are many algorithms that are developed for detecting PLSs and

finding the optimal algorithm will depend on the application and the type of images acquired [65].

3.1.2.1 Counting vs measuring intensity

In paper III, we design and perform a test to evaluate ways of quantifying the amount of fluorescent markers present in an image. The first approach consists of measuring the number of detected signals while the other approach consists of measuring the total intensity of the detected signals.

The enhancement of the signals was performed with a Laplace filter. The enhanced image was thresholded manually and each object after thresholding was counted as one signal. In the intensity measure, instead of counting the number of detected signals, the sum of all intensities in a neighborhood around the center of a detected signal is used as the measurement. A third method that could have been tested was to sum all intensities in the channel that contains the signals. This option would work in an ideal case, but in practice the presence of autofluorescence makes this measurement very unreliable.

We constructed artificial data with an increasing concentration of signals to evaluate the signal quantification using the two different methods. PLS were created by randomly distributing values in an image. To simulate a point spread function the signals were filtered with a Gaussian filter. In addition, Gaussian noise was added to the image.

The amount of signal was quantified with both measurements. In the low concentration region both methods give similar and reasonably accurate results. However, when reaching a high signal concentration, the intensity measurement gives better results. The results are in line with the logical assumption that, at high concentration, signals are difficult to separate due to clustering. For clustered signals, the intensity measurement will compensate for the missed detections by quantifying the intensity in a neighborhood around the detected signals. Furthermore, the intensity measurement still depends on the detection of a signal, and therefore at extremely high concentrations the measurement deviates from the true number of signals. This is an ideal case where none of the signals are saturated. If a lower bit depth was used many clustered signals would result in saturation, reducing the accuracy of the result from the intensity measure.

In conclusion, when there are many clustered signals present, using an intensity measurement provides more accurate results. On the other hand, when working with well separated signals, it is more suitable to count the number of signals. The intensity of the signals will differ in a sample and these differences will affect the intensity measure more than only counting the signals. Furthermore, it is easier to interpret and evaluate data that gives information on the number of detected signals rather than an intensity measure.

3.1.2.2 3DSWD

In many applications the accuracy of PLS detection is of high importance. The use of a Laplace filter, as in the previous section, is sometimes not good enough to get the desired detection accuracy. In Paper V we present a novel method for robust 3D signal detection in fluorescence images, called 3DSWD. The method is compared to three commonly used methods; TopHat, DoG and MP, all described in Sec. 2.4.2.4.

3DSWD

The idea of the 3DSWD is based on a method called Stable Wave Detector (SWD) that is used for landmark detection in 2D images [19]. The main idea of the 3DSWD is the combination of two steps in the detection of a signal: a detector and verifier/separator. The aim of the detector is to enhance regions with point-like source signals. The verifier/separator aims to verify if a high value from the detector is a true signal by examining the slope around the point. The verifier/separator is also used to separate signals that are close in space.

We performed several tests to evaluate the performance of the 3DSWD. Two tests were performed on artificial data: robustness to noise, and resolving power and sensitivity to signal intensity. The tests on robustness to noise showed that the 3DSWD had the highest number of True Positive (TP) and the lowest total amount of False Positive (FP) signals among all methods. From these tests the 3DSWD shows better robustness to noise than the other methods used in the comparison. We also showed that the 3DSWD is better at resolving signals that are in close proximity, and less sensitive to signal intensity differences, compared to the other methods.

Tests on artificial images provide a good tool to evaluate the differences in performance between the methods. However, using only artificial data is not enough to prove that they work in a real application; the methods need to be evaluated on real images. A simple GUI was constructed to allow experts from the application field to mark true signals in real fluorescence data. These manual signal detections were compared with the results from TopHat, MP, DoG and 3DSWD. *Precision*, *recall*, and *F-score* were used to quantify the performance of the different methods. *Precision* is the fraction of the detected signals that are actually true signals, while *recall* is the fraction of the true signals that are detected. The *F-score* is a weighted average of the two measures *precision* and *recall*. The *precision* for MP, TopHat and DoG are all higher than for the 3DSWD. However, when comparing the *recall* the 3DSWD outperforms the other methods. For the *F-score*, which is considered an overall measure of the method performance, the 3DSWD has the highest value among the tested methods. In addition, the *F-score* of the inter and intra-observer variability is in the same range as the 3DSWD.

Evaluation of verifier/separator step to other methods

Each of the conventional methods used in the evaluation could be used together with a verifier/separator for performance improvement. We have tested these methods on artificial data with and without the addition of a verifier/separator. We created 100 artificial images with increasing Gaussian noise, σ ranged from 0.04 to 1.24 and mean of 0. Similar to as we did in paper V, 3D signals were created with an anisotropic Gaussian. All parameters were optimized individually for each method. The strength of the verifier/separator is that it makes it possible to decrease the threshold to detect weaker signals without picking up additional false signals. If we use the same threshold for the methods with and without the verifier/separator no increase in TP will be seen so we decrease the threshold with 5% when using the verifier/separator. The results for all the methods can be seen in Fig 3.2.

If we look at the performance of all the methods without a verifier/separator we can see that the 3DSWD has the highest robustness to noise. The DoG has a slightly poorer performance than the 3DSWD. TopHat seems to be the least robust method to noise in these tests. The ranking order of performance of the methods is the same as in Paper V, where Poisson noise was used.

There is a performance enhancement for all methods when the verifier/separator step is added. The DoG has a small decrease in the FP and in addition an increase in the TP. The enhancement of the TopHat is similar to that of the DoG when the verifier/separator is added, with the difference that the FP has decreased more for the TopHat. The enhancement for the MP differs slightly from the other two methods. There is no significant increase in the number of TP, but there is a clear decrease of for the FP. With the addition of the verifier/separator to the DoG there is no significant difference between the signal detection performed by DoG and the 3DSWD. This is no surprise since the DoG is very similar to the detector used in 3DSWD. For these tests the combination of DoG-Verifier/separator and 3DSWD performed best. However, there is no method that will work perfectly on all types of images. In Paper VII we use a combination of MP and verifier/separator to detect neuron cells. We concluded from visual evaluation of result from different methods that this combination produced optimal signal detection performance. These neuron cells had a more flat peak than regular PLS and they varied in size, and that could be the reason why MP in combination with verifier/separator worked best in this case.

To conclude, the presented method, 3DSWD, has shown better performance than the conventional methods that were used for comparison. The concept of using both a detector and verifier/separator makes the signal detection more stable. The detector enhances regions with point-like source signals while the verifier/separator, by examining the direction of the slope around each detected point, verifies if a high value from the detector in fact is a true signal. We also show that the addition of a verifier/separator to other methods can improve their performance. Furthermore, the verifier/separator used in these

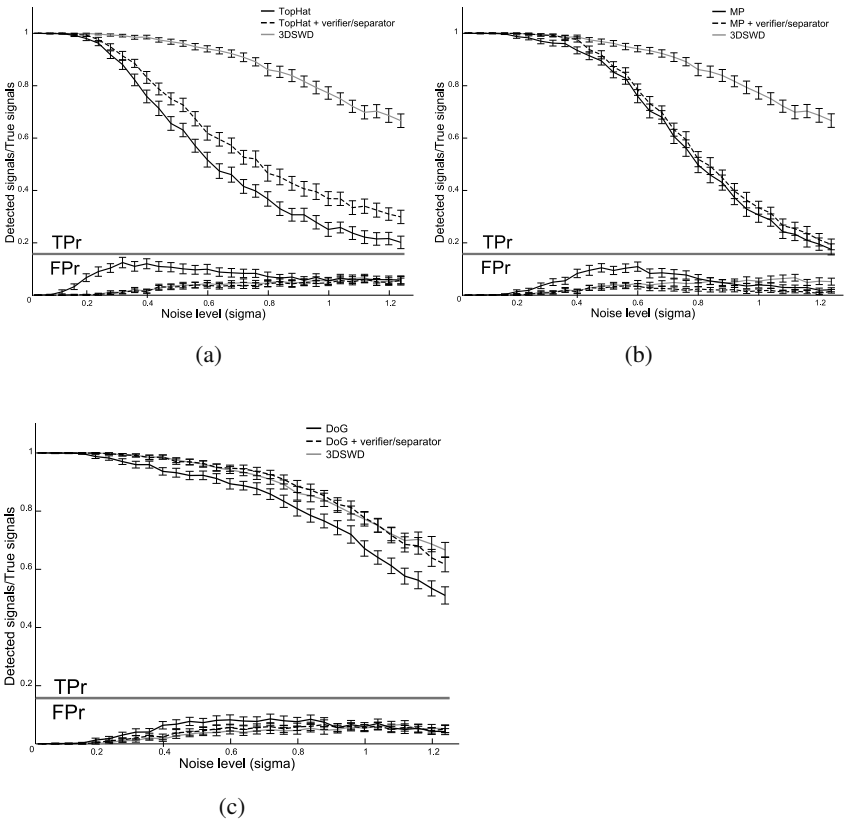


Figure 3.2: Comparison of signa detection using TopHat (a), MP (b) and DoG (c) with and without a verifier/separator. In each graph the result from the 3DSWD is also shown. The upper part of each diagram (TPr) represent the ratio of correct detected signals and the lower part of the each diagram (FPr) represent the ratio of false detected signals. Error bars represent a 99% confidence interval.

evaluations consist of a *sine* function, but could be substituted with any type of derivative filter. It should also be added that, if the signal concentration is very high and there are clustered and saturated signals the performance of the verifier/separator will be reduced.

3.1.2.3 Increasing the dynamic range

Sometimes the concentration of the biomolecular events detected by point-like signal is extremely high and the signals are saturated. Saturation means lost information in an image and accurate signal measurement is impossible, limiting the dynamic range of event detection. Changes can be made in the concentration of the detection probe, reporting on the analyte, to detect only a fraction of the occurrence. However, this will not work in heterogeneous

samples such as tissue sections, where the amount of a specific protein may vary greatly between neighboring cells.

In Paper VI we developed a method to increase the dynamic range of the PLA-based protein detection in order to allow detection of both abundant and scarce target molecules in the same sample. Reagents that give rise to three variants of the reporter DNA circles are added in a concentration ratio of 1:10:100. For any target molecule the probability of giving rise to each variant of Rolling Circle Product (RCP) is thus 1:10:100/111. The RCPs are detected with different fluorescent labels and the distinct signal concentrations can then be visualized in different colors. If the detection reaction is saturated in one color, then less abundant RCP detectable with another fluorophore can be used to quantify signals, thereby extending the dynamic range of an assay. The approach is not limited to *in situ* PLA, but can be applied to other RCA-based methods such as Immuno-RCA and padlock probes [26, 38, 62]. To evaluate this approach for increasing the dynamic range, we used the 3DSWD, described in Paper V. The cell and cytoplasm segmentation was done as previously described in Sec. 3.1.1. We performed the analysis on a cell line, on which we had produced a low to extremely high signal concentration for demonstrational purposes through a varying antibody concentration. Three oligonucleotides (A, B and C in Fig. 3.3) of the concentration ratio 100:10:1 were used together with three different fluorescent labels. The signal quantification in the different color channels can be seen in Fig. 3.3.

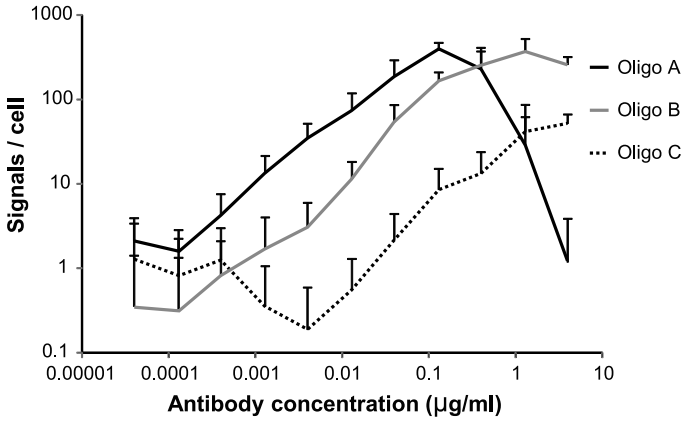


Figure 3.3: The number of signals detected for each oligonucleotide at different antibody concentrations (13-20 cells per data point; error bars show standard deviation).

We also analyzed a heterogeneous tissue sample, wherein the amount of target protein varied greatly. By using the proposed method of extending the dynamic range we could measure locally abundant and locally scarce target protein at the same time. Without the dynamic range extension, measuring

the abundant target molecule could only be done at the expense of loss of the scarce signals, and vice versa (see Fig. 3.4).

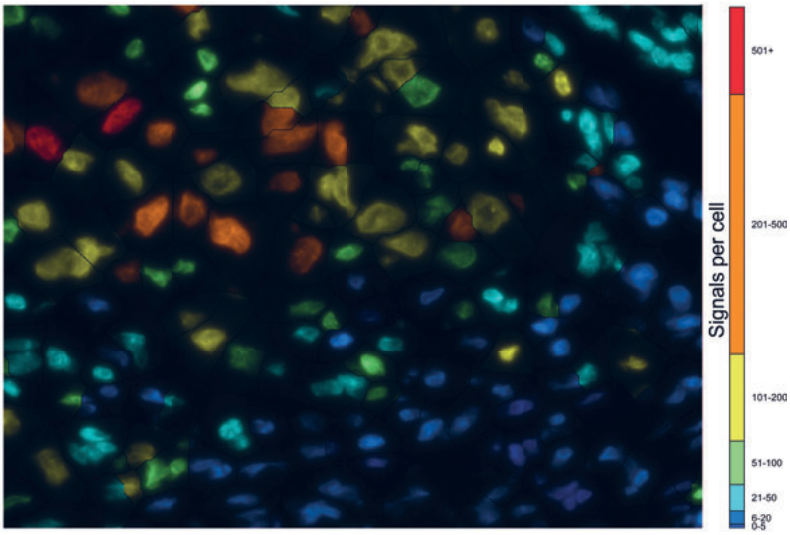


Figure 3.4: Result image from increased dynamic range showing the number of signals for each cell. The colors represent the amount of signals detected in each cell. An increased dynamic range enables the possibility of measuring both locally abundant (red nuclei) and locally scarce (blue nuclei) signals in the same sample.

3.1.3 Software

Development of new algorithms is the core of research in image analysis. The community that in the end will benefit from use of the novel methods, in this case the biomedical community, is not always able to use novel methods easily. The new algorithms developed are often just available as source code for some specific language, making it difficult for an external user to try them. Developing stand-alone applications that can be used by external users in the community provides a great opportunity to get several evaluators of the newly developed methods. The evaluation will be on real data, in large scale, and done by several independent users, providing a great feedback of the method's performance. As part of this thesis work, we have made some of the presented methods available in stand-alone software or made them available through existing software.

3.1.3.1 Visiopharm

For Paper I and II a plug-in was developed for the commercial software package VIS Image Analysis Software (Visiopharm, Denmark). The plug-in pro-

vided a user-friendly environment to perform single-cell analysis of fluorescence images. The cell segmentation for the plug-in is described in 3.1.1. The signal detection is performed with a Laplace filter and a threshold. The system can analyse one or a batch of images and export results as a spread sheet. The plug-in was implemented in C++ and incorporated in the VIS Image Analysis Software in order to utilize all the good features that already exist in the software. Incorporating the methods into the VIS software provides a user friendly application for performing single cell analysis. In addition, it provided us, the algorithm developers, a good feedback on the method's performance in "real life" and showed that there is a demand for such analysis.

3.1.3.2 BlobFinder

Based on request from our collaborators, we developed BlobFinder, a free software optimized for the analysis of images generated by the *in situ* PLA, described in section 2.3.1. The software can be used for other types of fluorescence microscopy images with point-source signals, such as data from, e.g., Fluorescent in Situ Hybridization (FISH). There exists a number of different image analysis software packages that can be programmed to perform a wide array of useful measurements, similar to the analysis in BlobFinder [1, 11]. However, the multi-application capability often compromises the simplicity of the tool. Therefore BlobFinder was developed for a specific task, making it easy to use, easy to learn and optimized for its task.

The software is implemented in MATLAB with a user-friendly GUI (graphical user interface) [46]. The Matlab code is compiled into a standalone application that can be run without MATLAB installed on the user's computer.

The algorithms in BlobFinder are similar to the algorithms implemented in VIS Image Analysis Software, see section 3.1.3.1. One difference from the plug-in to VIS, is that BlobFinder can import and analyze 3D images. The analysis is performed on a pre-filtered maximum projection of the 3D data. Even though the analysis is performed on a maximum projection, signals from different depths in the sample can be detected. All details on the software are described in Paper III.

3.1.3.3 BlobFinder Bright Field

A bright field version of BlobFinder was developed based on the paper by Zieba et al [77]. The RGB image is transformed to HSI colorspace [21]. The nuclei are segmented based on two thresholds: one threshold in the hue channel, defining everything that is blue (nuclei), and a threshold in the intensity channel to separate the nuclei from the background.

The signals are detected in the same way as in the original BlobFinder, but with an addition of a threshold in the hue channel. Blue spots inside the nucleus are sometimes detected as signals, but by using a color threshold we can discriminate the true red signals from the false blue.

BlobFinder Bright Field shares the same features as the original version of BlobFinder, such as single cell analysis, average analysis, signal count, signal intensity measure and batch process of large data sets.

3.1.3.4 Duolink ImageTool

BlobFinder has been widely used and the demand for more features of the software was high. The implementation in MATLAB was limited and therefore a software, purely implemented in C++, was developed. The commercial software, Duolink ImageTool, was developed together with Olink Bioscience (Uppsala Science Park, Sweden). Duolink ImageTool works on both bright field and fluorescence images. A large amount of time was spent on creating an intuitive GUI and enabling the user to get started with the software right away [2], see Fig. 3.5. All code for the image analysis was rewritten in C++ for optimal performance. Furthermore, the software contains direct file import from some of the most commonly used microscope manufacture file formats.

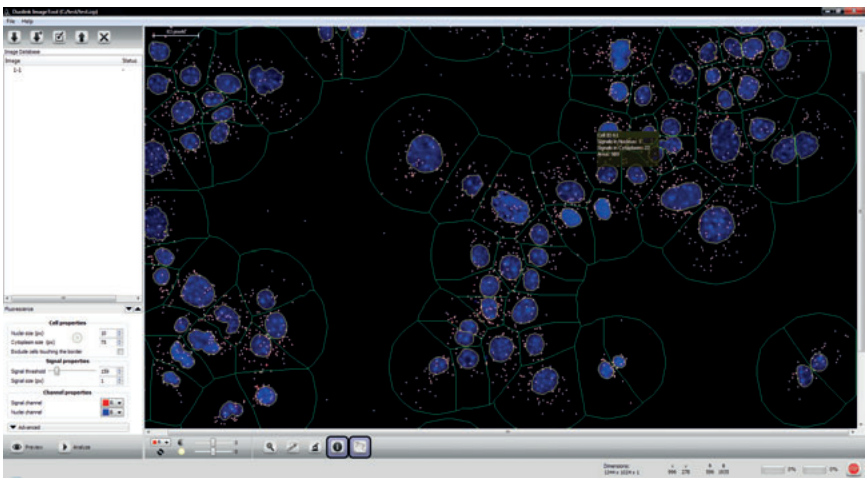


Figure 3.5: A screenshot of the GUI in Duolink ImageTool.

3.2 Zebrafish image analysis

High-Throughput screening (HTS) is used to find new treatments for disease, or better understand disease pathways through searching large libraries of chemical or genetic perturbants. Small size, optical transparency of complex organs and ease of culture makes the zebrafish larva an ideal organism for large-scale *in vivo* genetic and chemical screens not feasible in higher vertebrates such as mice. Zebrafish models of several human diseases have been developed [42, 48, 59, 64, 78] and lead compounds discovered by screening chemical libraries for efficacy in zebrafish disease models. These have been useful for pharmaceutical discovery owing to the high level of conservation of drug activity between humans and zebrafish [42, 48, 59, 64, 78].

However, visualizing most zebrafish organs requires manipulating and properly orienting the larvae. Even with confocal or two-photon microscopy, optical access is often impeded by pigmentation, by intervening organs such as eyes and heart, or by the highly autofluorescent yolk sac. Current methods to address these challenges involve treatment with the toxic chemical phenylthiourea to suppress pigmentation, and transferring zebrafish from multiwell plates or reservoirs and orienting them in viscous media such as agar using forceps.

To improve the throughput and complexity of zebrafish screens, the Yanik lab at the Research Laboratory of Electronics, MIT, developed the Vertebrate Automated Screening Technology (VAST) that automatically manipulates and images zebrafish at cellular resolution in three dimensions [54] (see Fig. 3.6).

Here we describe our contribution to the system, developed in close collaboration with the members of Yanik lab. One of the key features of the system is positioning of the fish before imaging. In Papers VII and VIII we present the automation of lateral and rotational positioning of zebrafish larvae using a 10x lens and bright field microscopy. The primary goal of accurate positioning is to subsequently acquire single cell resolution images from the brain of the animal by confocal microscopy, described in Paper VII.

In Paper IX we describe a further development of the system. Inspired by the rotation of the animal during positioning we developed a new system for high-throughput zebrafish tomography, both in bright field and fluorescence.

3.2.1 Vertebrate Automated Screening Technology (VAST)

The VAST system loads larvae from reservoirs or multiwell plates and drives them through the system using a syringe pump. A photodetector and a high-speed CCD are used to detect when a larva is entering the tube and the field of view of the microscopes. The tube is then rotated 360 degrees with the help of motors while a high-speed CCD camera acquires bright field images. An overview of the system is shown in Fig. 3.6. The rotational images are used to correctly position the fish and find the region of interest (ROI) that later will

be used for fluorescence acquisition. A dorsal and sagittal view of the fish can be seen in Fig. 3.7.

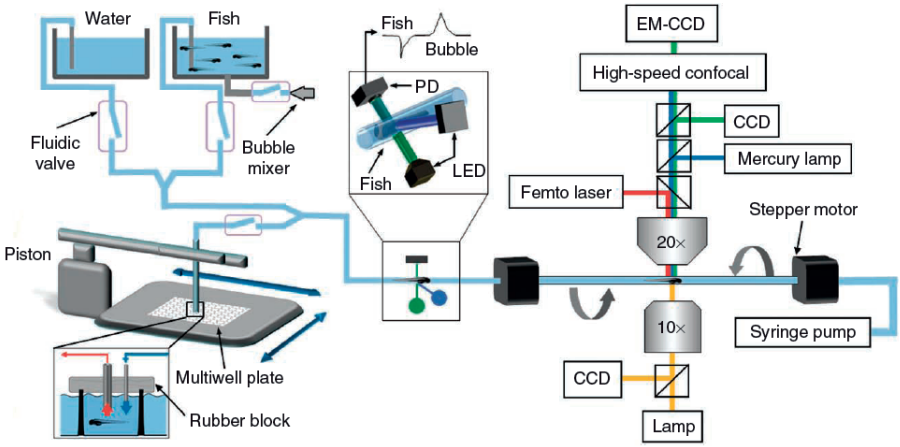


Figure 3.6: Figure from C. Pardo-Martin, Nature Methods 2009 [54], with permission from Nature Publishing Group (NPG). Schematic of the zebrafish manipulation and imaging platform Vertebrate Automated Screening Technology (VAST). The larvae are present in multiwell plates or reservoirs. They are loaded through a microfluidic system and a syringe pump at the other end. A photodetection system is used to distinguish the passage of a larva from air bubbles and debris. The larvae enters the field of view of the microscopes in a capillary that is hold in place by two stepper motors. The stepper motors enables rotation of the capillary and the larvae. There are two microscopes in the system, one upright confocal and an inverted bright field microscope. A femtosecond laser beam used for microsurgery is focused on the sample through the upper objective.

3.2.1.1 Zebrafish positioning in VAST

The larvae can enter the capillary either head- or tail-first, and this direction needs to be determined. By summing the inverse intensity values along axis orthogonal to the main tube axis we acquire a 1D profile along the fish $f_{dv}(x)$. There is a large intensity difference between the tip of the head and the empty tube in front of the fish. By taking the derivative of $f_{dv}(x)$ we obtain high absolute values in this region. The center of mass (CM) of $f_{dv}(x)$ is approximately at the center of the fish and by looking at the tip of the head and the CM it is possible to get the direction of the fish using

$$CM = \frac{\sum_{x=1}^{n} f_{dv}(x)x}{\sum_{x=1}^{n} f_{dv}(x)}. \tag{3.1}$$

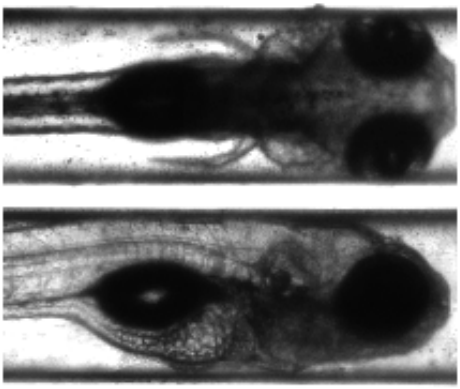


Figure 3.7: Top: A dorsal view of the fish inside the tube. Bottom: A sagitall view of the fish inside the tube.

$$D = \begin{cases} 1, & CM - x_0 > 0 \\ 0, & CM - x_0 = 0 \\ -1, & CM - x_0 < 0 \end{cases}, \tag{3.2}$$

where x_0 is a point where $|f'_{dv}(x)|$ attains a maximum. $D = 1$ means head to the left and $D = -1$ means head to the right.

To find the correct rotational orientation of the fish we calculate normalized cross-correlation [40] with two different templates of the larva, dorsal and sagittal, see Fig. 3.7. Normalized cross-correlation of image $f(x,y)$ with template t is defined as

$$Corr(u,v) = \frac{\sum_{x,y}[f(x,y) - \bar{f}_{u,v}][t(x-u,y-v) - \bar{t}\,]}{\left(\sum_{x,y}[f(x,y) - \bar{f}_{u,v}]^2 \sum_{x,y}[t(x-u,y-v) - \bar{t}\,]^2\right)^{\frac{1}{2}}}, \tag{3.3}$$

where $\bar{f}_{u,v}$ is the mean of $f(x,y)$ in the region under the template and $\bar{t}$ is the mean of the template.

Since all larvae, at approximately the same age, share similar gross morphology, we only need to construct the template once. A cross-correlation is calculated with both templates on all rotational images. Two vectors $Corr_D$ and $Corr_S$ containing the maximum correlation of the two templates, dorsal and sagittal, respectively, at different rotations are produced. The dorsal and ventral view share many similarities, hence we see two distinct peaks in $Corr_D$. In contrast, in $Corr_S$, only one clear peak can be seen. With the three

peaks and the known rotation direction, the rotational position of the fish can be found. The motors can then rotate the tube and larvae to the desired view.

The ROI is experiment specific and before the experiment the user draws the desired ROI on one of the templates. The information from the cross-correlation can be used to find the optimal position of the template on the desired view and from there the region of interest can be overlaid. Once the ROI is identified in the bright field view, the coordinates of this view are transformed to the confocal microscope and single cell resolution images can be acquired of the desired region. This minimizes the ROI to be captured by confocal microscopy, saving time, optimizing image quality, and minimizing strain for live animals.

The methods for zebrafish alignment and positioning are described in both Papers VII and VIII. Apart from the overlap in method description, Paper VII also present the result from a small-scale sample screen on the effect of drugs on neuron development as quantified by automated cell segmentation. Paper VIII provides a more thorough description of the system.

3.2.2 Zebrafish tomography

Conventional 2D microscopy is not always sufficient to capture all the complex 3D morphological details of an animal. Confocal microscopy and two-photon microscopy are routinely used to visualize larvae's organs in 3D. However, these techniques have a limited penetration depth and cannot be used to image the entire animal. In addition, they only work in combination with fluorescent probes. Optical projection tomography (OPT), described in section 2.2.5, has the penetration depth necessary to image an entire zebrafish and can be used both with bright field and fluorescently labeled fish [63]. Current optical projection tomography systems, however, have very low throughput and require the immersion of the sample in solid gels. This makes these techniques unsuitable for large-scale screens, and limits the possibilities with *in-vivo* samples. In Paper IX we present a new high-throughput optical projection tomography system cable of fast acquisition and reconstruction that can be used in large-scale *in vivo* or *in vitro* screens.

3.2.2.1 Tomography system setup

The tomography system is similar to the VAST system but optimized for tomography. The fish is loaded into the tube with a pump and detected as it enters the field of view. The fish is rotated, and bright field and fluorescent images are acquired throughout the rotation. The smoothness in the rotation is more important than in the VAST system, therefore a square tube surrounding the capillary is used to limit the movement of the tube. A low Numerical Aperture (NA) lens is used to get a large depth of field.

3.2.2.2 Alignment

The images acquired from multiple angles of the larva can be used to reconstruct a 3D image of the larva. To obtain a good reconstruction, the alignment of the rotational images is crucial. There are two major misalignments that need to be corrected; vertical movement of the tube and horizontal movement of the fish inside the tube. In addition, the tube must be in complete horizontal position and the center of rotation must be detected, see Fig 3.8.

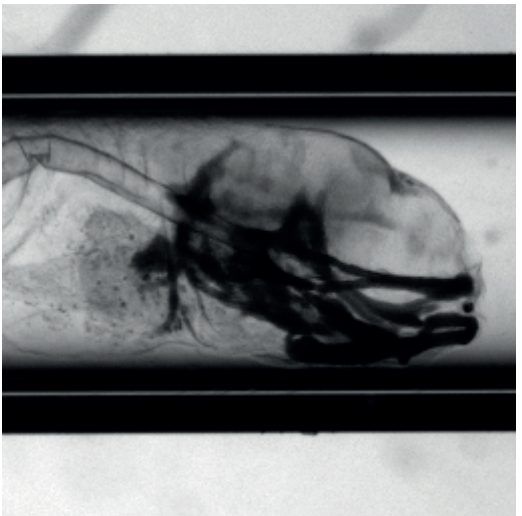


Figure 3.8: Red image channel of head of fish stained with Alcian blue, and acquired by bright field microscopy.

The system is intended to be fully automatic and the alignment therefore needs to be fully automated. To speed up the rest of the alignment the image is cropped to ensure that only the fish and the tube are in the image. The tube wall appears dark in the images, due to total internal reflection, and can therefore be easily detected with Otsu's thresholding method. A minimum intensity projection is performed on all rotational images to ensure that the cropped image will contain the whole tube in a all rotational images.

An edge detection method is applied on one of the rotational images followed by the Hough transform [18]. The angles of the detected lines are used to rotate the tube in a completely horizontal direction.

After rotating all images we need to correct for the vertical movement of the tube during the rotation. The center part of the tube wall contains a bright white line. We detect these lines in all rotational images with a filter that enhances thin lines. We perform a subpixel registration for the upper and the lower line of the tubewall separately with the initial rotational image [25]. During the rotation the tube is moving slightly up and down, towards and away from the microscope, this movement results in small size differences in

(a)

(b)

Figure 3.9: (a) Sinogram before (a) and after (b) vertical alignment and scaling of the tube. Note that both upper and lower thin white lines are straight after alignment and scaling of the tube.

the acquired images. By registering both sides of the tube separately we can get an estimate of the deviation in size during the rotation. The rotational images are then scaled to obtain the same diameter of the tube for all images. In addition, the lines are aligned accordingly to get a full vertical alignment, Fig. 3.9.

The tube is now vertically aligned and rotated in an exact horizontal position. The fish can enter the tube head first or tail first and the direction of the fish needs to be determined. The fish direction is detected in the same way as described in Section 3.2.1.1.

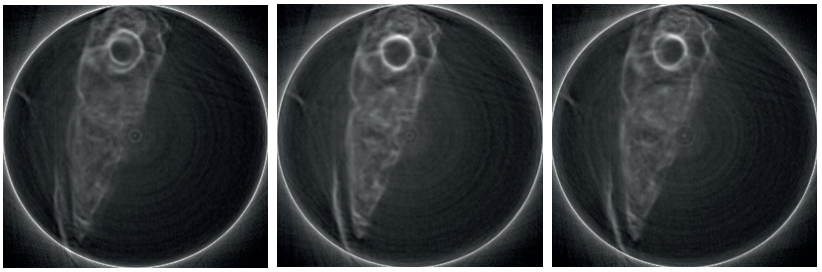
After detecting the direction of the fish, we need to compensate for the horizontal movement of the fish inside the tube. The fish is segmented in each rotational image. If the image is in grayscale, the segmentation is done with an intensity threshold and the largest object is considered to be the fish. If the image is a color image, a color classification is used to classify the image pixels as belonging to the fish or background. Training data for the classification is collected from one fish prior to the analysis. The user marks two training regions, inside the fish (object) and outside the fish (background). The largest object from the classification is considered to be the fish. The tip of the fish is adjusted in all rotational images to ensure that the fish is always in the same position.

Center of Rotation Correction

After the tube and fish is horizontally and vertically aligned, the Center of Rotation (COR) must be determined. This is a two-step process consisting of first

detecting the symmetry axis by cross correlation and second fine adjustment of the COR by performing several test reconstructions.

If all rotational images are summed together, the center of rotation should be the same as the horizontal symmetry axis. The symmetry axis is detected by looping over several rows in the image and cross correlating a region above the row with a mirrored region below that row. The row containing the maximum correlation is considered to be the axis of symmetry and the COR. This usually gives a good estimate of the COR but in some cases additional fine tuning is needed to get a better estimate. This is done by performing reconstructions with Filtered Backprojection (FBP) for a few slices of the fish with different assumptions for the positions of the rotational axis. We calculate the variance for each of the assumed positions of the COR. The position that has the highest variance is considered to be closest to the true COR [75], see Fig. 3.10.



(a) (b) (c)

Figure 3.10: Reconstructions made with different position of the Center of Rotation (COR). (b) has the clearest edges and can be considered as the best approximation of the true COR.

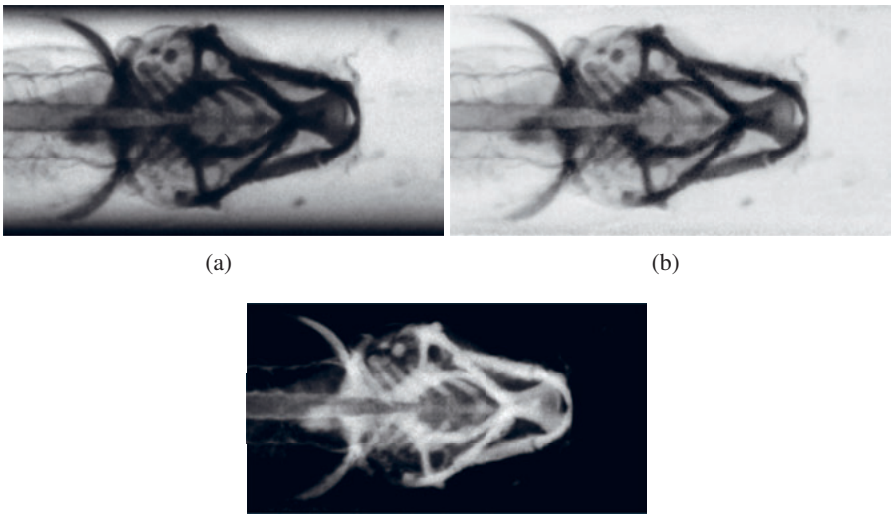
Illumination correction

Refraction between the tube wall and the surrounding fluid creates uneven illumination inside the tube. The part of the tube not occluded by the fish clearly shows how the illumination is not uniform, Fig. 3.11a. From previous alignment the tip of the fish has been detected and we therefore also know where the empty tube is located. By averaging this empty part of the tube, for all rotational images, we can get a cross-sectional profile of how much light passes through the tube at each position from the center of the tube. The aligned image is divide by the profile, resulting in an illumination corrected image. Furthermore, we take the logarithm of the data, Beer's law (3.4), to get a linear relationship between the intensities in the images, see Fig. 3.11b

$$A = \log_{10}(I/I_0), \tag{3.4}$$

where I_0 is the intensity of the light before it enters the sample, I is the light intensity after absorption in the sample and A is the absorbance. Before the

reconstruction the image is inverted and the background is set to zero, see Fig. 3.11c.



(a) (b)

(c)

Figure 3.11: (a) Aligned image without illumination correction. (b) Aligned image after illumination correction. (c) Aligned image after inversion and removal of background.

3.2.2.3 Light ray simulation

We know that there are some distortions in the light due to differences in refractive index between the tube wall and the surrounding fluid. By simulating the path of each of the rays through the tube it is possible to compensate for the distortions from the refraction.

In the simulation we trace each ray coming in to the detector in the microscope. We trace the rays from the light source to the detector and calculate the refractions at all boundaries of the tube, see Fig. 3.12. Each bin in the detector will collect light from an area inside the tube. The area for each bin can be simplified with a polygon containing four vertices.

The system matrix in the reconstruction, Sec. 2.4.3.2, describes the path that the light travels along, through the sample, to reach the detector bins. The system matrix is a large sparse matrix with $\theta N \times N^2$ elements, where θ is the number of acquisition angles and N is the number of detector bins. The polygon area that we calculated for each detector bin must be converted into a system matrix that can be used in the reconstruction. For each polygon we need to find the pixels that are contained inside the polygon. A pixel can fall between two regions and hence we need to use sub-pixel precision.

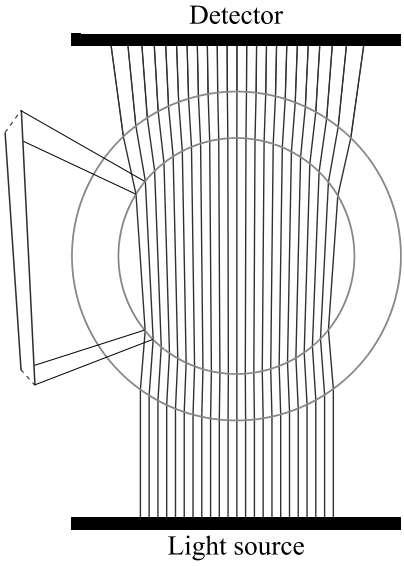


Figure 3.12: Ligh ray path through the tube. Each polygon consists of four lines.

The procedure for finding the sub-pixel coverage for a polygon described here is applied for all polygons and for all rotational angles used in the acquisition. First, we find the bounding box of the polygon, only pixels inside the bounding box are further used in determining the area coverage of the polygon. We then identify all pixels that are inside the polygon. Point p is inside a polygon if a ray starting from point p, in any direction, has an odd number of crossings with the polygon's edges [27]. This will determine if the center point of a pixel is inside a polygon. However, for pixels lying on the edge of the polygon we need to find the fraction of the pixel contained inside the polygon. Before we calculate the fractional pixel coverage we need to identify the pixels that are intersecting with the lines from the polygon. For each pixel we calculate the shortest distance to each of the line segments. A pixel p is intersected by a line segment l if the distance to the line is less than a threshold T_d (see Fig 3.13a),

$$T_d = \frac{1}{\sqrt{2}} \sin\left(\frac{\pi}{4} + \alpha\right), \tag{3.5}$$

where

$$\alpha = \begin{cases} \alpha_l & \text{if} \quad 0 \leq \alpha_l \leq \pi/4 \\ \pi/2 - \alpha_l & \text{if} \quad \alpha_l > \pi/4 \end{cases}, \tag{3.6}$$

and α_l is the angle of the line l, $0 \leq \alpha_l \leq \pi/2$.

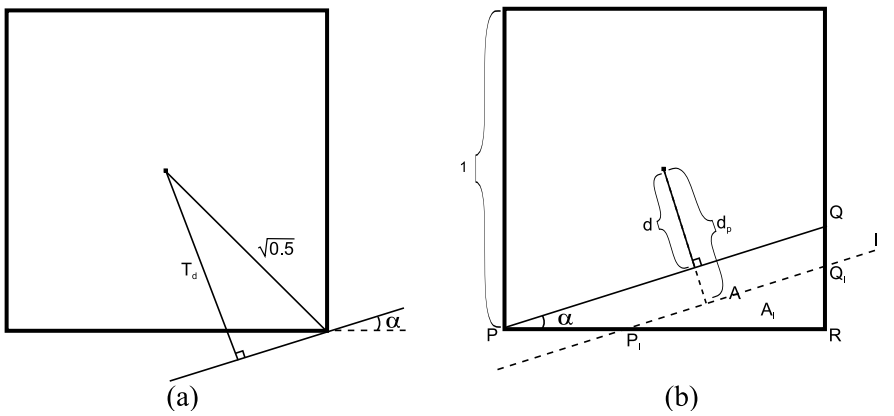


Figure 3.13: (a) If a line of angle α is intersecting the corner of a pixel, the shortest distance from the line to the center point of the pixel will be T_d. If an arbitrary line with the same angle α has a distance to the center of a pixel that is shorter than or equal to T_d the line is crossing that pixel.(b) Line l is crossing the pixel and divides it in two regions. To determine the area of the region A_l the area A for a triangle with known sides is used.

If the pixel is intersecting with the line l we need to find the fractional pixel coverage A_l (see $\triangle P_l Q_l R$ in Fig. 3.13b). The shortest distance between the center of pixel p and line l is defined as d_p. The length of $\overline{PR}$ and $\overline{QR}$ is known, and hence we can calculate the area A of $\triangle PQR$,

$$A = \frac{\tan(\alpha)}{2}. \tag{3.7}$$

The shortest distance between the center point of the pixel and the line segment $\overline{PQ}$ is defined as,

$$d = \frac{1}{\sqrt{2}} \sin\left(\frac{\pi}{4} - \alpha\right). \tag{3.8}$$

To get the area A_l we use the relationship $\triangle P_l Q_l R \sim \triangle PQR$. There are two cases we need to observe when calculating A_l:

(i) If $d_p > d$ then

$$A_l = \left(\frac{\tan(\alpha) - \overline{QQ_l}}{\tan(\alpha)}\right)^2 \cdot A, \tag{3.9}$$

where

$$\overline{QQ_l} = \frac{|d_p - d|}{\cos(\alpha)}. \tag{3.10}$$

(ii) If $d_p \leq d$ then

$$A_l = A + \overline{QQ_l}. \tag{3.11}$$

Previously we detected if the center point of the pixel is inside or outside the polygon. This, together with the calculated A_l, is used to get the sub-pixel coverage of the polygon and the pixel p:

$$A_l = \begin{cases} 1 - A_l & \text{if pixel center is inside polygon} \\ A_l & \text{if pixel center is outside polygon} \end{cases}. \tag{3.12}$$

There are a couple of special cases to consider when calculating the sub-pixel coverage of the polygons, see Fig 3.14.

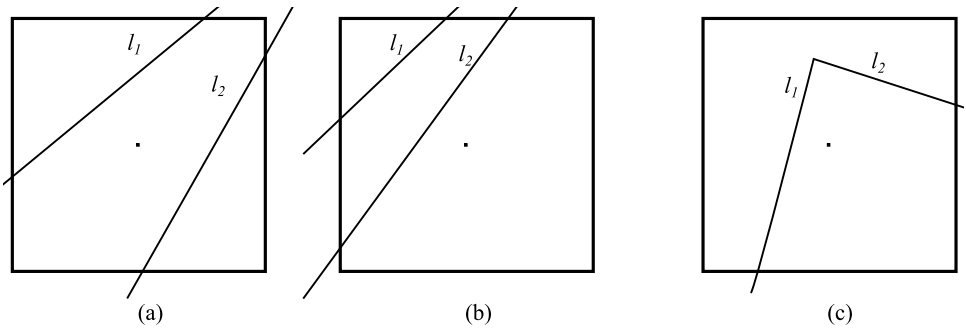


Figure 3.14: Special cases for pixel coverage. (a) Two lines crossing the same pixel and the center point of the pixel is between the lines. (b) Two lines crossing the same pixel and the center point of the pixel is not between the lines. (c) A vertex point of the polygon inside the pixel.

To solve case (a), for each line we calculate the part that is outside the polygon, i.e., the side without the center point of the pixel. This area for both of the lines are summed together and the total area of the coverage pixel is calculated by subtracting this value from the total area of a pixel. In case (b) the center point of the pixel is not between the lines. To get the area coverage we need to subtract the area A_{l_1} from A_{l_2}, calculated with (eq. 3.12). The area coverage in (c) is approximated by using sub-pixel precision for only one of the lines. The polygons in this case always have four lines, see Fig. 3.12. The dotted lines in Fig. 3.12 (upper and lower line segments) will be short and on the inner border of the tube, while the other two will be long and border adjacent polygons. The lines that are on the edge of the glass tube are assumed to be of less importance. To simplify the vertex problem we therefore chose to not have sub-pixel precision for the lines on the edge of the glass tube (dotted

lines); they will be either inside or outside the polygon, depending on the position of center point of the pixel.

Once the pixel coverage for all polygons at all the acquisition angles are calculated, the system matrix can be filled with the ray paths for each of the bins in the detector.

It is hard to evaluate the increase in accuracy of the reconstructions following this new light path algorithm. A good phantom object in zebrafish size that will fit inside the tube is yet to be designed to truly evaluate image quality improvements.

3.2.2.4 Tomographic reconstruction

As mentioned in section 3.2.2.2 the FBP is used to get a fast reconstruction for detecting the COR. When performing the reconstruction of the complete data set, the quality of the reconstruction is more important than the speed, we therefore use the OSEM algorithm together with the system matrix described in Sec. 3.2.2.3.

A gray-scale reconstruction of one fish, approximately 300 images per revolution and size of reconstruction 250x250x450 pixels, takes 50 sec with FPB. With OSEM, using 12 subsets and 2 iterations, the reconstruction time is around 3.5 min. These measurements were done on a Intel Core i7 2.67GHz with 6GB of RAM.

Bright field tomography

Tomographic reconstruction of bright field images is possible with stained or unstained fish. In paper IX an Alcian blue stain was used to enhance the bone of the fish, see Fig. 3.15. The alignment was done as described in Section 3.2.2.2. For color reconstruction, each channel is reconstructed individually and combined after reconstruction. An RGB reconstruction can be seen in Fig. 3.16.

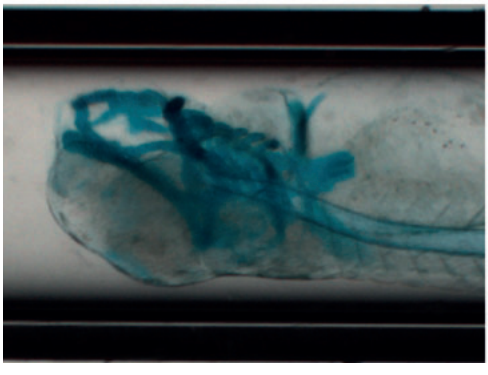


Figure 3.15: A single rotational image from a zebrafish larvae stained with Alcian blue

In paper IX, bone structure in the fish is segmented from the RGB reconstruction. The segmentation is done with a color classification. The user trains the classifier before segmenting all the fish by marking regions that belong to bone and not bone. Linear discriminant Analysis (LDA) is used to classify each pixel in the image as bone or not [20]. The feature vector consists of the three colors. Furthermore, a minimum threshold for the volume of bone structure is used to remove small objects. The result after color classification can be seen in Fig. 3.17. The segmentation provides a possibility to extract quantitative measurements of the bones.

Fluorescence tomography

When studying structures not easily targeted by light absorbing stains it is desirable to use fluorescence, and reconstruct 3D volumes by fluorescence tomography. Since both fluorescence and bright field images can be acquired at the same time, we only need to align one of them. We use the bright field alignment described in Section 3.2.2.2 and then use the same transformation for the fluorescence images. As for the bright field reconstruction, the fluorescent images are reconstructed with OSEM.

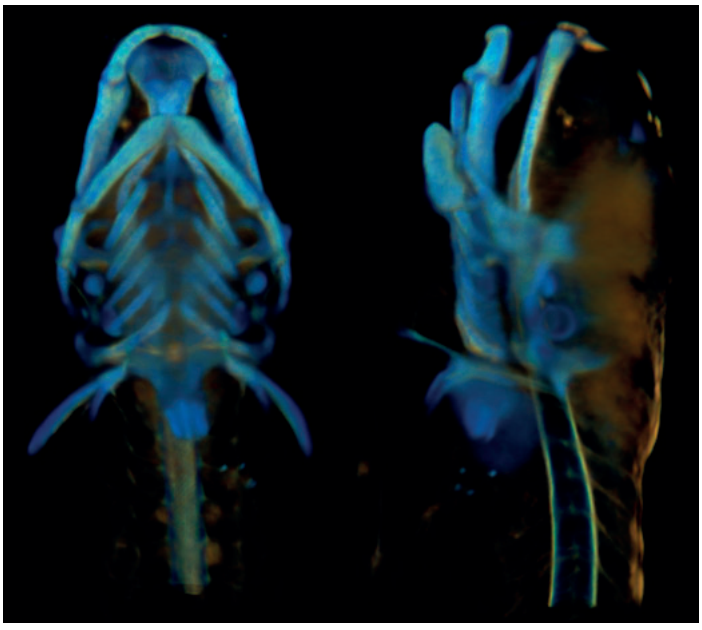


Figure 3.16: RGB reconstruction of zebrafish stained with Alcian blue.

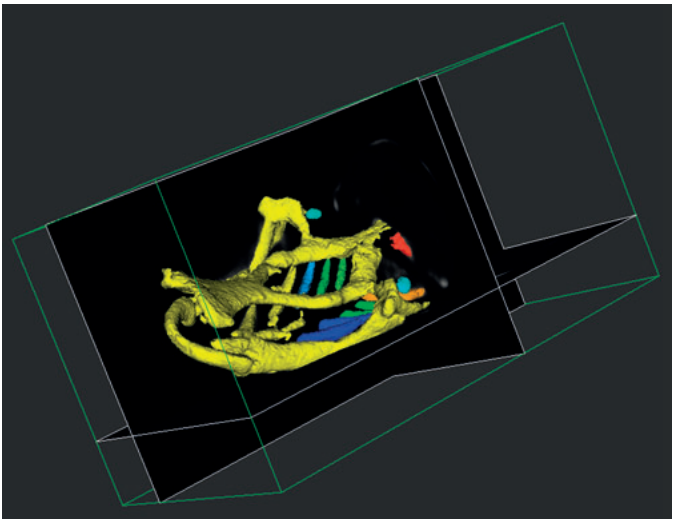


Figure 3.17: Color segmentation from RGB reconstruction of zebrafish stained with Alcian Blue.

4. Conclusion

4.1 Summary

Single cell analysis is crucial to capture variations between cells in the same population. Papers I and II present and evaluate a single cell analysis method for quantifying the number of signals per cell. This method provides sufficiently good results to discriminate between cells of different characteristics in the same sample. The algorithms where incorporated in the freely available software BlobFinder and BlobFinder Bright Field, which since the publication of Paper III in 2009 have been downloaded more than 500 times. This has provided good feedback of the single cell analysis method and shows that the demand for user-friendly applications of this type is high.

The detection of point-like signals in fluorescence microscopy images is an important field and many different algorithms have been developed to accurately detect these signals. In papers IV and V a novel method for signal detection is presented. The method shows better performance than the commonly used methods compared in the paper, both on synthetic and real data.

In samples with highly variable concentrations of signals, simultaneous detection of high and low concentrations is difficult and sometimes even impossible. Paper VI presents a method for extending the dynamic range of signal detection from PLA probes in order to allow detection of both abundant and scarce target molecules. The increased dynamic range provides a great tool for measuring signals in cell cultures with high variation in signal concentration.

High-throughput screening is an effective strategy for discovering disease pathways and new potential drugs. In Papers VII and VIII, improvements to the previously developed zebrafish high-throughput system VAST are presented. The throughput and utility of the system has been significantly increased by multi-threading the platform and adding fully automatic orientation and positioning of zebrafish. The improvements enable larger and more complex screening studies, using this model organism.

Paper IX presents a high-throughput optical projection tomography (OPT) system capable of acquiring 3D images of live zebrafish both in bright field and fluorescence. The combination of high-throughput and optical projection tomography provides a novel and useful tool for visualization and quantification of 3D features in large screens.

4.2 Concluding remarks

The increased use of digital images in microscopy, together with the development of new staining and acquisition techniques, has created a high demand for fast and accurate analysis of the data. Image analysis provides an invaluable tool for quantitative analysis of many types of microscopy images. Today, image analysis is an integral part in many biologists' everyday work. However, there are still many analyses performed manually that can be automated to increase accuracy and throughput. This thesis has presented several algorithms, methods, and software that will increase automation for many applications in the areas of biology and medicine.

The single cell analysis method presented and evaluated in this thesis is applicable on a wide range of different types of cells and cell cultures. The various software programs developed have been widely used, and the feedback from the users confirms the previously mentioned demand for these types of analysis. As the acquisition and staining techniques are continuously being improved, the software and algorithms also need to develop further to accommodate these changes. One such example is the increased dynamic range presented in this thesis. For biologists to fully utilize the benefits of multiple staining colors, a method and software that can automatically quantify large data sets of these types of images is required.

Detection of point source signals is application dependent, and the choice of detection algorithm might differ depending on the image data. In our evaluation, the 3DSWD performed better than the other methods used in the comparison. However, there are numerous other methods, not mentioned in this thesis, that can be used for detection of point-like signals. There are better or worse methods depending on the application, but no method will work perfectly on all types of images. Providing a library containing many different algorithms for signal detection would help users find the optimal algorithm for their specific type of images.

Information extracted from studying individual cells or cell cultures is limited when it comes to revealing information on the effects seen on a whole organism in response to the addition of potential drugs or genetic changes. There is a growing interest in studying model organisms as they provide a possibility to study effects on the organism as whole system. VAST and the hight throughput tomography system presented in this thesis not only provide a possibility to study the zebrafish as a whole organism, but also allow biologists to do this in a high throughput *in vivo* environment. It allows the power to extract information at the cellular scale by combining low resolution microscopy for positioning, and high resolution confocal microscopy for cell detection. This enables large-scale *in vivo* studies of biological processes such as organ development, neural degeneration and regeneration, pathogenesis, and drug toxicity.

Acknowledgement

This thesis is the result of the research conducted at the Centre for Image Analysis (CBA), Uppsala, Sweden. I would like to express my sincere gratitude and appreciation toward all those people who have contributed, directly or indirectly, to this thesis.

- My main supervisor Carolina Wählby, whose encouragement, guidance and support have been the absolute best possible. Thank you for the scientific expertise and all the great opportunities you have provided for me. I would also like to thank your wonderful family: Anders, Jonatan, Agnes and Matilda for welcoming me to your home during my stay in Boston, it was a great experience.
- My assistant supervisor Ida-Maria Sintorn for providing great scientific input and encouragement. I have known that your support has always been available for me, and that has been a well needed comfort.
- My assistant supervisor Ewert Bengtsson for your great leadership, assistance and for opening the gates to academia for me.
- My collaborators at MIT, Boston. Special thanks to Prof. Mehmet Yanik Fatih for providing me with the opportunity to come and work with your impressive group. It was an invaluable experience. Thanks to Carlos Pardo-Martin for a fruitful collaboration and always providing great image data as well as great company.
- My supervisors, Robin Strand, Cris Luengo, and Vladimir Ćurić for proof reading this thesis and providing me with great comments.
- Ingrid Carlbom, Ann-Catrin Andersson, Fredrik Hjelm and Carl-Magnus Clausson for providing feedback on parts of the thesis.
- Lena Wadelius for keeping CBA running smoothly and always reminding me of the things I forget. Olle Eriksson for keeping the machines running at CBA.
- My collaborators and co-authors at Rudbeck Laboratory, without your interesting projects and image data this thesis would not have been possible.
- Anne Carpenter and Broad Institute of Harvard and MIT for giving me the opportunity to come and visit you.
- Everyone at Olink Bioscinece for a great collaboration.
- The people at CBA who have been involved in my projects during this time, Amalka, Milan, and Vlad. I have learned a lot from working with you.

- Prof. Choi and his group at Inje University, Gimhae, Republic of Korea, thank you for your amazing hospitality.
- My collaborators and co-authors at Leiden University Medical Center for a fruitful collaboration and for taking good care of me in Leiden.
- Everyone at Visiopharm for a good collaboration.
- My colleagues and friends, past and present, at CBA, you have always been there for me and created a friendly environment.
- My roommate Patrik, for being the perfect officemate during this time. Thank you for providing the means for a great playground.
- BT, for all interesting and uninteresting discussions, and analysis of Habiben's issues.
- My Emma, for all the happiness and comfort you bring. I am looking forward to spending more time with you;) Sötis, du e bäst!
- My sister, thanks for believing in me, without your constant support and encouragement, I would not be the person I am today. Hope you enjoyed proof reading this thesis.
- Mom and Dad, for your love, support and belief in me.

Uppsala, October 5th 2011

Amin Allalou

Summary in Swedish

Datoriserad bildanalys är en teknik som används för att automatiskt analysera/extrahera information från digitala bilder. Datoriserad bildanalys omfattar många områden, t.ex. identifiering av fingeravtryck, automatiskt läsning av registreringsskyltar, eller räkning av celler i biologiska bilder. Med dagens kraftfulla datorer och nya bildalstrande tekniker har datoriserad bildanalys fått en betydande roll inom många forskningsområden.

Inom den medicinska forskningen används ofta ljusmikroskopi för att titta på celler eller organismer som är för små för att synas med blotta ögat. Man använder ofta odlade celler för att studera sjukdomar eller hur olika mediciner påverkar sjukdomstillstånd som kan modelleras på cellnivå. För att kunna se cellerna i ett mikroskop används ofta någon typ av infärgning av cellens olika delar. Genom att specifikt märka t.ex. DNA eller proteiner, går det att påvisa olika egenskaper hos celler. Mats Nilsson och Ulf Landegren vid Rudbeck-laboratoriet i Uppsala har tagit fram en metod som kallas 'pad-lock' eller 'proximity-ligation' som på ett mycket exakt sätt märker platser där en viss gen eller ett vist protein finns. Detta syns som små prickar, s.k. punktsignaler, runt eller i en cell i en fluorescensmikroskopibild. När bilderna sedan ska analyseras vill man räkna hur många DNA- eller proteinsignaler som existerar i de olika cellerna under olika förhållanden. Att analysera manuellt är extremt tidskrävande då det finns många prickar i varje cell, många celler i varje bild, och många bilder i varje experiment. Datoriserad bildanalys är då en bra lösning för en automatisk och snabb beräkning av celler och prickar.

Precis som människor så är ingen cell den andra lik, utan egenskaper varierar mellan olika celler. När det gäller analys av celler så är det viktigt att bildanalysmetoderna räknar prickar per cell och inte bara räknar den totala mängden prickar och celler i en bild. I den här avhandlingen presenteras och utvärderas en metod för att automatiskt analysera mängden DNA eller proteiner per cell i mikroskopibilder. Vi visar att med dessa metoder går det att på att ett robust sätt kvantifiera skillnader i celler som har olika egenskaper. Metoderna har utvärderats på artificiella bilder, sedan testats på flera olika typer av mikroskopibilder, och visat bra kvantifiering av antalet prickar per cell. Dessa metoder har sedan paketerats i olika mjukvaror för att både öka tillgängligheten av metoderna och användarvänligheten för forskar från andra områden än datoriserad bildanalys.

För att en automatisk analysmetod ska kunna räkna antalet punktsignaler per cell behövs det en robust metod för att detektera dessa signaler. I den här

avhandlingen presenteras en ny metod för robust detektion av punktsignaler i 3D-bilder; 3DSWD. Metoden har visat sig fungera bra genom tester på både artificiella och riktiga mikroskopibilder. 3DSWD har jämförts mot flera tidigare välkända metoder och visat bättre kvantifiering av signaler.

Människans kropp är ett komplext system uppbyggt av organ och mindre system som interagerar med varandra. Organfunktioner och denna sorts interaktioner är svåra att studera på cellnivå. För att lättare studera komplexa system används olika djurmodeller som t.ex. bananflugor, rundmaskar och zebrafiskar. Dessa djurmodeller är betydligt mindre och har ett mer lättöverskådligt system än människan. De har också en betydligt kortare generationstid (dagar-veckor) och det är oftast etiskt mer försvarbart att utföra storskaliga studier på enkla modellorganismer än på mer komplexa djur så som råttor och möss eller människor. Samtidigt är grundfunktionen för många organ och organinteraktioner, och därmed även sjukdomsmekanismer, väl bevarade genom evolutionen. Genom att utföra experiment på djurmodeller kan slutsatser dras om sjukdomsförlopp eller hur mediciner påverkar vissa sjukdomar hos människan. Av dessa enkla djurmodeller är zebrafisken en mycket bra modell för tidig utveckling hos ryggradsdjur. Dess embryon är enkelt uppbyggda och transparanta, vilket gör dem möjliga att studera med mikroskopi. Embryoutvecklingen är snabb och efter bara några dagar så har de flesta organ utvecklats. Efter 3-4 månader så kan en hona producera avkomma. Honan kan lägga 100-tals ägg under loppet av en vecka. Ett stort antal zebrafiskar kan alltså tas fram på mycket kort tid. Detta gör att zebrafisken är mycket lämplig att använda i storskaliga experiment.

Helautomatiserad hantering av zebrafiskar (eller celler) tillsamns med bildanalysteknik gör det möjligt att undersöka effekten av olika substanser på kort tid, så kallad 'High Throughput Screening' (HTS). HTS-system gör det möjligt att analysera effekten av en stor mängd substanser på ett stort antal djur. På Yanik lab, Research Laboratory of Electronics, MIT, har ett HTS system utvecklats för att titta på zebrafiskar. Levande fiskar transporteras i systemet genom små tunna rör tills de kommer in i ett mikroskops synfält där de sedan avbilds. Vid olika experiment studeras olika regioner av fisken. För att systemet automatiskt ska kunna ta bilder av en specifik region så behövs algoritmer för att detektera fiskens position. I den här avhandlingen så beskrivs bildanalysalgoritmer som är framtagna för att detektera och positionera fisken korrekt så att bilder av den önskade regionen kan tas automatiskt. Dessutom beskrivs ett nytt HTS system som har utvecklats för att ta tredimensionella (3D) volymsbilder av zebrafiskar. Fördelen med att använda 3D bilder är att man får ut mer och precisare information än som är möjligt att få från en 2D bild. Tillexempel volym, positioner i 3D rymden och vinklar mellan olika objekt mäts mer exakt i 3D än i 2D. 3D bilderna tas fram genom att fisken roteras och bilder tas under rotationen. Dessa bilder används sedan för att rekonstruera en 3D bild ungafär på samma sätt som i en röntgentomograf, men i detta fall används ett vanligt ljusmikroskop. Systemet är helt

automatiserat och medför nya möjligheter att göra storskaliga experiment på zebrafiskar. Avhandlingen visar bland annat att systemet kan användas för att se om varierande koncentration av olika läkemedel påverkar utvecklingen av zebrafiskars käkben, något som kan ge information om dess effekt på människan.

Bibliography

[1] M. D. Abramoff, P. J. Magelhaes, and S. J. Ram. Image Processing with ImageJ. *Biophotonics International*, 11(7):36–42, 2004.

[2] M. Adamek. Creating a Graphical User Interface Template for Izolde: The complete design process, focusing on usability and design. Master's thesis, Uppsala University, 2010.

[3] B. N. Alex. The concept of integrative levels and biology. *Science*, 101(2618):209–215, 1945.

[4] A. H. Andersen and A. C. Kak. Simultaneous algebraic reconstruction technique (SART): a superior implementation of the ART algorithm. *Ultrasonic Imaging*, 6(1):81–94, 1984.

[5] D. O. Antson, A. Isaksson, U. Landegren, and M. Nilsson. PCR-generated padlock probes detect single nucleotide variation in genomic DNA. *Nucleic Acids Research*, 28(12):E58, 2000.

[6] B. A. Barut and L. I. Zon. Realizing the potential of zebrafish as a model for human disease. *Physiological Genomics*, 2(2):49–51, 2000.

[7] S. Bolte and F. P. Cordelieres. A guided tour into subcellular colocalization analysis in light microscopy. *Journal of Microscopy*, 224(Pt 3):213–232, 2006.

[8] B. H. Brown. *Medical physics and biomedical engineering*. Series in medical physics. Institute of Physics Pub., 1999.

[9] I. B. Buchwalow and W. Böcker. *Immunohistochemistry: Basics and Methods*. Springer, 2010.

[10] A. E. Carpenter. Image-based chemical screening. *Nature Chemical Biology*, 3(8):461–465, 2007.

[11] A. E. Carpenter, T. R. Jones, M. R. Lamprecht, C. Clarke, I. H. Kang, O. Friman, D. A. Guertin, J. H. Chang, R. A. Lindquist, J. Moffat, P. Golland, and D. M. Sabatini. Cellprofiler: image analysis software for identifying and quantifying cell phenotypes. *Genome Biology*, 7(10):R100, 2006.

[12] V. Caselles, F. Catté, T. Coll, and F. Dibos. A geometric model for active contours in image processing. *Numerische Mathematik*, 66(1):1–31, 1993.

[13] T. F. Chan and L. A. Vese. Active contours without edges. *IEEE Transactions on Image Processing*, 10(2):266–277, 2001.

[14] K. R. Chi. How to reel in high-throughput results using worms and fish. *The Scientist*, 23(8):51, 2009.

[15] R. Cierniak. *X-Ray Computed Tomography in Biomedical Engineering*. Springer, 2011.

[16] M. Defrise. A short readers guide to 3D tomographic reconstruction. *Computerized medical imaging and graphics the official journal of the Computerized Medical Imaging Society*, 25(2):113–116, 2001.

[17] S. J. Du, V. Frenkel, G. Kindschi, and Y. Zohar. Visualizing normal and defective bone development in zebrafish embryos using the fluorescent chromophore calcein. *Developmental Biology*, 238(2):239–246, 2001.

[18] R. O. Duda and P. E. Hart. Use of the Hough transformation to detect lines and curves in pictures. *Communications of the ACM*, 15(1):11–15, 1972.

[19] J. Dupač and V. Hlaváč. Stable Wave Detector of Blobs in Images. In *Pattern Recognition*, volume 4174 of *Lecture Notes in Computer Science*, pages 760–769. Springer Berlin / Heidelberg, 2006.

[20] R. A. Fisher. The use of multiple measurements in taxonomic problems. *Annals of Eugenics*, 7(2):179–188, 1936.

[21] R. C. Gonzalez and R. E. Woods. *Digital Image Processing (3rd Edition)*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 2006.

[22] K. E. Gordon, B. Binas, R. S. Chapman, K. M. Kurian, R. W. E. Clarkson, A. J. Clark, E. B. Lane, and C. J. Watson. A novel cell culture model for studying differentiation and apoptosis in the mouse mammary gland. *Breast Cancer Research*, 2(3):222–235, 2000.

[23] R. Gordon, R. Bender, and G. T. Herman. Algebraic reconstruction techniques (ART) for three-dimensional electron microscopy and x-ray photography. *Journal of Theoretical Biology*, 29(3):471–481, 1970.

[24] P. Grangeat. *Tomography*. Digital signal and image processing series. John Wiley & Sons, 2009.

[25] M. Guizar-Sicairos, S. T. Thurman, and J. R. Fienup. Efficient subpixel image registration algorithms. *Optics Letters*, 33(2):156–158, 2008.

[26] Y. Gusev, J. Sparkowski, A. Raghunathan, H. Ferguson, J. Montano, N. Bogdan, B. Schweitzer, S. Wiltshire, S. F. Kingsmore, W. Maltzman, and V. Wheeler. Rolling circle amplification: a new approach to increase sensitivity for immunohistochemistry and flow cytometry. *The American journal of pathology*, 159(1):63–69, 2001.

[27] E. Haines. *Point in Polygon Strategies*, pages 24–46. Academic Press, 1994.

[28] M.A. Hayat. *Principles and techniques of electron microscopy: biological applications*. Cambridge University Press, 2000.

[29] G. T. Herman and A. Lent. Iterative reconstruction algorithms. *Computers in Biology and Medicine*, 6(4):273–294, 1976.

[30] H. M. Hudson and R. S. Larkin. Accelerated image reconstruction using ordered subsets of projection data. *IEEE Transactions on Medical Imaging*, 13(4):601–609, 1994.

[31] J. Inglese, R. L. Johnson, A.n Simeonov, M. Xia, W. Zheng, C. P. Austin, and D. S. Auld. High-throughput screening assays for the identification of chemical probes. *Nature Chemical Biology*, 3(8):466–479, 2007.

[32] A. C. Kak and M. Slaney. *Principles of Computerized Tomographic Imaging*. IEEE Press, New York, 1988.

[33] J. Kapuscinski. DAPI: a DNA-specific fluorescent probe. *Biotechnic histochemistry official publication of the Biological Stain Commission*, 70(5):220–233, 1995.

[34] M. Key. In immunohistochemical staining methods. *Dako pathology education guide*, pages 49–5, 2006.

[35] P. Kruizinga and N. Petkov. Computational model of dot-pattern selective cells. *Biological Cybernetics*, 83(4):313–325, 2000.

[36] K. Lange and R. Carson. EM reconstruction algorithms for emission and transmission tomography. *Journal Of Computer Assisted Tomography*, 8(2):306–316, 1984.

[37] C. Lantuéjoul and S. Beucher. On the use of geodesic metric in image analysis. *Journal of Microscopy*, 121:39–49, 1981.

[38] C. Larsson, I. Grundberg, O. Söderberg, and M. Nilsson. In situ detection and genotyping of individual mRNA molecules. *Nature Methods*, 7(5):395–397, 2010.

[39] K. Leuchowius, I. Weibrecht, and O. Söderberg. In situ proximity ligation assay for microscopy and flow cytometry. *Current Protocols in Cytometry*, Chapter 9, 2011.

[40] J. P. Lewis. Fast normalized cross-correlation. In *Vision Interface*, pages 120–123. Canadian Image Processing and Pattern Recognition Society, 1995.

[41] C. Li, C. Xu, C. Gui, and M. D. Fox. Level set evolution without re-initialization: A new variational formulation. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 1, pages 430–436, Washington, DC, USA, 2005. IEEE Computer Society.

[42] G.J. Lieschke and P.D. Currie. Animal models of human disease: zebrafish swim into view. *Nature Reviews Genetics*, 8:353–367, 2007.

[43] I. Lobo. Biological complexity and integrative levels of organization. *Nature Education*, 1(1), 2008.

[44] F. Madeo, S. Engelhardt, E. Herker, N. Lehmann, C. Maldener, A. Proksch, S. Wissing, and K. U. Fröhlich. Apoptosis in yeast: a new model system with applications in cell biology and medicine. *Current Genetics*, 41(4):208–216, 2002.

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

[46] MATLAB. *version 7.5.0 (R2007b)*. The MathWorks Inc., Natick, Massachusetts, 2007.

[47] C. R. Maurer, R. Qi, and V. Raghavan. A Linear Time Algorithm for Computing Exact Euclidean Distance Transforms of Binary Images in Arbitrary Dimensions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(2):265–270, 2003.

[48] P. McGrath and C. Li. Zebrafish: a predictive model for assessing drug-induced toxicity. *Drug Discovery Today*, 13:394–401, 2008.

[49] F. Meyer. Iterative image transformations for an automatic screening of cervical smears. *The journal of histochemistry and cytochemistry official journal of the Histochemistry Society*, 27(1):128–135, 1979.

[50] J. C. Olivo-Marin. Extraction of spots in biological images using multiscale products. *Pattern Recognition*, 35(9):1989–1996, 2002.

[51] S. Osher and J. A. Sethian. Fronts propagating with curvature-dependent speed: Algorithms based on Hamilton-Jacobi formulations. *Journal of Computational Physics*, 79(1):12–49, 1988.

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

[53] S.W. Paddock. *Confocal Microscopy: Methods and Protocols*. Methods in Molecular Biology Series. Humana Press, 2010.

[54] C. Pardo-Martin, T. Y. Chang, B.K. Koo, C. L. Gilleland, S.C. Wasserman, and M.F. Yanik. High-throughput *in vivo* vertebrate screening. *Nature Methods*, 7:634–636, 2010.

[55] P. Perona and J. Malik. Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(7):629–639, 1990.

[56] J. Radon. On the determination of functions from their integral values along certain manifolds. *IEEE Transactions on Medical Imaging*, 5(4):170–176, 1986.

[57] G. N. Ramachandran and A. V. Lakshminarayanan. Three-dimensional reconstruction from radiographs and electron micrographs: application of convolutions instead of Fourier transforms. *Proceedings of the National Academy of Sciences of the United States of America*, 68(9):2236–2240, 1971.

[58] J. Rittscher, R. Machiraju, and S.T.C. Wong. *Microscopic image analysis for life science applications*. Artech House bioinformatics & biomedical imaging series. Artech House, 2008.

[59] A.L. Rubinstein. Zebrafish: From disease modeling to drug discovery. *Current Opinion in Drug Discovery & Development*, 6:218–223, 2003.

[60] J.C. Russ. *The Image Processing Handbook*. Taylor & Francis, 2011.

[61] P. K. Sahoo, S. Soltani, A. K.C. Wong, and Y. C. Chen. A survey of thresholding techniques. *Computer Vision, Graphics, and Image Processing*, 41:233–260, 1988.

[62] B. Schweitzer, S. Wiltshire, J. Lambert, S O'Malley, K. Kukanskis, Z. Zhu, S. F. Kingsmore, P. M. Lizardi, and D. C. Ward. Inaugural article: immunoassays with rolling circle DNA amplification: a versatile platform for ultrasensitive antigen detection. *Proceedings of the National Academy of Sciences*, 97(18):10113–10119, 2000.

[63] J. Sharpe, U. Ahlgren, P. Perry, B. Hill, A. Ross, J. Hecksher-Sørensen, R. Baldock, and D. Davidson. Optical Projection Tomography as a Tool for 3D Microscopy and Gene Expression Studies. *Science*, 296(5567):541–545, 2002.

[64] J.T. Shin and M.C. Fishman. From zebrafish to human: modular medical models. *Annual Review of Genomics and Human Genetics*, 3:311–340, 2002.

[65] I. Smal, M. Loog, W. Niessen, and E. Meijering. Quantitative comparison of spot detection methods in fluorescence microscopy. *IEEE Transactions on Medical Imaging*, 29(2):282–301, 2010.

[66] O. Söderberg, M. Gullberg, M. Jarvius, K. Ridderstråle, K. J. Leuchowius, J. Jarvius, K. Wester, P. Hydbring, F. Bahram, L. G. Larsson, and U. Landegren. Direct observation of individual endogenous protein complexes *in situ* by proximity ligation. *Nature Methods*, 3(12):995–1000, 2006.

[67] P. Soille. *Morphological Image Analysis: Principles and Applications*. Springer-Verlag, 1999.

[68] T. Strachan and A.P. Read. *Human molecular genetics 3*. Garland Science, 2004.

[69] R. Y. Tsien. The green fluorescent protein. *Annual Review of Biochemistry*, 67(1):509–544, 1998.

[70] J. K. Udupa, V. R. LeBlanc, Y. Zhuge, C. Imielinska, H. Schmidt, L. M. Currie, B. E. Hirsch, and J. Woodburn. A framework for evaluating image segmentation algorithms. *Computerized Medical Imaging and Graphics*, 30:75–87, 2006.

[71] B. J. Vermolen, Y. Garini, I. T. Young, R. W. Dirks, and V. Raz. Segmentation and analysis of the three-dimensional redistribution of nuclear components in human mesenchymal stem cells. *Cytometry Part A*, 73(9):816–824, 2008.

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

[73] C. Wählby, I. M. Sintorn, F. Erlandsson, G. Borgefors, and E. Bengtsson. Combining intensity, edge and shape information for 2D and 3D segmentation of cell nuclei in tissue sections. *Journal of Microscopy*, 215(1):67–76, 2004.

[74] W. Wallace, L. H. Schaefer, and J. R. Swedlow. A working person's guide to deconvolution in light microscopy. *Biotechniques*, 31(5):1076–1078, 1080, 1082 passim, 2001.

[75] J. R. Walls, J. G. Sled, J. Sharpe, and R. M. Henkelman. Correction of artefacts in optical projection tomography. *Physics in medicine and biology*, 50(19):4645–4665, 2005.

[76] B. Zhang, J. Zerubia, and J.C. Olivo-Marin. Gaussian approximations of fluorescence microscope point-spread function models. *Applied Optics*, 46(10):1819–1829, 2007.

[77] A. Zieba, C. Wählby, F. Hjelm, L. Jordan, J. Berg, U. Landegren, and K. Pardali. Bright-field microscopy visualization of proteins and protein complexes by *in situ* proximity ligation with peroxidase detection. *Clinical chemistry*, 56(1):99–110, 2010.

[78] L. I. Zon and R. T. Peterson. In vivo drug discovery in the zebrafish. *Nature reviews. Drug discovery.*, 4:35–44, 2005.

Acta Universitatis Upsaliensis

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology 856

Editor: The Dean of the Faculty of Science and Technology

A doctoral dissertation from the Faculty of Science and Technology, Uppsala University, is usually a summary of a number of papers. A few copies of the complete dissertation are kept at major Swedish research libraries, while the summary alone is distributed internationally through the series Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology.

Distribution: publications.uu.se
urn:nbn:se:uu:diva-159196

ACTA
UNIVERSITATIS
UPSALIENSIS
UPPSALA
2011