Segmentation Methods and Shape Descriptions in Digital Images

Applications in 2D and 3D Microscopy

Ida-Maria Sintorn

Centre for Image Analysis
Uppsala

Doctoral Thesis
Swedish University of Agricultural Sciences
Uppsala 2005

Acta Universitatis Agriculturae Sueciae
2005:20

ISSN 1652-6880
ISBN 91-576-7019-6

Printed in Sweden by Universitetstryckeriet, Uppsala University, Uppsala, 2005

Abstract

Ida-Maria Sintorn. *Segmentation Methods and Shape Descriptions in Digital Images*. Doctoral Thesis

ISSN 1652-6880, ISBN 91-576-7019-6

Digital image analysis enables creating objective, fast, and reproducible analysis methods of objects or situations that can be imaged.

This thesis contains theoretical work regarding distance transforms for images digitized in elongated grids. Such images are the result of many, mainly 3D, imaging devices. Local weights appropriate for different elongation factors in 2D, as well as in 3D, are presented. Methods adapted to elongated grids save time and computer memory compared to increasing the image size by interpolating to a cubic grid.

A number of segmentation methods for images in specific applications are also included in the thesis. Distance information is used to segment individual pores in paper volume images. This opens the possibility to investigate how the pore network affects the paper quality. Stable and reliable segmentation methods for cell nuclei are necessary to enable studies of tumor morphology, as well as amounts of fluorescence marked substances in individual nuclei. Intensity, gradient magnitude, and shape information is combined in a method to segment cell nuclei in 2D fluorescence and 3D confocal microscopy images of tissue sections. Two match based segmentation methods are also presented. Three types of viral capsids are identified and described based on their radial intensity distribution in transmission electron micrographs of infected cells. This can be used to measure how a potential drug affects the relative amounts of the three capsids, and possibly, the viral maturation pathway. Proteins of a specific kind in transmission electron volume images of a protein solution are identified using a shape based match method. This method reduces the amount of visual inspection needed to identify proteins of interest in the images.

Two representation schemes, developed in order to simplify the analysis of individual proteins in volume images of proteins in solution, are presented. One divides a protein into subparts based on the internal intensity distribution and shape. The other represents the protein by the maximum intensity curve connecting the centers of the subparts of the protein. These representations can serve as tools for collecting information about how flexible a protein in solution is and how it interacts with other proteins or substances. This information is valuable for the pharmaceutical industry, when developing new drugs.

Key words: digital image analysis, volume images, microscopy images, elongated grid, distance transform, segmentation, shape description, grey-level, gradient magnitude, watershed, decomposition

Author's address: Centre for Image Analysis, Lägerhyddsvägen 3, SE-752 37 Uppsala, Sweden

Contents

Segmentation Methods and Shape Descriptions in Digital Images

Papers I – VIII

Papers appended to the thesis

The thesis is based on the following articles. All published papers are reproduced with permission from the publisher.

I Sintorn I-M., Borgefors G. (2001). Weighted distance transforms in rectangular grids. In Ardizzone, E. and Gesù, V., editors, *Proc. 11th International Conference on Image Analysis and Processing, Palermo, Italy*, pages 322–326. IEEE Computer Society.

II Sintorn I-M., Borgefors G. (2004). Weighted distance transforms for images digitized in elongated voxel grids. *Pattern Recognition Letters* 25:571–580.

III Sintorn I-M., Axelsson M., Svensson S., Borgefors G. (2005). Segmentation of individual pores in paper volume images. *Submitted for journal publication.*

IV Sintorn I-M., Homman-Loudiyi M., Söderberg-Nauclér C., Borgefors G. (2004). A refined circular template matching method for classification of human cytomegalovirus capsids in TEM images. *Computer Methods and Programs in Biomedicine*, 76(2):95–102.

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

VI Sintorn, I-M., Mata, S. (2004). Using grey-level and shape information for decomposing proteins in 3D images. In *Proc. IEEE International Symposium on Biomedical Imaging, Arlington, VA, USA*, pages 800–803. Mira Digital Publishing.

VII Sintorn, I-M., Gedda, M., Mata, S., Svensson S. (2005). Medial grey-level based representation for proteins in volume images. To appear in *Proc. of Iberian Conference on Pattern Recognition and Image Analysis, Estoril, Portugal, Lecture Notes in Computer Science*. Springer-Verlag.

VIII Sintorn, I-M., Borgefors, G. (2005). Shape based identification of proteins in volume images. *Submitted for publication.*

The work presented in Papers I and II were developed under close discussions with Gunilla Borgefors while the calculations and implementations (C, and C++) and most of the writing was performed by the author. The method presented in Paper III was developed and written mainly by the author but with help, comments and advice from the coauthors. Adapting existing methods and implementing new ones (C++) were performed by the author. The work in Paper IV was closely discussed with Mohammed Homman-Loudiyi, while the author performed the algorithm development, implementations (Matlab) and most of the writing herself. The work in Paper V was performed in close cooperation with Carolina Wählby. The method development, implementation and writing was split between her and the author. The work presented in Paper VI was performed mainly by the author, regarding method development, adapting existing implementations and creating new ones (C++), and writing the paper. Paper VII was produced in close cooperation with Stina Svensson, and the work with developing the method, implementing it and writing the paper was split mainly between her and the author. The method presented in Paper VIII was mainly performed by the author regarding method development, implementation (C++ and Matlab), and writing the paper.

The C and C++ implementations were either done as stand alone programs or as modules in the image processing platform IMP, developed at the Centre for Image Analysis originally by Bo Nordin and with substantial contributions by Joakim Lindblad.

1 Introduction and objectives

The number of areas using digital images as tools, for evaluation and analysis, is steadily increasing. Constantly improved technology for image generating systems, and cheaper and better computers, constitute the main underlying reasons. This growth, in turn, gives rise to a demand and wish for automatic analysis and extraction of information captured in the images. The idea of using computers for performing quantitative and objective studies of information present in images, has been around for quite some time. The first textbook on the subject was published already in the late sixties by Rosenfeld (1969), and since then, the demand and interest for digital image analysis has been ever increasing.

The work leading to this thesis was performed at the Centre for Image Analysis (CBA) in Uppsala, which is a joint institution between the Swedish University of Agricultural Sciences and Uppsala University. CBA was founded in 1988 and theoretical and application oriented image analysis research has since then been performed in the fields of discrete geometry, medical and biomedical image analysis, forestry and agriculture, and remote sensing. This thesis contains work related to several of the mentioned fields. Theoretical work on digital topology has been performed, regarding distance transforms, continuing the work by Professor Borgefors, supervisor of the author, and work presented in earlier CBA theses by Nyström (1997) assistant supervisor for the author, and Svensson (2001). One application of the theoretical results is the analysis of the structure of paper, related to the thesis of Aronsson (2002). A large part of this thesis is devoted to incorporating both shape and intensity information in methods developed for, and adapted to, specific biomedical applications. Many of these methods have their roots in work on binary shape analysis of 2D and 3D images by CBA researchers Borgefors, Nyström, and Svensson. Both intensity and shape information has also been used in the task of digital cell image analysis, connected to the thesis by Wählby (2003).

Objectives

The main objectives of this thesis have been to adopt established image analysis methods to work directly on elongated digitalization grids, and to incorporate intensity information in distance based shape analysis and representation schemes of 2D and 3D biomedical structures.

About this thesis

This is a thesis in digital image analysis, and the emphasis therefore lies on the developed methods, and not on the different applications. The concepts and methods are, hence, presented as generally as possible, although certain application specific circumstances and problems are discussed to motivate the choice of actions. The images studied in this thesis were all acquired with different microscopy techniques. A brief description of how the different microscopes work is given in Section 2. The author has, however, not acquired any of the images herself, and is not an expert on any of the systems. Section 3 contains image analysis concepts that constitute the foundation of this work. In Section 4, the methods in the appended Papers are described, together with discussions and ideas for further developments. Finally, conclusions and a summary of ideas for future work are presented in Section 5.

2 Microscopy techniques

The methods presented in this thesis were in most cases developed for either a specific application or a specific type of images. The images were all acquired by experts in the microscopy techniques briefly presented below. For comparative purposes, absorption light microscopy is also described, although no images studied in this thesis were acquired using that technique.

2.1 Light microscopy

In a common light microscope, the visible light that is not absorbed by the sample creates the magnified image. An illustration is shown in Figure 1 (left). Light from a light source is focused on the sample using a glass lens. The light that passes through the sample is magnified and focused on the detector using two lenses. The resolution limit in a light microscope is $0.2\mu m$, (Alberts et al., 1994), i.e., half the wavelength of blue light, which is the visible light with the shortest wavelength.

2.2 Fluorescence microscopy

In fluorescence microscopy, light emitted from fluorescent molecules in the sample is imaged. A fluorescent molecule absorbs light of certain wavelengths, and then emits light of longer wavelengths (longer wavelengths have less energy). Specific structures can be marked with fluorescent molecules, or sometimes a structure is fluorescent by nature. As this technique also uses light to magnify an object, the resolution limit is $0.2\mu m$, the same as in light microscopy. A dichroic, or beam splitting mirror, reflects light below a certain wavelength while longer wavelengths are transmitted. This is used in a fluorescence microscope, together with two filters, to ensure that only absorption wavelengths for the fluorescent marker to be imaged hit the sample, and at the same time only emitted wavelengths from that marker hit the detector. In Figure 1 (middle), a schematic representation of a fluorescence microscope is seen. Light from the light source of the absorption wavelengths is let through a filter, reflected in the dichroic mirror, and focused onto the sample. When the light reaches the fluorescent molecules they begin to emit light. Light from the sample is transmitted through the dichroic mirror, filtered to ensure that only light of certain wavelengths is present, and focused onto a detector where the image is formed.

2.3 Confocal microscopy

Just as in fluorescence microscopy, light emitted from fluorescent molecules is imaged with confocal microscopy. The technique is similar to ordinary fluorescence microscopy, but with the addition that a laser light source in combination with a blocking pinhole allows for imaging a specific spot of the sample at a time, see Figure 1 (right). The light source needs to be a well focused laser to make sure that the light, which is reflected in the dichroic mirror, is focused onto a small spot in the sample. Light emitted from this spot is transmitted through the dichroic mirror and

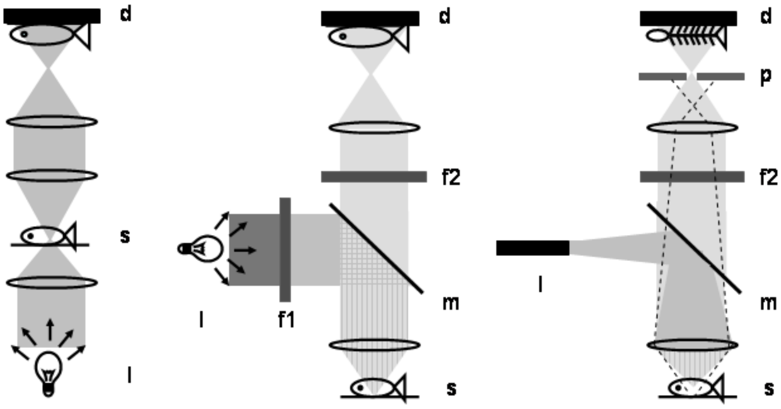


Figure 1: Simplified drawings of a light microscope (left), a fluorescence microscope (middle), and a confocal microscope (right). All three microscopes have a light source (l), a specimen (s), and a detector (d). The fluorescence microscope also has two filters (f1, f2) and a dichroic mirror (m) that control which wavelengths of the light that reach the specimen and the detector. The confocal microscope has a filter (f2) and a pinhole (p), in front of the detector, to ensure that light from out of focus objects (dashed lines) and light of undesired wavelengths do not reach the detector.

passes through a pinhole before it hits the detector. This pinhole blocks most of the light emitted from out of focus parts of the sample. The focal spot can scan the sample in x- and y-direction using a set of mirrors, and the sample can be moved in the z-direction. The confocal technique thereby allows for imaging of 3D structures, illustrated with the fish bone in Figure 1 (right). If a sampling corresponding to the best possible resolution is chosen for a specific setup, the resulting volume image will have elongated voxels, as light from out of focus objects limits the resolution in the z-direction more than in the x- and y-directions. It is very difficult to give any numbers for how thick objects that can be imaged with a confocal microscope as it depends on the sample itself as well as on the microscope setup. The amount of light reaching the detector decreases with depth, as emitted light from deeper into the sample will be absorbed and reflected by the sample on top of it and less light will, hence, reach the detector than from the top part of the sample. An example, to give a feeling for the order of the size it is possible to image, is that a good $50\mu m$ cell tissue section with a resolution of approximately $0.2\mu m$ in the x- and y-direction and $0.3\mu m$ in the z-direction can be achieved.

2.4 Transmission electron microscopy (TEM)

A transmission electron microscope, see Figure 2 (left) for a schematic drawing, images the amount of electrons that passes through the sample at different positions. The denser the material is in a spot of the sample, the more the electrons are scattered and less electrons will, hence, be transmitted through that spot. This is illustrated with the dense fish bone appearing in the resulting image in Figure 2 (left). Air molecules also scatter electrons and therefore the imaging takes place in vacuum. Electrons are focused on the sample by an electromagnetic lens and the electrons that are transmitted through the sample are magnified and projected onto a detector using two more electromagnetic lenses. The resolution limit in an electron microscope for biological samples is about $2nm$, (Alberts et al., 1994).

Electron tomography, or 3D imaging of an object using electron microscopy, can be performed by reconstruction from a series of images of the sample, acquired at different tilt angles. Sidec^{TM} Electron Tomography (SET), (SET webpage, 2004), uses the common reconstruction method filtered back projection, in combination with a refinement method called COMET, (Skoglund et al., 1996). This makes it possible to use a lower dose of electrons, which in turn, allows for acquisition of more images in a tilt series, without destroying the sample. With the combination of low dose electron microscopy and a refinement method, SET can reconstruct individual molecules down to a resolution of approximately $2nm$ (SET webpage, 2004).

2.5 Scanning electron microscopy (SEM)

In scanning electron microscopy (SEM), see Figure 2 (right), scattered electrons are detected instead of transmitted electrons. In backscatter mode, the electrons scattered back from the top part of the object are detected. The electron beam in a scanning electron microscope is focused by an electromagnetic lens and is bent using scan coils or a beam deflector to scan over the sample. At each spot of the sample, the backscattered electrons are detected and converted to an intensity value reflecting the density of the spot. Hence, a SEM image is a 2D image of the top part of the sample. The resolution limit for biological samples in SEM is about $10nm$, (Alberts et al., 1994). 3D SEM images can be produced as stacks of 2D images, each acquired using SEM, with a thin slice of the sample cut off between successive image acquisitions. The resolution in the x- and y-direction is usually higher than the thickness of the slices, and, hence, 3D images with elongated, box-like, voxels are the result.

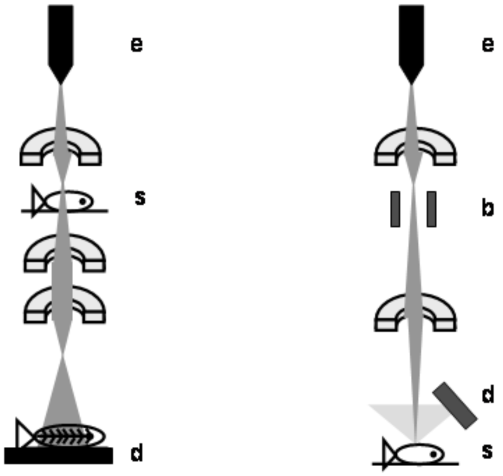


Figure 2: Simplified drawings of a transmission electron microscope (left), and a scanning electron microscope (right). Both microscopes have an electron gun (e), a specimen (s), and a detector (d). The scanning electron microscope also has a beam deflector (b).

3 Fundamental image analysis concepts

As stated in the title, image analysis methods for both 2D and 3D images have been developed. The images are, in this case, digital images (computerized discrete images). Such images are in a computer represented by grids, or matrices, with a value at each position. The location in the matrix corresponds to the spatial coordinates of the image, and the value at a location represents the grey-level, or intensity, of the corresponding position in the image. An element of a digital image, i.e., a position and value of the matrix, is in 2D denoted a pixel (picture element), and in 3D a voxel (volume picture element). Hereafter, image element or simply element will be used when no distinction between pixels and voxels is necessary. If only two intensity values are present in the image, it is a binary image. A grey-level image is an image having more than two intensity values. The range of values chosen when acquiring an image depends on what is required to solve the problem at hand and the amount of available memory. A common range is from 0 to 255, where 0 represents black and 255 white. All images used in this thesis were either acquired in that range or rescaled thereto.

A basic issue when it comes to digital image analysis is connectivity, i.e., how an image element is connected to other image elements (Kong and Rosenfeld, 1989). A pixel has two types of neighboring pixels: four edge neighbors, and four vertex neighbors, see Figure 3. A voxel has three types of neighbors; six face neighbors, twelve edge neighbors and eight vertex neighbors, see Figure 3. In 2D, depending on whether only the edge neighbors, or both the edge and vertex neighbors, are regarded as connected to a pixel, a method is based on 4-connectivity or 8-connectivity, respectively. In 3D, there are three different connectivities: 6-connectivity, where only the face neighbors are taken into consideration; 18-connectivity, where also edge neighbors are used; and 26-connectivity, where all neighbors are included.

If an image element, pixel or voxel, has the same size in all directions, all neighbors of the same type are at equal distance from the element. They should therefore be treated similarly in all processing and analysis steps. If the sampling rate is different along the different directions, all neighbors of the same type are not at equal distance from the element. This should be taken into consideration in all further processing of the image. As elongated image elements is one of the main issues in this thesis, it is discussed in more detail in Section 4.1.

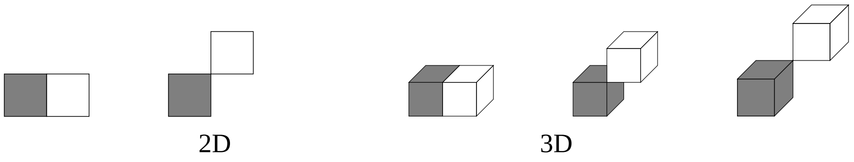


Figure 3: The types of neighbor relations in a 2D 3×3 neighborhood, and a 3D $3 \times 3 \times 3$ neighborhood.

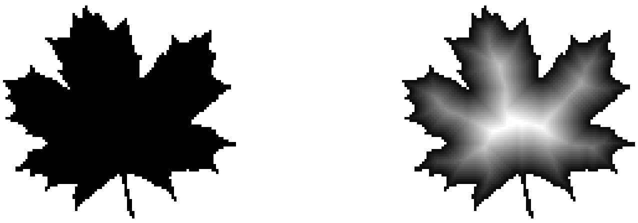


Figure 4: A binary image of a maple leaf (left), and its DT (right).

A group of connected elements that somehow is distinguished from the other elements is called an object. The complement to the object or objects is generally called the background. Usually 8-connectivity, in 2D, and 26-connectivity in 3D, is used for the object or objects. To avoid topological paradoxes (see Rosenfeld and Pfaltz (1966)), this implies that 4- and 6-connectivity has to be used for the background in 2D and 3D, respectively.

A survey of digital topology for 2D and 3D digital images can be found in Kong and Rosenfeld (1989). There, definitions for a $hole$ in 2D, and a $tunnel$ and $cavity$ in 3D are given. If an object completely surrounds a part of the background, it contains a $hole$ in 2D and a $cavity$ in 3D. If the background passes through the object, it contains a $tunnel$. A hollow ball is an example of a 3D object having a cavity, and a donut is an example of a 3D object having a tunnel.

In the following subsections fundamental image analysis methods, used or modified in the papers included in this thesis, are explained. Subsection *Digital distance transforms* covers how to approximate and calculate the closest distance from each image element in an object, or region, to the background. In Subsection *Segmentation* some methods for how to recognize and delineate objects in an image are explained. The Subsection *Object representation*, contains methods by which a segmented object can be represented to simplify further analysis.

3.1 Digital distance transforms

Measuring distances is useful in many image analysis applications, both as a descriptive measure in itself, and also as information for further processing and analysis. Distance information is used in all papers appended to this thesis. The pioneering work regarding digital distances was performed by Rosenfeld and Pfaltz (1966, 1968), and a brief overview of applications is found in Borgefors (1994).

A distance transformation is a procedure applied to an, in general, binary image, where it computes the distance from each image element in the object to the closest image element in the background. The result, the distance transform (DT), is, hence, a grey-level image where each object element has an intensity value equal to the closest distance to the background, see Figure 4.

In the earliest approaches (Rosenfeld and Pfaltz, 1966, 1968), the distance assigned to an object pixel in the DT is given by the minimum number of neighbor steps necessary to move from the pixel to the nearest pixel in the background. Depending on the connectivity used, 4 or 8 in 2D, and 6, 18, or 26 in 3D, the resulting distance values are different. The drawback with these DTs is that they are very rotation dependent. An object element can be assigned different distance values if the distance transformation is applied to the same object but rotated differently in the image.

To achieve a DT more stable under rotation, the different neighbor steps can be given different weights. The first idea (Montanari, 1968), was to approximate the length of each local neighbor step by the Euclidean distance of the step. However, this is not the best weights to use, no matter how intuitive it might sound. Borgefors (1984a) proved that other weights give a better approximation over longer distances, as well as a DT more stable under rotation. Good integer weights for the local steps in a 2D image are 3 and 4 for the edge and vertex neighbors, respectively, and for a 3D image good weights are 3, 4, and 5 for, the face, edge, and vertex neighbors respectively. Even better approximations to the Euclidean distances, and hence, a DT more stable under rotation, can be achieved if local distance information from a larger neighborhood is taken into account, see e.g. Borgefors (1986), and Svensson and Borgefors (2002a).

A DT can be calculated by propagating local distances in two passes over the image (Rosenfeld and Pfaltz, 1966). This is the case for both 2D and 3D images as well as for images of higher dimensionality. The elements in the object are set to infinity, and the elements in the background to 0, prior to running the two passes. During the first, forward pass, the image is processed from left to right and from top to bottom and in 3D, also from front to back. During the second, backward pass, the image is processed from right to left, bottom to top, and in 3D, also from back to front. The element under consideration, is given the minimum value of itself and the values of its already visited neighbors each increased by their respective local step weights. This process of propagating information over the image using local step weights is often referred to as chamfering, and weighted distance transforms (WDTs), are therefore often called chamfer distance transforms.

Once a DT has been computed, only a subset of the image elements of the object and their assigned distance values is needed to represent or reconstruct the original object. The distance values can be interpreted as radii of discs (2D) or balls (3D) totally enclosed in the object. If a disc or ball is completely covered by another disc or ball it is not needed in the representation of the original object. A disc or ball which is not completely covered by any other single disc or ball is called a maximal disc or ball. The center elements and corresponding distance values of such maximal discs, (CMDs), or balls, (CMBs), are all that is needed to represent the entire object. These CMDs/CMBs can be identified from a DT by inspecting the distance values in the local neighborhood, (Arcelli and Sanniti di Baja, 1988).

To reconstruct the original object from the CMDs/CMBs, a reverse distance transform, (Nyström and Borgefors, 1995), is computed from the distance values of the CMDs/CMBs. It is computed in the same manner as the WDT, i.e., by propagating local distance information in two passes over the image. The difference is that the maximum value of itself and the values of already visited neighbors each decreased by the respective local step weights, is assigned to the element under consideration.

It is sometimes useful to propagate distance information when obstacles are present in the image. This is performed in the same manner as in an ordinary DT with the exception that the obstacle elements are never considered. This constrained DT, CDT, (Piper and Granum, 1987), requires iteration of the two scans through the image until no more changes occur, to propagate the distance information around the obstacles. The number of needed iterations depends on the image contents (objects and obstacles).

Attempts have been made to calculate DTs on grey-level images, and thereby incorporate underlying grey-level information into the DT, (Rosin and West, 1995; Ikonen and Pekka, 2005; Saha et al., 2002; Svensson and Borgefors, 2002b). However, it is difficult to weigh together the importance of the distance and intensity information. Despite the problem of mixing modalities, grey-weighted DTs have shown to be useful in certain applications.

3.2 Segmentation

To separate the contents of an image into regions of interest, or objects, and background is called segmentation. This is a central process, as all further analysis and information extraction depend on the result of the segmentation. There exists very many different approaches to image segmentation, of which three general and common methods are explained below. These, as most image analysis methods, are usually modified to fit the segmentation task at hand.

Thresholding

Thresholding is a simple segmentation method useful in cases where regions of interest can be identified as image elements having similar grey-levels and, at the same time, the remaining image elements have different grey-levels. To find one (or more) suitable threshold values, it is often useful to study the grey-level histogram of the image. In the grey-level histogram, the grey-levels versus the number of elements of each grey-level, is plotted. In simple cases where an image contains bright objects in a dark background or vice versa, the histogram has a valley between two peaks, where the two peaks corresponds to the background and the object, respectively, see Figure 5 (top). The grey-level at the valley is then a suitable threshold. However, in most cases the histogram does not contain two well separated peaks, see Figure 5 (bottom), and choosing a suitable threshold is then not easy. Another common problem is a varying image background. In such cases it will not be possible to use a single threshold for the whole image, but more local approaches need to be considered. In Sahoo et al. (1988), an overview of different thresholding techniques is given.

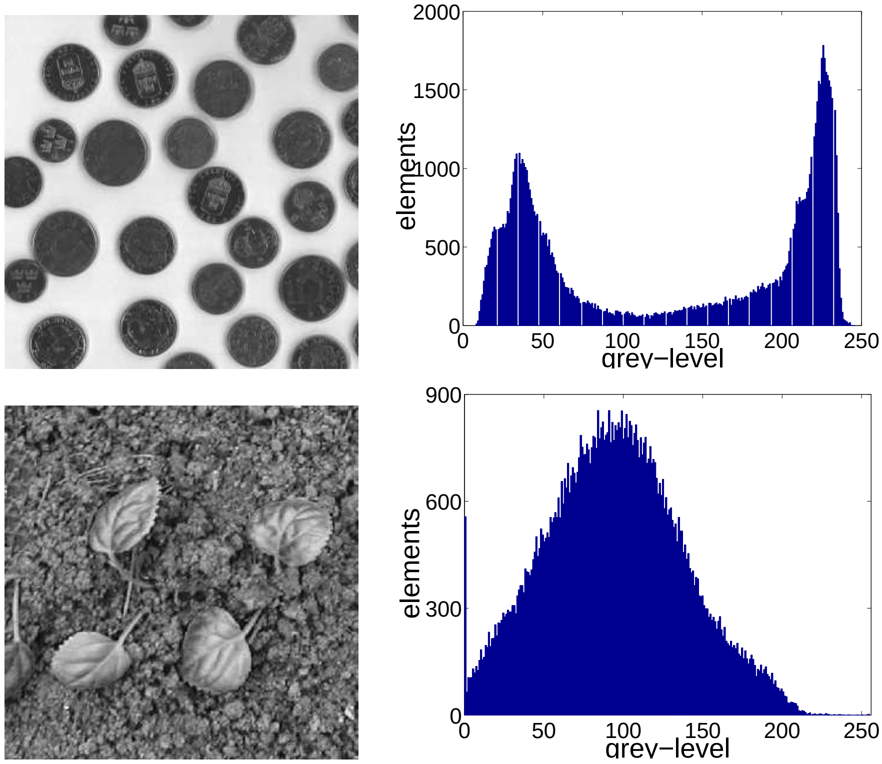


Figure 5: Images and corresponding grey-level histograms; coins (top), and leaves (bottom).

Watershed segmentation

A segmentation algorithm known as watershed, WS, was presented as a general contour detection algorithm by Beucher and Lantuéjoul (1979). The method can be applied to different kinds of image information, such as grey-level, distance or gradient magnitude (local contrast) information, to divide the image content into regions. Since the late seventies, the method has been used in a variety of different applications, see Meyer and Beucher (1990) and Vincent (1993) for overviews. A brief description of the method, applied to a 2D image, is given below, although it works similarly in other dimensionalities. The WS segmentation algorithm is easily understood by interpreting the intensity image as a landscape. High values (bright regions) in the image correspond to mountains or hills in the landscape, and low values (dark regions) correspond to lakes or valleys. Imagine drilling a small hole in every local hollow and slowly submerging the landscape in water. The deepest lakes or valleys will start to fill, and as the water level rises, they will eventually merge with other lakes or valleys. At places where different basins of water are about to meet, a watershed, or dam, is built. This inhibits water collections belonging to different local hollows to merge. When the whole landscape has been submerged in water, all image elements will have been assigned to a basin of water originating from a local hollow. Either the watersheds or the basins are referred to as the WS segmentation of the image. The latter is used in this thesis. In cases where separated mountains or hills are the desired output of the segmentation process, the image is inverted, i.e., the landscape is turned up-side-down, prior to applying the WS algorithm.

Depending on the type of information the WS algorithm is applied to, different regionalizations of the image can be achieved. One way is to apply the algorithm directly to the intensity information. Consider an image of tightly packed cells, see Figure 6(a). Since the objects in the image are bright, the image is inverted, Figure 6(b), prior to application of the WS algorithm. As the algorithm is explained above, a regionalization which completely covers the image will be created. If it instead is run only to a certain intensity, known a priori to be the background, it can create regions corresponding to well defined objects, as in Figure 6(c). The problem of deciding a good background level is, however, often as hard as it is to find a good thresholding level.

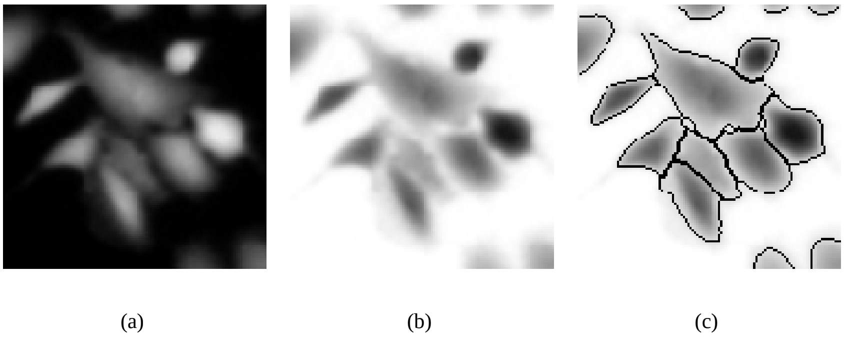


Figure 6: The watershed algorithm applied to intensity information. The original image (a), and inverted (b). The result,(c), of the algorithm applied to (b).

The WS algorithm is commonly applied to gradient magnitude information. Consider the image of cell nuclei in Figure 7(a). A gradient magnitude image can be calculated by filtering the image with one, or a set of local contrast detecting filters, see e.g. (Sonka et al., 1999), for a number of commonly used filters. Figure 7(b), is calculated by filtering the image with a set of Sobel filters which detects edges in different directions, and combining the filter responses in each position. Dark regions in the gradient magnitude image correspond to regions in the original image with low local contrast, whereas bright regions correspond to regions with high local contrast, i.e., sharp edges. When the WS algorithm is applied to the image in Figure 7(b), the result is, a regionalization where the watersheds will be situated where the gradient magnitude information is locally highest, see Figure 7(c). If the gradient magnitude image is interpreted as a landscape the borders will be placed along the crest line of the mountain range separating two valleys.

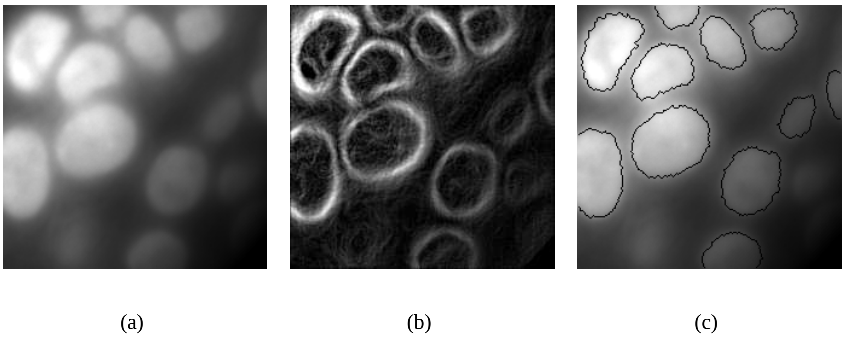


Figure 7: The watershed algorithm applied to gradient magnitude information. The original image (a), and the gradient magnitude image (b). The result, (c), of the algorithm applied to (b) and overlayed on (a).

Often, the regionalization resulting from the WS algorithm contains far more regions than desired. This is a problem often referred to as "over-segmentation", and can be handled in two different ways. One is to preprocess the image, e.g., by smoothing, in order to reduce the number of local minima, each giving rise to one region in the WS segmentation. The other is to merge the resulting regions into more significant ones, thereby obtaining the desired output. Most often a combination of both methods, adapted to the problem at hand, is needed to acquire a good segmentation result. Different criteria can be used to merge neighboring regions. One criterion is to set a threshold for the lowest allowed height, h, of the lowest point along the watershed between two regions, measured from the bottom of the more shallow region, see Figure 8. This means that the weakest point in the border decides whether two regions should be merged or not. Since this merging criterion only depends on one single image element, it is rather noise sensitive. A more robust method is to merge regions based on statistics from the whole separating border, e.g., the mean or median border value. Note that using statistics from the whole border is more computationally expensive than using only one point along the border. To use the median value is more computationally expensive than using the mean, as more values need to be kept during the complete merging process, in order to update border values correctly. If the mean value is used, only the number of border elements, together with the sum of the intensity values, need to be kept to enable a correct updating of the border during the merging. Merging does not necessarily need to be decided from intensity measures. It can also be based on regional measures such as size of a region, length of the border, or shape.

Instead of affecting all information in the image by smoothing the image to remove irrelevant local minima or by treating all WS regions as equally important when merging, seeds can be planted, automatically or manually, marking regions of interest. After seeding, the WS algorithm is run as described above with the exception that watersheds are only built where two seeded regions meet, (Meyer and Beucher, 1990). After a seeded WS there will, hence, be as many re-

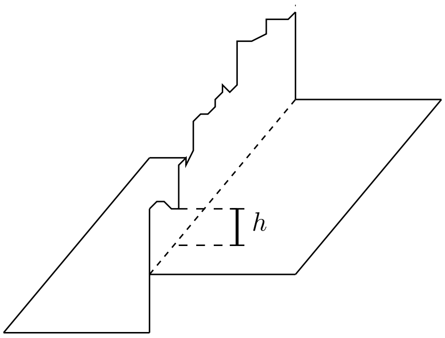


Figure 8: Illustration of the merging criterion where regions are merged based on the height, h, of the lowest point along the separating border, measured from the bottom of the more shallow region.

gions as there were seeds to start from. Seeding can serve as a useful way of combining different types of information. An example of this is illustrated on a small image of cell nuclei, Figure 9, where object seeds are found as local maxima in the original intensity image, and background seeds are local minima in the gradient magnitude image. The seeded WS algorithm is thereafter run on the gradient magnitude information. Naturally, also a seeded WS might need to be combined with preprocessing and/or a merging algorithm. The strategy of combining intensity and edge information in a seeded WS is used in the method developed in Papers V and VIII, and is further discussed in Sections 4.4 and 4.5.

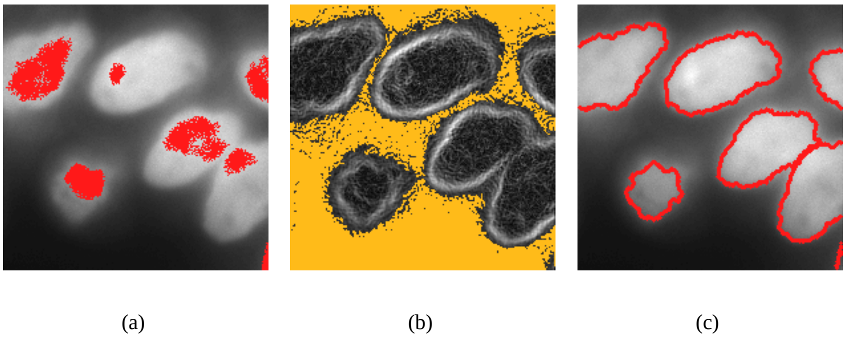


Figure 9: The seeded watershed algorithm applied to gradient magnitude information. The original image and object seeds (a), and the gradient magnitude image and the background seed (b). The result, (c), of the algorithm applied to (b) and overlayed on (a).

Once a segmentation of the image into objects and background has been performed, the WS algorithm can also be used to divide objects into smaller parts. This is usually desirable, since objects often are attached to each other in clusters. A WS algorithm applied to a distance transform of the cluster can serve as an efficient way of dividing the cluster into single objects, see Figure 10. This way of using the WS algorithm as a cluster separating method was used in Papers III and V, see Sections 4.2 and 4.4.

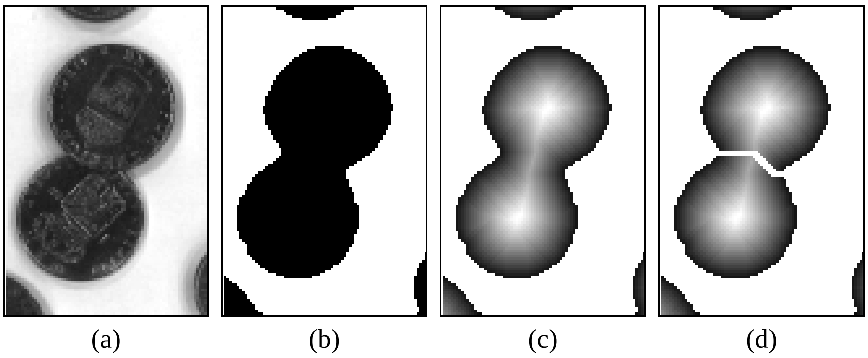


Figure 10: The watershed algorithm applied to a distance transform. An image of two overlapping coins (a), the regionalization after thresholding (b), a DT of the coins (c), the resulting regions after the watershed algorithm applied to the distance transform (d).

Yet another way of using the WS algorithm is to find suitable threshold levels for an image. If a grey-level histogram of an image is seen as a 1D image, the WS algorithm applied to the inverted histogram image, possibly with subsequent merging, can produce good threshold values for the original image. In Figure 11, an example of this way of finding two good threshold levels for a light microscopy image of a cell culture is shown. The peak furthest to the right in the histogram corresponds to the background, the central peak corresponds to healthy cells, and the peak to the left corresponds to dead cells due to lack of nutrition in the central part of the culture. WS segmentation of the inverted histogram detects good and image dependent threshold levels. This can be useful for segmenting objects in similar images but with varying contrast or overall brightness. Often more than one threshold level is desired, and as long as the peaks in the histogram correspond to the different regions of interest and are separated by the deepest valleys, WS based thresholding will supply the desired threshold levels.

Template matching

If specific structures in an image are sought, that can be distinguished by a certain shape and/or grey-level distribution, global approaches such as thresholding or WS segmentation, are often unsuccessful. An alternative way for finding the objects of interest is to locally compare, or match, the image information with a template describing the searched structure. Depending on what object feature the template represents, the similarity measure and matching procedure, i.e., how the image is searched, are usually somewhat different.

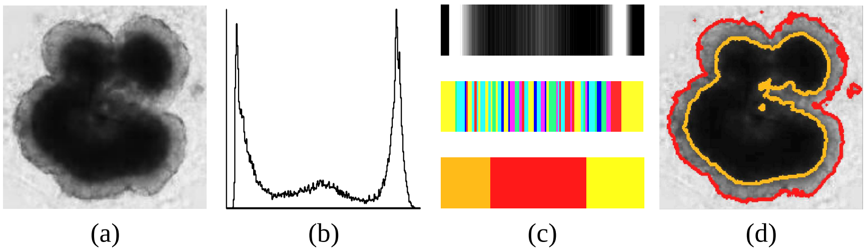


Figure 11: Using the watershed algorithm to automatically find grey-level thresholds. The original image (a). The grey-level histogram of the image (b). The histogram displayed as a 1D grey-level image (c) (top). The resulting regions of a watershed algorithm applied to the 1D image (c) (middle). The resulting regions after merging, using a border height criterion, until only three regions remain (c) (bottom). Borders of the resulting regions, (d), after thresholding the original image with the values 60 and 184, found as the values at the watersheds in (c) (bottom).

In cases where the template represents an intensity distribution, the similarity measure needs to weigh together the individual similarity between each position in the template and the its underlying image element. This is done by moving the template over the image and at each position measuring the similarity between the template and the subimage underneath. see Figure 12. The result from this matching is, hence, an image where the intensity of an element represents the similarity between the subimage and the template centered at that element. A common and simple similarity measure between the template and the subimage, is correlation (Gonzalez and Woods, 2002), which in its simplest form is the sum of every template element intensity multiplied by the intensity of the subimage element underneath it. Matching using this similarity measure can be performed very efficiently in the frequency domain. The disadvantage with this similarity measure is its sensitivity to intensity changes in the image. To compensate for this, the correlation coefficient can be used instead. The correlation coefficient, CC, at a position in the image is calculated as

$$CC = \frac{N\sum t_i s_i - \sum t_i \sum s_i}{\sqrt{N\sum t_i^2 - (\sum t_i)^2}\sqrt{N\sum s_i^2 - (\sum s_i)^2}}, \tag{1}$$

where N denotes the total number of elements in the template, and $i = 1...N$. The elements grey-levels in the template and subimage are denoted t_i and s_i, respectively. Matching using this similarity measure is usually performed in the spatial domain, as the formula for the CC is difficult to translate to the frequency domain. This approach for segmenting objects with a certain intensity pattern has been used in Paper IV, see Section 4.3.

A simpler measure, to compare the template and the subimage underneath it, can be used when the template instead represents a shape. The contour in 2D, and surface in 3D, of the shape then serves as the template, and the sum of values underneath the template, can be used as a similarity measure. If the image can be

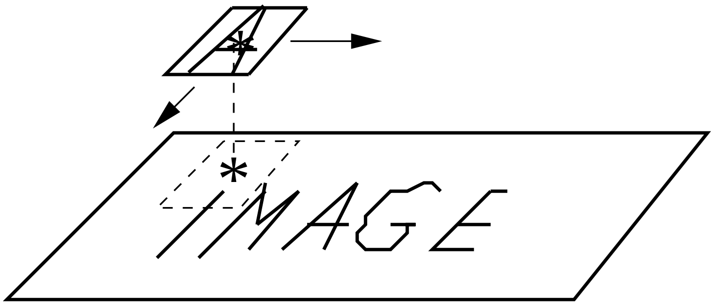


Figure 12: A template, centered at *, is moved over the image and at each position compared to the subimage underneath.

transformed so that the match score decreases (or increases) the closer to a perfect match the template is moved, a search algorithm can be used to speed up the matching procedure. This is the case for the chamfer matching algorithm first proposed by Barrow et al. (1977). The general idea is to calculate a DT from the contours of the objects in the image and then perform the matching in the distance image. The sum of the distance values hit by the template serves as the similarity measure. The lower the sum is, the better is the match. How the sum changes when the template is moved a small step in different directions is the key to finding the local best fits. Directions resulting in higher sums can be discarded, and only the directions giving lower sums should be further investigated. Note that a small step does not generally only mean translation, but also rotation, and in some cases scaling needs to be considered as well. Finding the best local matches is, hence, equivalent to finding local minima in an n-dimensional function, where n denotes the number of allowed geometrical transformations of the template. Since the search from a starting position may end in the closest local minimum (of course, depending on the search algorithm used), many starting positions are needed in order to find all objects of interest. For real images, this optimization problem is very large and therefore the original method is not suitable. In Borgefors (1988), a very important improvement to the method made it suitable for problems where no information about the approximate positions of the objects is available, a hierarchical chamfer matching algorithm (HCMA) was presented. The original chamfer matching algorithm is embedded in a resolution hierarchy which greatly reduces the computational cost of the matching, while also reducing the sensitivity to noise. A resolution pyramid of the contour image is calculated, see Figure 13, and the chamfer matching is started at the lowest resolution level. The good match positions are then used as starting positions when the matching is moved to the next resolution level. Another important improvement, (Borgefors, 1984b), is using the sum of the squared distance values hit by the template instead of just the sum. This similarity measure is better as it introduces fewer false minima. The HCMA algorithm, extended to 3D, is used for segmenting proteins in volume images in Paper VIII.

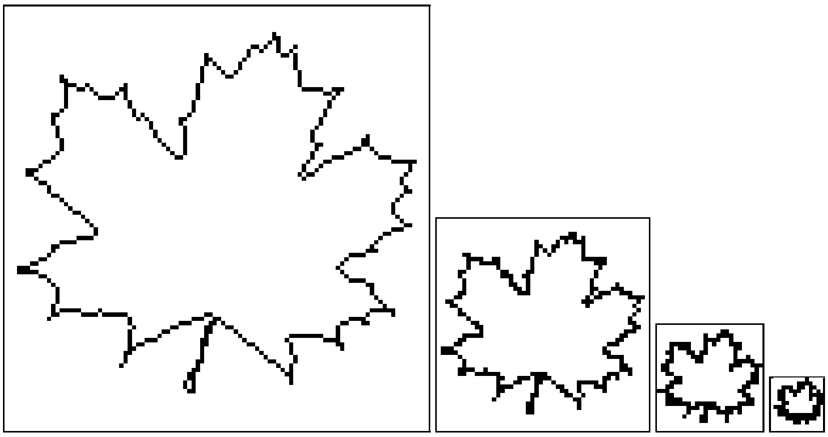


Figure 13: A resolution pyramid of the contours of a maple leaf. The resolution is halved between successive resolution levels.

3.3 Object representations

It is often useful to study the shape of a segmented object using different representation schemes, in order to facilitate recognition and analysis tasks. The aim is to represent an often complex object by a simpler structure, more easily analyzed. Two intuitive representations of an object will be explained here. The first is to represent the object by (simpler) subparts, and the second is to represent the object by a curve centrally located in the object.

Decomposing an object into subparts facilitates analysis as each part can be analyzed individually, as well as how the different parts are connected to compose the entire object. In Svensson and Sanniti di Baja (2002), a 3D method is presented which decomposes a binary object into simpler parts based on its shape. The method consists of three main steps, which all use information from a DT of the object. Seeds for the different subparts are identified in the DT of the object. Balls are then grown from each seed. How much each seed grows depends on its distance value. The remaining elements of the object are then each set to belong to the closest ball. A final merging process, used to remove non-significant parts that do not have much outer surface in the object, is often necessary. The identified seeds and the resulting decomposition of a binary "man" are shown in Figures 14(a) and 14(b). This method serves as the basis for the decomposition method in Paper VI, further described in Section 4.5, which incorporates intensity information in the decomposition.

As briefly mentioned at the end of Section 3.2, the WS algorithm can be used to decompose a region into subparts. The WS algorithm can be applied to the DT of an object resulting in a decomposition similar to the method described above. The difference is mainly that the WS will not necessarily result in dividing lines at positions where the surface bends. This is because the WS will only grow from local maxima, and elongated parts will therefore likely become part of a larger neighboring region. In Figure 14(c) the same object as in Figure 14(b) is decomposed with

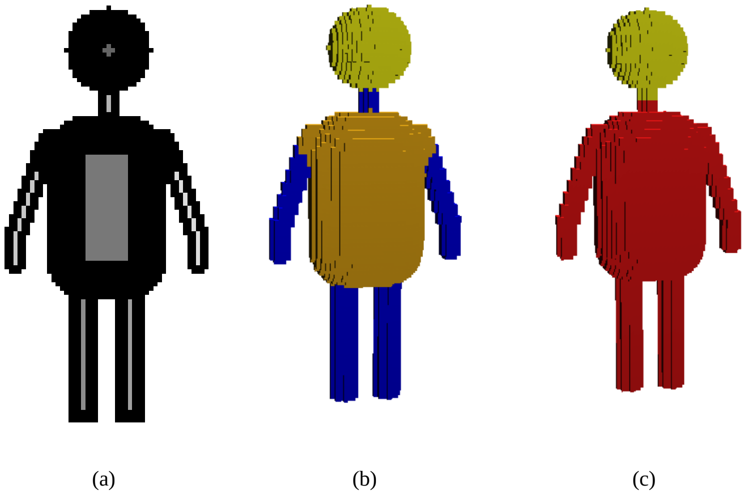


Figure 14: A cross section of a 3D object displaying the seeds identified in the binary decomposition method (a), and the resulting decomposition (b). The result of a watershed algorithm applied to a distance transform of the object (c).

the WS algorithm applied to a distance transform of the object. The WS algorithm can also be used to decompose a segmented object based on internal intensity information. The decomposition will then be strictly guided by the intensity distribution within the object, regardless of the shape of the object.

Instead of analyzing the entire object or its parts, a medial representation, such as its skeleton, can be useful. Analyzing the medial representation instead of the original object has proven successful in several 3D medical applications, see e.g. Nyström and Smedby (2001) and Fouard et al. (2004). Desirable features of a medial representation are that it is thin, centered within, and topologically equivalent to the original object. Topologically equivalent means that the number of holes in 2D, and the number of cavities, and the number of tunnels in 3D, should be the same for each connected component in original object and in the representation. For a 2D object, a medial representation possessing these features is a curve, while for 3D objects it can either be a surface, or further reduced to a curve, see Figure 15. For binary objects, the curve in 2D, or the surface in 3D, together with the distance from the background to each element in the representation, provides enough information to recover the original object, with an accuracy to the outermost layer. Therefore, the representation serves as an efficient way of storing binary objects. There are different approaches for calculating a medial representation of a binary object. A distance based approach can be performed by iteratively thinning the DT of an object. The elements in the outer distance layer are iteratively removed if they are not CMDs or CMBs, or needed for shape or topology preservation, (Svensson et al., 1999; Svensson, 2002). A more efficient 2D distance based method, using a path

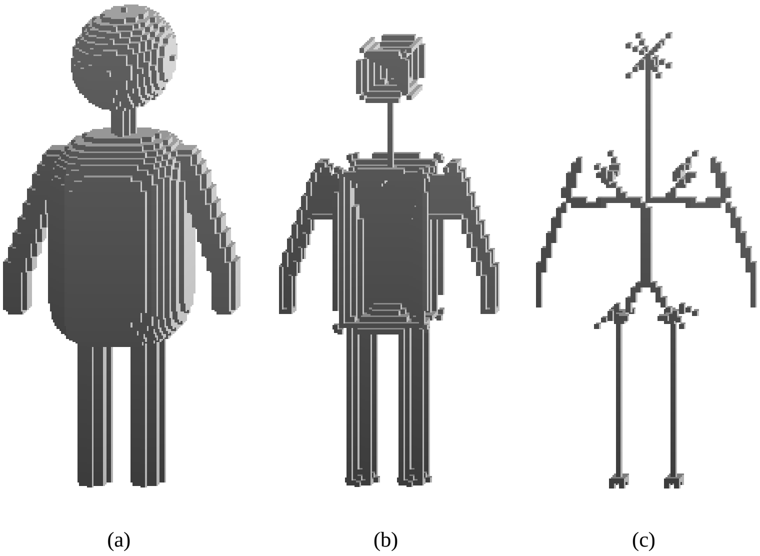


Figure 15: A surface skeleton (b), and a curve skeleton (c), of the 3D binary man (a).

growing approach, is found in (Sanniti di Baja, 1994).

The idea of representing a 3D shape by a centralized curve, to simplify analysis, can be transferred to the case of grey-level values inside the object, (Svensson et al., 2002). Instead of finding the most central curve, a curve following the maximum intensity is searched for, similar to finding crest lines in a mountain range. This can be achieved by thinning the shape based on intensity, under the constraint that the grey-level topology is not changed. This means, e.g., that a region within a 3D object with strictly lower grey-values than what is surrounding it will result in a closed surface in the grey-level medial representation. The idea by Svensson et al. (2002), of incorporating grey-level information in a medial representation was modified and used in Paper VII, and is, hence, further discussed in Section 4.5.

4 Contributions

In this Section, the methods and results thoroughly described in Papers I−VIII, are briefly presented. How distance transforms should be adopted to elongated 2D and 3D grids, Paper I and II, is described in Section 4.1. This is used in Paper III, described in Section 4.2, to segment individual pores in 3D SEM images of paper. Section 4.3, corresponding to Paper IV, presents how circular symmetric viral capsids can be described and identified in 2D TEM images based on distance and intensity information. Section 4.4, presents how different types of information are combined in Paper V to segment fluorescence labeled cell nuclei in 2D and 3D images. Finally, Section 4.5, describes two different representation schemes, Paper VI and VII, and a match based segmentation method, Paper VIII, for proteins imaged using SET.

4.1 Adapting distance transforms to elongated grids

In Section 3.1 the concept of a DT, and how to compute it by using a distance transformation, was explained. The local steps and weights were all given for images with square or cubic image elements. All images are, however, not sampled in such grids, especially not 3D images, where many tomographic and microscopic images have longer voxels in the z-direction. In 2D, elongated pixels are sometimes encountered in images produced with a line scanner or in satellite images. It is of interest to adapt methods to these elongated grids in order to avoid interpolation to cubic grids, as this makes the images larger without adding any information. In Papers I, and II, weighted DTs, WDTs, in images with elongated pixels and voxels are examined. This has also been investigated in 2D by Coquin and Bolon (1995); Bolon et al. (1992) and in 3D by Coquin et al. (1994); Fouard and Malandain (2005). Pixels in 2D images have a height equal to 1 and width equal to $\Lambda \geq 1$, and voxels in 3D images have height and width equal to 1 and length $\Lambda \geq 1$. An elongated pixel has three kinds of neighbors, in the following denoted by capital letters, while an elongated voxel has five different neighbors, denoted by lower case letters, see Figure 16. Optimal weights are calculated for local steps to the neighbors, i.e., to image elements in a 3×3, and $3 \times 3 \times 3$ neighborhood, for 2D and 3D, respectively.

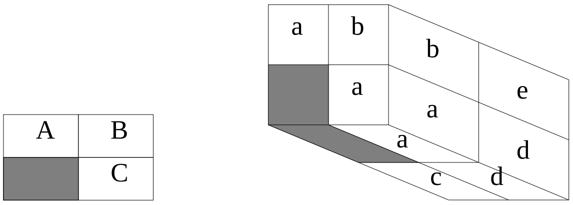


Figure 16: The different types of local steps for images digitized in an elongated 2D grid (left), and 3D grid (right).

The optimal weights for the local steps in the neighborhood were calculated by minimizing the maximum error between the WDT and the digital EDT, which is as rotation independent as possible, in square and cubic grids of sizes $\Lambda\ M\ \times\ M$ and $\Lambda\ M\ \times M\ \times\ M$, respectively. The optimal weights for the local steps are all dependent on Λ. The weights need to fulfill the *semi-regularity* criterion, which states that if two elements can be connected by only one type and direction of local step, this path should be the shortest, or equal to the shortest path between the two elements. In 2D, see Figure 16 (left), this leads to the following constraints:

$$A \leq B \quad \text{and} \quad C \leq B \quad \text{and} \quad B \leq A + C \ \ \forall \Lambda. \tag{2}$$

In 3D, see Figure 16 (right), the constraints implied by the *semi-regularity* criterion become:

$$a \leq b,\ \ a \leq c,\ \ b \leq d,\ \ c \leq d,\ \ d \leq e,\ \ 2e \leq b + 2d,\ \ b \leq 2a,\ \ d \leq a + c. \tag{3}$$

These constraints are not enough to generate unique expressions for the local distances in 3D. The points $(1,2,1)$ and $(1,1,2)$ can be reached from the origin in two ways, each fulfilling the constraints above. This gives rise to four different combinations of additional constraints. The combination chosen to give unique expressions for the local steps was:

$$e + a \leq d + b \quad \text{and} \quad e + c \leq 2d. \tag{4}$$

Due to symmetry, the optimization only needs to be performed in one quadrant in 2D, and one half of an octant in 3D. The maximal difference to the EDT is assumed to occur on the border of the image, The expressions for the differences between the WDT and the EDT will be different along the border according to the combination of steps needed to get there. To optimize the WDT, the local steps which minimize the maximum of these expressions is found by solving the system of equations the expressions give rise to. In 2D, the expressions for the local steps and the maximum difference to the Euclidean difference (scaled with M), denoted maxdiff are:

$$\begin{aligned}
&A(\Lambda) = \frac{3\Lambda^2 - \Lambda - (\Lambda - 1)\sqrt{\Lambda^2 + 1}}{5\Lambda^2 - 2\Lambda + 1} + \\
&\frac{2\Lambda\sqrt{(\Lambda + 1)(\sqrt{\Lambda^2 + 1} - 1)}}{5\Lambda^2 - 2\Lambda + 1}, \\
&C(\Lambda) = \Lambda A(\Lambda), \\
&B(\Lambda) = \Lambda A(\Lambda) + \sqrt{\Lambda^2 + 1} - \Lambda, \\
&\text{maxdiff}(\Lambda) = \mid \Lambda - \Lambda A(\Lambda) \mid .
\end{aligned} \tag{5}$$

The corresponding expressions in 3D are:

$$
\begin{aligned}
&a(\Lambda) = \frac{3\Lambda^2 + \Lambda(\sqrt{2} - 2 - \sqrt{\Lambda^2+2}) + \sqrt{\Lambda^2+2} + 1 - \sqrt{2}}{5\Lambda^2 - 2\Lambda + 1} + \\
&\frac{\sqrt{\Lambda^2(4\sqrt{\Lambda^2+2}(2\sqrt{2} - 1 + \Lambda) + 5\Lambda^2(2\sqrt{2} - 3) + 2\Lambda(3 - 4\sqrt{2}) + 6\sqrt{2} - 19)}}{5\Lambda^2 - 2\Lambda + 1}, \\
&b(\Lambda) = a(\Lambda) - 1 + \sqrt{2}, \\
&c(\Lambda) = \Lambda a(\Lambda), \\
&d(\Lambda) = \Lambda a(\Lambda) - \Lambda + \sqrt{1 + \Lambda^2}, \\
&e(\Lambda) = \Lambda a(\Lambda) - \Lambda + \sqrt{\Lambda^2 + 2}, \\
\\
&\text{maxdiff}(\Lambda) = \mid \Lambda - \Lambda a(\Lambda) \mid .
\end{aligned}
\tag{6}
$$

All local steps can be derived from one of the local steps. Here, the local steps are expressed as functions of the A, or a steps. The reason that all steps can be expressed as a function of one step, is that when optimizing on a square or cube, the error on the border of the cube, for all directions of local steps, should be equal. The WDT for elongated image elements underestimates the distances along the directions of the local steps and overestimates the distances in directions that are not well covered by the local distances.

Integer local steps are often used and they are acquired by multiplying the local steps with a scale factor and rounding off to the nearest integer. To make sure that the integer local steps are the best possible for each scale factor, an integer neighborhood for each rounded integer local step is examined. Since the best local steps depend on Λ, the local step equations are the main result of this work. The optimal and best integer local steps for 2D and 3D WDTs when Λ equals 3 is given in Tables 1 and 2, respectively. A good way two visualize the difference to the EDT is to show digital circles, 2D, or spheres, 3D, computed with the WDTs. The digital circle and sphere for $\Lambda = 3$ are shown i Figure 17.

Table 1: Local steps in 2D for $\Lambda = 3$.

scale factor	A	C	B	maxdiff
1	0.883	2.649	2.811	0.351
1.085	1	3	3	0.398
3.454	3	9	10	0.394
4.563	4	12	13	0.370
5.672	5	15	16	0.356
6.768	6	18	19	0.355
12.452	11	33	35	0.352
.	.	.	.	.
∞	A_{opt}	C_{opt}	B_{opt}	0.351

Note that the errors are very large for a Λ of 3. The WDTs for elongated grids are rather rotation dependent, see Figure 17. The larger Λ is, the more rotation de-

Table 2: Local steps in 3D for $\Lambda = 3$.

scale factor	a	b	c	d	e	maxdiff
1	0.865	1.279	2.595	2.757	2.911	0.406
1.940	2	3	5	5	5	0.739
2.347	2	3	6	7	7	0.444
6.909	6	9	18	19	20	0.422
9.281	8	12	24	26	27	0.414
17.205	15	22	45	48	50	0.411
19.570	17	25	51	54	57	0.410
.	.	.	.	.	.	
∞	a_{opt}	b_{opt}	c_{opt}	d_{opt}	e_{opt}	0.406

pendent the WDT becomes. To extend the neighborhood would reduce the error and the rotation dependency, but this would of course have the side effect of increasing the computational cost. Since the elements are elongated, an asymmetric neighborhood, 3×5 in 2D and $5 \times 5 \times 3$ in 3D, would be a good way to compensate for the largest errors without increasing the computational costs as much as with full 5×5 and $5 \times 5 \times 5$ neighborhoods. Weights for 3D WDTs in such asymmetric neighborhoods have been optimized over spheres in Coquin et al. (1994).

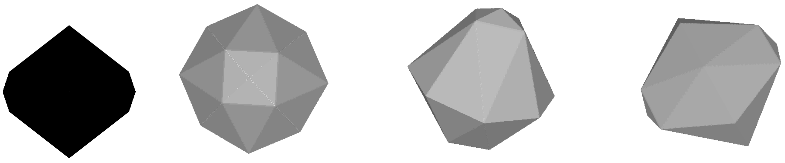


Figure 17: The disc and ball, 3 views, for $\Lambda = 3$, computed using the weights 3, 9, and 10 in 2D and 2, 3, 6, 7, and 7 in 3D.

4.2 Segmentation of paper pores

In Paper III, a method which segments individual pores in a paper volume image having elongated voxels is described. Information about the void space, the pores between the fiber network, is of interest for paper quality. It affects how the paper interacts with light and fluids, as well as the overall quality of the paper. Using 3D image analysis is the only way to extract individual pores for further feature extraction. The paper volume used to demonstrate the method in Paper III is a piece of milk carton, embedded in epoxy, and digitized as a stack of 2D images using SEM. Between successive 2D image acquisitions, the top part of the sample was cut off using a microtome. The voxels in the final aligned 3D volume are approximately seven times longer in the z-direction than in the x- and y-direction and correspond to true voxel sizes of $0.7 \times 0.7 \times 5.0 \mu m$. The volume was binarized into fibre network and void space using shading correction and grey-level thresholding, (Aronsson et al., 2002).

The algorithm for segmenting individual pores, can be divided into three steps: identification of the top and bottom surface of the paper; distance based WS segmentation of the void space; and merging of neighboring regions. All steps of the method use distance information and since the voxels in the volume are elongated, a WDT with weights adopted to the voxel size is used, see Section 4.1. Good integer approximations for the weights were calculated using the method in Paper II.

Prior to segmenting individual pores, the top and bottom surfaces of the paper need to be identified. This is performed using the rolling ball algorithm as described for this application in Svensson and Aronsson (2003). There, distance information is used to simulate a rolling ball with a user-provided radius. The paper surface is defined as being the points the surface of the ball traces. The remaining steps of the segmentation method are illustrated using a small part of a 2D slice from the volume, see Figure 18. A WDT is calculated for the void space, see Figure 18(b), where brighter values correspond to larger distance values. The WS algorithm is then applied to the distance information yielding as many regions as there were local maxima in the DT, see Figure 18(c). Merging of regions is necessary to reduce the over-segmentation to regions corresponding to individual pores. This merging is performed in two steps. First, all neighboring regions with a maximum distance value on the border, only slightly smaller than the largest distance value in one of the regions, are merged. Second, all neighboring regions with a maximum distance value on the border, larger than a certain size, are merged. The result after each of the two merging steps are shown in Figures 18(d) and 18(e), respectively. During the first merging step, regions separated due to a small contraction on the outer surface are merged, and small regions are merged to other small, or large, neighboring regions. During the second merging step, regions sharing a large border surface are merged. The first criterion is dependent on the sizes of the regions and the border between them, while the second is a fixed-size threshold.

In Figure 19, surface renderings of five individual pores identified in the volume are shown. The last slices of the fibre network are shown behind the pore regions to give a feeling for the complexity of the segmentation task. When the individual pores have been segmented features can easily be extracted and used to gather information about the paper structure. In Paper III, a few features are presented for the five pores visualized in Figure 19, as well as some scatter plots of features for all segmented pores.

4.3 Description and segmentation of viral capsids

When designing segmentation algorithms, it is important to incorporate as much information as possible about the objects to be found. If only objects of a certain kind are sought, a more specific segmentation, such as matching, is generally a good choice. In Paper IV, a matching method used to segment and classify three types of viral capsids in 2D is described. The capsids are imaged in host cell nuclei using TEM. This results in images with a severely textured background as can be seen in Figure 20. Matching based on intensity is computationally expensive. Therefore, less computationally demanding segmentation techniques, such as edge based segmentation methods and ring filters, were tried. They were, however, not successful

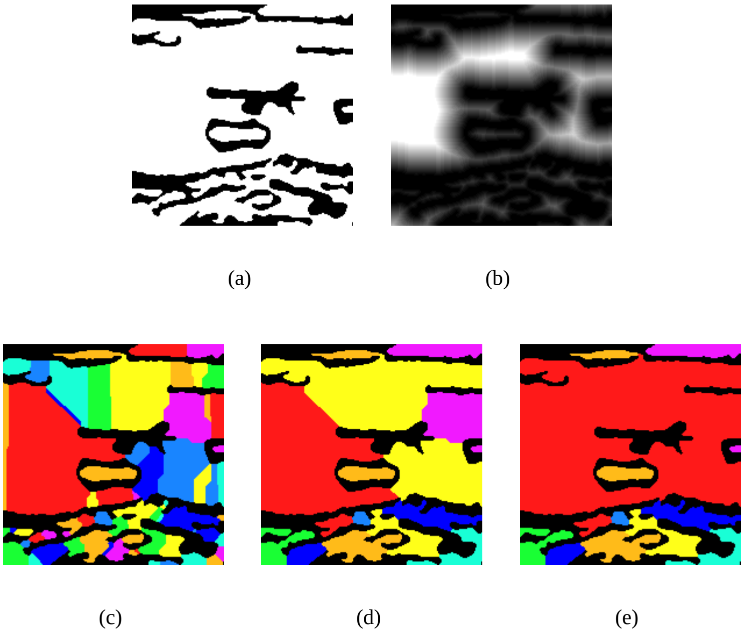


Figure 18: The pore segmentation algorithm illustrated on a part of a 2D slice from the volume. The original binary image (a), a distance transform of the void space (b). The result of watershed algorithm (c), when applied to (b). The result after each of the two merging steps (d) and e), respectively.

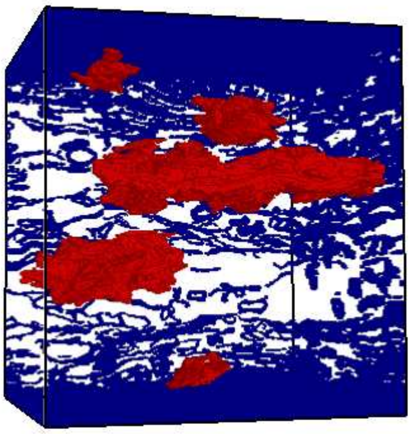


Figure 19: A surface rendering of five individual pores identified by the method described in Paper III.

on these images due to the texture in the background.

The three capsid types are all radially symmetric, but their radial density, or grey-level, distribution differ, see Figure 20. The radial grey-level distribution for each type is used to create a unique description, or signature, which, in turn, is used to create a template for the matching method.

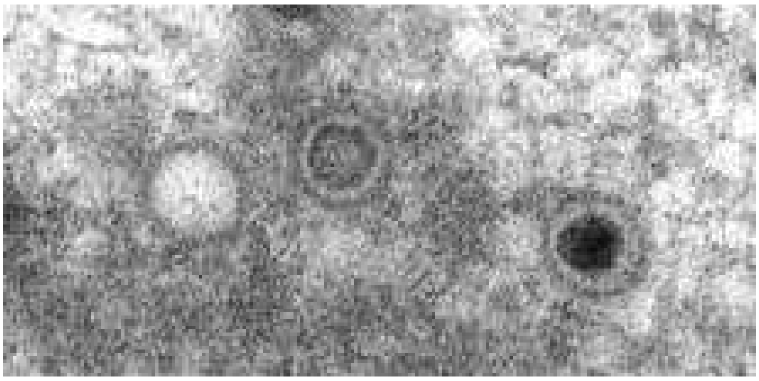


Figure 20: Part of a TEM image showing the three different capsid types.

A circular symmetric object in an image can be described by a 1D function of r, where the value at position r represents the grey-value at distance r from the center of the object. The objects in consideration here are approximately circular and can therefore be described by such a 1D function. The value at each r is calculated as the mean of all values at distance r from the center. This representation of a circular object will be referred to as the density profile of an object. A profile representing each class is created by averaging eight manually picked capsids of each class after size and grey-level normalization. A template for each class is created from the density profiles by doing the opposite to a profile calculation, i.e., all pixels at distance r from the center in the template image are assigned the value at position r in the profile. Density profiles along with their corresponding templates are shown in Figure 21.

The templates are used one at a time to calculate three correlation images, where the value in each position shows the similarity between the underlying image and the template centered at that position. The similarity measure is the correlation coefficient (CC) and the matching is performed in the spatial domain, see Section 3.2. The positions with the highest percentage of CCs for each template were matched to templates of smaller and larger sizes. The size of the template that produced the best CC, along with the CC itself, were kept for further analysis and classification. Starting from the highest CC for each class and moving on to lower values, the circular subimage giving rise to the score was transformed into polar coordinates and placed as rows in a matrix. Each row was then aligned to the class profile and the matrix was matched against a similar matrix produced from the class profile. To classify the subimage as a true capsid, the CC for the aligned matrix needs to be

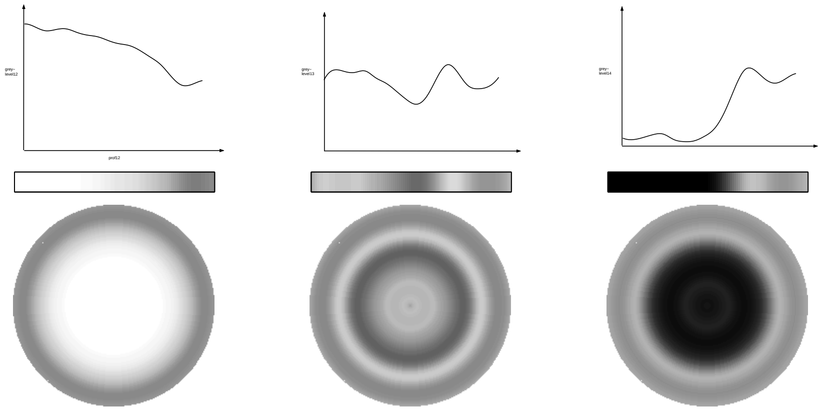


Figure 21: Density profiles as functions and images, and corresponding templates for the three different capsid types.

higher than a threshold, $T1$. This threshold is set as a percentage of the highest CC in the image. In addition, the CC for the circular subimage needs to be higher than a threshold, $T2$, set as a percentage of $T1$. For one of the capsid types, the middle type in Figure 20, the variance of the sums of the rows in the aligned matrix also needs to be lower than a threshold. These analysis steps are performed for positions of each class until a number of positions not fulfilling the criteria are found.

This segmentation and classification method was applied to micrographs from three different preparations. The results on two images, from which the eight capsids used to construct the three class profiles were picked, are shown in Figure 22. The method was also applied to 10 micrographs not used for developing the method. The result was compared to visual classification performed by an expert. In total 185 (82%) of the 226 capsids were correctly classified, 41 (18%) were missed, and 38 (17%) false objects were found. The results are very satisfactory considering the severely textured background and the low contrast between objects and background. Using the two thresholds, T1 and T2, allows for excluding false capsids as they are generally not as radially symmetric as true ones. It also allows for including true capsids with lower original match scores due to not perfectly circular shape or a somewhat distorted interior. As the method is described, it needs very little human interaction. The thresholds can be tuned for a small image and then used for all images produced under the same condition.

These results can be further improved if a human expert would make the final decision. The thresholds could then be slightly lowered and the iterative examination of the best position could be run farther. The subimages classified as true capsids can be presented in a table on the screen where the human expert can mark the false or true capsids. The false positives can thereby be eliminated, and as more objects are kept, more missed objects can be found.

The density profile is a compact way to describe circular symmetric objects. Besides being a good and useful description, certain measurements can easily be

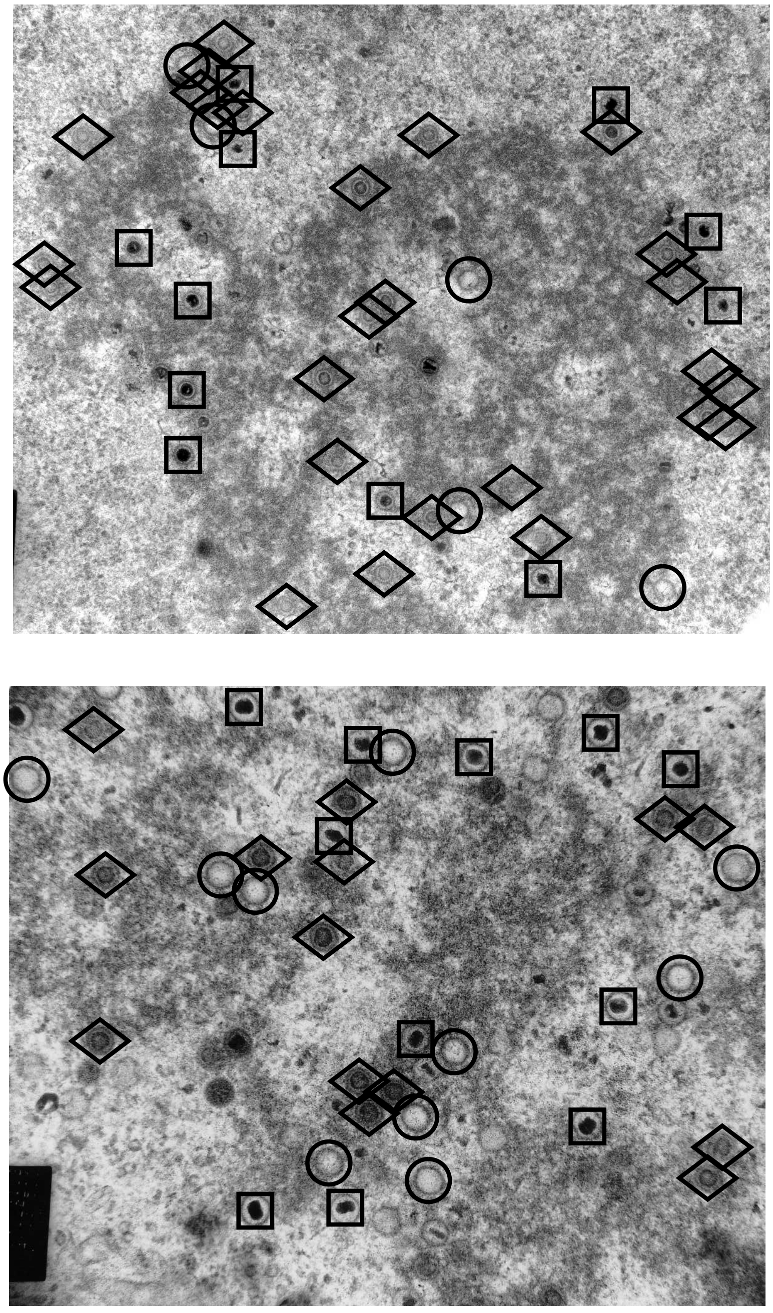


Figure 22: Result of the segmentation and classification method applied to two micrographs. Positions classified as each of the three capsid types are marked with circles, diamonds and squares, respectively.

extracted from the profiles and used to distinguish objects of different types. In the viral capsid case, it is of interest to measure the width of the outer shell. The reason is that when a viral capsid matures, proteins are attached to the shell, and it is of interest to know the thickness of this layer and where it is acquired. The width can be calculated as the length between the two zero-crossings, at an approximate radius from the center, in the second derivative of a smoothed profile, see Figure 23.

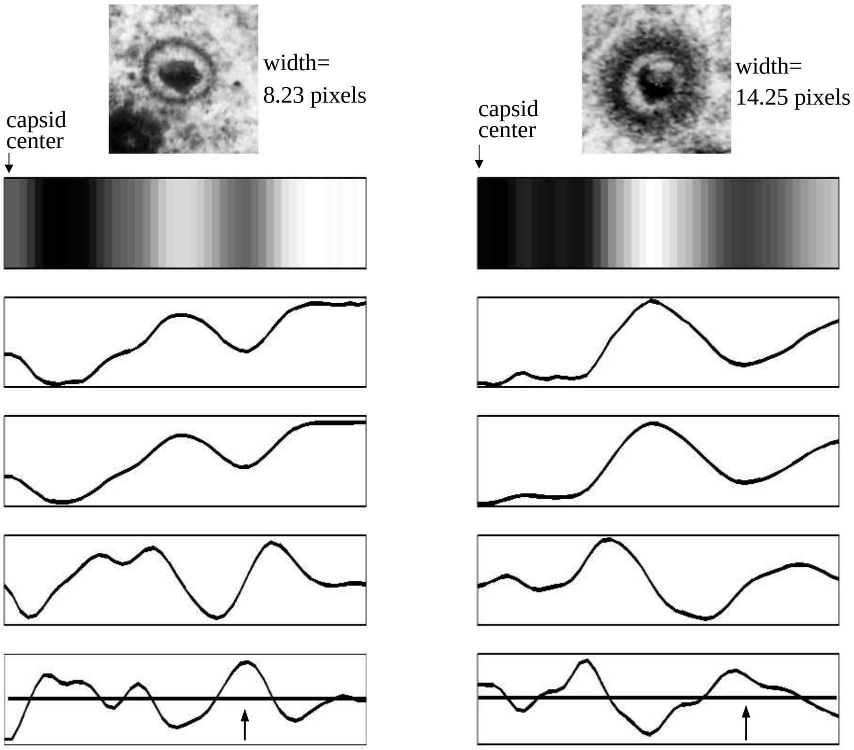


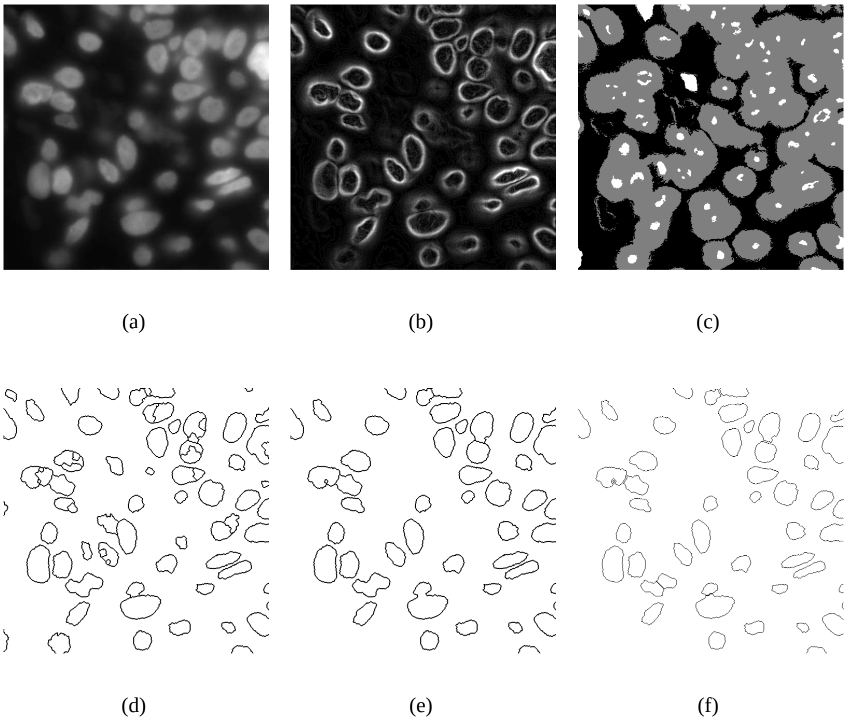
Figure 23: Calculation of the width of the outer capsid shell for a capsid without attached proteins (left), and for a capsid with attached proteins (right). From top to bottom: capsid image, density profile as image, profile as function, smoothed profile, first derivate of the smoothed profile, and second derivative of the smoothed profile. The straight line in the second derivative of the smoothed curve marks the 0 level.

4.4 Segmentation of cell nuclei

The aim of the method presented in Paper V was to develop a segmentation method which incorporates several different types of information, to be used for automatic segmentation of 2D and 3D images of fluorescence stained cell nuclei in tissue. This allows for studying cells in their natural context. The benefits of automatic segmentation compared to manual delineation of the nuclei using a mouse, are that it is more objective, easily reproduced, and faster.

Mainly three problems are encountered when fluorescence images of stained cell nuclei are studied. The background intensity is generally not flat but varies, due to autofluorescence from the tissue and fluorescence from out of focus objects. Usually, there are also intensity variations within each object and the cell nuclei are often clustered, see Figure 24(a). The two last issues complicate the task of finding borders between all nuclei, while not finding false borders within the nuclei. The problems with varying background, presence of false borders and weak true borders are encountered in many other types of imagery as well. Therefore, this method, or part of it, is likely to be useful in other situations.

The method consists of four main steps: automatic marking of region seeds; WS segmentation; merging of neighboring regions with weak borders; and splitting of regions based on a shape criterion. The different steps of the method are illustrated in a small 2D image of cell nuclei in Figure 24. The extended h-maxima transform, (Soille, 1999), was used to place seeds in local maxima higher than h, corresponding to objects or parts of objects in the intensity image. The same transform but applied to the inverted gradient magnitude image, see Figure 24(c), was used to find background seeds. This will, of course, result in background seeds also inside the objects. These are removed based on their small size. The object seeds and remaining background seeds are shown in Figure 24(c). Using local maxima in the intensity image as object seeds, and large local minima in the gradient magnitude image as background seeds, removes to a great extent the problem with varying background. The borders between the seeds were then found with the WS algorithm applied to the gradient magnitude image, see Figure 24(d). Due to the intensity variation within the objects, many objects will consist of several regions, each originating from a separate seed. These should be joined without joining regions from neighboring objects. This is a difficult problem, as borders separating different objects are not always stronger than borders between regions within an object. The assumption made in this method is that the average border between objects is stronger than the average border between regions within an object. The average gradient magnitude strengths between all regions in the image are, hence, calculated. All borders weaker than a certain threshold are then merged, see Figure 24(e), one by one, starting from the weakest border. It is important that the borders are merged in order, as the average border strength for regions affected by a merging needs to be updated. The assumption made in the previous step is not always valid for clustered objects. They are sometimes so tightly packed, that there is no clear border between them. The last step of the method uses distance based shape information to split regions that have deep concavities. A distance transform is calculated for each region, and the WS algorithm is applied to the distance information. Neighboring regions are merged if the border between them have high enough distance values, compared to the distance values inside the regions. The result of this shape based WS will be a splitting of regions that are connected by a narrow waist, see e.g. the two objects in the lower left quadrant of the result in Figure 24(f). This method will, hence, segment objects that contain at least one seed, have borders to neighboring objects that have sufficiently strong gradient magnitude average and have a reasonable cell nucleus shape.


(a)

(b)

(c)

(d)

(e)

(f)

Figure 24: Illustration of the method combining intensity, edge, and shape information, in the segmentation of cell nuclei. A small 2D image (a), and its gradient magnitude (b). The object and background seeds are shown in white and black, respectively (c). The result after a seeded WS (d), and after merging based on mean edge strength (e). The final result after splitting regions based on shape information (f).

A small smoothing filter was applied to both the 2D images and the 3D image to speed up the WS segmentation. Prior to that, the z-slices of the volume were compensated for light attenuation and the volume was subsampled to cubic voxels. To compensate for light attenuation, the mean of all pixel intensities higher than 10 was calculated for each slice. This value is plotted against slice number as a solid curve in Figure 25. A straight line was fit to the linear part of the plot, representing the slices inside the tissue sample, in a least square sense, and used to compensate for the attenuation. The intensities in a z-slice were compensated by multiplication with $m/(k \times z + m)$, where k is the slope of the line and m is the offset. The mean of all pixel intensities above 10 for each slice after the compensation is seen as the dotted curve in Figure 25. A multiplicative compensation is important to make sure that values equal to, or close to, zero are not increased. As can be seen in the plot, some slight attenuation still remains, but the contrast in all slices is now sufficient for segmentation. Each voxel in the original volume was 98nm in the x-

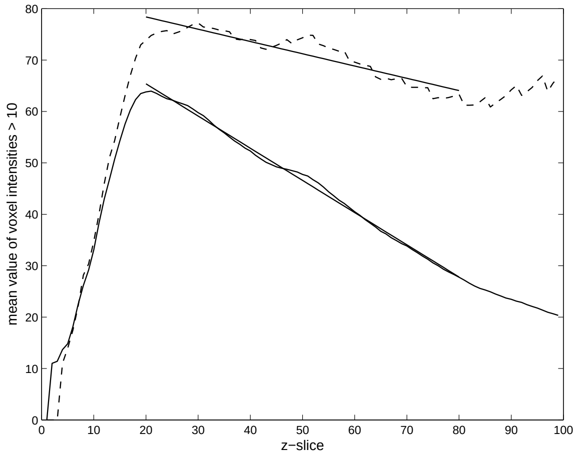


Figure 25: A plot of the mean intensity for all element values > 10 in each z-slice in the confocal microscopy image versus slice number. The solid line represents the values prior to light attenuation compensation and the dashed line represents the values after compensation. The compensation was performed as described in the text.

and y-directions and $163nm$ in the z-direction. The image was subsampled to cubic voxels by nearest neighbor interpolation, as this simple interpolation method proved to be sufficient.

The method was evaluated on six 2D images and one 3D image of cell nuclei in tissue. There was a total of 689 nuclei in the 2D images, and 90 nuclei in the 3D image, of which 91% were correctly segmented for both the 2D images and the 3D image. A surface rendering of the 3D result and a close-up are shown in Figure 26. Apart from the nuclei there were fluorescent objects in the images not at all resembling proper nuclei. These were also segmented but if necessary they can easily be removed based on their different size. This would improve the percentage of correctly segmented nuclei.

The mean intensity in the z-slices still decrease with depth, even after the compensation has been performed, see the dashed curve and the fit straight line in Figure 25. A better compensation would probably have been achieved by fitting the line to the mean of the highest percentage of intensities in each slice, instead of the mean of all pixel intensities higher than 10 in each slice. The line should then only be fit to the slices inside the tissue sample. The slices before the peak in Figure 25 originate from when the focal plane was outside the sample and should, hence, not be used for fitting the line. Calculating the mean of all pixel values higher than a low threshold in each slice might, however, be necessary to decide where sample begins to be in focus.

To test whether the decrease in signal through the 3D sample was only due to

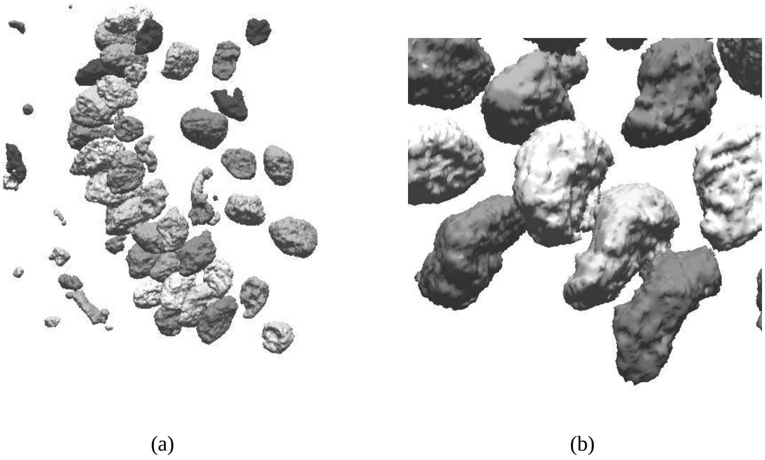


Figure 26: The result of the segmentation method applied to a volume image of cell nuclei (a), and a close-up showing that the method manages to separate touching nuclei (b).

light attenuation or also due to less staining in the deeper slices some additional experiments were performed. New tissue sections were prepared as in Paper V. In addition, one section was stained from both sides, instead of just letting the stain diffuse into the sample from one side, which is more common. Volumes were acquired with different objectives (and from both sides of the section which was stained from both sides). These volumes revealed that staining from one side reaches through the whole volume without decrease and that the decrease in signal was only due to light attenuation.

4.5 Identification and representation of molecules imaged with SET

As mentioned in Section 2, SET is a protein imaging method producing 3D images with a resolution down to $2nm$. This is good enough to give coarse structural information of proteins, together with information on how they interact with other proteins or molecules. SET volumes contain thousands of objects of which only a handful are the objects of interest. These are, so far, found by visual inspection of objects having approximately the proper size after suitable thresholding of the image. Developing methods for segmentation and analysis of objects in the volumes would reduce the amount of human interaction needed to find the objects of interest. The objects of interest are very small in the volume images, and therefore the amount of available shape information is limited. The reconstructions also suffer from noise and varying background. In Papers VI and VII, two representations of proteins imaged with SET are presented. They are both based on the internal density distribution of the object, which make them insensitive to changes on the border. These methods can serve as valuable tools when analyzing structural differences of objects of a certain kind and how they interact with other objects. They can also be used in the segmentation process to help distinguish true objects of interest. In Paper VIII, a segmentation method which combines intensity, edge, and shape information is presented. This creates a method which reduces the mentioned problems of varying background, touching objects, and need for large amount of human interaction, in cases where the shape of the objects of interest is known.

Volume images of proteins created from atom positions deposited in the protein data bank, (PDB), (Berman et al., 2000), were created and used for developing and evaluating the methods. These volumes were created by representing each atom by a Gauss kernel with a magnitude corresponding to the weight of the atom. The sum of the contributions from all atoms less than a resolution dependent distance away from each element, constitutes the intensity value assigned to that image element.

Object decomposition using internal intensity and shape

In Paper VI, a method which decomposes an object into its constituent parts is described. Measurements of the individual parts and how they are connected in the object can be used to describe, and distinguish between, different objects. The method is an extension of a decomposition method for binary objects presented by Svensson and Sanniti di Baja (2002) to incorporate also the internal intensity distribution of an object. The intensity information decides how many parts an object will be decomposed into, while the shape decides where the borders between the parts will be placed. This makes the method robust to small changes on the border of an object. The choice of grey-level threshold used for segmenting the image is therefore not critical, as long as an object stays connected. The threshold will, hence, not affect the number of parts an object is decomposed into, but it will still affect the sizes, and to a certain extent, the shape of the parts.

The decomposition method is applied to a grey-level object already segmented from the background. It consists of two main steps: identification of seeds from the internal grey-level distribution, and shape based growth from these seeds. Local grey-level maxima are identified and marked as seeds. Since the images are rather noisy, using all local maxima would create too many seeds, and consequently, too many regions in the final decomposed object. Therefore, the image is smoothed by applying a Gaussian filter prior to detection of local maxima. All maxima detected on the smoothed image are considered as significant and used as seeds. A sphere, centered at each seed is then grown until it somewhere reaches the border of the object. A second growth step consists of assigning the remaining object elements to the closest sphere. The two growth steps are both based on distance information. A distance transform is calculated inside the object, and the distance values assigned to the seeds are used in a reverse DT, RDT. This RDT corresponds to the spheres, centered at the seeds. In the second growth step a constrained DT from the borders of the spheres with the background as obstacle, is calculated. The spheres are then grown distance layer by distance layer until all image elements are assigned to a region. The purpose of distance based growth in two steps instead of one, is to stop seeds in protruding or peripheral parts of the object, to conquer elements from larger regions. The result of the decomposition method applied to some protein images created from PDB are shown in Figure 27, and to some proteins (antibody IgG) imaged with SET in Figure 28. The first columns of the two Figures contain identifications, the second columns contain volume renderings, and the third contain the decomposition results. The content of the fourth columns, a grey-level based medial representation, will be described below. In the last row of Figure 27, the star in the identification denotes that the molecule presented is an assumed biological structure, which in this case consists of three units of the protein with PDB identification 1eo8.

In some cases, e.g. for 1igt (which consist of two Fab arms and a Fc stem), quantitative information (number of amino acids or atoms) about the structural subunits (domains) of a protein can be found in PDB. The number of subparts in the decomposition result depends on the resolution, as this affects how many local intensity maxima that can be distinguished. In a volume image with a resolution of $2nm$ produced from PDB, the protein 1igt is decomposed into 7 subparts, see Figure 27. Each of the two Fab arms are decomposed into 2 regions and the Fc stem is decomposed into 3. To enable comparison of the subparts of the decomposition and domain information, in some cases available in PDB, the image can be smoothed until it decomposes into the correct number of subparts. If this is done for 1igt, each Fab arm and the Fc stem make up 34, 32, and 34 % of the total volume, respectively. The corresponding percentages for the number of amino acids in each domain is 33 for each Fab arm, and 34 for the Fc stem.

Grey-level based medial representation

A medial representation of objects in volume images is suggested in Paper VII. This representation facilitates analysis and comparison, as it is compact. Moreover, it enhances and reveals internal grey-level features that are difficult to discover by inspecting the original objects. The representation is based entirely on grey-level distribution and no specific segmentation of the object is necessary. Seeds for the representation are identified as in the decomposition method described above, i.e., as local maxima after Gaussian smoothing. Some sort of segmentation is, hence, necessary to decide which seeds belong to the object and which do not. The representation scheme can roughly be described as the maximum intensity path connecting these seeds of the different parts. It is developed from the method presented in Svensson et al. (2002), and adapted to the application of studying proteins in volume images.

After identification of the local maxima, any possible grey-level cavities are set to belong to the background. The grey-level cavities are identified as connected components of only one grey-level, with neighbors of strictly higher grey-levels. Iterative thinning guided by the grey-level distribution in the image is then performed. A distance label is assigned to each voxel, denoting the minimum constrained distance to a region with a lower grey-level, when all regions with higher grey-levels are obstacles. Each voxel will, hence, have two values, a grey-value, and a distance value. Thinning is then performed per grey-level, and within each grey-level, per distance value. Voxels are removed if their removal does not change the topology of the background or the object, (Bertrand and Malandain, 1994; Saha and Chaudhuri,

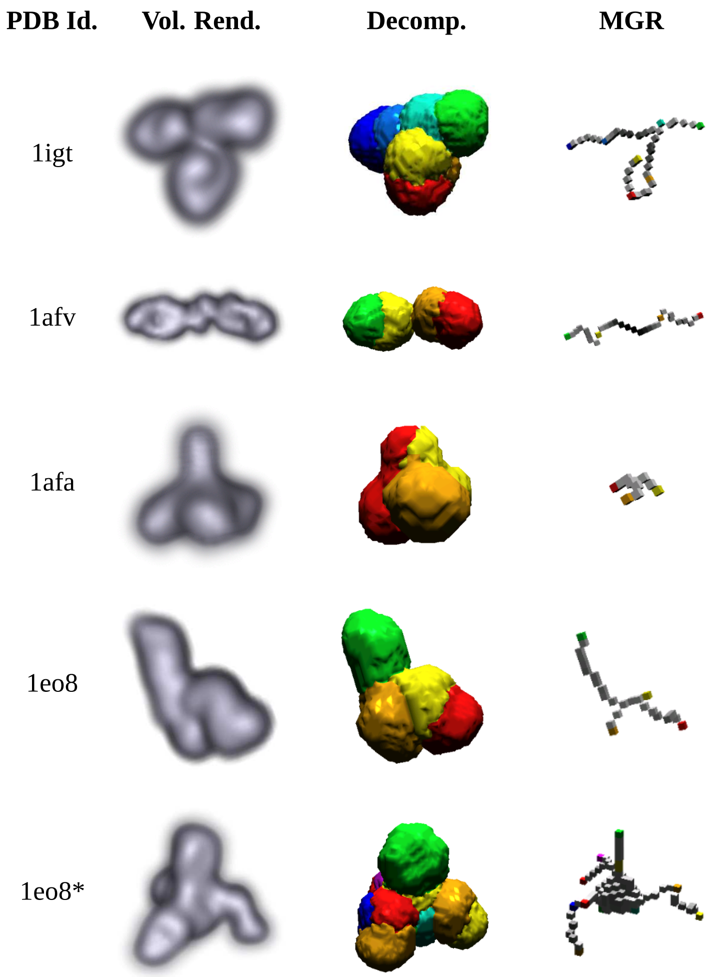


Figure 27: From left to right: PDB identification; volume rendered original object; grey-level decomposition; medial grey-level based representation, MGR, for constructed PDB volumes.

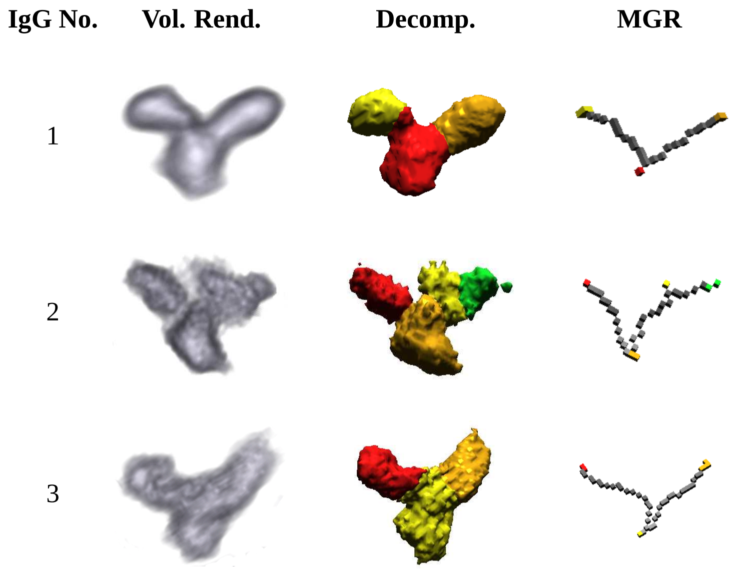


Figure 28: From left to right: volume rendered original object; grey-level decomposition; and medial grey-level based representation, MGR, for three IgG molecules from SET.

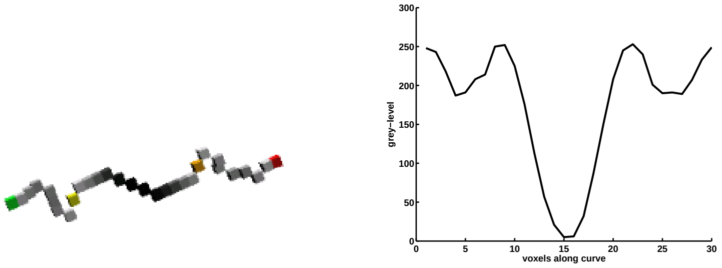


Figure 29: Left: Medial grey-level based representation of PDB ID 1afv. Right: Grey-levels along the representation.

1994). The local maxima serve as endpoints in the representation, and are therefore considered as unremovable during thinning. The resulting representation is denoted Medial Grey-level Representation (MGR). MGR for some proteins constructed from PDB are shown in the fourth column of Figure 27, and for some proteins imaged with SET in the fourth column of Figure 28. The grey-values of the voxels that are part of the MGR are kept in the representation, as they reveal information about the grey-level distribution along the path, or paths, connecting the center points. As the grey-levels depict density, this could indicate how tightly different parts are connected. An illustration of this is shown in Figure 29, for a protein constructed from the PDB ID 1afv. The grey-levels between the two middle local maxima are low, indicating that they are only weakly connected, while the higher grey-levels between the middle and the outer local maxima indicate that they are probably more tightly connected.

Shape based identification

The segmentation method used so far for segmenting objects in SET volumes consists of grey-level thresholding together with size discrimination and visual inspection. This method has several drawbacks. The background varies in the volumes, and using one grey-level threshold for the whole volume therefore means that the size of an object will be dependent on its position in the volume. To be sure not to throw away true objects of interest, the range of accepted sizes, hence, has to be rather generous. This in turn, increases the amount of objects that need to be visually inspected. As the visual inspection is very time consuming, reducing the number of objects that need visual inspection is desirable. Another problem with the size discrimination step is that objects touching other objects or objects that have been split into parts are easily discarded as they will be interpreted as one object being too large or several objects being too small. In Paper VIII, a method is presented which focuses on the described problems. It combines intensity and gradient magnitude information to extract stable borders of the objects. A template is then used to search for borders corresponding to objects of interest.

The method uses the hierarchical chamfer matching algorithm, (HCMA), (Borgefors, 1988), which matches a binary edge template to a DT of a binary edge image. The matching is embedded in a resolution pyramid to speed up the calculations and avoid getting stuck in false local minima. The sum of the squared distance values hit by the template is used as the comparison measure, see Section 3.2. A low sum will, hence, correspond to a good match. The template can, in the case of SET volume segmentation, be either a volume image of a protein of the kind of interest, constructed from PDB, or a manually identified object in a SET volume. The edge of the template is found through grey-level thresholding and extracting the border of the object. The edges of the image are found by using a seeded WS algorithm. Object seeds are identified as voxels having an intensity higher than a threshold, known to definitely correspond to the interior of possible proteins. A background seed is identified as the largest connected component of elements having a gradient magnitude lower than a threshold. This produces a background seed spread over the whole volume even if the background is varying. A seeded WS is then applied to the gradient magnitude image, and the borders between the resulting regions will be placed along the gradient magnitude maxima. This is similar to the method used in Paper V, with the exception that the object seeds are all given the same label, as the borders, and not the individual objects, are sought. The border of the object region is extracted and used in the HCMA algorithm. The positions generating low sums are the result of the method. Objects at these positions should be presented, in order of increasing sum, to an expert for final visual inspection.

In Figure 30, the objects at the two best positions in each of three volumes with the antibody IgG as objects of interest are presented. The volumes each contain one visually identified IgG antibody. The method manages to find the correct objects as the best position in volume one and three and as the second best position in volume

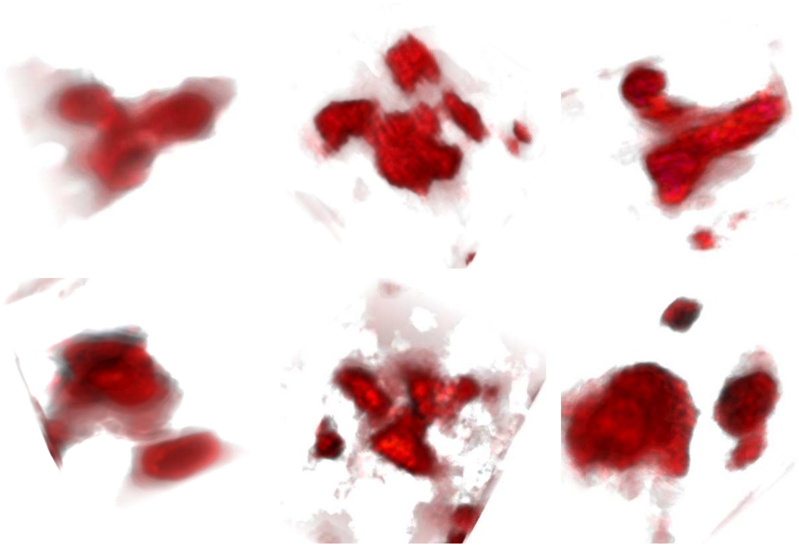


Figure 30: Volume renderings of subvolumes corresponding to the the best (top row) and second best (bottom row) match positions in three SET volume images containing the antibody IgG in solution. The visually judged true IgG proteins are at the best position for the first and third volume and at the second best position in the second volume.

two, even though the template and the objects differ quite a lot. The structure at the best position in volume two is rather similar to the template and might possibly be an antibody not found by the former segmentation method.

As a postprocessing step, the mean intensity of the objects at the best positions can be investigated to make sure that the position corresponds to a position with high intensity values, and not a part of the background between two objects. This would possibly have revealed that the second best positions in volumes one and three correspond to false proteins.

5 Conclusions and future work

This thesis includes image analysis research essentially within three different fields. In the digital geometry field, theoretical results regarding optimization of the local weights in distance transforms (DTs) adapted to elongated grids in 2D and 3D have been described.

In the segmentation field, several methods developed for different applications have been presented. A DT with weights optimized for an elongated voxel grid, in combination with the watershed (WS) algorithm was used to segment individual pores in a paper volume. A match based segmentation algorithm was developed to identify and classify three types of viral capsids in 2D images of infected cell nuclei. An average radial grey-level profile for each capsid type was used to create the templates used in match based segmentation. Intensity, edge strength, and shape information were combined in a WS based method to segment cell nuclei in 2D and 3D images. The last segmentation algorithm presented in this thesis combines intensity, edge, and shape information to segment proteins of a specific kind in volume images.

In the field of object representation, some description schemes have been presented. The radial grey-level profiles used for identifying and classifying capsids serve as an efficient way of describing and comparing different capsids. Two methods developed to describe and simplify analysis of proteins in volume images have been presented. The first decomposes a segmented object into subparts based on internal intensity distribution and shape. The second creates a medial representation of an object solely from its internal intensity distribution.

Further investigation on how to adapt image analysis methods to elongated grids would be useful. Whether, e.g., connectivity and edge detection filters should be treated differently in such grids, compared to square or cubic grids, needs investigation. Using information about elements in a neighborhood of the element under consideration can perhaps reveal information about its connectivity to other elements and how steep an edge is and where it should be positioned.

In relation to the analysis of pores in paper, as well as to the decomposition of proteins into subparts, the question of how to find the main part of an object arose. The largest inscribed ball is often used, but this is a measure that in most cases is not representative, neither of the shape, nor of the volume. The largest inscribed convex object would be a more representative feature to investigate. However, it is so far not clear how to identify such an object, or a good approximation thereof.

Another idea, not yet investigated, is whether it is possible to use the presented decomposition method to estimate the resolution achieved in Sidec Electron Tomography (SET) volumes (or in other electron tomography reconstructions). If a protein of the same kind as in the reconstruction is deposited in the protein data bank (PDB), volume images of it can be constructed at different resolutions. Comparing the decomposition result of the PDB protein in images of different resolutions with the decomposed object in the SET image would probably reveal the approximate resolution obtained in the SET image. There are two circumstances that are likely to limit the accuracy of the estimated resolution. The first is that the SET images contain noise, which influence the detection and localization of local maxima. The second is that the protein structures deposited in PDB are likely to be tighter than the same protein imaged with SET, as SET images proteins moving freely in solution while most protein structures deposited in PDB are the structures proteins have in crystal lattices.

In the hierarchical chamfer matching (HCMA) algorithm, DTs are calculated from extracted edges in the image to be searched. Some work has been performed by Rosin and West (1995) on how edge strength information (or other information) can be included in the DT and subsequently used in the matching. Rosin and West (1995), however, start from already extracted binary edges. A slightly different approach would be to calculate a DT directly from the gradient magnitude information. This can be achieved by calculating a reverse DT (or a DT from the inverted gradient magnitude image) from the values already present in the gradient magnitude information. The benefit of this strategy is that no decision of what is to be counted as an edge or not need to be taken. This edge-weighted DT has been somewhat investigated by the author with promising results. The problem that occurs, though, is how to weigh different types of information, a problem present in all attempts to combine modalities. This needs further investigation.

A tunnel is defined as the background passing through the object, and this definition is also used in the medial grey-level based representation method. Considering volume images of proteins and protein complexes, it could easily be imagined that *grey-level tunnels* exist and are of interest. A grey-level tunnel would be a path of grey-levels higher than the background but lower than the voxels surrounding them, passing through the object. Identifying such grey-level tunnels is a topologically interesting task, and incorporating them into the medial grey-level based representation would make it even better suited for the protein analysis application.

Information from the decomposition result and the medial grey-level based representation could most likely be used in the segmentation process. It could preferably be added at the last step of the presented match based segmentation method to discard false objects and strengthen the overall rating of true objects of interest.

The idea of using the WS algorithm on the grey-level histogram of an image to automatically find threshold levels, illustrated in Figure 11, should also be further investigated and evaluated on different types of images.

In connection to the description and discussion of each method in the contribution section, some additional ideas or suggestions for improving the methods were mentioned. Investigating these ideas further, together with the briefly described unexplored ideas above, is left for future work.

References

Alberts, B., Bray, D., Lewis, J., Raff, M., Roberts, K., and Watson, J. D. (1994). *Molecular Biology of the Cell*, chapter 4, pages 139–156. Garland Publishing, Inc.

Arcelli, C. and Sanniti di Baja, G. (1988). Finding local maxima in a pseudo-Euclidean distance transform. *Computer Vision, Graphics, and Image Processing*, 43:361–367.

Aronsson, M. (2002). *On 3D Fibre Measurements of Digitized Paper*. PhD thesis, Swedish University of Agricultural Sciences. Silvestria, No. 254.

Aronsson, M., Sävborg, Ö., and Borgefors, G. (2002). Minimizing scanning electron microscope artefacts by filter design. *Journal of Microsocpy*, 206(1):84–92.

Barrow, H. G., Tenenbaum, J. M., Bolles, R., and Wolf, H. C. (1977). Parametric correspondence and chamfer matching: Two new techniques for image matching. In *Proceedings 5th International Joint Conference on Artificial Intelligence*, pages 659–663, Cambridge, Massachusetts.

Berman, H., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T., Weissig, H., Shindyalova, I., and Bourne, P. (2000). The protein data bank. *Nucleic Acids Research*, 28:235–242.

Bertrand, G. and Malandain, G. (1994). A new characterization of three-dimensional simple points. *Pattern Recognition Letters*, 15:169–175.

Beucher, S. and Lantuéjoul, C. (1979). Use of watersheds in contour detection. In *International Workshop on Image Processing: Real-time and Motion Detection/Estimation*, pages 2.1–2.12.

Bolon, P., Vila, J. L., and Auzepy, T. (1992). Opérateur local de distance en maillage rectangulaire. In *Proc. 2ème Colloque de Géométrie Discrète en Imagerie: Fondements et Applications*, Grenoble, France, pages 45–56. In French.

Borgefors, G. (1984a). Distance transformation in arbitrary dimensions. *Computer Vision, Graphics, and Image Processing*, 27:321–345.

Borgefors, G. (1984b). An improved version of the chamfer matching algorithm. In *7th International Conference on Pattern Recognition*, pages 1175–1177, Montreal, Canada.

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

Borgefors, G. (1988). Hierarchical chamfer matching: A parametric edge matching algorithm. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 10(6):849–865.

Borgefors, G. (1994). Applications using distance transforms. In Arcelli, C., Cordella, L. P., and Sanniti di Baja, G., editors, *Aspects of Visual Form Processing*, pages 83–108. World Scientific Publishing Co. Pte. Ltd.

Coquin, D. and Bolon, P. (1995). Discrete distance operator on rectangular grids. *Pattern Recognition Letters*, 16:911–923.

Coquin, D., Chehadeh, Y., and Bolon, P. (1994). 3D local distance operator on parallelepipedic grids. In *Proc. 4th Discrete Geometry for Computer Imagery*, Grenoble, France, pages 147–156.

Fouard, C. and Malandain, G. (2005). 3-D chamfer distances and norms in anisotropic grids. *Image Vision and Computing*, 23(2):143–158.

Fouard, C., Malandain, G., Prohaska, S., Westerhoff, M., Cassot, F., Marc-Vergnes, J.-P., Mazel, C., and Asselot, D. (2004). Skeletonization by blocks for large 3D datasets: Application to brain microcirculation. In *2nd. IEEE International Symposium on Biomedical Imaging*, pages 89–92.

Gonzalez, R. C. and Woods, R. E. (2002). *Digital Image Processing*, chapter 12, pages 698–704. Prentice-Hall, 2nd edition.

Ikonen, L. and Pekka, T. (2005). Shortest routes on varying height surfaces using gray-level distance transforms. *Image and Vision Computing*, 23(2):133–141.

Kong, T. Y. and Rosenfeld, A. (1989). Digital topology: Introduction and survey. *Computer Vision, Graphics, and Image Processing*, 48:357–393.

Meyer, F. and Beucher, S. (1990). Morphological segmentation. *Journal of Visual Communication and Image Representation*, 1:21–46.

Montanari, U. (1968). A method for obtaining skeletons using a quasi-Euclidean distance. *Journal of the Association for Computing Machinery*, 15(4):600–624.

Nyström, I. (1997). *On Quantitative Shape Analysis of Digital Volume Images*. PhD thesis, Uppsala University. Faculty of Science and Technology, No. 288.

Nyström, I. and Smedby, Ö. (2001). Skeletonization of volumetric vascular images – distance information utilized for visualization. *Journal of Combinatorial Optimization*, 5(1):27–41. Special Issue on Optimization Problems in Medical Applications.

Nyström, I. and Borgefors, G. (1995). Synthesising objects and scenes using the reverse distance transformation in 2D and 3D. In Braccini, C., De Floriani, L., and Vernazza, G., editors, *Image Analysis and Processing*, volume 974 of *Lecture Notes in Computer Science*, pages 441–446. Springer-Verlag.

Piper, J. and Granum, E. (1987). Computing distance transformations in convex and non-convex domains. *Pattern Recognition*, 20(6):599–615.

Rosenfeld, A. (1969). *Picture Processing by Computer*. Academic Press, New York.

Rosenfeld, A. and Pfaltz, J. L. (1966). Sequential operations in digital picture processing. *Journal of the Association for Computing Machinery*, 13(4):471–494.

Rosenfeld, A. and Pfaltz, J. L. (1968). Distance functions on digital pictures. *Pattern Recognition*, 1:33–61.

Rosin, P. L. and West, G. A. W. (1995). Salience distance transforms. *Graphical Models and Image Processing*, 57(6):483–521.

Saha, P. K. and Chaudhuri, B. B. (1994). Detection of 3-D simple points for topology preserving transformations with applications to thinning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(10):1028–1032.

Saha, P. K., Wehrli, F. W., and Gomberg, B. R. (2002). Fuzzy distance transform: Theory, algorithms, and applications. *Computer Vision and Image Understanding*, 86:171–190.

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

Sanniti di Baja, G. (1994). Well-shaped, stable, and reversible skeletons from the (3,4)-distance transform. *Journal of Visual Communication and Image Representation*, 5(1):107–115.

SET webpage (2004). Sidec Technologies. http://www.sidectech.se. visited 2004-12-10.

Skoglund, U., Öfverstedt, L.-G., Burnett, R., and Bricogne, G. (1996). Maximum-entropy three-dimensional reconstruction with deconvolution of the contrast transfer function: A test application with adenovirus. *Journal of Structural Biology*, 117:173–188.

Soille, P. (1999). *Morphological Image Analysis: Principles and Applications*. Springer-Verlag.

Sonka, M., Hlavac, V., and Boyle, R. (1999). *Image Processing, Analysis, and Machine Vision*, chapter 4, pages 77–88. Brooks/Cole Publishing Company, 2nd edition.

Svensson, S. (2001). *Representing and Analyzing 3D Digital Shape Using Distance Information*. PhD thesis, Swedish University of Agricultural Sciences. Silvestria, No. 211.

Svensson, S. (2002). Reversible surface skeletons of 3D objects by iterative thinning of distance transforms. In Bertrand, G., Imiya, A., and Klette, R., editors, *Digital and Image Geometry*, volume 2243 of *Lecture Notes in Computer Science*, pages 395–406. Springer-Verlag.

Svensson, S. and Aronsson, M. (2003). Using distance transform based algorithms for extracting measures of the fibre network in volume images of paper. *IEEE Transactions on Systems, Man, and Cybernetics*, 33(4):562–571. Special issue on 3-D Image Analysis and Modeling.

Svensson, S. and Borgefors, G. (2002a). Digital distance transforms in 3D images using information from neighbourhoods up to $5 \times 5 \times 5$. *Computer Vision and Image Understanding*, 88(1):24–53.

Svensson, S. and Borgefors, G. (2002b). Fuzzy border distance transforms and their use in 2D skeletonization. In Kasturi, R., Laurendeau, D., and Suen, C., editors, *16th International Conference on Pattern Recognition (ICPR 2002)*, volume I, pages 180–183. IEEE Computer Society.

Svensson, S., Borgefors, G., and Nyström, I. (1999). On reversible skeletonization using anchor-points from distance transforms. *Journal on Visual Communication and Image Representation*, 10(4):379–397.

Svensson, S., Nyström, I., Arcelli, C., and Sanniti di Baja, G. (2002). Using grey-level and distance information for medial surface representation of volume images. In Kasturi, R., Laurendeau, D., and Suen, C., editors, *Proceedings 16th International Conference on Pattern Recognition (ICPR 2002)*, volume 2, pages 324–327. IEEE Computer Society.

Svensson, S. and Sanniti di Baja, G. (2002). Using distance transforms to decompose 3D discrete objects. *Image and Vision Computing*, 20(8):529–540.

Vincent, L. (1993). Morphological grayscale reconstruction in image analysis: Applications and efficient algorithms. *IEEE Transactions om Image Processing*, 2(2):176–201.

Wählby, C. (2003). *Algorithms for Applied Digital Image Cytometry*. PhD thesis, Uppsala University. Faculty of Science and Technology, No. 896.

Other publications and conferences

The author has also been author or co-author to the following papers.

- Höglund S., Su J., Sandin Reneby S., Végvári Á., Hjertén S., Sintorn I.-M., foster H., Wu Y., Nyström I., Vahlne A. (2002). Tripeptide Interference with Human Immunodeficiency Virus Type 1 Morphogenesis. *Antimicrobial Agents and Chemotherapy*, 46(11):3597–3605.
- Sintorn I.-M., Borgefors G. (2002). Weighted distance transforms for images using elongated voxel grids. In Braquelaire, A., Lachaud, J., Vialard A., editors, *Discrete Geometry for Computer Imagery*, volume 2301 of *Lecture Notes in Computer Science*, pages 244-254. Springer-Verlag.
- Aronsson M., Sintorn I.-M. (2002). Ring Shaped Object Detector for Non-Isotropic 2D Images using Optimized Distance Transform Weights. In *Proc. IEEE International Conference on Image Processing*, pages 985–988.
- Sintorn I.-M., Homman M. (2002). Description, Segmentation and Classification of Human Cytomegalovirus Capsids. In Åström K., editor, *Proc. SSAB'02, Symposium on Image Analysis*, pages 21–24.
- Sintorn, I.-M. (2002). Automatic identification and classification of Cytomegalovirus capsids in electron micrographs. Abstract in *Proc. VINNOVA 2002*.
- Homman, M., Sintorn, I.-M., Hultenby, K., Borgefors, G., Söderberg-Naucler, C. (2002). Nuclear egress of human Cytomegalovirus capsids by budding through the nuclear membrane. Extended abstract in *Proc. International conference on Electron Microscopy*.
- Sintorn I.-M. (2004). Object Decomposition Based on Grey-level and Shape Information. In Bengtsson E., Eriksson M., editors, *Proc. SSBA'04, Symposium on Image Analysis*, pages 25–28.
- Sintorn, I.-M., and Weistrand, Ola. (2002). A report on the first French-Nordic summer school in mathematics. In *Centre for Image Analysis Internal report*, No. 25.

The following international conferences have been attended.

- 9th International Conference on Discrete Geometry for Computer Imagery (DGCI), Uppsala, Sweden, Dec. 2000.
- 11th International Conference on Image Analysis and Processing (ICIAP), Palermo, Italy, Sept. 2001.
- 10th International conference on Discrete Geometry for Computer Imagery (DGCI), Bourdeaux, France, April 2002.

- 11th International Conference on Discrete Geometry for Computer Imagery (DGCI), Naples, Italy, Nov. 2003.
- IEEE International Symposium on Biomedical Imaging (ISBI), Arlington, VA, USA, April 2004.

Acknowledegments

Special thanks to the following people for contributing to this thesis in some way or the other.

- Gunilla Borgefors and Ingela Nyström, my supervisors, for friendship, help, and support during these years, and for having confidence in me and encouraging me to develop my own ideas.
- Lena Wadelius for help with EVERYTHING and keeping things at CBA running smooth.
- Ewert Bengtsson for introducing me to Sidec and several skilled researchers.
- The more permanent staff at CBA: Gunilla, Ewert, Tommy, Fredrik, Olle, Lena, Ingela, and Bosse for creating a great atmosphere at CBA.
- The less permanent staff at CBA: current and former PhD student colleagues Anna, Ola, Robin, Mats, Maria, Petra, Lina, Amalka, Magnus, Patrick, Stina, Roger, Felix, Xavier, Erik, Jocke, Mattias A, for lunches, fika, friday pubs, wednesday pubs, and any other day pubs.
- Olle for help with the computer system, giving me more memory, and for restoring, by me mistakenly deleted files.
- Bosse, Jocke, Erik, for not loosing patience while helping me with my never ending questions about computers and programming.
- Gunilla, Ingela, Stina, Lina, and Bosse, for proofreading this thesis.
- Collaborators and co-authors from outside CBA: Sidec & Sara Sandin, Örjan Sävborg, Fredrik Erlandsson, Susana Mata, and especially Mohammed Homman-Loudiyi, optimistic, enthusiastic, and funny; simply the best collaborator one can imagine.
- Collaborators and co-authors at CBA: Gunilla Borgefors, Maria Axelsson, Magnus Gedda, Ewert Bengtsson, and most of all Stina Svensson and Carolina Wählby.
- Erik Vidholm for great support and help when this thesis and the last Papers were produced, and panic was close.
- My office roommates: Roger Lundqvist, Nataša Sladoje, and a bunch of master thesis students Erik, Per, Kristin and Mathias, for making our room the best working place.
- Petra, Stina, Anna, Lina, Nataša, and Erik for a friendship started at work that I hope will last forever!
- The other members of CBA's golf-section, Tommy, Roger, and Erik, for not always working on sunny summer afternoons.

- All my friends from Uppsala and elsewhere!
- Old and new members of my old and new family: Mamma Kerstin & Pappa Jalle, Kalle & Nadine, Anna-Karin, Jo-Hanna, Barbro & Roger, Magda & Viktor, I love you all!
- Mathias, you are the best! Jag älskar dig tusen miljarders!

Uppsala, February 2005

Ida-Maria Sintorn