

Abstract

The reconstruction of biomedical images is of a great importance. In molecular *pathology** numerical and structural *chromosome** aberrations have been found to be an important clue for the diagnosis of *tumors**. The existence or nonexistence of such aberrations can be detected by *interphase** *fluorescence in situ hybridization (FISH)** [38, 54]. This method, which is based on tracing the *FISH signals** in the nucleus, helps to find out, whether there are any gene irregularities, like gain (e. g. *trisomy**) or loss (e. g. *monosomy**) of certain *base sequences** in the *deoxyribonucleic acid (DNA)**. The correct evaluation of such events is a necessary condition for a prospective use in diagnosis of certain genetic diseases. To guarantee an unbiased signal count per nucleus, a whole nucleus has to be inspected. For this reason a correct segmentation of each cell is very important and therefore accurate segmentation technique is required.

In this work the overview of methods concerning cell boundary extraction from images acquired from microscope is given. All the techniques can be divided into two main groups: manual and semi-automatic. Because of the simplicity of maintenance and speed the great emphasis is put on the latter one. At the present time, the research has been focused on the development of semi-automatic approaches called deformable models. These techniques have deep mathematical and physical background, are able to use *a priori* knowledge about the inspected object, and are as self-sufficient as possible.

As a contribution a new method based on the common segmentation principles has been designed, presented and tested. This method is capable of correct tissue reconstruction with a minimum degree of user interaction.

Keywords

Biomedical image segmentation • Simplex mesh • Dual simplex mesh • Deformable model

*For explanation of starred emphasized words in the text see glossary on page 115.

Contents

1	Introduction	1
1.1	Image origin	1
1.2	Image preprocessing & analysis	1
1.3	Objectives to be reached	4
2	Background	6
2.1	Manual segmentation	8
2.2	Unsupervised segmentation	9
2.2.1	Thresholding	9
2.2.2	Edge tracing based	12
2.2.3	Region based	15
2.2.4	Template based	18
2.2.5	Texture based	18
2.2.6	Deformable models – fundamentals	19
2.2.7	Deformable models – alternative approaches	23
2.3	Recent methods summary	26
3	Coarse segmentation	28
3.1	Hough Transform based	29
3.2	Multilevel adaptive thresholding	32
3.2.1	Threshold setting	32
3.2.2	Image division	33
4	Fine segmentation	37
4.1	General simplex mesh	37
4.1.1	Definition	37
4.1.2	Topological properties	38
4.1.3	Geometrical properties	39
4.1.4	Simplex mesh as a deformable model	41
4.1.5	Simplex mesh generation	45
4.2	Star-shaped simplex mesh	48
4.2.1	Definition	48
4.2.2	Topological properties	49

4.2.3	Star-shaped simplex mesh as a deformable model	49
4.3	Dual simplex mesh	52
4.3.1	Definition	52
4.3.2	Dual simplex mesh as a deformable model	53
5	Application	58
5.1	Materials and methods	58
5.1.1	Sample preparation	58
5.1.2	Hardware description	59
5.1.3	Image acquisition	60
5.2	Cell segmentation using dual simplex meshes	60
5.2.1	Application tool	65
6	Results	66
6.1	Coarse segmentation	66
6.1.1	HT based marker localization	69
6.1.2	MAT based marker localization	73
6.2	Fine segmentation	73
6.2.1	Star-shaped simplex meshes	73
6.2.2	Dual simplex meshes	77
7	Conclusion	91
7.1	Coarse segmentation	91
7.2	Fine segmentation	93
A	Input images	95
B	Toolbox	108
B.1	Required packages	108
B.2	HT utility	109
B.3	MAT utility	110
B.4	viewer3d utility	110
B.5	DSM utility	110
C	Fourier Analysis	113
C.1	Basic definitions	113
C.2	Convolution theorem	114
	Glossary	114
	Bibliography	117

Chapter 1

Introduction

1.1 Image origin

*Biomedical** image analysis and recognition is an interesting and non-trivial problem. In order to measure certain cell properties, some of its components must be visualized by means of one of the available visualization methods.

The aim of our interest is the nucleus architecture. Since the nucleus contains mainly the DNA molecules, we will focus on the visualization of the DNA. The procedure of the DNA visualization is called FISH. For FISH we can use specific DNA probes modified by *digoxigenin**, *biotin** (nondirect labeling) or fluorescent dyes (direct labeling). *Hybridization** and post-hybridization washing are performed according to the instructions of the manufacturers. The principle of the hybridization method is based on incorporating the specific stained DNA sequences to the denaturated DNA matrix with the complementary sequences of the bases (see Fig. 1.1).

In the past, the probes were stained with radioisotopes (ISH method). The drawback of this method stemmed from a low resolution. Recently, the visualization has been performed using the fluorescence technique – hence the FISH method.

1.2 Image preprocessing & analysis

Before starting the own analysis, the images have to be submitted to the preprocessing step. It deals with some undesirable phenomena which appear in the scanned images and enhances the quality of them. Sometimes the preprocessing step can coincide with the own analysis, but anyway it should not be omitted:

Noise suppression When acquiring the images from the microscope, the noise is naturally present. It is not possible to avoid it entirely but we are able to reduce its amount. The typical methods for noise reduction are Gaussian or Median filtering [60]. There is an important difference between them: while the Median filter removes spikes and preserves edges, the Gaussian filter just blurs them. Therefore, the Gaussian filter

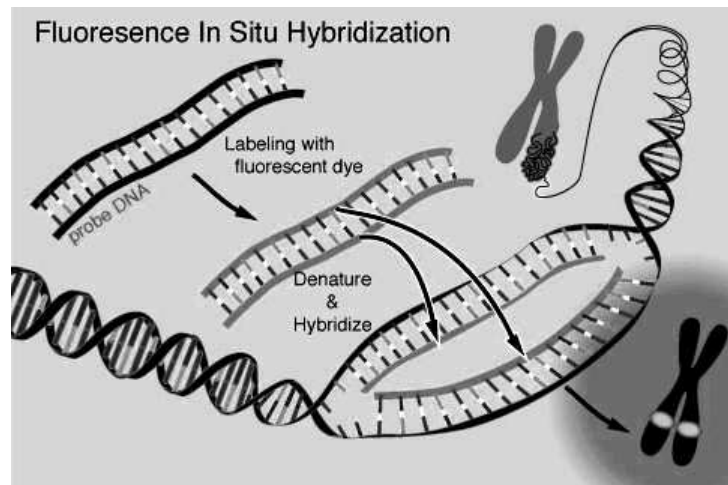


Figure 1.1: The process of FISH (Fluorescence In Situ Hybridization)^a

^asource: National Human Genome Research Institute, Division of Intramural Research

is usually used only if the noise level is low or as a preprocessing step for a certain image analysis algorithm that requires a smoothed image.

Deconvolution Deconvolution is a process of obtaining a deblurred image from a blurred one [60, 75]. The blurring phenomenon is caused by optical laws valid for the imaging devices. Theoretically, if there is no noise present it is possible to compute the original (unblurred) image data from any scanned image data. Unfortunately, the blurring function (PSF) is never precisely known, and moreover it changes dramatically with the position in the image (especially along the axial direction). It is important to know, that good deconvolution algorithms take hours per image.

These two principles are not the only ones which are used to improve the quality of the images. A great accent is put on the choice of scanning instrument – whether to use *wide-field (conventional)* or *confocal* microscope [38]. Although conventional microscopy provides whole range of high-quality tools for industry as well as for research, it has a drawback which stems from the difficulty with three-dimensional (3D) observations. Therefore, only 2D studies are usually performed. Reconstruction of 3D information from 2D slices acquired at different axial positions is possible using deconvolution techniques, but either the result is of poor quality or the time consumption of the deconvolution algorithm is unacceptable. The confocal effect can be easily obtained using a scanning optical microscope with a small pinhole in front of the detector. The main advantage of confocal microscopy is the possibility of obtaining confocal images even for faint fluorescent objects and, consequently, 3D images of the objects.

Segmentation When processing the biomedical images, besides some other demands, one usually requires image segmentation. The segmentation process is seen as an effort to recognize the image contents – the identification of regions within the image. In our case, the regions are assumed to be *cell** clusters (such as in *tissues**), cells or cell *nuclei**, for example.

Many papers concerning this field of image processing have been published and a lot of methods and algorithms have been developed. According to the degree of user interaction, these can be divided into three main groups: manual, semi-automatic and automatic.

Manual methods, based on the user interaction, e. g. drawing some sketches or region boundaries with a mouse, are the gold standard. These give the best results. On the contrary, the amount of the interaction required is too high. Therefore, these methods are looked at as quite cumbersome and time consuming. In addition, each expert has his own unique knowledge concerning the analysed problem. This typically leads to user specific results, which are even unrepeatable.

Nowadays, image analysis methods are required to process a large amount of image data. The examples of such suppliers producing large sets of huge three-dimensional images in the capacities of hundreds of megabytes per one image are the *Magnetic Resonance Imaging (MRI)** or *Confocal Laser Scanning Microscopy (CLSM)**. This fact drives the development of the image analysis methods towards automatic approaches to abandon the manual ones, that are time-consuming.

As mentioned above, the main advantage of the automatic methods is their self sufficiency. In general, they are faster and easier to maintain. In spite of this fact, practically their application is not as easy as it looks. The usage of them brings another computation step – verification of the result. In manual methods, we do not need this testing because they are driven by the user during the evaluation process and therefore the results are consistent with the user request. This is not the case of the automatic ones.

Because of the above reasons the semi-automatic methods are likely to be the best choice. They are capable of driving the computation on their own with a small amount of user interaction. Of course, this interaction is reduced to the minimum in order to reduce the time which is spent by the user.

The degree of the interaction is not the only aspect that one might be interested in. Another usual question is whether the processed data are planar (2D) or spatial (3D). This influences the computation time of the methods processing such images and the quality of their results. As a matter of fact, if the precise segmentation of the cells in image is achieved, inaccuracy in feature measurement can persist due to the usage of 2D images. The two-dimensional images do not represent complete cells in its entirety. Some features such as size, shape, chromosome density, volume, etc., cannot be measured precisely using 2D images. Most of these problems can be reduced by the use of volumetric (3D) images. Use of thick-tissue specimens and the three-dimensional (3D) imaging results in complete and detailed representation of the cells. The examination of spatial distribution of the cells, tissue architectures, tumor grading,

structural and geometric feature measurements, counting FISH signals, etc., can be done more accurately by three-dimensional image analysis.

In this field of study the segmentation is applied to two different types of image data. The first one is the scan of cells that are freely distributed in space. These can be blood cells or cell cultures, for example. The second one concerns the cells, that are part of the cell clusters such as in tissues.

In the first case, the segmentation process includes extraction of the particular isolated cells and separation of the touching or overlapping ones. If the cells do not touch each other, the computation is really easy, otherwise the division has to be solved. The quality of the result depends on the degree of overlap or touch of studied cell pairs. Usually, in one image there are a lot of cells. Therefore, if the division process tends to give ambiguous results, the problematic cells can be omitted and cast out of the further manipulation.

If the segmentation of the *histological** images is performed, the computation process is a bit more complex. There is hardly any cell in the image, which can be clearly separated. Nearly all of the studied cells are touching each other or overlapping. In this case more sophisticated segmentation methods have to be used to get the proper results. For this purpose the approaches using *a priori** knowledge about the processed data have been developed.

The basic segmentation techniques such as *thresholding* [58] or *watershed* [67] do not use *a priori* knowledge about the studied objects. These methods are general and are capable of segmentation any type of objects. However, they give really poor results if the images are noisy or contain partial objects. In case we know the likely shape of the segmented objects we can apply some specific methods, which cooperate with this knowledge. The *active contour model* [32] which stems from the principle of *deformable models* [66] is an example of such a method. For instance, it might be known that the shape of the cells is topologically equivalent to a sphere. This technique and the similar ones are usually handtailored for a specific problem and therefore are not so general as the previous ones. On the other hand, these perform their task very well.

1.3 Objectives to be reached

The area of image segmentation has already been highly developed and therefore it is easy to find a large set of approaches which solve this task. In spite of this fact, there are several specific problems, such as reconstruction of biomedical images acquired from microscopes, which have not been satisfactory solved yet.

In general, there are only a few basic principles applied to solve these problems (see Chapter 2). In all the methods recently used these common origins can be traced up.

In this work the most important methods used in biomedical applications are summarized. The text follows searching for their advantages, drawbacks and hence their suitability for practical use. The main contribution of this thesis is design of a method

suitable for tissue image reconstruction. A great emphasis is placed upon the minimization of user interaction, i. e. the method is self-sufficient as much as possible and requires minimal amount of inputs. As mentioned above, the automatic approaches need verification of their results. This study concerns this problem as well. In this case the particular objects are analyzed according to their shape and volume. The consecutive verification of the results evaluates the suitability of the reconstructed objects for the further processing. Hence, it helps an operator when observing the quality of the results.