Extraction of thin color pattern from images for histology investigation

A.Nedzved and V. Starovoitov, *Member, IEEE*

***Abstract*—Quantitative measurement of membrane characteristics in a histological section is an important step in cancer diagnostic. We present an algorithm of colour image processing of a histological sample for membrane analysis. The algorithm includes a unique color thinning operation that allows to extract a membrane pattern. Colour properties of the detected membrane allow to describe physiological processes in histological tissue for medical analysis**

I. Introduction

An objective analysis of cytological and histological images has been a subject of research for many years. One of the most difficult fields in biomedical image analysis is automated extraction and classification of cells and nucleus.

Basic characteristic of objects in histological images is colour of cells and nuclear membrane. Membrane is not just a passive skin. It is actually a fluid bilayer of lipids with hydrophilic ends on the outside. This membrane is thus active, being penetrated by protein complexes that determine what substances are allowed in and what is allowed out. Also, the outer layer is occupied by glycoproteins that have attached sugars, and these act as a kind of finger-print of the cell, allowing it to be recognised by other cells and important compounds like hormones. Therefore membrane pay an important role for controlling physiological process in tissue. Analysis of this structure is very important for pathology diagnoses.

Membrane pattern in a histology image is characterized by colour, elongated shape and inhomogeneity body. For medical diagnose some analysis of physiological features of tissues and properties of membranes may be useful. There are some algorithms for membrane pattern extraction. Most simple membrane segmentation techniques are combinations of brightness segmentation and thinning [1]. For a good processing these methods require a high quality image and well prepared cells. It is very rarely happens. Some improvement can be done by usage of mathematical morphology [2]. Some efficient algorithms are based on deformable models (snakes) [3-5]. Classical snake models are edge-oriented and work well if the target objects have distinct gradient values. This is not always true in biomedical imagery, which makes the models very dependent on conditions of initial image preparation. This approach was applied for biomedical image segmentation and has positive results but suffer from the well-known drawbacks of initialisation and minimisation.

In paper [3] authors show a Viterbi search-based dual active contour algorithm that is able to overcome many of these problems and achieve over 99% accurate segmentation on a database of 20130 Pap stained cell images. In the literature many variants of deformable models have been proposed. In these methods only one object may be indicated (or initialized) manually that is not always convenient [6]. Results of using deformable models are better than for thresholding-based approaches. But they individually process every cell and require a lot of time to work interactively. Recently new methods of clustering were developed for membrane segmentation. They are based on shape [7], color [8-10] and some fuzzy properties [11]. The main drawback of these methods is a long time to work.

Thus the task of the membrane pattern extraction from a colour histological image is very important. Thresholding-based approaches with application of mathematical morphology allow to get a satisfactory result for separation of homogeneous structure. The main drawback of these methods is processing of a whole object, but not its boundary.

In this article we propose a new algorithm that is a combination of a new colour image thinning method and colour-based clustering.

.

II. Thinning of Colour Images

For elongated object analysis one of the basic operations is thinning. There are many algorithms for binary or gray image thinning. In binary images white colour presets objects and black presents background. In image processing thinning means a transformation of the original "thick" objects into lines of one-pixel thickness. Due to clearance of object boundaries thinning is very well developed for binary images [1].

In gray-scale images object borders are not so sharp and the result of thinning operation is variable. During last decades, there appeared many algorithms for thinning of grey-scale images [2,3] but there are not practically known algorithms for thinning of colour images. However, colour always gives specifics in image processing and object recognition [4,5] and in our task colour is an important feature of the object of interest and should be used.

Manuscript received May 2, 2010. (Write the date on which you submitted your paper for review.) This work was supported in part by the ISTC Grant B1489 and partly by ISTC project #B-1682 and #B-1636.

A.Nedzved and V. Starovoitov are with the United Institute of Informatics Problems of the National Academy of Sciences of Belarus, Surganova str. 6, Minsk Belarus, NedzvedA@newman.bas-net.by.

In gray-scale images each pixel is represented by three coordinates (x,y and brightness), a color image is exemplified by five coordinates: three characteristics of color (hue, luminance, saturation), and two coordinates of the spatial location of the pixel (x,y). Mutual influence of these five parameters is very complex and their analysis is usually time-consuming. Based on color coordinates, we may consider that in color image processing we deal with three gray-scale values. Correspondingly it is possible to adapt gray-scale algorithms for color image processing.

Here, we consider processing of histological images that usually contains cells, kernels, etc. Cells may be classified by using their shape and structure. The cell shape can be treated as almost round-shaped (circular-based primitive) and not elongated. The cell structure is characterized by its kernels and kernels' position. Another important cell feature is its position in a tissue.

It is practically impossible to select or develop automatic segmentation methods that could be applied for these images. That is why most of investigators consider particular features of cell images and methods of their segmentation [7,8]. These results also depend on a cell image quality and when difference between cell or kernel and background is small, most of the methods do not work properly. We have analyzed a wide range of cell images and think that cells` shape can be extracted through edge detection and thinning approach.

In this paper, we propose a new algorithm for thinning of colour histological images that is based on thinning of pseudo gray-scale images. To extract accurately gray-scale levels, we propose a new coordinate system for color representation and propose an algorithm to thin separately gray-scale images.

It is possible to define three main approaches that are mainly used for processing of color images [5,6]:

1) Separate processing of colours. Each layer, corresponded to R,G,B (or another system of coordinates) is processed independently like gray-scale image. After processing, results grouped again in some way into one colour image.
2) Process only one component of a converted colour image. For example, RGB image may be converted into the HSV space, than it is sufficient to process only V component (which contains information about brightness), combine the result with the original components H and S, and convert back to RGB representation.

Simultaneous processing of pixels presented as points in a colour space like vectors. In this technique, we process color features along the vectors in the space, where image processing is performed.

A. Color Coordinate Systems

We construct a transformation of a colour image into a pseudo-colors description by introduction of a special coordinate system. From other side, the new coordinate system should reflect specificity of images and algorithms. Such as we are oriented on mathematical morphology algorithms [9] applied to colour histological images that have some special features, we can build a corresponding coordinate system for the image representation.

One of the most important characteristics of an object in a histological color image is its chromaticity. If image fragments have an abrupt jump in chromaticity, they are either objects or their parts. However, chromaticity almost does not influence the object topology. From other side, if we process one coordinate instead of two, we get a twofold gain in the speed. In most operations of thinning, the image processing consists of much more iterations, hence this gain is essential.

Thus, processing of vector of color distance between the origin of coordinates and the de-sired point can be an advantageous. It equals the sum of the vectors of the luminance and saturation or the sum of the basis RGB vectors which directly specify the color. Therefore, it contains information about both the luminance and saturation and is a gray-scale value, which is most appropriate for the thinning on color images. That is why we propose to introduce the co-ordinate system where one of the axes is the vector of the color distance, and another, the quantity which features chromaticity.

Let us consider a cylindrical coordinate system whose central axis is a result of the vector summation of axes of RGB system and has a meaning of the monochromatic component. First of all, it should be noted that chromaticity, which is characterized by the angle of rotation, is an independent feature of the object, whereas saturation and luminance depend upon the external-conditions. Thus, the most efficient for analysis will be employment of the vector of the color distance, which is equal to the sum of vectors of luminance and saturation or the sum of the basis RGB vectors, which directly exemplify the color. It remains to introduce a quantity which will characterize a correlation of the luminance and saturation. Such a quantity is the angle between the vector of the color distance and central axis of the frame of reference. This quantity has a meaning of the relative saturation and should not vary during image processing. The purpose of this quantity is to reconstruct the relation of the luminance and saturation in inverse transformations. Thus, in the following, we shall call this spherical system of coordinates as PHS, where P is the vector of the color distance, H is hue (chromaticity), S is relative saturation.

In order to obtain a desired system, it is necessary to carry out two rotations of the coordinate axes about 45°. We have carried out a rotation in the plane RB, and then in the plane RG. This resulted in the following transformations of the coordinates

$$\begin{cases} Z = \frac{1}{2}(R + \sqrt{2}G + B) \\ Y = -\frac{\sqrt{2}}{2}(R - B) \\ X = \frac{1}{2}(R - \sqrt{2}G + B) \end{cases}.$$

Passing from the Cartesian coordinates ZYX to the spherical ones, we obtain the PHS coordinate system,

$$\begin{cases} P = \sqrt{R^2 + G^2 + B^2} \\ H = \arctan(\frac{\sqrt{2}(B - R)}{\sqrt{2}G - R - B}). \\ S = \arctan(\frac{\sqrt{2}G + R + B}{2P}) \end{cases}$$

Relations between three color coordinate systems RGB, ZYX and PHS are presented in Fig.1.

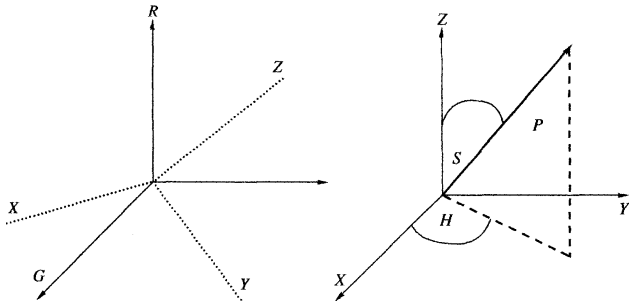

Fig. 1. Relations between coordinate systems a) RGB and ZYX, b) ZYX and PHS.

For gray-scale morphology operations we can use P component. For separation, hue H is employed. The component S is always invariable, because it corresponds to the relation of luminance and saturation. Thus, we have obtained a new system of color coordinates, PHS, which can be profitably employed for thinning of color images.

B. Color Coordinate System Oriented for Thinning

We propose to use an image described in the new PHS colour space, each pixel of the image has two main features: length of the color vector P and colour value as an angle H. Both components of the colour image may be investigated as gray-scale image, and we can apply an algorithm of the gray-scale thinning for P component of our color image description.

Membrane is an elongated object. In our opinion the Zhang-Suen algorithm [12] is optimal for thinning of elongated objects with small width. This algorithm was developed for binary image processing. Therefore the basic task in our research was an adaptation of binary conditions of the Zhang-Suen algorithm to a color image description. In the binary case each pixel has a Boolean value. We propose to change the binary value of a pixel in thinning conditions by more complex logical expression described below that includes some new pixel's coordinates from the PHS space.

Let us consider the hue component H. By the Grassmann's definitions of color it has continuous values, so it is not a gray scale. When comparing the characteristics of pixels for morphological features description, the hue value H do not defines a color distance. Difference between hue angles plays a fundamental role in our approach. For example, the hue difference between 0 and 10 is equal to difference between 0 and 350. This feature can be used as an additional condition for determination of the neighboring properties of a pixel.

Thus we have a new colour coordinate system PHS, which can be effectively used for operations of gray-scale morphology. The values P and |ΔH| are the main features for our analysis.

– Connectivity between pixels in a color space depends on brightness, saturation and hue values in a small neighborhood.

For brightness and saturation in determining the connectedness of a pixel color, using a vector P, which includes properties of both gray-scale components. This value is gray and analyzed by traditionally methods. But for the hue became important features of human perception.

Human distinguishes four basic color tones. Edges of the main sector of hue circle are perceived as sharp contrast regions. In the sector with slightly sharp contrast, we can take 1/8 of the hue circle. It can to use the difference in color tone of not more than 90° for additional condition of connectedness of pixels. But if saturation decrease, perception of the hue difference increases for close colors. Thus we introduce a color pseudoangle A, it will be included in a new condition of connectedness,

$$A = H\, S_{max}/S,$$

where A is a color pseudoangle, H – the hue value of a pixels, S_{max} – the maximal saturation value of the colour image, S – the saturation value of the pixel.

Now we define a new expression of description of connectivity of two pixels in a colour image as follows:

$$(P_j >= P_x \ \text{AND} \ |A_i - A_x| < 90°) \qquad (1)$$

where P_j - the length of the pixel j spatially neighboring to the pixel x, P_x - the brightness value of pixel x, $|A_i - A_x|$ - color pseudoangle difference between two neighboring pixels. We analyze P, H and S values of 8-connected neighboring pixels in the spatial plane.

Thus pixel contains only one value for binary and grayscale image. We change the binary result in the Zhang-Suen algorithm [12] by the result of logical variable (1) saving the rest ideas of the algorithm. Condition of pixels changing are involved for a color image. In this case such

operation corresponds to shift color pixels properties from the minimum neighborhood to analyzed pixel. If properly organized iteration it even allows you to speed up processing.

After such improvement of Zhang-Suen thinning algorithm [12] the pixel values are changed by described above condition of color connectedness.

In result, the Zhang-Suen algorithm can be used to process color images.

C. Figures

All tables and figures will be processed as images. **However, IEEE cannot extract the tables and figures embedded in your document.** (The figures and tables you insert in your document are only to help you gauge the size of your paper, for the convenience of the referees, and to make it easy for you to distribute preprints.) Therefore, **submit, on separate sheets of paper, enlarged versions of the tables and figures that appear in your document.** These are the images IEEE will scan and publish with your paper.

D. Color Thinning Results

The proposed definition of connectivity provides a high-quality results of the Zhang-Suen thinning algorithm and allows to define a color skeleton of membrane in colour histological images. Testing conducted by the operation of thinning by on an image consisting of colored lines. Results was comparing with thinning components in the RGB color system, brightness component of HLS and vector of Bit-Mix transformation [13].

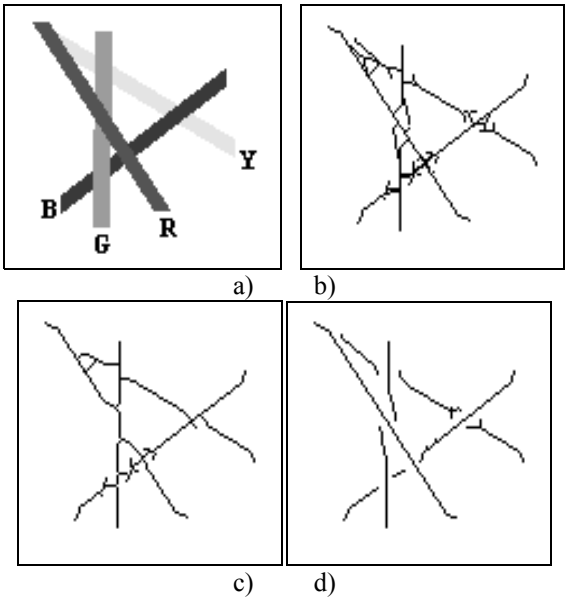

Fig. 2 Test results: a) original image, b) thinning of the merger in RGB, c) the thinning of the brightness in HLS system and method Bit-Mix, d) thinning of the brightness in PHS color space.

The skeleton of the image in the RGB system has connection for yellow color. On place mixed colors and formed irregular false branches that are joined to the middle line of objects.

Results of thinning in the HLS colour space and by the Bit-Mix method [13] are similar (Fig.2c). Different colors lines are defined as a single one complex line of a connected object with variable brightness. Proposed in this article algorithm allow to take best result.

The developed algorithm has been applied for detection membrane pattern from a histological image (Fig.3). Practical verification has demonstrated that the algorithm allows one to obtain a high-quality skeleton of color image with cell membrane (Fig.5). However, employment of the harmonic functions decelerates the process of the preparation of the image for processing. Therefore, subsequent refinement of the algorithm through the optimization of the coordinate transformation is possible.

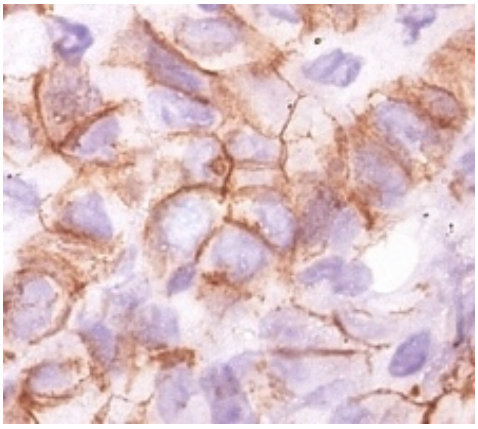

Fig. 3 Original histological image with colored membrane (brown ridges)

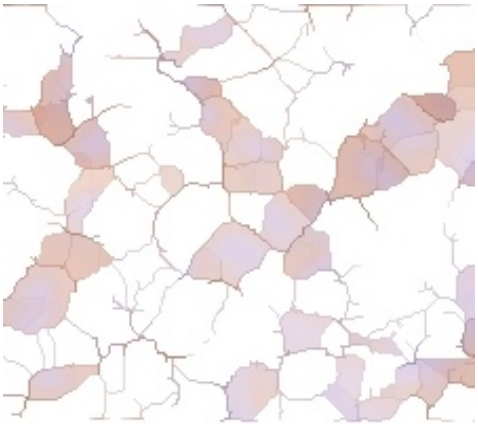

Fig. 4 Result of thinning of previous color histological image by the proposed algorithm.

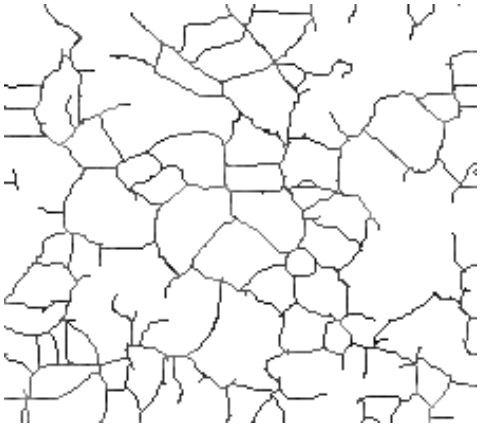

Fig. 5 Skeletons extracted from the oroginal histological image (Fig.3) with colored membrane using condition (1).

III. Clustering

Each color plane is here clustered independently by splitting its (gray-scale) histogram in several sections corresponding to representative classes of pixels in the image. We assume that the number of classes of pixels is a priori known. Furthermore, we suppose that the representative sections (or clusters) of each color plane histogram are located around its main modes. It is thus reduced to find the threshold values dissecting the histograms in a finite number of clusters.

For clustering skeleton pixels the HLS color space was used. The above presented tinning algorithm allows to decrease the number of pixels as result Hue-180° histogram was changed (Fig.6-7).

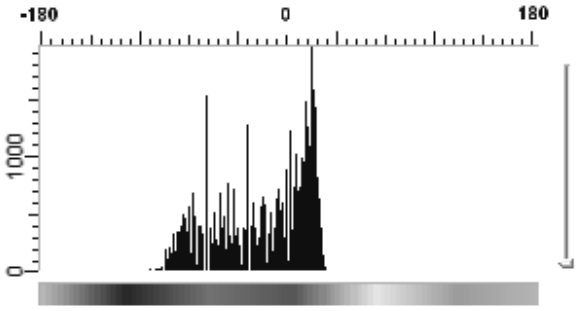

Fig. 6 The Hue histogram of an original color image (Fig.3) in the range [-180°, +180°].

After thinning of the original image and skeletonization, the hue histogram has only one main peak close to brown-red colors. In the original colour image it corresponds to main colours of membrane. This pick corresponds to membrane regions. Therefore it can be used as a center of a basic cluster corresponding to the membrane pattern. Thresholds may be selected by shifting from the mode of the Hue histogram based on a skeleton on the standard error value.

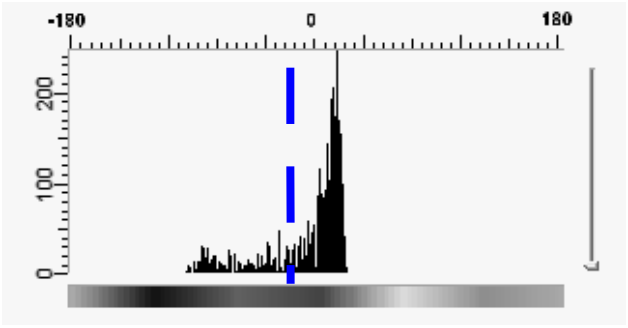

Fig. 7 The Hue histogram of skeleton pixels from Fig.5 used as a mask for the original histological color image presented in Fig.3 and a binarization threshold depected by the dashed line.

Values of brightness and saturation in the HLS color space were clustered by k-mean method into tree classes:

1) stoma;
2) membrane and nuclear;
3) all other objects.

In result we define regions for nuclear and membrane. Clustering of color skeleton by hue and consolidation of this results with the results of brightness and saturation clustering allow us to select regions containing only membrane (Fig 8-9).

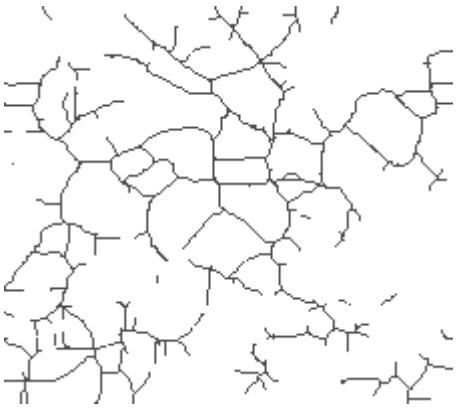

Fig. 8 Corrected skeleton of membrane after clustering.

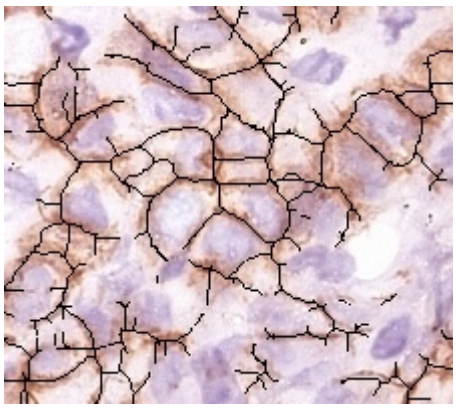

Fig. 9 Result of membrane pattern extraction that represented as layer in the original image from Fig.3.

Common scheme of a membrane pattern extraction algorithm is represented in Fig.10.

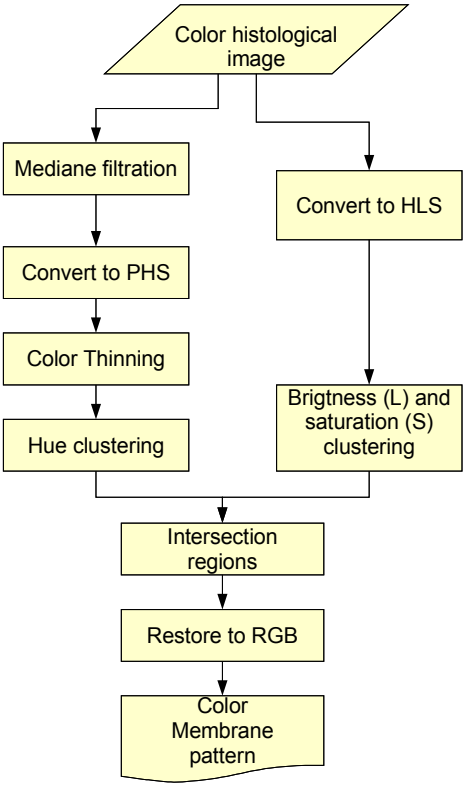

Fig. 10 Result of membrane pattern extraction that represented as layer on source image.

IV. Analysis of membrane pattern

Serving of color properties of membrane allows to investigate feathers of membrane construction and properties of the exchange of substances between cells.

Topological features of membrane can be described by a color skeleton. The basic parameters for a membrane topological structure description are full area, skeleton, nodes and tails (Fig. 11, 12). A skeleton is a mean line of a geometrical dendrites object. Nodes are points of branching of dendrite. Tails are the last fingers of branches.

Fig. 11. Topological features of dendrites: gray – membrane body, light gray – skeleton, black – nodes.

Fig. 12. Tails of membrane: by black – a dendrites body, by gray – tails.

For description of topological features of dendrite we can calculate some characteristics: branchness, curliness, membrane length, mean membrane width, tailness, tails curliness, tails ratio.

Branchness and curliness describe complexity of dendrite. They are defined by ratio of count of nodes to length of skeleton and count of skeletons segment correspondingly. The skeletons segment is fragment of skeleton between points of nodes or skeleton ends.

$$\text{Branchness} = \text{count (nodes)} / \text{length (skeleton)}$$
$$\text{Curliness} = \text{count (nodes)} / \text{count (segments)}$$

Full length of membrane corresponds to length of skeletons.

$$\text{Dendrite length} = \text{Length (skeleton)}$$

Definition of tails allows to produce characteristics for description correlation termination properties with complexity of membrane. They are tailness, tails curliness, tails ratio.

$$\text{Tailness} = \text{count (tails)} / \text{length (skeleton)}$$
$$\text{Tails curliness} = \text{count (tails)} / \text{count (segments)}$$
$$\text{Tails ratio} = \text{length (tail)} / \text{length (skeleton)}$$

These parameters may be used for description of complexity of the membrane structure and medical diagnostics.

V. Conclusion

An new algorithm for membrane structure extraction is presented. Some characteristics for membrane shape and complexity were elaborated.

An algorithm for thinning of objects in color images has been proposed. A special color coordinate system PHS has been proposed that allows to take into account specifics of histological images and the thinning algorithm. The proposed algorithm was tested on images of complex medical objects. A comparison with other coordinate color systems shows that image thinning in the PHS colour system produces a rather high-quality skeleton of the elongated objects in color biomedical images. The algorithm was

tested on synthesized and real histological images. For all cases results demonstrated good quality.

Usage of simple image processing operations in the developed algorithm allows us to realize such processing very fast.

Acknowledgment

This work is supported by ISTC project #B-14892 and and partly by ISTC project #B-1682 and #B-1636.

References

[1] [1] W. Bocker, W.-U. Muller, & C. Streffer, Comparison of different automatic threshold algorithms for image segmentation in microscope images, Proc. SPIE Conference on Applications of Digital Image Processing XVIII, San Diego, CA, vol.2564, 1995, 230-241.

[2] [2] A. Nedzved, S. Ablameyko, & I. Pitas, Morphological segmentation of histology cell images, Proc. Intern. Conference on Pattern Recognition, Barcelona, Spain, 2000, 500-503.

[3] [3] P. Bamford, & B. Lovell, Unsupervised cell nucleus segmentation with active contour, Signal Processing, 71(2), 1998. 203-213.

[4] [4] F. Sadeghian, Z. Seman, A-R. Ramli, H. Kahar, & M-I. Saripan, A Framework for white blood cell segmentation in microscopic blood images using digital image processing, Biological Procedures Online, 11(1), 2009, 196-206.

[5] [5] J. Park, & J.M. Keller, Snakes on the watershed, IEEE Transactions on Pattern Analysis and Machine Intelligence, 23(10), 2001, 1201–1205.

[6] [6] A.Garrido, & N. Perez de la Blanca, Applying deformable templates for cell image segmentation, Pattern Recognition, 33(5), 2000, 821-32.

[7] [7] C. Changming Sun, P. Vallotton, D. Wang, J. Lopez, Y. Ng, & D. James, Membrane boundary extraction using circular multiple paths. Pattern Recognition, 42(4), 2009, 523-530.

[8] [8] S. Acton, Biomedical image analysis at the cellular level, Proc. Int. Conference on Machine Vision and Image Processing, Portrush, Northern Ireland, 2008, 27.

[9] [9] K.Z. Mao, P. Zhao, & P.H. Tan, Supervised learning-based cell image segmentation for p53 immunohistochemistry, IEEE Transactions on Biomedical Engineering, 53(6), 2006, 1153-1163.

[10] [10] C. Ortiz De Solorzano , R. Malladi , S. A. Lelièvre & S. J. Lockett, Segmentation of nuclei and cells using membrane related protein markers. Journal of Microscopy, 201(3), 2001, 404 – 415.

[11] [11] M.Hu, X. Ping, & Y. Ding, Applying fuzzy growing snake to segment cell nuclei in color biopsy images, Lecture Notes in Computer Science, 3314, 2005, 672-677.

[12] [12] R. Zhang, & R. Z. Suen, A fast parallel algorithm for thinning digital patterns, Communications of the ACM, 27(3), 1984, 236-239

[13] [13] J.Chanussot, & P.Lambert, Total ordering based on space filling curves for multivalued morphology, Proc. 4th Int. Symposium on Mathematical Morphology and its Applications, Amsterdam, The Netherlands, 1998, 51-58.