**Proceedings of the Fifth IASTED International Conference**
**BIOMEDICAL ENGINEERING**
**February 14–16, 2007, Innsbruck, Austria**
**ISBN Hardcopy: 978-0-88986-648-5 / CD: 978-0-88986-649-2**

# MORPHOMETRICAL FEATURE EXTRACTION ON COLOR HISTOLOGICAL IMAGES FOR ONCOLOGICAL DIAGNOSTICS

A. Nedzved[1], A. Belotserkovsky[1], T.M. Lehmann[2], S. Ablameyko[1]
[1]United Institute of Informatics Problems, National Academy of Sciences of Belarus,
Surganova str., 6, 220012 Minsk, Belarus
{NedzvedA, abelotser}@newman.bas-net.by
[2]Department of Medical Informatics, Aachen University of Technology (RWTH)
Pauwelsstr. 30, 52057 Aachen, Germany
lehmann@computer.org

**ABSTRACT**
The diagnostics of oncological diseases is based on histological specimens in hematoxilin-eosin staining. Since manual evaluation of microscopy images is time consuming and depends on the human expert, several approaches to automatic image analysis and classification have been published. In such systems, feature extraction usually relies on a fixed resolution and a small number of numerical features. Contrarily, this framework is based on a morphometric study using two levels of optical magnification (50 and 200 times, correspondingly). In this paper, we propose a principle scheme for automation of the oncological diagnostics, and an algorithm of morphometric feature extraction of tissue fragment at low magnification. In particular, patterns of cells, vessels, and fragments of tissue are considered individually and combined for correct identification of objects extracted from the specimen. The fact of invasion is established automatically after this procedure as well as polymorphism, polychromism and anaplasia. Using this method, diagnostics of 86 out of 100 patients was confirmed.

**KEY WORDS**
Image processing, histological specimens, morphometric feature analysis, classification.

## 1. Introduction

Modern computer support and facilities in microscopy bring new perspectives of studying of cells structures. At the same time, the most commonly used method for tissue analysis is still the oldest one, i.e., the morphological method, which gives reasonable biological conclusions.
The group of morphological features, which are used for extraction of similar types of cells and for organ analysis and tissue fragments, is noticeably extending. Usually, there is no any relation between different types of features. Therefore, types of histological tissue fragment are separated from each other at their morphological features. A systematization of histological objects is very important in order to provide a morphological analysis and oncological diagnosis.
There are various approaches to segment biomedical images. One of the most popular of them is based on mathematical morphology. Many morphology-based algorithms for cell segmentation have been proposed through the years [1-3]. The initial image segmentation is determined by classifying the image local variation information obtained with dilation and erosion operations. A median filter is then can be used to smooth the initially segmented image. A median filter is applied subsequently to correct possible classification errors inside the cells. The modified median filter can clear small regions of misclassified pixels while avoiding significant changes to the cell profiles. An erosion operation is finally used to restore the cell regions.
Edge-based segmentation can be divided into two independent stages: edge detection and edge linking. Obtained edges are used to determine the cell location and contour model is further used to select the set of edges involved in this cell location. In paper [4], authors propose an edge-based potential aimed at the elimination of local minima due to undesired edges. This approach integrates knowledge about the features of the desired boundaries apart from gradient strength and eliminates local minima, which makes the segmentation less sensitive to initial contours.
Color is an important feature in histological image segmentation. There are several effective algorithms to automatically detect cells and other histological objects. For segmentation, RGB space and Lab space are combined to segment nuclei. The candidate cancer cells are selected using some morphological features of nuclei. Due to specifics of specimen preparation and painting a method for nonlinear color quantization based on human color perception [5] can be applied for segmentation here.
This paper is organized as follows. In Section 2, we discuss the classification of objects in histological specimens. Section 3 is devoted to principle scheme of morphological analysis of histological specimens, which are studied at more than one magnification. There are plenty of methods analyzing cells and their features at

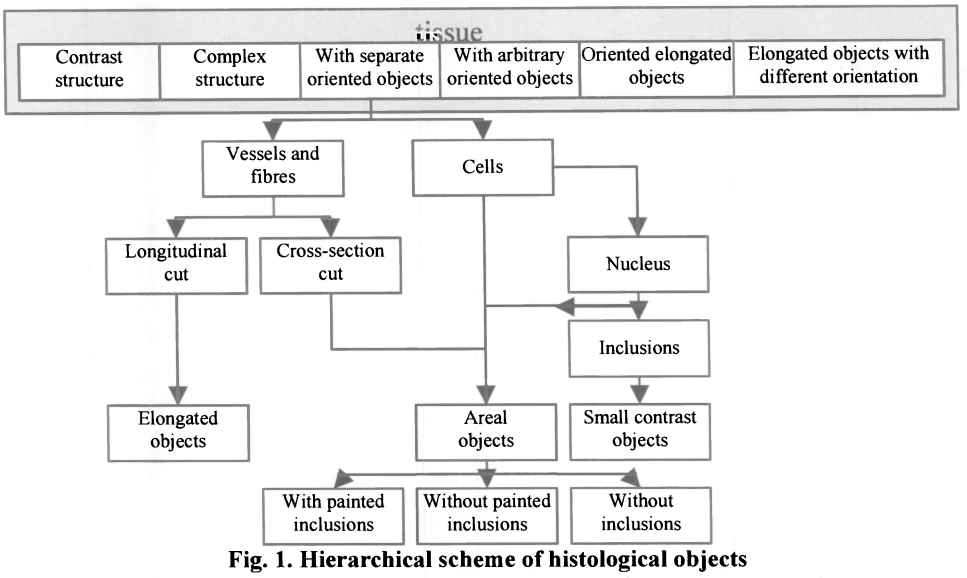


**Fig. 1. Hierarchical scheme of histological objects**

high magnification. However, these features are insufficient to make a truthful diagnostic decision. As an initial step, we propose a new approach to image analysis and feature extraction for morphological analysis of specimens at low magnification (Section 4). In particular, the processing of objects of interests is adapted to cells and vessels (Section 5) and tissue fragments (Section 6). Classification results of the experiments that are based on 100 patients are given in Section 7, followed by a discussion (Section 8) and conclusion (Section 9).

## 2. Classification of objects of histological specimen images

Usually, an image is decomposed to analyze the specimen. Therefore, the segmentation process (i.e., extraction of homogeneous regions) is considered as the basic step for formal scene description. It is necessary to define a correct set of features and feature characteristics for a suitable choice of segmentation methods.

The histological objects that are to be extracted are determined according to tasks to be solved. Automated histological specimen analysis is based on topological features of images. It allows to define the whole procedure of study for object extraction. However, automatic analysis of histological specimen depends on the optical magnification of the image. There are a certain group of topological features of tissue and its components for each magnification. This fact has prompted to consider histological objects over magnification of histological specimens.

General scheme of hierarchical analysis of histological objects is presented in Fig.1. Different tissue fragments, which are composed of group of homogeneous cells and fibers, form an entire image of histological specimen. Commonly, these fragments being objects of interest show a certain texture. Therefore, a region growing approach is used for object extraction.

In accordance with two-stage processing of histological specimens, the first levels of object hierarchy are tissues at low optical magnification (50–100x). The image processing of these objects makes it possible to gain a pattern of tissue.

The second level is composed of cells, fibers and vessels as objects of interest that form fragments of tissue [6]. A correlation of grey-level background and objects characteristics is specific for these images. The background is covered mostly with hollows of very small cells, fibers, and other particles.

Furthermore, the background usually contains different clutter and noise, which may appear at the step of registering an image. Hence, the background is characterized by mostly homogeneous pixel intensity excepting single pulses. However, the pixel intensity of the background depends on the tissue density distribution, the quality of the section, illumination, and electronic noise [7].

Two types of vessels and fibers are distinguished for longitudinal and cross-sectional cuts. Cross-sectional cut vessels and fibers are ring-like objects which gird the region with other densimetric characteristics. Inside the fibers, pixel intensity is always uniform, but particles of substance such as blood might appear in vessels. Segmentation of these images is the same as of cells when it carried out by geometrical parameters [8]. In longitudinal cuts, vessels and fibers are shown as elongated dendritic objects. It is quite complicated to determine such an objects because of their inconstant intensity, which varies with the intersection thickness or based on object overlapping. Therefore, a centerline or skeleton optimally describes longitudinal cut objects.

Cells are more complex objects than vessels. As for the cross-sectional cuts of vessels, width and shape are very

important for cells, too. Furthermore, it is essential for tissue identification to take into account the localization of cells, and different geometries characterize the cell. Methods for segmentation depend on grey-level characteristics of images of cells, background and of it correlations [9]. So, cell images are classified as follows:

1. Individual cells of the same type and any other objects differ from background with respect to the pixel's intensity. Therefore, thresholding methods are used for segmentation.
2. Individual cells of the same type and the object's background may vary uniformly. Therefore, mathematical morphology is used here.
3. Individual cells may show pixel's intensity close to background values, or other objects or noise is present in the images. Therefore, a region merging approach is applied for segmentation.

A classification of topological and geometrical characteristics such as shape, size, presence of nucleolus, or inclusion is performed to determine the appropriate the type of cells. Accordingly, the cell's nucleolus as the next type in hierarchy is also divided into these three groups.

## 3. The principle scheme of morphological analysis of histological specimens

The proposed scheme of histological objects and structures plays a key part in automated analysis of tissue and its components when coping with inter-individual variations of histological objects. In general, the geometric characteristics, which are important for preliminary diagnostics, are changing. Therefore, overall tissue characteristics are studied at a low magnification. In particular, these are the tissue structure and shape, the presence of polyferation, tissue uniformity, etc. Characteristics of cells and its surrounding are determined at high magnification (500–2000x). At low magnification, tissue fragments are visible (Fig. 2).

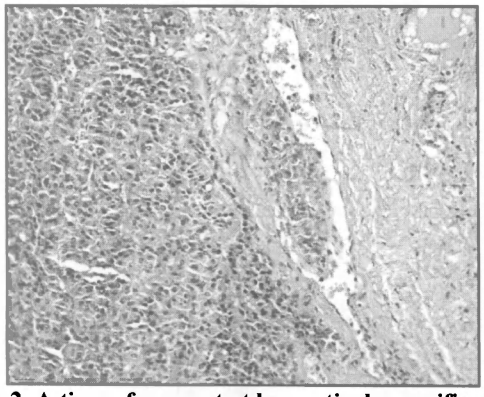

**Fig. 2. A tissue fragment at low optical magnification**

The cells presented here form small contrasted objects. Clusters of cells form tissue regions, which are functionally different from each other. Hence, the features of cell groups mainly characterize the tissue structure. It is necessary to carry out the analysis of geometric, topological and texture characteristics of cells to appreciate tissue fragments at oncological abnormalities, such as polymorphism, polychromism, and anaplasia.
Summarizing all aforesaid, we describe a morphological analysis, which is discovered in general as follows:

1. On a low optical magnification (50–100x), a presence of tumor is determined based on these characteristics and features of tissue and cell clusters:

- Uniformity of tissue layer:
  - Smoothness of layer edge;
  - Form-factor of layer edge;
- Cells organization of layer in tissue:
  - Distribution of cells by its area in layer;
  - Orientation of cells in layer;
  - Intercellular distance.

2. On a high optical magnification (500–2000x), a diagnosis is verified based on the cell characteristics:

- Morphometry of the cell:
  - Size;
  - Shape;
  - Nucleocytoplasmic ratio.
- Morphometry of the nucleus:
  - Size;
  - Shape;
  - Inner topological structure:
    (a) sulci;
    (b) inclusions;
    (c) mitoses (pathology in nucleolus organization).

## 4. Automated processing of histological images at low magnification

Based on the classification of histological images and the principle scheme of morphological analysis, a novel algorithm for the analysis of tissue images is proposed. This algorithm consists of the following steps, where each step is classified by objects of interest in histological images (Fig. 3):

1. Color conversion (RGB to CIELAB);
2. Stepwise decomposition of image;
3. Thresholding of binary patterns of cells and functional fragments of tissue;
4. Edge detection of vessel owned regions, identification of vessels;
5. Invasion determination by image comparing.

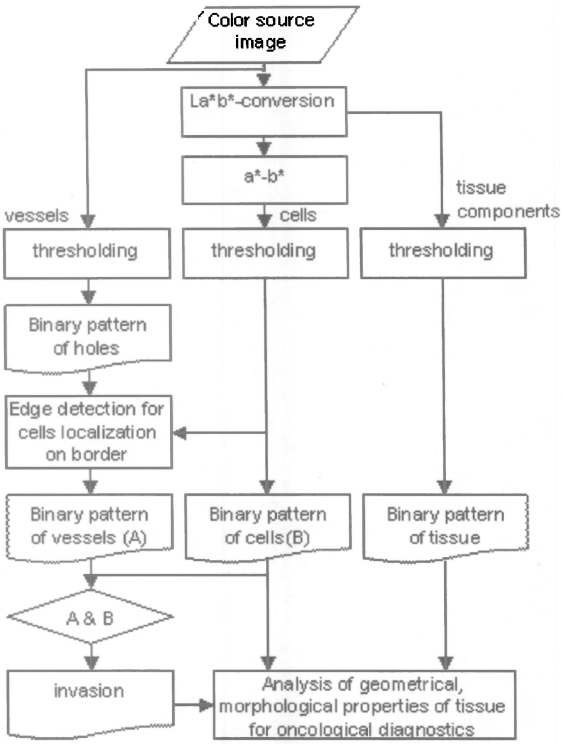


**Fig. 2. A tissue fragment at low optical magnification**

## 5. Morphological feature analysis of cells and vessels at low magnification

The first stage of histological image segmentation at low magnification extracts cells. This region extraction using intensity and color is based on a coloration of specimens. There are several types of staining and red, violet, blue, green, orange and brown are prevalent historically in histology. Straightforwardly, region extraction is done by decomposing the image on color components. Since these colors are the nearest to orthogonal components in opponent color models, the most effective way for the cell extraction is to provide a decomposition in CIELAB color system (Fig. 4).

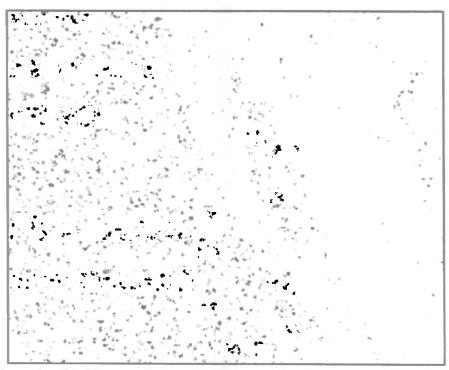

**Fig. 4. Cell nuclei at low optical magnification**

The bulk of a vessel is an unfilled space (Fig. 5), which may actually represent ruptures or artifact. Hence, thresholding itself is insufficient for vessel extraction. Additional edge analysis of these unfilled spaces is performed by mathematical difference of erosion and dilatation.

**Fig. 5. Unfilled space of vessel (capsule)**

According to the scheme presented in Figure 3, it is significant to localize cells which are present on the tissue surrounding the vessels (Fig. 6). If such cells are present, a vessel is identified by the unfilled area. Else, this area is regarded as an artifact, and it is impossible to determine an oncological disease.

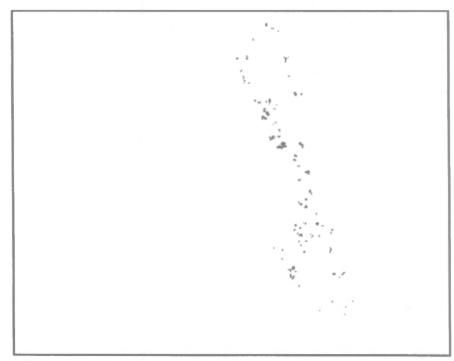

**Fig. 6. Boundary cells of vessel's capsule**

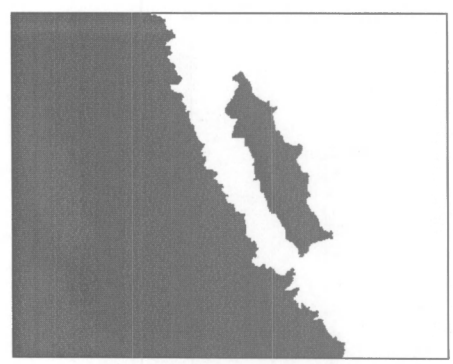

**Fig. 7. Cell cluster's area**

The next step is to compare two patterns of vessels and cells to localize a cells relatively vessels with logical

"and" operator. Pathology (invasion) is the case when cells are inside of the vessel. Another feature of oncological disease is a tumor encapsulation into space (expansion). For classification of this pathology, a cluster of cells is matched to the unfilled area of a vessel (Fig. 7).

## 6. Morphological feature analysis of functional tissue fragments at low magnification

Usually, the extraction of functional areas of tissue is based on common tissue characteristics such as staining, size, density, and texture. A tissue fragment that is labeled as background is represented by one of the color components in Lab color system (CIELAB color space). For hematoxilin-eosin staining, tissue is colored in red. Thresholding the b - component (Fig. 8) allows to gain a binary tissue pattern (Fig. 9).

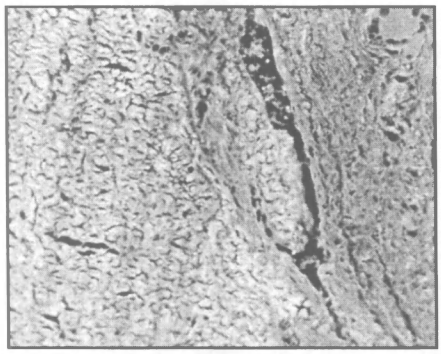

**Fig. 8. Tissue image in La*b* (b*-component)**

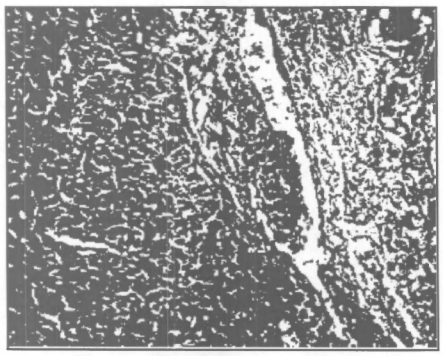

**Fig. 9. Binary image of tissue**

The distance between extracted objects is the most important texture parameter [10], and several ways of its calculation can be applied to functional tissue fragments at low magnification (Fig. 10).
The following texture characteristics are based on the distance and used to specify the cell interactions in tissue, which is very important for tumor diagnostics:

- intersection number of horizontal and vertical grid imposed on image with object boundaries;
- stereological parameter equal to overall length of horizontal and vertical lines of the grid;
- stereological parameter equal, to overall length of horizontal chords imposed on grid, which are inside of objects;
- total anisotropy of binary image relatively to coordinate axes.

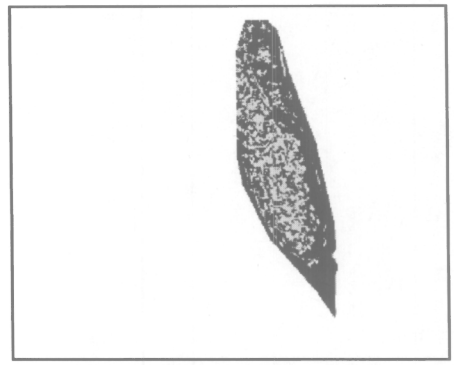

**Fig. 10. Detected cells regions**

## 7. Results

To describe a texture of cells, common statistical characteristics of object sample are used which allow to provide classification of tissue fragment and to define preliminary factors of disease. These are like statistical parameters of sizes and shape of nuclei as well as its orientation. Statistical deviations from normal distribution testify to atypical localization of cells in the tissue, which is abnormal. Testing was performed on binary patterns of nuclei of tumor's tissue areas. A binary image and interactively extracted regions of tissue was matched. The proposed approach is compared with morphological segmentation [1], edge detection, snakes [4] and clasterization in color space [5]. Closeness to initial image was defined with respect to Hausdorff metrics and root mean square error (Table 1).

**Table 1. Quality and speed measures for tissues using different approaches**

| | Hausdorff | RMS error | Processor's ticks |
|---|---|---|---|
| Morphological segmentation | 0,22 | 0,21 | 398 |
| Edge detection | 0,38 | 0,19 | 162 |
| Active contours | 0,31 | 0,18 | 1056 |
| Color clasterization | 0,20 | 0,17 | 1022 |
| Proposed approach | 0,21 | 0,18 | 454 |

For the medical eligibility, the proposed approach was checked on images of histological specimens of different kinds of tumor. The rate of recognition was determined as the ratio of automatically extracted cells to number of cells which were extracted interactively. A difference in results has appeared for different specimens of one

diagnosis. It explains why ranges of recognition rates are presented in the table bellow (table 2).

**Table 2. Quality and speed measures for cells using different approaches**

| | Carcinoma | Struma | Adenoma |
|---|---|---|---|
| Morphological segmentation | 0,6-0,7 | 0,7-0,9 | 0,7-0,8 |
| Edge detection | 0,4-0,7 | 0,6-0,8 | 0,5-0,7 |
| Color clasterization | 0,5-0,7 | 0,6-1,0 | 0,6-0,8 |
| Proposed approach | 0,6-0,7 | 0,8-1,0 | 0,7-0,8 |

The success of recognition strongly depends on quality of histological specimen preparation. It explains a wide dispersion of recognition rate however a comparative analysis of results allows to affirm preferences in time and quality of proposed algorithm.
On the basis of defined features and characteristics scheme for automated morphological analysis is worked through. In total, 100 verified patients were tested so far, and 86 times the result was confirmed. Misclassification occurred due to poor specimen preparations and optical leakages of the microscopes.

## 8. Discussion

Regarding a qualitative analysis of specimens, a single-scale definition is sufficient for most features. As shown, the second level of analysis of histological specimen is for accurate definition of diagnosis and an expert decision is reached after that. Segmentation is a key procedure to extract all objects preserving it topology and geometry, but no matter how truthful the result of image segmentation there are a plenty of features outside of consideration if we study one magnification only. Moreover, it is very important to consider patterns of cells, vessels and fragments of tissue together for correct identification of object extracted from image.

## 9. Conclusion

In this paper, we classified a histological objects by it topological properties, which has allowed to define a set of tumor characteristics in terms of computational geometry. We proposed the algorithm to extract cells and vessels for oncological diagnostics. A malignancy of detected pathology is diagnosed performing an analysis of the specimens at high magnification. Therefore, further investigations will be guided to adapt and to elaborate a methods for image processing of such an images.

## Acknowledgement

This work is partially supported by INTAS project No 04-77-7067.

## References

[1] A. Nedzved, S. Ablameyko, I. Pitas, Morphological segmentation of histology cell images, *Proc. Int. Conf. on Pattern Recognition*, 1, 2000, 500-503.
[2] J. Costa, N. Mascarenhas, M. de Andrade Netto, Cell nuclei segmentation in noisy images using morphological watersheds, *SPIE, 3*(164), 1997, 314-338.
[3] Wu HaiShan, J. Barba, J. Gil, Morphological segmentation of textured cell images, *Imaging Science Technology, 40*(3), 1996, 265-335.
[4] X.M. Pardo D. Cabello Biomedical active segmentation guided by edge saliency Pattern Recognition Letters, 2000, Vol. 21(67): 559– 572.
[5] A. Belotserkovsky, S. Ablameyko, Adapted Posterization of Color Image in HSB Space *International Journal of Pattern Recognition and Image Analysis, 15*(1), 2005, 125-129.
[6] A. Nedzved, S. Ablameyko, E. Oczeretko, Extraction of nerve cells in images with herpetic infections. *Annual Proceedings of Medical Sciences, 50*, 2005, 284-288.
[7] P. Sahoo, S. Soltani, A. Wong, Y. Chen, A survey of thresholding techniques. *Computer Vision, Graphics and Image Processing, 41*, 1988, 233-260.
[8] A. Pitiot, H. Delingette, P. Thompson, Automated image segmentation: Issues and applications, *Medical Imaging Systems Technology, 3*, 2005, 1-47.
[9] S. Petushi, C. Katsinis, C. Coward, F. Garcia, A. Tozeren, Automated identification of microstructures on histology slides. *Proc. IEEE Int. Symposium on Biomedical Imaging: Macro to Nano,* 1, 2004, 424-427.
[10] V.V. Starovoytov, *Local Geometrical Methods of Digital Processing and Image Analysis* (Minsk: IEC NAS of Belarus, 1997).