

*Проведено дослідження ознак областей, контурів та ознак текстури для сегментації біомедичних зображень. Розроблено метод і алгоритм сегментації цитологічних зображень на основі попередніх розміток. Розроблено метод і алгоритми текстурної сегментації гістологічних зображень на основі просторових моментів. Розроблено алгоритми оцінки якості сегментації на основі метричного підходу. Наведено результати порівняння якості сегментації, проведеної розробленими алгоритмами*

*Ключові слова: гістологічні та цитологічні зображення, сегментація, попередні розмітки, просторові моменти, оцінка сегментації*

*Проведено исследование признаков областей, контуров и признаков текстуры для сегментации биомедицинских изображений. Разработан метод и алгоритм сегментации цитологических изображений на основе предыдущих разметок. Разработан метод и алгоритмы текстурной сегментации гистологических изображений на основе пространственных моментов. Разработаны алгоритмы оценки качества сегментации на основе метрического подхода. Приведены результаты сравнения качества сегментации, проведенной разработанными алгоритмами*

*Ключевые слова: гистологические и цитологические изображения, сегментация, предыдущие разметки, пространственные моменты, оценка сегментации*

# DEVELOPMENT OF ALGORITHMS FOR BIOMEDICAL IMAGE SEGMENTATION BASED ON PRELIMINARY MARKUP AND TEXTURE ATTRIBUTES

**Yu. Batko**

PhD, Senior Lecturer\*

E-mail: programmer\_tang@ukr.net

**N. Batryn**

PhD, Associate Professor

Department of International Economic Relations\*\*

E-mail: nbatryn@tneu.edu.ua

**G. Melnyk**

PhD, Associate Professor\*

E-mail: mgm@tneu.edu.ua

**S. Verbovyi**

Lecturer\*

E-mail: vso@tneu.edu.ua

**T. Datsko**

PhD, Associate Professor\*\*\*

E-mail: datsko\_t@tdmu.edu.ua

**P. Selskyy**

MD, Professor\*\*\*

E-mail: selskyy@tdmu.edu.ua

\*Department of Computer Engineering\*\*

\*\*Ternopil National Economic University

Lvivska str., 11, Ternopil, Ukraine, 46020

\*\*\*Department of Pathological Anatomy

with section course and Forensic medicine

I. Horbachevsky Ternopil State Medical University

Voli sq., 1, Ternopil, Ukraine, 46001

## 1. Introduction

Biomedical images are used for diagnosis and treatment of malignant neoplasms. Biomedical images of normal and pathological cells and tissues are obtained by using light microscopes. These images are the objects of study in histology and cytology.

In order to automate analysis of histological images, the systems of automated microscopy (SAM) are applied. SAM include hardware and software parts [1]. One of the most important stages in the automated measurement of optical and geometrical parameters of cells and tissues is the segmentation of micro objects on histological images [2, 3]. The complexity of the analysis lies in the high variability of parameters and low contrast of micro objects.

Micro objects on histological images are the cells of tissues of organs. Tissues are composed of rounded cells, which

are arranged in layers. The size of the cells ranges from 0.5 to 1.2  $\mu\text{m}$ . Micro objects on the cytological images are the cells that are randomly arranged.

An analysis of a histological image involves the following stages: obtaining an image, manual or automated segmentation, measurement of the size, shape, position and optical parameters of micro objects, classification and statistical processing of measurements results.

Histological images have the following features [4, 5]:

- weak contrast due to the use of low-resolution cameras;
- images contain micro objects surrounded by the background that is complex in geometric and optical characteristics;
- non-even background is the result of incorrect configuration of the microscope illumination module;
- attributes of brightness of micro objects are the same as in the background;

- a contrast of micro objects depends on the degree of optical zoom of a microscope;
- images contain regions with repeated pattern;
- stable color palette of the image is enabled by methods of chemical dyeing. Using the specialized color spaces, it is possible to improve quality of morphological operations and speed up processing rate compared to traditional RGB color space.

Image segmentation leads to the splitting of images into regions with similar characteristics. Among the main characteristics of image segmentation is the brightness for black and white images and color component for color images. The segmentation also involves attributes of contour and texture. The process of segmentation splits the image but does not identify separate segments or their ratios [6].

At present, there are no universal methods for segmentation. In practice, a set of specialized methods are employed that are the most effective for this class of tasks. Paper [2] proposed the following requirements for segmentation:

- regions should be homogeneous for the characteristics of brightness or texture;
- part of the region should be simple in shape and have small “openings”;
- adjacent regions should significantly differ from each other by some preset characteristics;
- contours of each region should be simple, continuous and spatially precise.

Improving the segmentation of cytological images with a low contrast and non-uniform background requires evaluation of attributes of points from neighboring homogeneous regions. Such an assessment will make it possible to form the rules for building up regions following the preliminary image markup. Because the nuclei of cells are significantly darker than the surrounding background, then in the process of detection of groups of cells in histological images it is required to consider the distribution of levels of gray on the plane. We propose to estimate this distribution using spatial moments calculated for each point of the image. Application of spatial first-order moments will enable segmentation of blood vessels and ducts in the image.

---

## 2. Literature review and problem statement

---

There are several approaches to the classification of segmentation algorithms [8]. The approaches are based on the following criteria [9]: properties of points, regions, contours, *a priori* knowledge about micro objects, etc. Additional criteria for categorization of segmentation algorithms [10] are: image type (color, grayscale, binary), the nature of the segmentation process (parallel or sequential processing) [13–15]. However, these characteristics are rather ambiguous. For example, threshold segmentation may take place in parallel or under sequential mode and process both binary and grayscale images [11]. This leads to ambiguity in the algorithms of segmentation classification [12]. The criteria specified below will make it possible to better categorize segmentation algorithms.

*Algorithms based on the properties of the points.* A decision on the inclusion of a point in a homogeneous segment is made by analyzing characteristics of the point [16, 17] (brightness, color components characteristics). The advantage of this class of algorithms is speed, since the decision

on the inclusion of a point in a homogeneous region is made for each separate point. The drawback is the complexity of selection (evaluation) of preliminary information for images in a large number of homogeneous regions with similar characteristics.

*Algorithms based on texture properties [4].* A decision on the inclusion of a point in a homogeneous region is made based on the similarity of texture attributes in a given point. This type of algorithms is recommended for images with duplicate patterns.

A texture segmentation process [14–17] can be divided into 3 stages:

1) formation of a set of characteristics (attributes) for the original image, which creates a multidimensional feature space. Each vector describes attributes of a texture [7];

2) stage of the classification, where each point of the image corresponds to a certain class based on the estimation of the vector of attributes;

3) stage of the original image segmentation is based on the information obtained after classification.

The advantage of this class of algorithms is the choice of segments with the same texture, but the drawback is the complexity of the segmentation algorithms.

*Algorithms based on the selection of the contour [5].* A decision on the inclusion of a point in a homogeneous region is made by analyzing characteristics of the point of the contour. The algorithms are recommended for images with characteristic changes in brightness along the region's contour. The advantage of this class of algorithms is speed. Disadvantages are the presence of discontinuities at the edges (for the algorithms of filtering) and the difficulty of identifying the *a priori* information (algorithms of active contour).

Systems for analysis of biomedical images contain tools for manual detection of micro objects by selecting parameters for the segmentation algorithms [17]. In turn, academic studies have not addressed a problem on the automation of setting the algorithms. The results of algorithms work during research are controlled visually. Minimizing manual operations in the selection of segmentation parameters and the estimation of quality of the performed segmentation would allow automated segmentation of particular types of images.

---

## 3. The aim and objectives of the study

---

The aim of present research is the construction of algorithms for the segmentation of cytological and histological images, which would improve quality, when compared with known algorithms, under conditions of low-contrast images, non-uniformity of the background.

To achieve the set aim, the following tasks have been solved:

- to analyze relations between neighboring points and to devise rules for combining adjacent points into homogeneous regions to highlight cells;

- to analyze textural features of complex micro objects on histological images for the development of methods and algorithms for highlighting the layers of cells;

- to construct algorithms for segmentation quality evaluation and to estimate segmentation algorithms compared to the splitting into homogeneous regions by experts.

#### 4. A method of image segmentation based on preliminary markup

We shall consider characteristics of separate points in an image and the relationship between them. We shall denote:  $I$  – original image;  $I_s_i$  – original image marked up using the  $i$ -th type of markup;  $V_j$  – the  $j$ -th homogeneous region in the original image marked up using the  $i$ -th type of markup;

$M_k(x, y, z)$ ,  $x=1..l$ ,  $y=1..m$ ,  $z=1..8$  is the array of coefficients of correlations for the  $k$ -th markup,  $l$  is the width of the original image,  $m$  is the height of the original image,  $z$  is the number of neighboring pixels.

An array of summary coefficients of correlations  $M_{sum}$  is equal to (1):

$$M_{sum} = \sum_{k=1}^n M_k, \quad (1)$$

where  $n$  is the number of preliminary markups used in the process of segmentation.

*Definition 1.* We shall denote markup as the process of splitting original image  $I$  into an array of homogeneous regions  $V_j$  based on the homogeneity criterion  $KO$ . The homogeneity criterion is determined preliminary based on the analysis of original image  $I$ :

$$I \xrightarrow{KO} \{V_j\}.$$

*Definition 2.* If two adjacent points  $I(x_1, y_1)$  and  $I(x_2, y_2)$  are located in a homogeneous region, the relationship between them is equal to 1:

$$R = 1, P(I(x_1, y_1)) = P(I(x_2, y_2)),$$

where  $P(I(x_1, y_1))$  is the identifier of a homogeneous region to which point  $I(x_1, y_1)$ , belongs,  $P(I(x_2, y_2))$  is the identifier of a homogeneous region to which point  $I(x_2, y_2)$ , belongs,  $R$  is the coefficient of relationship between two neighboring points.

*Definition 3.* If two adjacent points  $I(x_1, y_1)$  and  $I(x_2, y_2)$  are located in different homogeneous regions, the relationship between them is equal to 0:

$$R = 0, P(I(x_1, y_1)) \neq P(I(x_2, y_2)).$$

*Definition 4.* The total coefficient of relationship between two pixels  $I(x_1, y_1)$  and  $I(x_2, y_2)$  is defined as the amount of relations at  $n$  markups (2):

$$Rsum = \sum R_i, i = 1..n, \quad (2)$$

where  $R$  is the total coefficient of relationship between two neighboring pixels  $I(x_1, y_1)$  та  $I(x_2, y_2)$ .

##### 4.1. Segmentation algorithm based on previous markups

The segmentation algorithm is as follows:

1) perform preliminary markup of the original image  $I$  using  $n$  markups;

2) create an array of coefficients of correlations  $M_k$  between adjacent points for each of the  $n$  markups;

3) create an array of total correlation coefficients  $M_{sum}$  between adjacent points for each of the  $n$  markups;

4) group the points of original image  $I$  into homogeneous regions based on the array of total relation coefficients  $M_{sum}$ .

Preliminary markup can be performed in three ways.

*Manual markup.* The markup of the images is performed by several users independently, by manually selecting the homogeneous regions. This approach is labor-intensive and subjective in character because preliminary markup is performed under the influence of the human factor. The advantage of the approach is that the quantity of preliminary markups can be minimal.

*Automated markup.* Marking is carried out using the known methods of segmentation, but the user assigns the input parameters independently. The advantage of this approach is: high accuracy and performance speed, improvement of the objectivity of the preliminary markup.

*Automated markup.* Preliminary markup is performed based on the automated analysis of the original image, for example, an analysis of the histogram of distribution of brightness and defining different thresholds of belonging to a homogeneous region.

Since the algorithm was designed for the segmentation of color images during preliminary automated tagging, it enables a transition from a three-dimensional representation of color to the one-dimensional. Representation of the image in a 1-dimensional space allows for the automated analysis of histograms of color distribution using known algorithms for determining the thresholds.

Preliminary markups can be applied in different color spaces.

The following rules were devised to finalize the process of segmentation in order to categorize points of the original image in homogeneous regions based on the relationships between adjacent points:

1) if a connection between two adjacent points  $I(x_1, y_1)$  and  $I(x_2, y_2)$  is maximum  $M_{sum} \rightarrow \max$  for the original image, then such points are merged into homogeneous region  $V_j$  (Fig. 1, a);

2) if the total coefficient of connection of point  $I(x_1, y_1)$  with neighboring point  $I(x_2, y_2)$  is larger than that with other neighboring points, these points are merged into homogeneous region  $V_j$ ;

3) if point  $I(x_1, y_1)$  has the same total coefficient of connection with two (or more) neighboring points

$$M_{sum}(x_1, y_1, z) = M_{sum}(x_2, y_2, z) = M_{sum}(x_3, y_3, z),$$

which are combined in homogeneous region  $I(x_2, y_2) \in V_j$ ,  $I(x_3, y_3) \in V_j$ , then it is added to the respective homogeneous region  $I(x_1, y_1) \in V_j$  (Fig. 1, b);

4) if point  $I(x_1, y_1)$  has the same total coefficient of connection with two (or more) neighboring points

$$M_{sum}(x_1, y_1, z) = M_{sum}(x_2, y_2, z) = M_{sum}(x_3, y_3, z),$$

which do not belong to one homogeneous region  $I(x_2, y_2) \in V_j$ ,  $I(x_3, y_3) \in V_j$ ,  $i \neq j$ , then the point is added to the region with which it has more neighboring points (Fig. 1, c).

The result of work of the method is a set of homogeneous regions. Since the objects in an image typically consist of a group of homogeneous regions, then it is necessary, in order to select the objects in an image, to perform a procedure for additional combination of homogeneous regions.

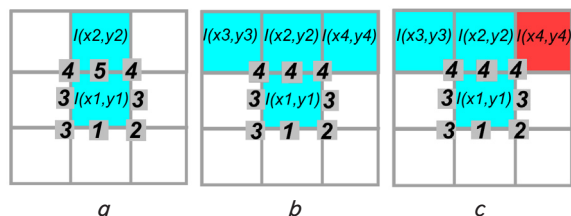


Fig. 1. Examples of merging the points into homogeneous regions: *a* – maximum connection between points; *b* – the same total connection coefficient with neighboring points; *c* – the same total connection coefficient with adjacent points that do not belong to one homogeneous region

### 5. Method of texture segmentation

A texture segmentation method involves the following stages [19]:

- a) computation of texture attributes for each point in an image within a sliding window the size of  $W \times W$ ;
- b) segmentation of the created texture field.

We shall consider textural attributes based on the spatial moments of a region and matrices of the distribution of levels of gray.

#### 5.1. Computation of spatial moments in a region

A texture of the image can be quantitatively described using simple statistical characteristics, such as mathematical expectation, variance and moments of the highest order [19]. The term of spatial moments (SM) originates from mechanics. SM, when applied to images, reflect distribution of the levels of gray image along its axes. Based on SM, it is possible to calculate region attributes that are invariant to rotation, transfer and scale.

Spatial moments of region at point with coordinates  $(x, y)$  and the value of function of levels of gray  $f(x, y)$  are calculated from (3):

$$m_{p,q} = \iint x^p y^q f(x,y) dx dy, \tag{3}$$

where  $p+q=0,1,2,3$ .

An image is considered to be a function of two variables  $f(x, y)$ ; we calculate a series of points of the lower order for each pixel in image ( $p+q \leq 2$ ). The moments are calculated within local windows the size of  $W \times W$  around each pixel.

In a discrete variant, SM within a window with a center in pixel  $(i, j)$  are calculated as the sum at normalized coordinates  $(x_m, y_n)$ :

$$m_{p,q} = \sum_{-W/2}^{W/2} \sum_{-W/2}^{W/2} f(m,n) x_m^p y_n^q, \tag{4}$$

where  $m, n$  are the coordinates of point relative to the window.

To estimate a distribution of the levels of gray, we using row-column moment of inertia  $m_{1,1}$ .

#### 5.2. Computing the matrices of distribution of the levels of gray

In the matrix of distribution of the levels of gray (MDLG) [19]  $P_d$  for displacement vector  $d = (dx, dy)$ , the value of element  $p_{i,j}$  is the number of occurrences of the pair of values for levels of gray  $i$  and  $j$ , which are located at distance  $d$ . Thus, each point in the image  $f(x, y)$  can be matched with

matrix  $P_d$ , which describes the distribution of brightness in a window the size of  $W \times W$  centered at the point with coordinates  $(x, y)$ . The elements of matrix  $P_d$  are determined from:

$$P_d(i, j) = \sum_{(m,n) \in D} f_{i,j}(x_{m,n}; x_{m+d,n+d}), \tag{5}$$

where  $D$  is the square the size of  $W \times W$  ( $W$  is an odd number);  $i, j = 0..255$  are the values of brightness of points;  $x_{m,n}$  is the brightness of point with coordinates  $(m, n)$ .

Function  $f_{i,j}(x_{m,n}; x_{m+d,n+d})$  is derived from:

$$f_{i,j}(x_{m,n}; x_{m+d,n+d}) = \begin{cases} 1, & (x_{m,n} = i \text{ and } x_{m+d,n+d} = j) \\ & \text{or } (x_{m,n} = j \text{ and } x_{m+d,n+d} = i); \\ 0, & \text{otherwise.} \end{cases} \tag{6}$$

This function is an indicator that the points that are located at a given distance have a certain level of brightness. Parameter  $d$  determines the distance at which the analysis of neighboring points is performed. Based on MDLG, we define such textural attributes as energy, entropy, contrast, homogeneity, correlation, etc.

By applying matrix  $P_d(i, j)$ , which describes the distribution of brightness inside a region centered at point  $(x, y)$ , we compute attributes of textures. After processing the entire image, we form for each of the attributes a matrix that stores a value for the attribute in all processed points – a field of texture attributes. To describe the attributes, we shall employ auxiliary magnitudes:

$$m_j = \sum_j j P(i, j), \quad p_j = \sum_j P(i, j), \tag{7}$$

mathematical expectation

$$Mx = \sum_i i p_i.$$

Below are the formulae for computing the attributes.

1. Total mean value:

$$F_1 = \sum_i m_i p_i. \tag{8}$$

2. Inertia:

$$F_2 = \sum_i \sum_j (i-j)^2 P(i, j). \tag{9}$$

3. Second angular moment:

$$F_3 = \sum_i \sum_j P^2(i, j). \tag{10}$$

4. Total correlation:

$$F_4 = \sum_i \sum_j m_i m_j. \tag{11}$$

5. Entropy:

$$F_5 = - \sum_i \sum_j \ln(P(i, j)) P(i, j). \tag{12}$$

6. Correlation:

$$F_6 = \sum_i \sum_j (i - Mx)(j - Mx) P(i, j). \tag{13}$$

The complexity of applying MDLG for the segmentation of biomedical images is in the large number of the enumer-

ated attributes, need to choose values for  $d, m, n$ . Existing SAM do not possess means to search for optimal parameters of the texture segmentation algorithm.

### 5. 3. Texture segmentation algorithm

The proposed texture field segmentation algorithm involves the following stages:

1. Construction of texture field  $G$  whose every point  $g \in G$  in accordance with (1):

$$g(i, j) = m_{i,1} = \sum_{-W/2}^{W/2} \sum_{-W/2}^{W/2} f(m, n) x_m y_n.$$

2. Normalization of  $G, g \in [0, 255]$ .

3. Search for thresholds  $t_1, t_2, \dots, t_n$  using the following steps:

a) Setting interval  $R=[a, b]; a=0$  and  $b=255$ .

b) Finding mathematical expectation  $\mu$  root-mean-square deviation  $\sigma$  of all pixels in interval  $R$ .

c) Computing thresholds  $t_1$  and  $t_2$  as  $t_1 = \mu - k\sigma$  and  $t_2 = \mu + k\sigma$ . Parameter  $k$  affects the magnitude of interval  $R$ .

d) Computing new intervals  $a=t_1+1, b=t_2-1$ .

e) Repeat steps a–d  $n/2$  times ( $n$  is the number of thresholds), setting new bounds of intervals  $a=t_1+1, b=\mu$  and  $a=\mu+1, b=t_2-1$ .

4. Segmentation of textural field  $G$ , resulting in obtaining  $n+1$  binary masks  $s_i, i=\{1, \dots, n+1\}$ :

$$s_i(x, y) = \begin{cases} 1, & \text{if } t_{i-1} \leq g(x, y) \leq t_i, \\ 0, & \text{otherwise,} \end{cases}$$

where  $t_0=0, t_{n+1}=255$ .

5. Segmentation of the original image in order to obtain  $n+1$  images  $\hat{s}_i, i=\{1, \dots, n+1\}$ :

$$\hat{s}_i(x, y) = \begin{cases} f(x, y), & \text{if } s_i(x, y) = 1, \\ 0, & \text{otherwise.} \end{cases}$$

To test MDLG (5) as a texture's attribute, it is necessary to execute a given algorithm by constructing at step 1  $g(i, j)=F_5$ . Parameter  $k$  is used to control the interval between the lowest and the largest threshold.

The optimal threshold number  $n$  (accordingly, the number of iterations of the algorithm) can be assigned *a priori*, based on actual application, or chosen based on a change in the value of signal/noise ratio  $\rho$ . Values of  $\rho$  can be calculated between the original image and the segmented image with averaged values of pixels in the middle of the segments.

## 6. Quantitative estimation of segmentation algorithms

In order to quantitatively estimate algorithms for the segmentation of cytological and histological images, we shall employ a metric approach. It implies using the Fréchet and Hausdorff metrics. The original image is segmented by an expert, that is, it is split into a set of homogeneous regions that do not intersect. The same image is processed by the constructed segmentation algorithms. The expert's set of homogeneous regions and a set after algorithm's work in a general case do not match. The level of similarity between these sets defines the quality of segmentation. To compare the contours of images, we shall apply the Fréchet metric while the Hausdorff metric will be employed to compare

regions. In order to find the smallest distances between contours, we shall use a metric by Gromov-Fréchet; between regions – by Gromov-Hausdorff. Quantitative assessment of the quality of segmentation algorithms is based on the following algorithms [22].

Consider polygonal regions.

Fig. 2 shows polygonal regions after being segmented by an expert and after segmentation executed by the algorithm.

Consider algorithms for determining the Fréchet distance and the Hausdorff distance.

### 6. 1. Algorithm for computing the discrete Fréchet distance

Let two contours  $C$  and  $R$  be assigned (Fig. 2).

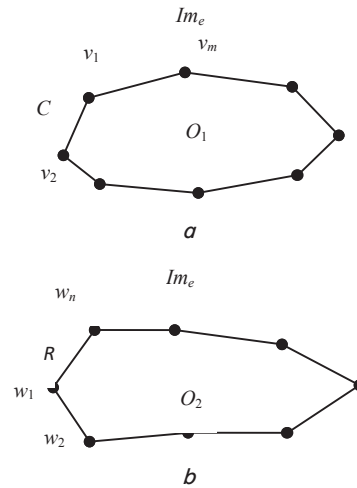


Fig. 2. Polygons: *a* – expert segmentation; *b* – algorithm segmentation

1. Represent contours  $C$  and  $R$  by sets:

$$\sigma(C) = (v_1, \dots, v_r),$$

$$\sigma(R) = (w_1, \dots, w_s),$$

(14)

where  $r, s$  is the number of segments.

2. Introduce set  $L$  between  $C$  and  $R$ :

$$L = (v_{a_1}, w_{b_1}), (v_{a_2}, w_{b_2}), \dots, (v_{a_m}, w_{b_m}),$$

$$a_1=1, b_1=1, a_m=r, b_m=s.$$

3. Then finding the Euclidean norm

$$\|L\| = \max_{i, j=1, \dots, m} (v_{a_i}, w_{b_j}),$$

will involve the following steps.

3. 1. If  $i=1$  and  $j=1$ , then the Euclidean distance is equal to:

$$d = \sqrt{(w_{b_j} - v_{a_i})^2}. \quad (15)$$

3. 2. If  $i>1$  and  $j=1$ , then the Euclidean distance is equal to:

$$\max\{d(v_{a_{i-1}}, w_{b_1}), d(v_{a_i}, w_{b_1})\}. \quad (16)$$

3. 3. If  $i=1$  and  $j>1$ , then the Euclidean distance is equal to:

$$\max\{d(v_{a_1}, w_{b_{j-1}}), d(v_{a_1}, w_{b_j})\}. \quad (17)$$

3. 4. If  $i > 1$  and  $j > 1$ , then the distance is equal to:

$$\max\{\min(d(v_{ai-1}, w_{bi}), d(v_{ai-1}, w_{bi-1}), d(v_{ai}, w_{bi-1}), d(v_{ai}, w_{bi}))\}. \tag{18}$$

Before finding distances between contours, we perform segmentation of a biomedical image and linear approximation of contours of the homogeneous regions.

**6. 2. Algorithm for computing the Hausdorff distance**

Thus, we have two polygonal convex regions  $O_1 = (v_1, v_2, \dots, v_m)$  and  $O_2 = (w_1, w_2, \dots, w_n)$ , where  $v_i (i = 1, m)$ ,  $w_i (i = 1, n)$ . Then the Hausdorff distance between regions  $O_1$  and  $O_2$  is equal to:

$$d_H(O_1, O_2) = \max\{\max_{i=1, \dots, m} d_{O_2}(a_i^1, b_i^1), \max_{i=1, \dots, n} d_{O_1}(a_i^2, b_i^2)\}, \tag{19}$$

where  $d_{O_2}(a_i^1, b_i^1)$  are the projections of vertices of region  $O_1$  onto region  $O_2$ ;  $d_{O_1}(a_i^2, b_i^2)$  are the projections of vertices of region  $O_2$  onto region  $O_1$  [24].

Projections  $d_{O_l} (l = 1, 2)$  are computed from:

$$d_{O_l}(v, w) = \begin{cases} \|(v, w) - Proj_{O_l}(v, w)\|, & (v, w) \notin Interior(O_l), \\ 0 & \text{in other cases,} \end{cases} \tag{20}$$

where  $Proj_{O_l}(v, w)$  is the minimum Euclidean distance from point  $P(v, w)$  to region  $O_l$  [22].

Then the Hausdorff distance is calculated based on steps:

1) We obtain polygonal regions  $O_1 = (v_1, v_2, \dots, v_m)$  and  $O_2 = (w_1, w_2, \dots, w_n)$ . Find distances  $d_{O_l} (l = 1, 2)$  of regions  $O_1$  and  $O_2$  according to expression (20).

2) The Hausdorff distance  $d_H$  is to be found according to expression (19).

Computing the Gromov-Fréchet distance and the Gromov-Hausdorff distance implies conducting a number of isometric transformations. That is why we shall consider the algorithm of isometric transformations of polygonal regions.

**6. 3. Algorithm of isometric transformations**

Isometric transformations are assigned by set

$$T_i = \{P, R, S\},$$

where  $P$  is the parallel carry,  $R$  is rotation,  $S$  is mapping relative to the axis (Fig. 3). Matrix of transformations in affine space is represented in the following way:

$$T_A = \begin{pmatrix} a & b & p \\ c & d & q \\ m & n & s \end{pmatrix}.$$

For the isometric transformations, the matrix is represented as:

$$T_A = \begin{pmatrix} a & 0 & 0 \\ 0 & d & 0 \\ m & n & 1 \end{pmatrix}, m, n \in R, a, d \in Z, a = d = \{-1, 0, 1\}.$$

We shall represent the algorithm of isometric transformations in the following steps:

1. Compute centers of masses  $M_1(x_{c_1}, y_{c_1})$  and  $M_2(x_{c_2}, y_{c_2})$  for regions  $O_1$  and  $O_2$ .
2. Build straight lines  $m_1$  and  $m_2$  for regions  $O_1$  and  $O_2$ , which intersect them at points  $A'(x'_1, y'_1)$ ,  $B'(x'_2, y'_2)$  and  $A''(x''_1, y''_1)$ ,  $B''(x''_2, y''_2)$ .
3. Compute angles  $\theta_1, \theta_2$  between straight lines  $m_1$  and  $m_2$  and the  $OX$  axis.
4. Perform parallel carry  $P$  of field  $O_2$  onto region  $O_1$  to match the centers of masses  $M_1$  and  $M_2$ .
5. Perform rotation  $R$  for the difference of angles between  $\theta_1$  and  $\theta_2$ .
6. Calculate the Hausdorff distance based on expression:

$$d_H = \max(\max_j \min_i d_H(O_{1j}, O_{ei}), \max_i \min_j d_H(O_{1j}, O_{ei})).$$

7. Compute the discrete Fréchet distance according to expression

$$d_{dF}(C, R) = \min\{\|L\|\}.$$

8. Compute the least distance between the polygons based on expression

$$D_i = \alpha d_{dFi} + \beta d_{Hi},$$

where  $\alpha$  and  $\beta$  are the weight coefficients.

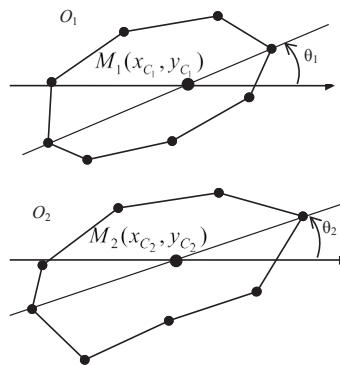


Fig. 3. Isometric transformations of polygons

Estimation of the quality of segmentation for two images will consist of the following stages: splitting the image into homogeneous regions, approximation of contours, finding the Hausdorff distance, finding the Fréchet distance, calculating the weighted sum of distances for two metrics.

**7. Experimental research into segmentation algorithms**

In order to perform computer experiments, software in the Java programming language was developed using the OpenCV computer vision library. The software is designed for preliminary processing, image segmentation and segmentation quality evaluation. Preprocessing algorithm presented in detail in [24]. Segmentation algorithms are estimated based on the devised metric approach [22].

The main purpose of the algorithm is to adapt automatically to different types of images that can greatly simplify operation compared to improving an image manually. The devised algorithm involves the following stages:

1. Determining the level of brightness, contrast, mean value of red, green, and blue channels of the RGB original image.

2. Selecting parameters for improving quality of an image depends on the input parameters obtained at stage 1. As a rule, the levels of brightness of cytological and histological images are not always uniform, which is why they require correction. The result of correcting brightness and contrast of cytological images is shown in Fig. 4.

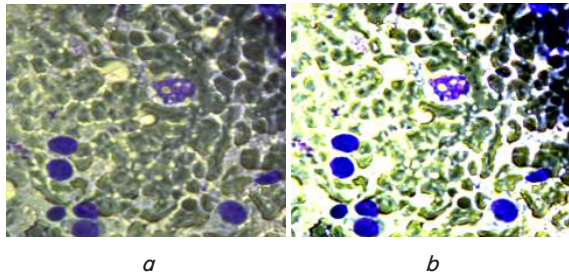


Fig. 4. Result of the correction of contrast and brightness of the image: *a* – original image; *b* – adjustment of contrast and brightness

3. Filtration of images. Results of the experiments based on a sample of 100 cytological images from database [25] have shown that the averaging filter core with  $p=3$  is the most appropriate to reduce the level of noise. Fig. 5 shows an example of the image filtration.

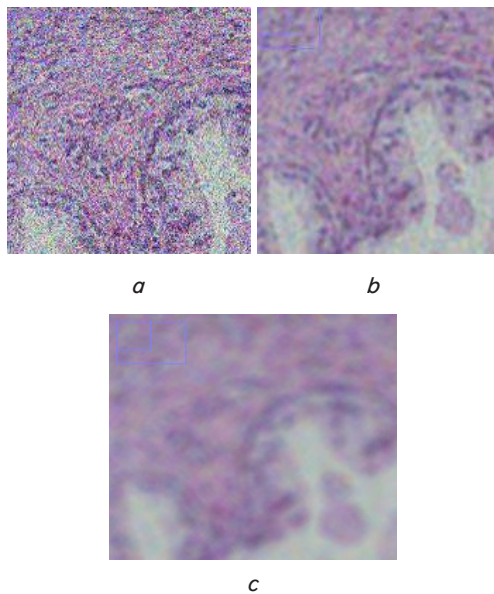


Fig. 5. Result of image filtration: *a* – noisy image; *b* –  $p=3$ ; *c* –  $p=7$

Fig. 5 shows that filter with a core of  $p=3$  reduced the noise level of the original image without distorting the image visually. Filter with a core of  $p=7$  reduced the noise level, but it significantly distorted the image.

4. Obtaining the mean values of RGB channels, which are used to perform morphological operations (“dilation”, “erosion”). Morphological operation “dilation” implies blurring the images using a filter. The image is generated from the local minimum, that is, dark regions increase. Operation “dilation” implies enlargement of the image using a filter. The image is generated from the local maximum, that is, bright regions increase. Examples of performing operations of dilation and erosion are shown in Fig. 6.

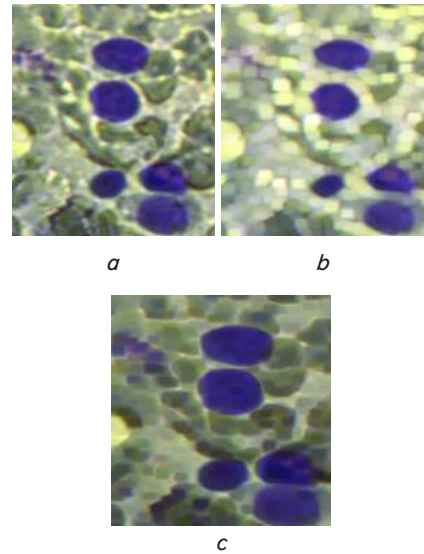


Fig. 6. Morphological operations: *a* – original image; *b* – dilation; *c* – erosion

5. Image segmentation based on the algorithm of preliminary markup.

6. Application of threshold segmentation. This segmentation algorithm is employed to convert an image into format “objects – background”.

7. Evaluation of segmentation algorithms.

The study has shown that the algorithm that is based on the previous markup identifies homogeneous regions very clearly. However, it is necessary to perform additional processing to identify micro objects in the image.

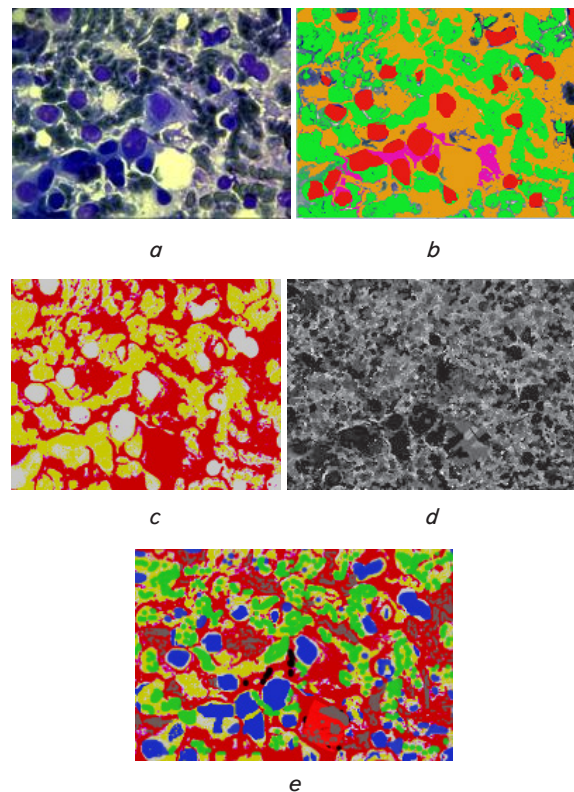


Fig. 7. Images segmented by using different algorithms: *a* – original image; *b* – expert segmentation; *c* –  $k$ -means algorithm; *d* – waterfront algorithm; *e* – algorithm based on preliminary markup

Histological images contain such complex micro objects as cross-sections of glands and ducts. Fig. 8, *a* shows a fragment of the histological image of breast tissue. To study texture segmentation algorithms experimentally, we employed a base of 100 histological images [25]. We conducted segmentation based on SM algorithm. Fig. 8, *b* shows space of texture features.

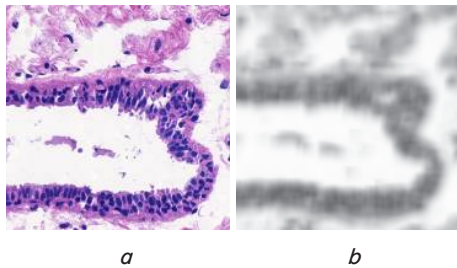


Fig. 8. Fragment of the original image and space of texture attributes: *a* – original image; *b* – space of texture attributes

Space of texture attributes is split into regions by the calculation of thresholds. The numbered regions form markup (Fig. 9, *a*). One of the homogeneous regions corresponds to a breast duct shown in Fig. 9, *b*.

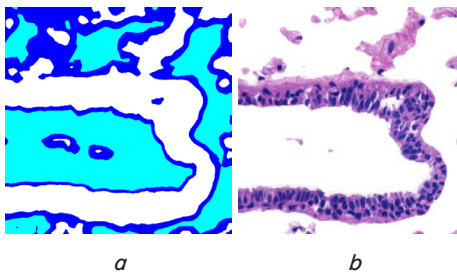


Fig. 9. Image segmentation and selection of a homogeneous region: *a* – marked up image; *b* – region of the breast duct

Table 1 gives comparison of segmentation algorithms of the cells. We applied the Hausdorff, Fréchet, Gromov-Hausdorff, Gromov-Fréchet metrics for comparison.

Table 1

Comparison of segmentation algorithms

Algorithm	Hausdorff metric	Gromov-Hausdorff metric	Fréchet metric	Gromov-Fréchet metric
Algorithm of <i>k</i> -means	67.89	64.63	67.89	64.63
Waterfront algorithm	130.9	125.7	132.76	126.1
Algorithm based on preliminary markup	54.38	53.27	54.38	53.27

Results of the texture segmentation were evaluated using a metric approach. Table 2 gives results of the comparison of texture segmentation algorithms to separate the tissue in a histological image. The third column of the metric denotes the following: Hausdorff – 1; Gromov-Hausdorff – 2; Fréchet – 3; Gromov-Fréchet – 4.

Table 2

Evaluation of texture segmentation algorithms

Expert segmentation	Texture segmentation	Distance value		
			SM	MDLG
		1	122.14	158.78
		2	119.31	173.00
		3	240.5	288.60
		4	135.11	175.64
		1	335.62	436.31
		2	323.27	468.74
		3	335.62	402.74
		4	322.94	549.00
		1	111	155.40
		2	103.56	150.16
		3	405.26	486.31
		4	389.07	505.79
		1	422.9	549.77
		2	408.3	592.04
		3	420.31	630.47
		4	405.36	672.90
		1	170.35	221.46
		2	170.35	247.01
		3	231.13	277.36
		4	231.13	300.47

The results of comparison show that the application of segmentation algorithm based on SM made it possible to improve the quality of segmentation by 18 % on average in comparison with the algorithm based on MDLG.

### 6. Discussion of results of studying the effect of properties of the points as texture attributes on the quality of segmentation

The use of connections between points for the segmentation of cytological images has the following advantages: processing images of any type, splitting micro objects that are in contact, insensitivity to noise.

The devised algorithm for texture segmentation based on spatial moments makes it possible to segment a group of cells (tissue) in a histological image. In addition to the better quality of segmentation, another benefit of the constructed algorithm is a smaller number of parameters. The study in the future should be conducted in the direction of development of an algorithm for calculating the size of a sliding window to compute spatial moments. This will allow automatic setting of parameters of the algorithm according to the multiplicity of magnification of a microscope or the scale of an image. However, the automation of processing biomedical images is hampered by the Ukrainian reality of applying manual microscopes.

The estimation of segmentation quality based on the Hausdorff distance involves convex regions, which is why a shortcoming is the need to additionally convert the non-convex regions into convex regions. This causes enhanced computational complexity. A promising direction for the



estimation of quality of segmentation is to search for metrics to compare the non-convex regions. This path implies overcoming a problem of the dependence of computing cost on the complexity of contours of objects.

The developed algorithms and software tools make it possible to automate the operations of preliminary segmentation in the systems of automated microscopy and to reduce the time of image processing. Automating manual operations leads to improved efficiency of medical diagnosis based on cytological and histological images.

---

## 9. Conclusions

---

1. We devised the method and constructed the algorithm for segmentation based on preliminary markup, which made it possible to improve the quality of cytological image segmentation by 21 % on average. A method of segmentation implies splitting a color image into homo-

geneous regions, computing a coefficient of connection between adjacent points, and merging the points into homogeneous regions. The algorithm allows for the automated segmentation.

2. We devised the method and built the algorithms for texture segmentation of layers of cells, which made it possible to improve the quality of histologic image segmentation by 18 % on average. The method involves computing the values of spatial moments for each point of the image. A feature space, obtained in this way, is segmented by the algorithm. The algorithm calculates thresholds based on mathematical expectation.

3. The algorithms of segmentation quality assessment based on a metric approach were developed. The application of the Fréchet and Hausdorff distances made it possible to find similarity between the contours and regions in the quantitative form. The use of the Gromov-Fréchet and Gromov-Hausdorff metrics allowed us to find the shortest distances between the contours and regions, respectively.

---

## References

1. Argyris, G. 3D Terrain Modelling of the Amyntaio – Ptolemais Basin [Text] / G. Argyris, I. Kapageridis, A. Triantafyllou // 2nd International Workshop in “Geoenvironment and Geotechnics”. – Milos island, 2008.
2. Berezsky, O. M. Method of image segmentation based on previous layouts images [Text] / O. M. Berezsky, Y. M. Batko, G. M. Melnyk // Proceedings of the 4th International Scientific-Technical Conference “Computer Science and Information Technology 2009”. – Lviv: “Tower and Co”, 2009. – P. 48–52.
3. Berezsky, O. M. Vision-based medical expert system [Text] / O. M. Berezsky, K. M. Berezska, Y. M. Batko, G. M. Melnyk // 6th International Scientific and Technical Conference “Computer Sciences and Information Technologies”. – Lviv, 2011. – P. 49–50.
4. Bieri, M. Quantitative analysis of Alzheimer plaques in mice using virtual microscopy [Text] / M. Bieri, A. Wethmar, N. Wey // First European Workshop on Tissue Imaging and Analysis. – 2009. – P. 31–38.
5. Muralidhar, G. S. Snakules: A Model-Based Active Contour Algorithm for the Annotation of Spicules on Mammography [Text] / G. S. Muralidhar, A. C. Bovik, J. D. Giese, M. P. Sampat, G. J. Whitman, T. M. Haygood et. al. // IEEE Transactions on Medical Imaging. – 2010. – Vol. 29, Issue 10. – P. 1768–1780. doi: 10.1109/tmi.2010.2052064
6. Berezsky, O. M. Texture segmentation of biomedical images based on spatial moments [Text] / O. M. Berezsky, Y. M. Batko, G. M. Melnyk // Proceedings of the 4th International Scientific-Technical Conference “Computer Science and Information Technology 2009”. – Lviv: “Tower and Co”, 2009. – P. 42–45.
7. Berezsky, O. Segmentation of cytological and histological images of breast cancer cells [Text] / O. Berezsky, Y. Batko, G. Melnyk, S. Verbovy, L. Haida // 2015 IEEE 8th International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS). – 2015. doi: 10.1109/idaacs.2015.7340745
8. Fu, K. S. A survey on image segmentation [Text] / K. S. Fu, J. K. Mui // Pattern Recognition. – 1981. – Vol. 13, Issue 1. – P. 3–16. doi: 10.1016/0031-3203(81)90028-5
9. Pal, N. R. A review on image segmentation techniques [Text] / N. R. Pal, S. K. Pal // Pattern Recognition. – 1993. – Vol. 26, Issue 9. – P. 1277–1294. doi: 10.1016/0031-3203(93)90135-j
10. Skarbek, W. Color Image Segmentation – A Survey [Text] / W. Skarbek, A. Koschan. – Berlin, 1994. – 81 p.
11. Lucchese, L. Color Image Segmentation: A State-of-the-Art Survey, Image Processing, Vision, and Pattern Recognition [Text] / L. Lucchese, S. Mitra // Proceedings of the Indian National Science Academy (INSA-A). – New Delhi, India, 2001. – P. 207–221.
12. Jipkate, B. R. A Comparative Analysis of Fuzzy C-means Clustering and K-means Clustering Algorithms [Text] / B. R. Jipkate, V. V. Gohokar // International Journal of computational Engineering Research. – 2012. – Vol. 2, Issue 3. – P. 737–739.
13. Thilagamani, S. A Survey on image segmentation through clustering [Text] / S. Thilagamani, N. Shanthi // International journal of research and Information sciences. – 2011. – Vol. 1, Issue 1.
14. Sharma, N. Colour image segmentation techniques and issues: an approach [Text] / N. Sharma, M. Mishra, M. Shrivastava // International Journal of Scientific & Technology Research. – 2012. – Vol. 1, Issue 4. – P. 9–12.
15. Saini, S. A Study Analysis on the Different Image Segmentation Techniques [Text] / S. Saini, K. Arora // International Journal of Information & Computation Technology. – 2014. – Vol. 4, Issue 14. – P. 1445–1452.

16. Belaid, L. J. Image segmentation: a watershed transformation algorithm [Text] / L. J. Belaid, W. Mourou // Image Analysis & Stereology. – 2011. – Vol. 28, Issue 2. – P. 93. doi: 10.5566/ias.v28.p93-102
17. Abirami, M. S. Analysis of Image Segmentation Techniques for Medical Images [Text] / M. S. Abirami, T. Sheela // Proceedings of International Conference on Emerging Research in Computing, Information, Communication and Applications (ERCICA-14). – 2014.
18. Li, C. A Level Set Method for Image Segmentation in the Presence of Intensity Inhomogeneities With Application to MRI [Text] / C. Li, R. Huang, Z. Ding, J. C. Gatenby, D. N. Metaxas, J. C. Gore // IEEE Transactions on Image Processing. – 2011. – Vol. 20, Issue 7. – P. 2007–2016. doi: 10.1109/tip.2011.2146190
19. Berezsky, O. M. Design of computer systems for biomedical image analysis [Text] / O. M. Berezsky, K. M. Berezska, Y. M. Batko, G. M. Melnyk // Proceedings of the X-th International Conference “The Experience of Designing and Application of CAD Systems in Microelectronics” CADSM'2009. – Lviv-Polyana, 2009. – P. 186–192.
20. Arbeláez, P. Contour Detection and Hierarchical Image Segmentation [Text] / P. Arbeláez, M. Maire, C. Fowlkes, J. Malik // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2011. – Vol. 33, Issue 5. – P. 898–916. doi: 10.1109/tpami.2010.161
21. Joshi, V. S. A Review of an Enhanced Algorithm for Color Image Segmentation [Text] / V. S. Joshi, A. N. Shire // International Journal of Advanced Research in Computer Science and Software Engineering. – 2013. – Vol. 3, Issue 3. – P. 435–441.
22. Berezsky, O. N. Quantitative evaluation of the quality of the segmentation of images based on metrics [Text] / O. N. Berezsky, E. N. Berezskaya // Upravlyayushchie Sistemy i Mashiny. – 2015. – Vol. 6. – P. 59–65.
23. Atallah, M. J. A linear time algorithm for the computation of some distance functions between convex polygons [Text] / M. J. Atallah, C. C. Ribeiro, S. Lifschitz // RAIRO – Operations Research. – 1991. – Vol. 25, Issue 4. – P. 413–424. doi: 10.1051/ro/1991250404131
24. Berezsky, O. Automated Processing of Cytological and Histological Images [Text] / O. Berezsky, O. Pitsun // 2016 XII International Conference on Perspective Technologies and Methods in MEMS Design (MEMSTECH). – Lviv-Polyana, 2016. – P. 51–53. doi: 10.1109/memstech.2016.7507518
25. Berezky, O. M. Relational database of intelligent automated microscopy system [Text] / O. M. Berezky, O. Yo. Pitsun, S. O. Verbovyi, T. V. Datsko // Scientific Bulletin of UNFU. – 2017. – Vol. 27, Issue 5. – P. 125–129. doi: 10.15421/40270525