UDC 004.93

Berezsky O. M.[1], Pitsun O. Y.[2]

[1]Professor, Head of Computer engineering department, Ternopil National Economic University, Ternopil, Ukraine
[2]Post-graduate student, Ternopil National Economic University, Ternopil, Ukraine

EVALUATION METHODS OF IMAGE SEGMENTATION QUALITY

Context. The basic methods of quantitative evaluation of image segmentation quality are explored. They are used to select segmentation algorithms for specific image classes. The object of the study is cytological and histological images that are used in diagnosing the pathological processes in oncology. The subject of the study is quantitative methods for segmentation algorithms' quality evaluation.

Objective. The purpose of the work is to introduce the Gromov-Fréchet metric and develop a metric-based method for quantitative evaluation of segmentation quality for image segmentation algorithms' comparison.

Method. The quantitative evaluation criteria, which are based on comparison with etalon image and without the comparison with etalon image, are analyzed. The algorithms for measuring the distances between images based on the Fréchet, Hausdorff, and Gromov-Hausdorff metrics are analyzed.

To calculate the distance between the contours of images, the Gromov-Fréchet distance was introduced. The condition of identity, symmetry and triangle is proved, and it is shown that the Gromov-Fréchet distance is a metric.

The metric-based method of quantitative evaluation of segmentation quality is developed. It is based on the use of the Gromov-Hausdorff and Gromov-Fréchet metrics. The method is based on the algorithms for non-convex-into-convex polygon transformation, weighted chord algorithm, and algorithms for calculating the Fréchet and Hausdorff distances. To calculate the Hausdorff distance between convex regions, the Atalah's algorithm was used. The Thierry and Manillo algorithm was used to find the discrete Fréchet distance. These algorithms have the lowest computational complexity among their class of algorithms.

Results. The Gromov-Fréchet metric was introduced and the metric-based method of quantitative evaluation of segmentation quality was developed.

Conclusions. The conducted experiments on the basis of cytological images confirmed the performance of software for evaluation the distances between images. The developed method showed a high accuracy of estimation the distances between images. The developed software module was used in intelligence systems for diagnosing the breast precancerous and cancerous conditions. The software can be used in various software systems of computer vision. Promising areas for further research are search for new metrics to evaluate the distances between images.

Keywords: segmentation, quantitative segmentation evaluation, Fréchet metric, Hausdorff metric, Gromov-Hausdorff metric, Gromov-Fréchet metric, polygon, cytological images.

NOMENCLATURE

AMS – automated microscopy system;

AR – area relation;

FOM – figure of merit;

NR FOM – number relation figure of merit;

AUMA – absolute ultimate measurement accuracy;

RUMA – relative ultimate measurement accuracy;

FOC – figure of certainty;

FRAG– fragmentation;

CSP – correctly selected pixels;

PDE – pixel distance error;

I – a predefined image;

I^e – an expert (etalon) image;

I^{A_j} – an image, segmented by algorithm A_j;

A_j – a segmentation algorithm;

$d(I^e, I^{Aj}) = d_j$ – a distance between etalon image I^e and segmented image I^{A_j} by algorithm A_j;

$d(x, y)$ – an Euclidean distance between points x and y;

d_H – a Hausdorff distance;

$d_2(V_2, W_2)$ – a distance between contours;

d_F – a Frйchet distance;

d_{GH} – a Gromov-Hausdorff distance;

d_{GF} – a Gromov-Fréchet distance;

A, B – compact sets;

f, g – continuous curves;

α, β – arbitrary continuous non-decreasing function;

(X, d) – metric space X with metric d ;

γ_i – parameterized curves in metric spaces,

f, g – isometric embeddings into a space;

Z – a set of isometric transformations;

P, Q, P_i, Q_j – polygons;

C_{P_i} – a contour (external boundary) of convex polygon P_i;

O_{P_i} – an internal boundary of convex polygon P_i;

$d_1(V_1, W_1)$ – a distance between regions;

D – a combined distance;

φ, λ – weighted coefficients;

V – initial set of polygon vertices;

n – a number of vertices in polygon;

m – a number of vertices corresponding to condition $\alpha_i < 180°$;

B – a set of polygon vertices after convex regions formation;

α_i – interior angles of the polygon;

O – algorithm computational complexity.

INTRODUCTION

Image processing and analysis have been widely used in computed tomography, magnetic resonance tomography,

X-ray imaging (digital radiography), histology and cytology, etc. [1]. For diagnostics in oncology, automated microscopic system (AMS) are used to process and analyze cytological and histological images.

The object of this research is cytological and histological image segmentation.

Cytological image is a microscopic image of preparations containing cells and their components (nucleus, cytoplasm) [2]. Histological image is a microscopic image of preparation thin sections of fixed tissues that reflect their structure [3].

The main disadvantages of most histological and cytological images are low quality, non-uniform illumination of areas, presence of noise, lack of clear contours between the microscopic objects and the background. It is known that image segmentation is crucial for the average computer vision. There are many algorithms for image segmentation, such as threshold, watershed distribution, *k*-means, and others. Therefore, the choice of segmentation algorithms and their parameters is of great significance. To select the optimal parameters for segmentation algorithms, it is necessary to evaluate the results of segmentation.

The subject of the research is methods of segmentation quality evaluation.

There are the following segmentation evaluation criteria: non-standard segmentation criteria that do not require standard (etalon) segmentation and standard (etalon-based) segmentation criteria.

One of the main criteria of the first group is homogeneity of segments. This approach is based on calculation of value dispersion of a certain image feature used for segmentation [4]. Another criterion for evaluating segmentation is contrast between segments. Besides, a shape of a segment is also a criterion that can help evaluate segmentation quality. This criterion can be effectively used for a specific image analysis. The advantage of using non-standard criteria is simplicity and speed.

The most famous algorithms belonging to the second group are AR, FOM, NR FOM, AUMA, RUMA, FOC, Baddeley, Average Distance, Variance distance, FRAG [5]. This group of algorithms is based on the use of metrics [6]. The results of etalon-based segmentation are compared with the results of specific segmentation algorithm. The main advantage of the second group of algorithms is high accuracy of evaluation.

There is no single unified theory of image segmentation. Most algorithms are heuristic in nature. Therefore, the actual problem is objective quantitative evaluation of segmentation quality.

The purpose of this article is to analyze and compare the current evaluation methods of image segmentation, develop a metric-based method and algorithms for quantitative evaluation of image segmentation using histological and cytological images.

1 PROBLEM STATEMENT

Let I be a predefined image (fig. 1). Expert (etalon) image I^e is obtained after expert's segmentation. Let image I^{A_j} be an image segmented by algorithm A_j. Then the distance
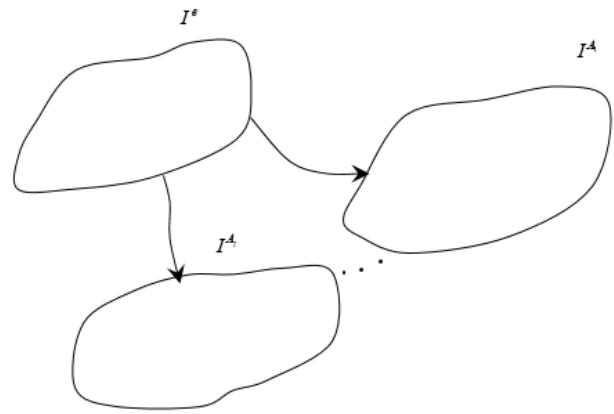

Figure 1 – Expert image and image segmented by algorithms

between an etalon image and an image segmented by algorithms A_i will be equal to $d(I^e, I^{A_i})$, $i = 1,..., l$.

It necessary to find out the algorithm A_j that meets the requirement of distance minimum $d_j = \min(d_1, d_2,..., d_l)$.

It is necessary to find out the algorithm.

2 LITERATURE REVIEW

Segmentation quality evaluation has become a research focus of many researchers since the late 80s – early 90s. Zhang, Mattana, and Huo [5] stated that it is accuracy of individual objects' selection that can serve as image segmentation evaluation. The authors considered individual characteristics of objects. To evaluate the threshold segmentation methods, Lee, Chung, and Park introduced segmentation error probability criterion based on counting wrongly classified pixels. Yasnoff and Mui [7] introduced pixel distance error (PDE) to find out the distances between pixels in the target and segmented images. The longer the distance is, the higher the segmentation error will be. Gerbrands [8] introduced FOM criterion to determine the distance between the segmented pixel and "correct" pixel location. Criterion RUMA, offered by Zhang [9], uses geometric object parameters to evaluate segmentation quality.

The analyzed criteria mainly evaluate the distances between individual objects in etalon and segmented images. In practice, it is necessary to evaluate the quality of segmentation for a group of objects.

The algorithms for comparing segmentation results by means of metrics are based on known Frйchet and Hasdorf metrics.

Thus, Lopez and Reisner [10] developed an algorithm to reduce a number of vertices of the convex polygon for a given error ξ in the Hausdorff metric. The algorithm can be used only for convex polygons. Alt and Scharfz [11] calculated the Hausdorff distance between algebraic plane curves using Voronoi diagrams. The algorithm is used for partial cases with algebraic curves and has a high computational complexity. Chew and Kedem [12] developed an algorithm for finding the minimum Hausdorff distance in metrics L_i and L_∞. The resulting computational complexity

is $O(n^2 \cdot \log^2 n)$. Knauer and Scherfenberg [13] developed a search method by a given image pattern, which has the least distance in the metric of Hausdorff. In this case, translation of the specified pattern is used to the search image. The algorithm has a high computational complexity. Alvarez and Seidel [14] developed a method for finding the minimum weight spanning tree based on the Hausdorff metric for d-dimensional space. The problem of approximation of such a tree is solved in polynomial time. Atallah [15] developed an algorithm for finding the Hausdorff distance between convex polygons. The computational complexity of this algorithm is $O(m \cdot n)$, where m and n refer to a number of vertices of the first and second polygons, respectively.

A number of publications are devoted to the development of algorithms for finding the Frйchet distance between curves. Alt and Godau [16] developed algorithms for finding the distance between parametric curves. The computational complexity is $O(mn \cdot \log mn)$. Mosig and Clausen [17] developed an algorithm for calculating the Frйchet discrete distance for polygonal curves. They used groups of rigid motion. The computational complexity is $O(m^2 \cdot n^2)$, where m and n refer to a number of fragments on the first and second curves. K. Buchin, M. Buchin, and Wenk [18] developed an algorithm for calculating the Frйchet distance for surfaces that are represented by simple polygons. The algorithm has a polynomial complexity. Rote [19] developed the algorithm for calculating the Frйchet distance between two curves given by the set of m and linearly approximated segments. The computational complexity is $O(m \cdot n)$. Schlesinger, Vodolazskiy, and Yakovenko [20] obtained the computational complexity $O(m \cdot n)$ for the Frйchet metric for closed polygonal curves. Ahn et al [21] developed an algorithm for calculating the Frйchet distance with inaccurately given vertices. For d-dimensional space they obtained the computational complexity of $O(d \cdot m \cdot n)$. Cook, Driemel, Sherette, and Wenk [22] developed an algorithm for calculating the Frйchet distance between non-flat surfaces. The authors reached the polynomial time in the L_∞ metric. Gudmundsson and Smid [23] developed a fast algorithm for finding the similarity of polygonal trees in the Frйchet metric. The algorithm has a polynomial complexity.

Eiter and Mannila [24] developed an algorithm for finding the Frйchet discrete distance between polygons. The computational complexity of the algorithm is $O(m \cdot n)$. The best algorithms for finding the Frйchet distance for plane closed curves have the computational complexity of $O(m \cdot n)$. The best algorithms for finding the Hausdorff distance for convex regions have computational complexity $O(m \cdot n)$.

The mentioned algorithms calculate the distances separately between the curves (contours) of the images in the Fréchet metric and between the regions of the images in the Hausdorff metric.

Therefore, it is necessary to develop a metric, method, and software for evaluating segmentation quality for complex images with many objects.

3 MATERIALS AND METHOD

In order to compute distances between images, we use Hausdorff and Frйchet metrics. To compute the shortest distances between images, we use Gromov-Hausdorff and Gromov-Frйchet metrics.

We present the basic metrics below.

Frйchet metric. Let X be a metric space with a metric d. For two curves $f:[a,b] \to X$ and $g:[a',b'] \to X$, the Frйchet distance between them is equal to [10]:

$$d_F = \inf_{\alpha,\beta} \sup_{t \in [0,1]} d\big(f(\alpha(t)), g(\beta(t))\big),$$

where α and β are arbitrary non-decreasing continuous functions from the interval $[0,1]$ on intervals $[a,b]$ and $[a',b']$, respectively. The value of function is $\alpha(0) = 0$ and $\alpha(1) = 1$, and it is similar for function β.

Hausdorff metric. For a metric space (X,d), the Hausdorff metric d_H will be called a metric on the set $\mathfrak{I}$ of all non-empty compact subsets X, which is defined by means of the formula [25]:

$$d_H^X(A,B) := \max\left\{\max_{x \in A} \min_{y \in B} d(x,y), \max_{y \in B} \min_{x \in A} d(x,y)\right\}.$$

Gromov-Hausdorff metric. The distance between two compact sets A and B is equal to [26]:

$$d_{GH}(A,B) := \inf_{X,f,g} d_H^X(f(A), g(B)),$$

where $f:A \to X$, $g:B \to X$ denote isometric embeddings into a metric space (X,d).

Gromov-Frйchet metric. To measure the distance between two curves, we use Gromov-Fréchet metric:

$$d_{GF}(S,Q) := \inf_{X,S,Q} d_F^X(f(S), g(Q)),$$

where $f:S \to X$, $g:Q \to X$ – isometric embeddings into a metric space (X,d). Suppose there are two parameterized curves in metric spaces $\gamma_i:[0,1] \to X_i$, $i=1,2$. (X_i, d_i) is a metric, $t \to \gamma_i(t) \in X_i$ (fig. 2). We embed isometrically $j_i : X_i \to Z$, $i=1,2$, (Z,d).

Then we take a Fréchet distance between $j_1\gamma_1$ and $j_2\gamma_2$,

$$d_F(j_1\gamma_1, j_2\gamma_2). \tag{1}$$

Finally, infimum of such (1) along all isometric embeddings j_1, j_2 will be a Gromov-Fréchet distance:

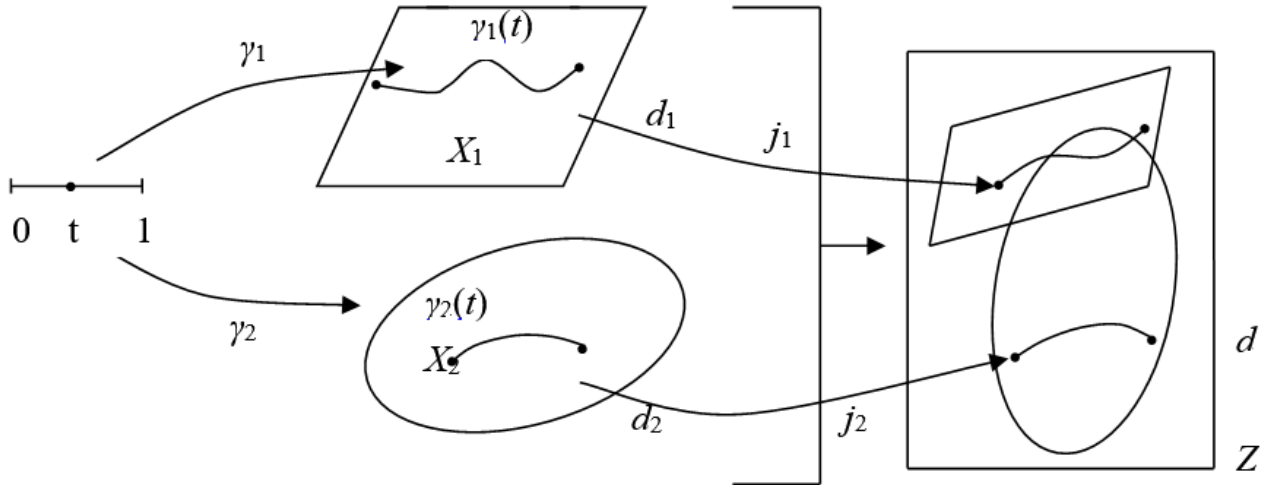$$d_{GF}(\gamma_1, \gamma_2) = \inf_{j_1, j_2, Z} d_F(j_1\gamma_1, j_2\gamma_2).$$

Figure 2 – Gromov-Fréchet metric between two parameterized curves

Theorem. The function d_{GF} is a metric.

Proof. First we check if the definition is correct. Having two spaces X_1 and X_2 we consider their bouquet $Z = X_1 \vee X_2$, sticking together two arbitrary points x_1^0 and x_2^0.

Let us define a metric d in a bouquet by means of the formula:

$$d(x,y) = \begin{cases} d_i(x,y), \text{if } x,y \in X_i. \\ d_1(x,x_1^0) + d_2(x_2^0,y), \text{if } x \in X_1, y \in X_2. \end{cases}$$

Then we check the definition of a metric for d_{GF}.

Obviously, $d_{GF}(\gamma_1,\gamma_2) \geq 0$.

Besides, $d_{GF}(\gamma,\gamma) = 0$ for each γ.

Let us denote that the definition implies that $d_{GF}(\gamma_1,\gamma_2) \geq d_{GF}(\gamma_1([0,1]),\gamma_2([0,1]))$.

If $d_{GF}(\gamma_1,\gamma_2) = 0$, then it implies that $d_{GF}(\gamma_1([0,1]),\gamma_2([0,1])) = 0$, therefore, we can assume that $\gamma_1([0,1]) = \gamma_2([0,1])$.

The last equality in case of embedded curves gives equality $\gamma_1 = \gamma_2$ exactly to parameterization. The symmetry of function d_{GF} is obvious.

Let us verify the triangle inequality. Let $\gamma_i : [0,1] \to X_i$, $i = 1,2,3$ be parameterized curves; $d_{GF}(\gamma_1,\gamma_2) = a$, $d_{GF}(\gamma_2,\gamma_3) = b$; and $\varepsilon > 0$. There are metric spaces Z_{12} and Z_{23}, and such isometric embeddings

$$j_i : X_i \to Z_{12}, i = 1,2,$$

$$k_i : X_i \to Z_{23}, i = 2,3,$$

so that

$$d_F(j_1\gamma_1,j_2\gamma_2) < a + \varepsilon,$$

$$d_F(k_2\gamma_2,k_3\gamma_3) < b + \varepsilon.$$

Let $Z = Z_{12} \cup_{k_2 j_2^{-1}} Z_{23}$. In other words, Z is derived from $Z_{12} \cup Z_{23}$ by identification of each point $j_2(x)$ with point $k_2(x)$, $x \in X_2$.

Metric d on Z is defined by the formula:

$$d(x,y) = \begin{cases} d'(x,y), \text{ if } x,y \in Z_{12}. \\ d''(x,y), \text{ if } x,y \in Z_{23}. \\ \inf\{d'(x,j_2(z)) + d''(k_2(z),y) \,|\, z \in X_2\}, \text{if } x \in Z_{12}, y \in Z_{23}. \end{cases}$$

(Here d' is a metric on Z_{12}, d'' is a metric on Z_{23}).

Then from the triangle inequality for d_F it derives that

$$d_{GF}(\gamma_1,\gamma_2) \leq d_F(j_1\gamma_1,k_3\gamma_3) \leq d_F(j_1\gamma_1,j_2\gamma_2) +$$
$$+ d_F(k_2\gamma_2,k_3\gamma_3) < a + b + 2\varepsilon =$$
$$= d_{GF}(\gamma_1,\gamma_2) + d_{GF}(\gamma_2,\gamma_3) + 2\varepsilon$$

Since $\varepsilon > 0$ is arbitrary, we obtain the required inequality.

To evaluate segmentation quality, the researchers developed quantitative evaluation method of segmentation quality (QEMSQ), which is built on metric-based measuring the distance between images.

After segmentation, we obtain a set of segments that we approximate linearly and get a set of polygons. In general, they are not convex. Thus, the task is to compare two non-convex polygons after a specific algorithm segmentation and expert segmentation.

Let P and Q be two non-convex polygons (fig. 3).

Figure 3 – Polygons P and Q

Let us divide polygons P and Q into sets of convex polygons, that is, $P = P_1 \cup ... \cup P_i \cup ... \cup P_n$, $Q = Q_1 \cup ... \cup Q_j \cup ... \cup Q_m$. Each convex polygon P_i and Q_j is represented as $P_i = C_{P_i} \cup O_{P_i}$, where C_{P_i} and O_{P_i} are contour (external boundary) and internal region of convex polygon P_i, respectively. Similarly, polygon $Q : Q_j = C_{Q_j} \cup O_{Q_j}$.

Then we receive

$$P = (C_{P_1} \cup O_{P_1}) \cup (C_{P_i} \cup O_{P_i}) \cup ... \cup (C_{P_n} \cup O_{P_n})\} \mid V,$$

$$Q = (C_{Q_1} \cup O_{Q_1}) \cup (C_{Q_{ji}} \cup O_{Q_{ji}}) \cup ... \cup (C_{Q_m} \cup O_{Q_m})\} \mid W.$$

Let us represent the latter expressions in the following form:

$$P = (C_{P_1} \cup ... \cup C_{P_i} \cup ... \cup C_{P_n}) \cup (O_{P_1} \cup ... \cup O_{P_i} \cup ... \cup O_{P_n}),$$

$$Q = (C_{Q_1} \cup ... \cup C_{Q_j} \cup ... \cup C_{Q_m}) \cup (O_{Q_1} \cup ... \cup O_{Q_j} \cup ... \cup O_{Q_m}).$$

Denote:

$$C_{P_1} \cup ... \cup C_{P_i} \cup ... \cup C_{P_n} = V_1,$$

$$O_{P_1} \cup ... \cup O_{P_i} \cup ... \cup O_{P_n} = V_2,$$

$$C_{Q_1} \cup ... \cup C_{Q_j} \cup ... \cup C_{Q_m} = W_1,$$

$$O_{Q_1} \cup ... \cup O_{Q_j} \cup ... \cup O_{Q_m} = W_2.$$

Then the distance between polygons P and Q is equal to a sum of distances between contours and internal regions of convex polygons P_i and Q_j. The distance between regions equals:

$$d_1(V_1, W_1) = \inf \{\varepsilon_1 > 0 \mid \forall i = \overline{1, n}, \exists j = \overline{1, m}, d_H(O_i, O_j) \le \varepsilon_1$$

and vice versa $\forall j = \overline{1, m}, \exists i = \overline{1, n}, $ and $d_H(O_i, O_j) \le \varepsilon_1\}$,
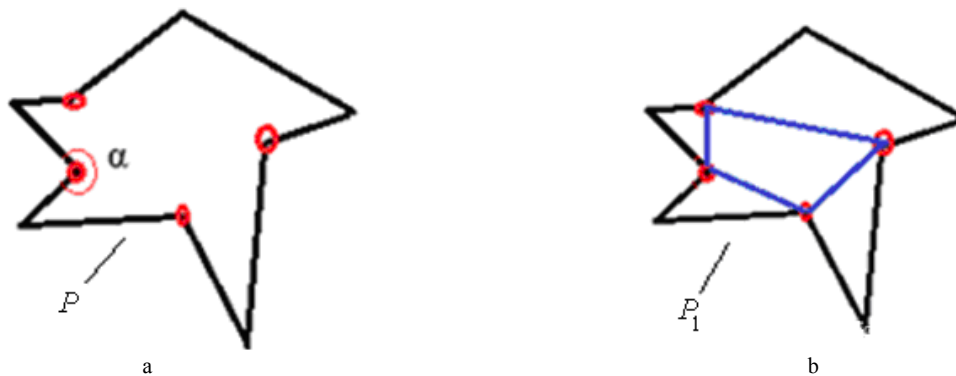
d_H − a Hausdorff distance.

Similarly, we calculate the distances between contours:

$$d_2(V_2, W_2) = \inf \{\varepsilon_2 > 0 \mid \forall i = \overline{1, n}, \exists j = \overline{1, m}, d_F(C_i, C_j) \le \varepsilon_2$$

and vice versa $\forall j = \overline{1, m}, \exists i = \overline{1, n}, $ and $d_F(C_i, C_j) \le \varepsilon_2\}$,

d_F − a Fréchet distance.

Quantitative evaluation method of segmentation quality is based on a combination of algorithms that ensure finding the shortest distances between images. This combination includes a set of algorithms: algorithm for non-convex-into-convex polygon transformation, weighted chords algorithm, Hausdorff distance algorithm, and discrete Fréchet distance algorithm.

Stepwise quantitative evaluation method of segmentation quality can be represented as follows:

1. Formation of a convex polygons' set.
2. Conducting isometric transformations for embedding convex polygons with a maximum cross section.
3. Computing a Fréchet distance for convex polygons.
4. Computing a Hausdorff distance for convex polygons.
5. Finding the shortest distance based on weighted metrics (Fréchet and Hausdorff metrics) between polygons P and Q according to $D = \varphi d_H + \lambda d_F$.

Let us describe the basic algorithms that underpin this method.

Algorithm for non-convex-into-convex polygon transformation:

Let P be a non-convex polygon with vertices $V = \{v_0, v_1, ..., v_{n-1}\}$. Algorithm for convex polygon formation consists of the following steps:

1. Beginning with an upper vertex, we select the vertices with internal angles α_i larger than 180°. If such angles do not exist, then the polygon is convex and the algorithm is completed. Otherwise, we get an array of vertices $B = \{b_0, b_1, ..., b_m\}$.

2. We connect consistently the received vertices b_1 beginning with the top vertex and get a polygon P_1.

3. We repeat step 1 with a received polygon P_1.

4. We repeat steps 1 and 2 with polygons P_1 until the angle of each vertex meets the condition $\alpha < 180°$.

The algorithm is graphically demonstrated in fig. 4

Figure 4 – Partitioning non-convex polygon into convex polygons: a – polygon P, with internal angles larger than $180°$, b – polygon P_1, convex region formation

Algorithm computational complexity is $O(n+m)$, where n is a number of vertices of initial non-convex polygon and m is a number of vertices of non-convex polygon.

Weighted chords algorithm is described in the article of Berezsky, Melnyk, Batko, and Pitsun [27] Computational complexity of the algorithm is $O(n \cdot m)$, where n is a number of weighted chords of the first polygon and m is a number of weighted chords of the second polygon. For convex regions' comparison, the Atallah's algorithm [15] was used. For contour comparison, we used the Fréchet discrete distance algorithm developed by Eiter and Mannila [25]. So, the developed QEMSQ algorithms have the least computational complexity.

4 EXPERIMENTS

For computer experiments, we used cytological images [28].

To compare segmentation quality of individual micro objects, we use etalon image shown in fig. 5.

Other images for these experiments were generated randomly. In this case, we use the following methods to evaluate segmentation quality: CSP, WSP, PDE, FRAG, AUMA, and RUMA.
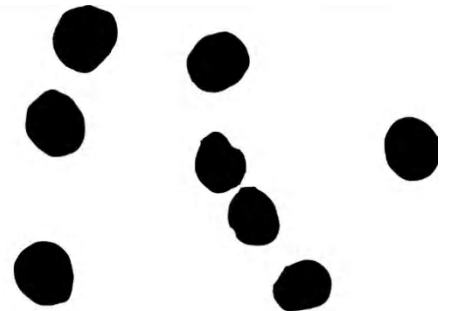
To evaluate segmentation quality of micro objects' groups, we use cytological images. Fig. 6 shows the etalon image and images segmented by means of thresholding, k-means, and watershed distribution algorithms [29]. We used Hausdorff, Fréchet, Gromov-Hausdorff, and Gromov-Fréchet metrics.
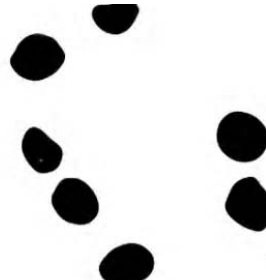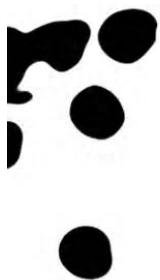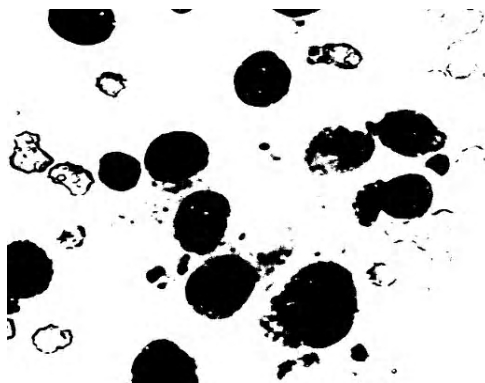

Figure 5 – Etalon image
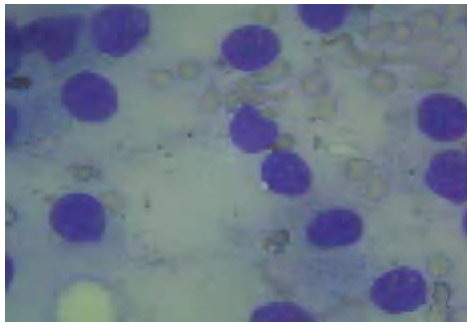
a

b

c

d

e

Figure 6 – Image segmentation: a – output image, b – etalon image, c – segmentation results by k-means + threshold algorithms, d – segmentation results by watershed + threshold algorithms, e – segmentation results by threshold algorithm

5 RESULTS

Comparative analysis of image segmentation is demonstrated in Table 1.

Thus, the described methods of segmentation quality evaluation rated the most similar to etalon images and the most dissimilar images. These methods should be used to evaluate the segmentation quality of individual microscopic objects rather than the entire image as a whole. Table 2 shows the results of segmentation quality evaluation of images illustrated in fig. 6.

Analysis of data in Table 2 demonstrates that Gromov-Fréchet and Gromov-Hausdorff metrics showed better results than others. The combination of segmentation algorithms k-means and threshold optimally suits the segmentation of cytological images.

6 DISCUSSION

The analyzed methods of quantitative evaluation of segmentation quality, such as CSP, WSP, PDE, FRAG, AUMA, RUMA provide evaluation for only individual microscopic objects.

To evaluate the quality of segmentation of micro objects' groups, it is necessary to apply metrics. The most common metrics are the classical metrics of Fréchet and Hausdorff. At present, the best known algorithms that implement the Fréchet metric for contours (flat curves) have the computational complexity of $O(m \cdot n)$. Known algorithms for calculating the Hausdorff distance between convex regions (polygons) have the computational complexity of $O(m \cdot n)$. The Hausdorff and Fréchet metrics allow separately measuring the distances between regions and contours of images. To find the shortest distances between regions, the Gromov-Hausdorff metric is used.

The disadvantages of the known algorithms are the following:

1. Algorithms based on the Fréchet metric calculate the distance only between the contours of images;

2. Algorithms based on the Hausdorff metric calculate the distance only between convex regions of images;

3. There are no algorithms that calculate the shortest distance between contours.

The advantages of the developed algorithms are the following:

1. The proposed Gromov-Fréchet metric allows estimating the shortest distance between the contours of images;

2. The use of a combined metric based on the metrics of Gromov-Hausdorff and Gromov-Freche provided the possibility to calculate the shortest distances between contours and non-convex regions of images.

3. The developed EMISQ, which is based on the best known algorithms for calculating the Fréchet and Hausdorff distances, automatically estimates the shortest distances between groups of micro objects.

CONCLUSIONS

In the article, the authors introduced the Gromov-Fréchet distance and proved that distance is a metric. The method of quantitative evaluation of image segmentation quality is developed, on the basis of which a program module is designed and implemented, which allows calculating the shortest distance between images in an automatic mode.

The scientific novelty of the results is the following:

– for the first time, a Gromov-Fréchet metric was proposed for measuring the shortest distance between the contours of images;

– for the first time, quantitative evaluation method of segmentation quality based on the integrated use of Gromov-Fréchet and Gromov-Hausdorff metrics was applied allowing to evaluate the shortest distances between images.

The practical significance of the results is in the development of software to evaluate the shortest distances between the images. Computer experiments that were conducted on the example of cytological and histological images showed high efficiency of software that was used in image automatic segmentation algorithms.
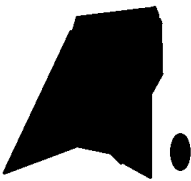
Table 1 – Comparative analysis of image segmentation

Methods of segmentation quality evaluation				
CSP, %	80.96	93.52	70.5	97.84
WSP, %	53.72	47.9	58.26	44.01
FRAG, %	98	90	70	97
AUMA, -pixels	4347.0	1165.0	5113.0	1105.0
RUMA, %	15.22	4.08	17.9	3.87

Table 2 – Comparative analysis of image segmentation based on metrics

Metrics	b) → b)	b) → c)	b) → d)	b) → e)
Hausdorff	0	19.20	21.37	46.57
Gromov-Hausdorff	0	19.20	21.37	46.57
Fréchet	0	36.76	26.4	52.77
Gromov-Fréchet	0	36.76	24.18	52.77
Weighted metrics	0	27.98	22.78	49.67

Further areas of research embrace the development of algorithms of de-paralleling the metric quantitative evaluation method of segmentation quality, which will help speed up the process of segmentation quality evaluation and segmentation algorithm optimization. Besides, a promising area for further investigations is development of a metric for evaluation the similarities of non-convex polygons.

REFERENCES

1. Yinpeng J. Contrast Enhancement by Multi-scale Adaptive Histogram Equalization / J. Yinpeng, L.Fayadb, A. Laine // Proceedings of SPIE. – 2001. – Vol. 4478. – P. 206–213. DOI: 10.7916/D8QZ2M29

2. Baron T. H. A Prospective Comparison of Digital Image Analysis and Routine Cytology for the Identification of Malignancy in Biliary Tract Strictures / T. H. Baron, G. Harewood, A. Rumalla // Clinical Gastroenterology and Hepatology. – 2004. – Vol. 2, Issue 3. – P. 214-219. DOI: 10.1016/S1542-3565(04)00006-0

3. Petushi S. Large-scale computations on histology images reveal grade-differentiating parameters for breast cancer / S. Petushi, F. U. Garcia, M. M. Haber // BMC Medical Imaging. – 2006. – Vol. 6, Issue 14. – P. 496–499. DOI: 10.1186/1471-2342-6-14

4. Levine M. D. Dynamic measurement of computer generated image segmentations / M. D. Levine, A. Nazif // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 1985. – Vol. 7, Issue 2. – P. 155–164. DOI: 10.1109/TPAMI.1985.4767640

5. Zhang Y. J. A review of recent evaluation methods for image segmentation / Y. J. Zhang // Signal Processing and its Applications (ISSPA) : Sixth International Symposium, Kuala Lumpur Malaysia, Aug 13–16, 2001 : proceedings. – Kuala Lumpur, 2001. – P. 148–151.

6. Berezsky O. Methods of quantitative evaluation of image segmentation quality / O. Berezsky // Signal/Image Processing and Pattern Recognition (UkrOBRAZ'2014) : XIIIth All-Ukraine International Conference, Kyiv, 3–7 November 2014 : proceedings. – Kyiv, 2014. – P. 51–54.

7. Lee S. U. A comparative performance study of several global thresholding techniques for segmentation / S. U. Lee, S. Y. Chung, R. H. Park // Computer Vision, Graphics, and Image Processing. – 1990. – Vol. 52, Issue. 2. – P. 171–190.

8. Zhang Y. J. Segmentation evaluation using ultimate measurement accuracy / Y. J. Zhang, J. J. Gerbrands // Image Processing Algorithms and Technique. – 1992. – Vol. 1657. – P. 449-460.

9. Zhang Y. J. Objective and quantitative segmentation evaluation and comparison / Y. J. Zhang, J. J. Gerbrands // Signal Processing. – 1994. – Vol. 39, Issue.2. – P. 43-54. DOI:10.1016/0165-1684(94)90122-8

10. Lopez M. Hausdorff approximation of convex polygons / M. A. Lopez, S. Reisner // Computational Geometry. – 2005. – Vol. 32, Issue 2. – P. 139–158. DOI: 10.1016/j.comgeo.2005.02.002

11. Alt H. Computing the Hausdorff distance between curved objects / H. Alt, L. Scharfz // International Journal of Computational Geometry. – 2008. – Vol. 18. – P. 307–320. DOI: 10.1142/S0218195908002647

12. Chew L. P. Getting around a lower bound for the minimum Hausdorff distance / L. P. Chew, K. Kedem // Computational Geometry. – 1998. – Vol. 10, Issue 3. – P. 197–202. DOI: S0925-7721(97)00032-1

13. Knauer C. Approximate nearest neighbor search under translation invariant hausdorff distance / C. Knauer, M. Scherfenberg // International Journal of Computational Geometry. – 2011. – Vol. 21, Issue 3. – P. 369–381. DOI: S0218195911003706

14. Alvarez V. Approximating the minimum weight spanning tree of a set of points in the Hausdorff metric / V. Alvarez, R. Seidel // Computational Geometry. – 2010. – Vol. 43. – P. 94–98.

15. Atallah M. J. Computing Some Distance Functions Between Polygons / M. J. Atallah, C. Celso // Computer Science Technical Reports. – 1990. – Vol. 9. – P. 1–10.

16. Alt H. Computing the Fréchet distance between two polygonal curves / H. Alt, M. Godau // International Journal of Computational Geometry and Applications. – 1995. – Vol. 5. – P. 75–91.

17. Mosig A. Approximately matching polygonal curves with respect to the Fréchet distance / A. Mosig, M. Clausen // Computational Geometry. – 2005. – Vol. 30, Issue 2. – P. 113–127. DOI: 10.1016/j.comgeo.2004.05.004

18. Buchin K. Computing the Fréchet distance between simple polygons / K. Buchin, M. Buchin, C. Wenk // Computational Geometry. – 2008. – Vol. 441, Issue 1–2. – P. 2–20. DOI: 10.1145/1137856.1137870

19. Rote G. Computing the Fréchet distance between piecewise smooth curves / G. Rote // Computational Geometry. – 2007. – Vol. 37. – P. 162–174. DOI: 10.1016/j.comgeo.2005.01.004

20. Schlesinger M. I. Frechet Similarity of Closed Polygonal Curves / M. I. Schlesinger, E. V. Vodolazskiy, V. M. Yakovenko // International Journal of Computational Geometry. – 2016. – Vol. 26. – P. 53–66. DOI: 10.1142/S0218195916500035

21. Computing the discrete Fréchet distance with imprecise impute / [H.-K. Ahn, C. Knauer, M. Scherfenberg et al.] // International Journal of Computational Geometry. – 2016. – Vol. 22. – P. 27–44. DOI: 10.1142/S0218195912600023

22. Computing the Fréchet distance between folded polygons / [A. F. Cook, Anne Driemel, Jessica Sherette et al.] // Computational Geometry. – 2015. – Vol. 50. – P. 1–16.

23. Gudmundsson J. Fast algorithms for approximate Fréchet matching queries in geometric trees / J. Gudmundsson, M. Smid // Computational Geometry. – 2015. – Vol. 48. – P. 479–494. DOI:10.1016/j.comgeo.2015.02.003

24. Computing discrete Fréchet distance: Technical Report: CD-TR 94/64 / T. Eiter, H. Mannila // Information Systems Department, Technical University of Vienna. – Vienna, 1994. – 7 p.

25. Deza M. M. Encyclopedia of Distances / M. M. Deza. – Berlin : Springer-Verlag, 2009. – 590 p.

26. Gromov M. Metric Structures for Riemannian and Non-Riemannian Spaces / M. Gromov. – Boston, MA, Progress in Mathematics, 1999. – 1041 p.

27. Regions Matching Algorithms Analysis to Quantify the Image Segmentation Results / [O. Berezsky, Y. Batko, O. Pitsun et al.] // Sensors & Transducers. – 2017. – Vol. 208, Issue 1. – P. 44–49. DOI: 10.1109/STC-CSIT.2016.7589862

28. Fuzzy system diagnosing of precancerous and cancerous conditions of the breast / [O. Berezsky, S. Verbovyy, L. Dubchak et al.] // Computer Sciences and Information Technologies (CSIT) : XI th International Scientific and Technical Conference, Lviv, 6–10 September 2016 : proceedings. – Lviv, 2016. – P. 200–203. DOI: 10.1109/STC-CSIT.2016.7589906

29. Berezsky O. Automated Processing of Cytological and Histological Images / Oleh Berezsky, Oleh Pitsun // Perspective Technologies and Methods in MEMS Design (MEMSTECH'2016) : XII th International Conference, Lviv-Polyana, 20–24 April 2016 : proceedings. – Lviv, 2016. – P. 51–53. DOI:10.1109/MEMSTECH.2016.7507518

Березький О. М.[1], Піцун О. Й.[2]

[1]Д-р техн. наук, професор, завідувач кафедри комп'ютерної інженерії, Тернопільський національний економічний університет, Тернопіль, Україна

[2]Аспірант, Тернопільський національний економічний університет, Тернопіль, Україна

МЕТОДИ ОЦІНКИ ЯКОСТІ СЕГМЕНТАЦІЇ ЗОБРАЖЕНЬ

Актуальність. Розглянуто основні методи кількісної оцінки якості сегментації зображень. Вони використовуються для вибору алгоритмів сегментації для конкретних класів зображень. Об'єктом дослідження є цитологічні та гістологічні зображення, які використовуються при діагностуванні патологічних процесів в онкології. Предметом дослідження є кількісні методи оцінки якості алгоритмів сегментації.

Мета роботи – введення метрики Громова-Фреше та розроблення метричного методу кількісної оцінки якості сегментації для порівняння алгоритмів сегментації зображень.

Метод. Проаналізовані критерії кількісної оцінки, які базуються на порівнянні та без порівняння з еталонним зображенням. Проаналізовано алгоритми знаходження відстані між зображеннями на основі метрик Фреше, Хаусдорфа, Громова-Хаусдорфа.

Для знаходження відстані між контурами зображень введено відстань Громова-Фреше. Доведено умову тотожності, симетричності та трикутника і показано, що відстань Громова-Фреше є метрикою.

Розроблено метричний метод кількісної оцінки якості сегментації. Він базується на використанні метрик Громова-Хаусдорфа та Громова-Фреше. Метод побудований на основі алгоритмів перетворення неопуклих полігонів у опуклі, алгоритму зважених хорд, алгоритмів обчислення відстаней Фреше та Хаусдорфа. Для знаходження відстані Хаусдорфа між опуклими областями використано алгоритм М. Аталаха. Для знаходження дискретної відстані Фреше використано алгоритм Томаса Ейтера і Хейкі Манілли. Дані алгоритми мають найнижчу обчислювальну складність серед свого класу алгоритмів.

Результати. Введено метрику Громова-Фреше та розроблено метричний метод кількісної оцінки якості сегментації.

Висновки. Проведені експерименти на основі цитологічних зображень підтвердили працездатність програмного забезпечення оцінки відстані між зображеннями. Розроблений метод показав високу точність оцінки відстаней між зображеннями. Розроблений програмний модуль використаний у інтелектуальній системі діагностування передракових і ракових станів молочної залози. Програмне забезпечення може бути використане в різних програмних системах комп'ютерного зору. Перспективними напрямами досліджень є пошук нових метрик для оцінки відстані між зображеннями.

Ключові слова: сегментація, кількісна оцінка сегментації, метрика Фреше, метрика Хаусдорфа, метрика Громова-Хаусдорфа, метрика Громова-Фреше, полігон, цитологічні зображення.

Березский О. Н.[1], Пицун О. Й.[2]

[1]Д-р техн. наук, профессор, заведующий кафедрой компьютерной инженерии, Тернопольский национальный экономический университет, Тернополь, Украина

[2]Аспирант, Тернопольский национальный экономический университет, Тернополь, Украина

МЕТОДЫ ОЦЕНКИ КАЧЕСТВА СЕГМЕНТАЦИИ ИЗОБРАЖЕНИЙ

Актуальность. Рассмотрены основные методы количественной оценки качества сегментации изображений. Они используются для выбора алгоритмов сегментации для конкретных классов изображений. Объектом исследования являются цитологические и гистологические изображения, используемые при диагностировании патологических процессов в онкологии. Предметом исследования являются количественные методы оценки качества алгоритмов сегментации.

Цель работы – введение метрики Громова-Фреше и разработка метрического метода количественной оценки качества сегментации для сравнения алгоритмов сегментации изображений.

Метод. Проанализированы критерии количественной оценки, основанные на сравнении и без сравнения с эталонным изображением. Проанализированы алгоритмы нахождения расстояния между изображениями на основе метрик Фреше, Хаусдорфа, Громова-Хаусдорфа.

Для нахождения расстояния между контурами изображений введено расстояние Громова-Фреше. Доказано условие тождественности, симметричности и треугольника и показано, что расстояние Громова-Фреше является метрикой.

Разработан метрический метод количественной оценки качества сегментации. Он базируется на использовании метрик Громова-Хаусдорфа и Громова-Фреше. Метод построен на основе алгоритмов преобразования невыпуклых полигонов в выпуклые, алгоритма взвешенных хорд, алгоритмов вычисления расстояний Фреше и Хаусдорфа. Для нахождения расстояния Хаусдорфа между выпуклыми областями использован алгоритм М. Аталаха. Для нахождения дискретного расстояния Фреше использован алгоритм Томаса Рейтера и Хейкки Маниллы. Данные алгоритмы имеют самую низкую вычислительную сложность среди своего класса алгоритмов.

Результаты. Введено метрику Громова-Фреше и разработано метрический метод количественной оценки качества сегментации.

Выводы. Проведенные эксперименты на основе цитологических изображений подтвердили работоспособность программного обеспечения оценки расстояния между изображениями. Разработанный метод показал высокую точность оценки расстояний между изображениями. Разработанный программный модуль использован в интеллектуальные системе диагностирования предраковых и раковых состояний молочной железы. Программное обеспечение может быть использовано в различных программных системах компьютерного зрения. Перспективными направлениями исследований является поиск новых метрик для оценки расстояния между изображениями.

Ключевые слова: сегментация, количественная оценка сегментации, метрика Громова-Хаусдорфа, метрика Громова-Фреше, полигон, цитологические изображения.

REFERENCES

1. Yinpeng J., Fayadb L., Laine A. Contrast Enhancement by Multi-scale Adaptive Histogram Equalization, *Proceedings of SPIE,* 2001, Vol. 4478, pp. 206–213. DOI: 10.7916/D8QZ2M29
2. Baron T. H., Harewood G., Rumalla A. A Prospective Comparison of Digital Image Analysis and Routine Cytology for the Identification of Malignancy in Biliary Tract Strictures, *Clinical Gastroenterology and Hepatology*, 2004, Vol. 2, Issue 3, pp. 214–219. DOI: 10.1016/S1542-3565(04)00006-0
3. Petushi S., Garcia F. U., Haber M. M. Large-scale computations on histology images reveal grade-differentiating parameters for breast cancer, *BMC Medical Imaging*, 2006, Vol. 6, Issue 14, pp. 496–499. DOI: 10.1186/1471-2342-6-14

4. Levine M. D., Nazif A. Dynamic measurement of computer generated image segmentations, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1985, Vol. 7, Issue 2, pp. 155–164. DOI: 10.1109/TPAMI.1985.4767640

5. Zhang Y. J. A review of recent evaluation methods for image segmentation, *Signal Processing and its Applications (ISSPA), Sixth International Symposium, Kuala Lumpur Malaysia, Aug 13–16, 2001, proceedings*. Kuala Lumpur, 2001, pp. 148–151.

6. Berezsky O. Methods of quantitative evaluation of image segmentation quality, *Signal/Image Processing and Pattern Recognition (UkrOBRAZ'2014) : XIIIth All-Ukraine International Conference, Kyiv, 3–7 November 2014 : proceedings*. Kyiv, 2014, pp. 51–54.

7. Lee S. U., Chung S. Y., Park R. H. A comparative performance study of several global thresholding techniques for segmentation, *Computer Vision, Graphics, and Image Processing*,1990, Vol. 52, Issue. 2, pp. 171–190.

8. Zhang Y. J., Gerbrands J. J. Segmentation evaluation using ultimate measurement accuracy, *Image Processing Algorithms and Technique,* 1992, Vol. 1657, pp. 449–460.

9. Zhang Y. J., Gerbrands J. J. Objective and quantitative segmentation evaluation and comparison *Signal Processing,* 1994, Vol. 39, Issue.2, pp. 43–54. DOI:10.1016/0165-1684(94)90122-8

10. Lopez M. A., Reisner S. Hausdorff approximation of convex polygons, *Computational Geometry*, 2005, Vol. 32, Issue 2, pp. 139–158. DOI: 10.1016/j.comgeo.2005.02.002

11. Alt H., Scharfz L. Computing the Hausdorff distance between curved objects, *International Journal of Computational Geometry*, 2008, Vol. 18, pp. 307–320. DOI: 10.1142/S0218195908002647

12. Chew L. P., Kedem K. Getting around a lower bound for the minimum Hausdorff distance, *Computational Geometry,* 1998, Vol. 10, Issue 3, pp. 197–202. DOI: S0925-7721(97)00032-1

13. Knauer C., Scherfenberg M. Approximate nearest neighbor search under translation invariant hausdorff distance, *International Journal of Computational Geometry*, 2011, Vol. 21, Issue 3, pp. 369–381. DOI: S0218195911003706

14. Alvarez V., Seidel R. Approximating the minimum weight spanning tree of a set of points in the Hausdorff metric, *Computational Geometry*, 2010, Vol. 43, pp. 94–98.

15. Atallah M. J., Celso C. Computing Some Distance Functions Between Polygons, *Computer Science Technical Reports*, 1990, Vol. 9, pp. 1–10.

16. Alt H., Godau M. Computing the Frechet distance between two polygonal curves, *International Journal of Computational Geometry and Applications*, 1995,Vol. 5, pp. 75–91.

17. Mosig A., Clausen M. Approximately matching polygonal curves with respect to the Frechet distance, *Computational Geometry*, 2005, Vol. 30, Issue 2, pp. 113–127. DOI: 10.1016/j.comgeo.2004.05.004

18. Buchin K., Buchin M., Wenk C. Computing the Frйchet distance between simple polygons, *Computational Geometry*, 2008, Vol. 441, Issue 1–2, pp. 2–20. DOI: 10.1145/1137856.1137870

19. Rote G. Computing the Frechet distance between piecewise smooth curves, *Computational Geometry*, 2007, Vol. 37, pp. 162–174. DOI: 10.1016/j.comgeo.2005.01.004

20. Schlesinger M. I., Vodolazskiy E. V., Yakovenko V. M. Frechet Similarity of Closed Polygonal Curves, *International Journal of Computational Geometry*, 2016, Vol. 26, pp. 53–66. DOI: 10.1142/S0218195916500035

21. Ahn H.-K., Knauer C., Scherfenberg M. et al. Computing the discrete Frechet distance with imprecise impute, *International Journal of Computational Geometry,* 2016, Vol. 22, pp. 27–44. DOI: 10.1142/S0218195912600023

22. Cook A. F., Driemel Anne, Sherette Jessica et al. Computing the Fréchet distance between folded polygons, *Computational Geometry*, 2015, Vol. 50, pp. 1–16.

23. Gudmundsson J., Smid M. Fast algorithms for approximate Fréchet matching queries in geometric trees, *Computational Geometry,* 2015, Vol. 48, pp. 479–494. DOI:10.1016/j.comgeo.2015.02.003

24. Eiter T., Mannila H. Computing discrete Fréchet distance: Technical Report: CD-TR 94/64 / Information Systems Department, Technical University of Vienna. Vienna, 1994, 7 p.

25. Deza M. M. Encyclopedia of Distances. Berlin, Springer-Verlag, 2009, 590 p.

26. Gromov M. Metric Structures for Riemannian and Non-Riemannian Spaces. Boston, MA, Progress in Mathematics, 1999, 1041 p.

27. Berezsky O., Batko Y., Pitsun O. et al. Regions Matching Algorithms Analysis to Quantify the Image Segmentation Results, *Sensors & Transducers,* 2017, Vol. 208, Issue 1, pp. 44–49. DOI: 10.1109/STC-CSIT.2016.7589862

28. Berezsky O., Verbovyy S., Dubchak L. et al. Fuzzy system diagnosing of precancerous and cancerous conditions of the breast, *Computer Sciences and Information Technologies (CSIT) : XI th International Scientific and Technical Conference, Lviv, 6–10 September 2016 : proceedings*. Lviv, 2016, pp. 200–203. DOI: 10.1109/STC-CSIT.2016.7589906

29. Berezsky O., Pitsun O. Automated Processing of Cytological and Histological Images, *Perspective Technologies and Methods in MEMS Design (MEMSTECH'2016) : XII th International Conference, Lviv-Polyana, 20–24 April 2016 : proceedings*. Lviv, 2016, pp. 51–53. DOI: 10.1109/MEMSTECH.2016.7507518