Sensors & Transducers, Vol. 208, Issue 1, January 2017, pp. 44-49

Sensors & Transducers
Published by IFSA Publishing, S. L., 2017
http://www.sensorsportal.com

Regions Matching Algorithms Analysis to Quantify the Image Segmentation Results

Oleh BEREZSKY, Grygory MELNYK, Yuriy BATKO and Oleh PITSUN
Ternopil National Economic University, Lvivska str., 11, Ternopil, 46020, Ukraine
Tel.: +380352475051, fax: +380352475051
E-mail: ob@tneu.edu.ua, mgm@tneu.edu.ua, bum@tneu.edu.ua, o.pitsun@tneu.edu.ua

Received: 20 September 2016 /Accepted: 30 December 2016 /Published: 31 January 2017

Abstract: In the article the matching algorithms of region images are analyzed. The work also presents their advantages and disadvantages. The comparing images algorithm of regions is developed on the basis of the measured chords. The comparison regions algorithms are used to evaluate segmentation algorithms in the Gromov–Hausdorff metric. The algorithm of metric evaluation is developed as an example of biomedical images segmentation.

Keywords: Image, Segmentation quality, Matching, Segmentation, Gromov-Hausdorff metric.

---

1. Introduction

Image segmentation is one of the fundamental processes in many software applications with image or video processing and computer vision systems. It is often used to divide the image into separate areas that correspond to different real objects. This is an important step in the content analysis stage or in the process of images "understanding". Modern scientists have developed many approaches to solve problems of images dividing into separate areas, but most of the proposed algorithms can solve the problem of segmentation only with a specific type of image [1]. Therefore, the automatic selection of effective segmentation algorithms is a complex task, since the comparison of different algorithms must consider both the type of the algorithm, and the parameters of its work. The main advantages of using segmentation analysis results are: a) objective selection of the best partitioning of the image; b) the possibility of quality control during the segmentation processes of images and correcting the work of the algorithms in real time; c) the ability to attract the most efficient algorithms for segmentation in the work program of image processing [2].

As a result of segmentation, the following two types of errors occur: the segmented image point is marked in such way as it belongs to a homogeneous region, and it doesn’t belong to the field in the model image; in the segmented image the point is not marked as it belongs to homogeneous area, but it is such in the model image [3]. Therefore, two criteria are often used to quantify the quality of segmentations: the level of regions similarity, the level of regions difference [4].

Among the segmentation results test we distinguish the following algorithms: the algorithm based on the Rand index that calculates the plurality of pixels pairs that are marked as tags of the same type and as both the segmentation algorithm and the model splitting [5]. The global consistency error measures the degree if the segmentation can be seen as another refinement of the segmentation and the error is 0, if two segments overlap completely and a certain value is nonzero otherwise. The boundary points error is determined by the algorithm that measures the average deviation of one contour point from the nearest contour point on the reference segmentation. The accuracy of segmentation is defined as the percentage of correctly segmented points to the number of pixels in an image [6]. The average absolute error characterizes the average difference between the model and actual results obtained for all tests [7].

In addition to the quantitative estimation there is another qualitative (subjective) one. The subjective criteria are the criteria of visual perception, valued during the examination by some groups of observers (experts). One of the approaches to quantify the quality of the segmentation approach is based on metrics usage [8-9].

2. Problem Statement

Let Im be an image. Let's segment this image using a particular algorithm or a set of algorithms for the segmentation, i.e. $S = \{A_1, A_2, ..., A_k\}$, where A_i is the segmentation algorithm. The result is $S_1(Im) = Im_1$. An expert manually segments the original image. Then we get the image Im_e. Let's present the image data in the forms of (1) and (2).

$$Im_e = \bigcup_{i=1}^{m_e} C_{ei} \cup O_{ei} , \qquad (1)$$

where C_{ei}, O_{ei} is the shape and the region marked by the expert.

$$Im_1 = \bigcup_{i=1}^{m_1} C_{1i} \cup O_{1i} , \qquad (2)$$

where C_{1i}, O_{1i} is the shape and the region marked by the algorithm.

We use the Hausdorff metric (3) to evaluate the proximity between regions O_1 and O_2.

$$d_H^X(O_1,O_2) := \max\left\{ \max_{x\in O_1} \min_{y\in O_2} d(x,y), \max_{y\in O_2} \min_{x\in O_1} d(x,y) \right\}, \qquad (3)$$

where d_H^X is the Hausdorff distance in the metric space (X,d).

We evaluate the shortest distance between the regions O_1 and O_2 using the Gromov - Hausdorff metrics (4):

$$d_{GH}(O_1,O_2) := \inf_{X,f,g} d_H^X(f(O_1), g(O_2)), \qquad (4)$$

where $f: O_1 \to X$, $g: O_2 \to X$ are the isometric embedding of metric space (X,d).

Thus, to determine the minimum distance between the regions O_1 and O_2 we must perform isometric transformation in order to find the maximum matching (section), that $S = O_1 \cap O_2 \to \max$.

3. Solution

After segmenting we obtain two homogeneous areas (Fig. 1), which are generally nonconvex polygons.

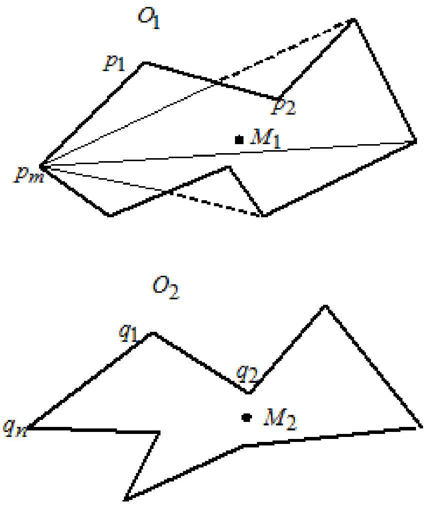


Fig. 1. The homogeneous region obtained after segmentation.

Thus, the problem of the maximum two-section polygons is reduced for the isometric transformation, translation and rotation. There are a number of heuristic algorithms for isometric transformations, such as the use of mass centers, ranges of their subsequent imposition, limiting use of a rectangle and extreme peaks [10].

The isometric transformation calculation requires the development of a number of heuristic algorithms.

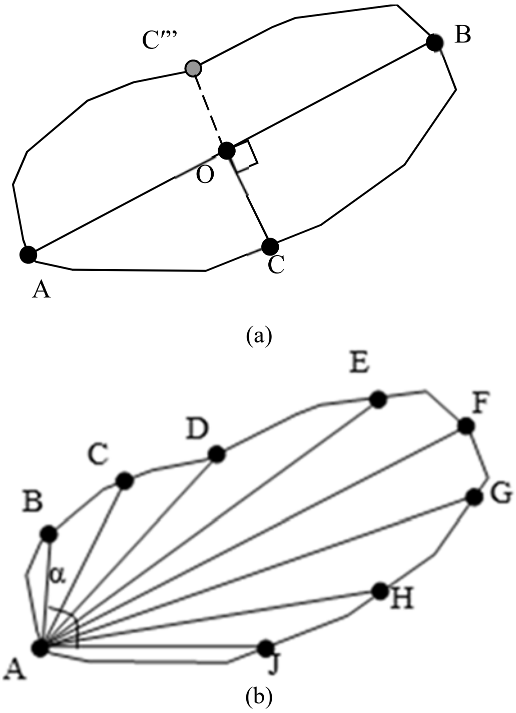
The longest chord algorithm is to find two of the longest chords on the polygons. After that we define maximum chord angles slope to the horizontal axis OX. In case of mapping one polygon rotates relative to the other with the corners (5) and (6). The computational complexity of the algorithm is directly proportional to the number of vertices of the polygon n.

$$\beta_1 = |\alpha_1 - \alpha_2|, \beta_2 = 180 + |\alpha_1 - \alpha_2|, \qquad (5)$$

where α_1 and α_2 are chord angles to the horizontal axis OX.

The previous algorithm can be modified by calculating the coordinate of the third point. The algorithm of the three points [11] on the contour polygon determines its two most distant points. For this we conduct the segment maximum length $AB = L_{\max}$, which includes the contour. The third point is defined as one that belongs to a median perpendicular and most distant $CO > C'O$ from the middle of the segment AB (point O). The example of the three points is shown in Fig. 2(a). The matching algorithm performs overlapping centers of polygons and bends in such way that the third point lies on one and the same half-plane. The computational complexity of the algorithm of the three points is also linear and proportional to the number of vertices of the polygon.

The example of a combinatorial algorithm of the three points on the circuit is the section lines algorithm [12]. The algorithm of section lines is the direct construction of lines, crossing the path of some of the starting point A (the farthest point from the center of the polygon mass) through a specified angle α (Fig. 2(b)). At the intersection of a contour n with lines we get n points of the intersection. Then we choose three points for the polygons mapping from the resulting set of points of the intersection. The computational complexity of this algorithm is proportional to n^3.



(a)

(b)

Fig. 2. (a) Algorithm based on three points; (b) secant lines algorithm.

4. Algorithm of Matching Polygons Based on Measured Chords

For the polygons $O_1 = (p_1, p_2, ..., p_m)$ and $O_2 = (q_1, q_2, ..., q_n)$, given that vertices conduct chord. For the polygon O_1 given by m vertices, we get a set of chords $\{h_1, h_2, ..., h_k\}$, where $k = \frac{m(m-3)}{2}$. For the polygon O_2, which is given by n vertices we get a set of chords $\{l_1, l_2, ..., l_p\}$, where $p = \frac{n(n-3)}{2}$. Thus, in the case of full search, we obtain the following computational complexity $O(n^2 \cdot m^2)$.

In order to reduce computational complexity let's sort the chords by both measuring and using the following coefficients:

a) The ratio of the relative chord length is (6)

$$\delta_{l_i} = \frac{l_i}{l_{\max}}, \tag{6}$$

where l_i is the length of i^{th} chord in the polygon, $l_{\max}$ is the maximum chord length in the polygon;

b) The overlapping polygon chord coefficient (7)

$$\delta_{O_i} = \frac{l_{Oi}}{l_i}, \tag{7}$$

where l_{Oi} is the length of i-th chord in the polygon, belonging to the inner region of the polygon, l_i is the length of i-th chord in the polygon.

Based on the given coefficients δ_{l_i} and δ_{O_i}, the coefficients are obtained according to (8)

$$W_i = \alpha\delta_{l_i} + \beta\delta_{O_i}, \tag{8}$$

where α and β are the measure coefficients chosen from the set of the values $[0..1]$ and $\alpha + \beta = 1$.

Then the array of the suspended chords for the polygon O_1 equals to $\{h_{W_1}, h_{W_2}, ..., h_{W_k}\}$, and for O_2 – $\{l_{W_1}, l_{W_2}, ..., l_{W_p}\}$.

Then the algorithm to find the Gromov - Hausdorff distance is:

1. We obtain the expert regions O_1 and O_2 using both the algorithm segmentation and the segmentation.
2. For the regions O_1 and O_2 we conduct piecewise linear approximation image edges (9):

$$O_i = \bigcup_{j=1}^{l} \{(x, a_j x + b_j) \mid x \in [c_j; d_j]\}, \tag{9}$$

where a_j, b_j, c_j, $d_j \in R$, $i = \overline{1,2}$ and we obtain the polygons.

3. On the basis of the Formulas (6), (7), (8) we form the arrays of the measured chords for the polygon O_1 $\{h_{W_1}, h_{W_2}, ..., h_{W_k}\}$, and for O_2 – $\{l_{W_1}, l_{W_2}, ..., l_{W_p}\}$.

4. We give the threshold Δ and sort arrays of the chords for the regions O_1 and O_2 according to the conditions $W_i \le \Delta$.

5. We calculate the centers of the masses $M_1(x_{C_1}, y_{C_1})$ and $M_2(x_{C_2}, y_{C_2})$ for the regions O_1 and O_2.

6. We perform parallel translation P of the region O_2 to O_1 so that the masses M_1 and M_2 centers match.

7. We perform rotation of R with the difference of rotations θ_1 and θ_2, found on the base of the sorted chords.

8. Then, we find the Hausdorff distance using the following Formula (10):

$$d_H^X(O_1,O_2) := \max\left\{\max_{x\in O_1}\min_{y\in O_2} d(x,y), \max_{y\in O_2}\min_{x\in O_1} d(x,y)\right\} \quad (10)$$

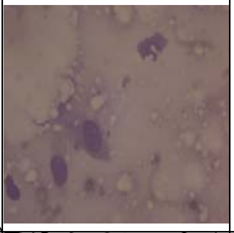
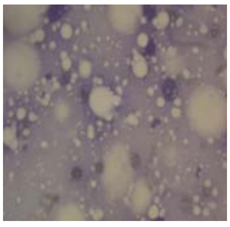
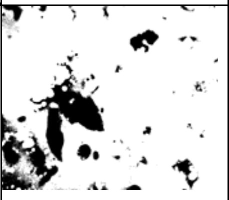
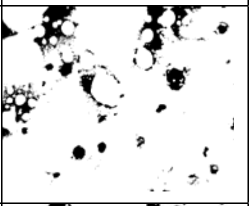
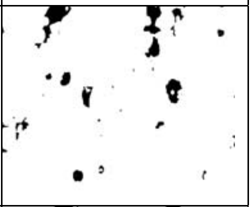
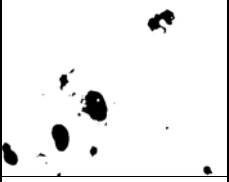
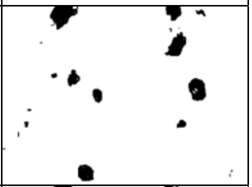
5. Experimental Results

The database of histological and cytological images of precancerous and cancerous conditions of the mammary gland was used for the experiments and it was also developed by the authors [13]. Generally, cytological and histological images are characterized by high complexity processing. The main disadvantages are the presence of different kinds of noise and high definition. The image preprocessing module is based on the library OPENCV [14] as well as it is designed to be adapted to different types of images and to improve their quality. Its work includes the following stages:

1. Image loading;
2. Determining the level of brightness, the average value of red, green and blue channels of RGB image input;
3. Options selection in order to improve the picture quality depending on the input parameters obtained in step 2. In this case, there are the following operations: brightness adjusting, contrast, morphological operations (dilate, erode);
4. Filtering by averaging filter, window size 7×7. This allows the image filter part;
5. Determining the peak signal to noise ratio. We have then the input data as the original and filtered image;
6. Image filtering algorithms (median, Gaussian, bilateral and other adaptive filters); selecting the type of filter and its parameters depends on the peak signal to noise ratio;
7. Image segmentation with k-means algorithm. As a result of the experiments for cytological image, k-means algorithm has shown the best results;
8. Threshold segmentation. Threshold segmentation is used to obtain images of the "object - background".

Table 1 compares the results of image segmentation algorithms with and without preprocessing. Image preprocessing is principle in order to improve the segmentation quality, including the adjustment of brightness and contrast. Noise reduction improves the segmentation quality significantly. To assess the quality of segmentation we use the Hausdorff metric, the Gromov - Hausdorff metric and the Gromov - Freshe metric.

Table 1. Results of cytological image segmentation with and without preprocessing.

Images / **Type of pro-cessing**		
Thres-holding without prepro-cessing		
K-means without prepro-cessing		
Thresholding with prepro-cessing		
K-means with prepro-cessing		
Expert		

The experiments were conducted on a test sample of 100 images. Table 2 presents the mean values of distances in the investigated metrics for different segmentation algorithms with and without pre-processing of images.

For experimental studies of the algorithm we have created the arbitrary set of test images of nonconvex areas. The images examples are shown in Fig. 3.

We have performed a pairwise matching images of nonconvex regions with each other and in each case the Gromov-Hausdorff distance is calculated. Fig. 4 shows the difference (deviation) found within each experiment, the distance found full-force, which is accepted as the reference. Each line of the graph represents a different algorithm. The algorithms are marked as: the longest chord, the algorithm of three points calculation – 3 points, the secant method, the developed algorithm – the list of chords. As we see, the developed algorithm determines the distance between the regions with precision.

Table 2. Evaluation of segmentation.

Metrics / **Type of processing**	**Hausdorff**	**Gromov-Hausdorff**	**Frechet**	**Gromov-Frechet**
Thresholding without preprocessing	304.53	304.53	307.56	305.7
K-means without preprocessing	89.4	89.23	89.4	89.23
Thresholding with preprocessing	13.45	12.36	13.45	12.36
K-means with preprocessing	10.21	10.02	10.21	10.02

Fig. 3. Example of test images.

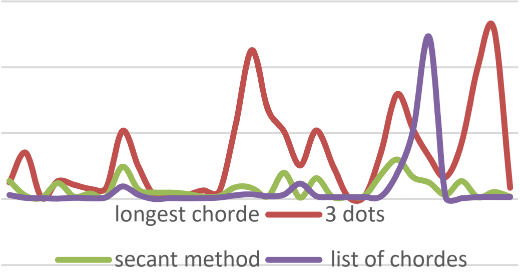


Fig. 4. Deviations of the Gromov-Hausdorff distance for some experiments on the outcome of full search.

Fig. 5 shows the evaluation of the execution time of each algorithm. If we match the accuracy at the level of the exhaustive algorithm search the developed algorithm shows the time at other algorithms based on contour features.

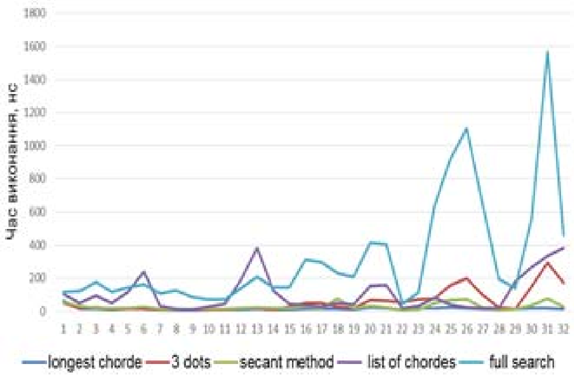


Fig. 5. Execution time for each algorithm.

We have worked out the algorithm to find the smallest distance between regions using the coefficient of relative chord length and the chord coefficient of an overlapping polygon. To assess the region in which the coefficients make the least difference, we have built the graph of the distance from the coefficients (Fig. 6). In the figure the relative chord length coefficient is pointed as Chorde distCoef, and the coefficient of the polygon chord overlapping is pointed as Chorde overlapCoef.

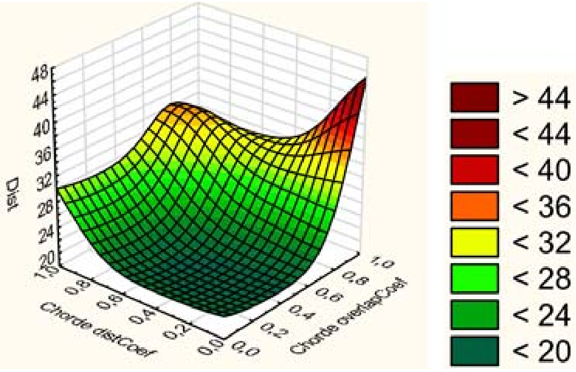


Fig. 6. Dependence of the Gromov-Hausdorff distance on the chord length coefficients and the overlap region.

5. Conclusions

Thus, the previous image processing significantly affects the quality of image segmentation cytology. The distances in metrics of previously processed images are much smaller than the images without preprocessing. Moreover, the k-means segmentation algorithm has shown better results in all metrics than the threshold segmentation.

Finally, we have worked out the algorithms, which are the most accurate as for time and commensurate expenses. They are compared for nonconvex regions in case of the Gromov-Hausdorff distance calculation among the following algorithms: the longest chord of three points, the secant, the measured chords algorithm.

Acknowledgements

The proposed research has been developed within the state budget project "Hybrid Intelligent Information Technology Diagnosing of Precancerous Breast Cancer Based on Image Analysis" (state registration number 1016U002500).

References

[1]. Zhang Y. J., An Overview of Image and Video Segmentation in the Last 40 Years, in Advances in Image and Video Segmentation, *IRM Press*, 2006, pp. 1-15.

[2]. Sharma M., Chouhan V., Objective Evaluation Parameters of Image Segmentation Algorithms, *International Journal of Engineering and Advanced Technology (IJEAT)*, Vol. 2, Issue 2, December 2012, pp. 84-87.

[3]. Christensen H. I., Phillips P. J. (editors), Empirical evaluation methods in computer vision, *World Scientific Publishing Company*, July 2002.

[4]. Philipp-Foliguet S., Guigues L., Multi-scale criteria for the evaluation of image segmentation algorithms, *Journal of Multimedia*, Vol. 3, No. 5, December 2008, pp. 42-56.

[5]. Chinnadurai V., Chandrashekhar G., Improvised levelset method for segmentation and grading of brain tumors in dynamic contrast susceptibility and apparent diffusion coefficient magnetic resonance images, *International Journal of Engineering Science and Technology*, Vol. 2, No. 5, 2010, pp. 1461-1472.

[6]. Zhang Y. J., Image segmentation evaluation in this century, in Encyclopedia of Information Science and Technology, ed. M. Khosrow-Pour, 2nd ed., *IGI Global*, 2009, pp. 1812-1817.

[7]. Padamavati S., Subashini P., Sumi A., Empirical Evaluation of suitable segmentation algorithm for IR Images, *IJCSI*, Vol. 7, Issue 4, No. 2, July 2010, pp. 22-29.

[8]. Berezsky O. N., Berezskaya E. N., Quantitative evaluation of the quality of the segmentation of images based on metrics, *Upravlyayushchie Sistemy i Mashiny*, Vol. 6, 2015, pp. 59-65.

[9]. Berezsky O., Batko Yu., Melnyk G., Verbovyy S., Haida L., Segmentation of Cytological and Histological Images of Breast Cancer Cells, in *Proceedings of the 8th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS)*, 24-26 September 2015, pp. 287-292.

[10]. Skiena S. S., The Algorithm Design Manual, *Springer*, 2008, 730 p.

[11]. Berezsky O. N., The algorithm of analysis and synthesis of biomedical images, *Journal of Automation and Information Sciences*, Vol. 39, Issue 4, 2007, pp. 69-80.

[12]. Melnyk G., Algorithm of matching of microobjects with different shapes, in *Proceedings of the Information Technologies in Innovation Business Conference (ITIB)*, 2015, pp. 31-34.

[13]. Berezsky O., An Intelligent System for Cytological and Histological Image Analysis, in *Proceedings of the 13th International Conference 'The Experience of Designing and Application of CAD Systems in Microelectronics' (CADSM'15)*, Polyana-Svalyava (Zakarpattya), Ukraine, 24-27 February 2015, pp. 28-31.

[14]. Berezsky O., Automated Processing of Cytological and Histological Images, in *Proceedings of the XIIth International Conference «Perspective Technologies and Methods in MEMS Design» (MEMSTECH'16)*, Lviv-Polyana, 20-24 April 2016, pp. 51-53.