UDC 519.7

V.I. Lytvynenko

APPLICATION OF CLONAL SELECTION ALGORITHM FOR SOLVING THE PROBLEM OF IMAGE SEGMENTATION CELLS

У статті пропонується підхід до рішення задачі сегментації зображень кліток при високому рівні шуму. Для цієї мети використовується методика динамічного об'єднання в кластери та алгоритм клонального відбору. Запропонований метод припускає наявність апріорних відомостей про форму клітини. Для опису границь клітки використовується модель побудови контуру клітки.

Ключові слова: *клітина, модель, кластеризація, клональний відбір, сегментація*

Introduction

There are three central tasks in the analysis of images: detection, segmentation and classification of object images. Detection of localization of objects as well as their boundaries of segmentation is the first step in many image analysis problems, especially in problems of quantitative analysis of objects.

For example, in medical imaging, detection and segmentation of cells, organs, etc. play an important role for the diagnosis and prognosis. Unfortunately, even with the help of image analysis software, traditional manual analysis is tedious and time consuming, especially in cases which should be determined by a large number of objects [1]. Thus, the development of effective and sustainable methods for automatic and rapid detection and accurate segmentation of their boundary is a very important goal.

Image Segmentation - is the most important and difficult stage of the analysis of medical images. This process is a natural and logical extension of functionalities of the digital image processing because it allows for visual analysis of objects and their yaskravistnyh geometrical characteristics [2]. Segmentation divides an image into its constituent area and facilities. That level of detail, which proves this division depends on the problem being solved. Image segmentation is not trivial, is one of the most complex image processing tasks.

The quality of segmentation depends on the accuracy of morphological calculation characteristics of biological objects and

therefore the accuracy of the classification and diagnosis.

Detection and segmentation of the object does not mean its identification. Many applications of image analysis require that the object was classified. For example, in [3] presented an algorithm that allows to extract outlines five types of vehicles moving. They are also segment the contours of the vehicle using their templates to further classify the vehicle in question. In this and other studies [4,5], segmentation is used to obtain information about the form [6,7,8].

In [4] the segmentation of cells images of the thyroid gland is used with a genetic algorithm. In [5] the same author used Tabu-search. In our work for the first time to solve the problem of segmentation of gray-scale images of blood cells used the immune clonal selection algorithm [10].

The suggested method consists of the following components: a) obtaining a gray-scale images; b) use gray-scale images to obtain of points that perhaps belong to the cell boundaries; c) application of Clonal algorithm to tune the ellipsoid model parameters to establish compliance its boundaries with the boundaries of the real cell.

1. Statement of the problem

Pathologists often make diagnostic decisions according to the results observation of individual cells, including their geometrical parameters such as area, radius, circumference, perimeter of cage, its compact size, the bump coefficients, Fourier descriptors and others [9].

However, to perform an accurate segmentation it is necessary the shape facilities modeling. Over recent years, has been developed a set of methods for image segmentation of cells [11-16]. As examples of such methods are methods based on regions, different boundary methods and so on. Methods based on the regions to perform image segmentation use the real increasing of region, its division and fusion. Boundary methods - they are a simple methods based on the classification of individual pixels. According to such methods with each pixel is associated with any property, such as brightness level; this property is compared with a threshold and the pixel is classified as part of the object or image, or as a body. The limiting conversion can be considered as a transaction in which the comparison is made with the function T that has the form [1]:

$$T = T(x, y, p(x, y), f), \quad (1)$$

 ISSN 1562-9945

where f - is an image and $p(x,y)$ indicates a some local characteristics of image points (x,y) ; for example, the average brightness in the vicinity of the center at this point. The image $g(x,y)$, obtained as a result of limit transformation, defined as follows:

$$g(x,y)=\begin{cases} 1, if\ f(\mathrm{x},\mathrm{y})>T \\ 0, if\ f(x,y)\le T \end{cases} \quad (2)$$

Thus, the pixels which appropriated the value 1 correspond to objects and pixels with value 0 corresponds to the background. If the value T depends only on f, that is the same for all image points, then the threshold is called global. If the threshold T depends on the spatial coordinates x and y, that it is called local or dynamic. If the threshold is depends on $p(x,y)$, such threshold is called adaptive. Thus, the main difficulty in this method consist in determining of threshold. The most simple method - a selection of the threshold based on the principal value in the histogram [1]. More sophisticated version of this approach is given in [18, 19]. The problem of these methods is that they use only local information (for individual pixels) and do not use any information about the shape of the object as a whole, although such knowledge can greatly improve segmentation especially in the presence of noise and other image distortions that lead to fuzzy distinction between object and background. Images of cells are characterized by the following features.

1. **Feeble** contrast, that is, objects (cells) in their brightness levels can be similar to the background.

2. **Multiple overlay of objects** on the considered area of the image. Multiple overlay makes segmentation very difficult.

3. **Poor quality.** The staining methods of cell preparations make a lot of irregularities in the image and lead to the fact that not all parts of the investigated object are equally colored.

In the presence of noise, interference, and overlays, the image segmentation the cell is a difficult problem to be solved [19]. Therefore, for the solution of such problems, are used constraints imposed a priori knowledge. This means that the efficiency of segmentation can be greatly improved by using a priori knowledge of the shape of cells. In fact, one of the most common problems in medical image segmentation is an extension of traditional approaches to segmentation and classification of objects with inclusion of information about the form,

and not only about the intensity of objects. In this paper, for accurate image segmentation the cell is not only used the information about the border, but also information about its shape.

For image segmentation of cell, we propose to use the immune clonal selection algorithm [10,23].

The proposed method consists of the following three parts:

1) identification of possible boundaries of cells;

2) an approximate determination of the location of the cells and the detection of image points that are most probably are belong to boundaries of cells;

3) creation a model the cell circuit, characterized by five parameters that accurately allocate circuit the cell and eliminate the noise impact.

By setting the parameters of the model it is possible to accurately describe the contour of cells in the image. Thus, the problem of image segmentation is transformed into optimization problem.

2. Contoured model of cell picture

The most cells of the human body are generally of ellipsoid shape as shown in fig. 1a and fig. 2a. It can be seen that although the cells themselves brightness is lower than the brightness of body, image contrast remains quite low. In addition, there is the presence of noise and overlapping cells in the image. As mentioned earlier, this type of problem can be solved by overlay restrictions on the parameters in the form of a priori information. The approach based on modeling of the cell contour, can be represented as a problem of parameters optimization. If the contour of the cell model defined, you can restore the segmented image for the necessary values of the geometrical parameters. To describe the contour of the cell will use the equation of an ellipse:

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1 \qquad (3)$$

where a and b – respectively with the length of the major and minor axis.

Given the possible movements and rotation, equation (3) can be written as:

$$\frac{\left[(x-x_0)\cos\theta+(y-y_0)\sin\theta\right]^2}{a^2}+\frac{\left[(x-x_0)\sin\theta+(y-y_0)\cos\theta\right]^2}{b^2}=1 \quad , \qquad (4)$$

 ISSN 1562-9945

where x_0, y_0 - the center of the ellipse, θ indicates its orientation.

Thus, we obtain the five parameters: x_0, y_0, a, b and θ that describe the model.

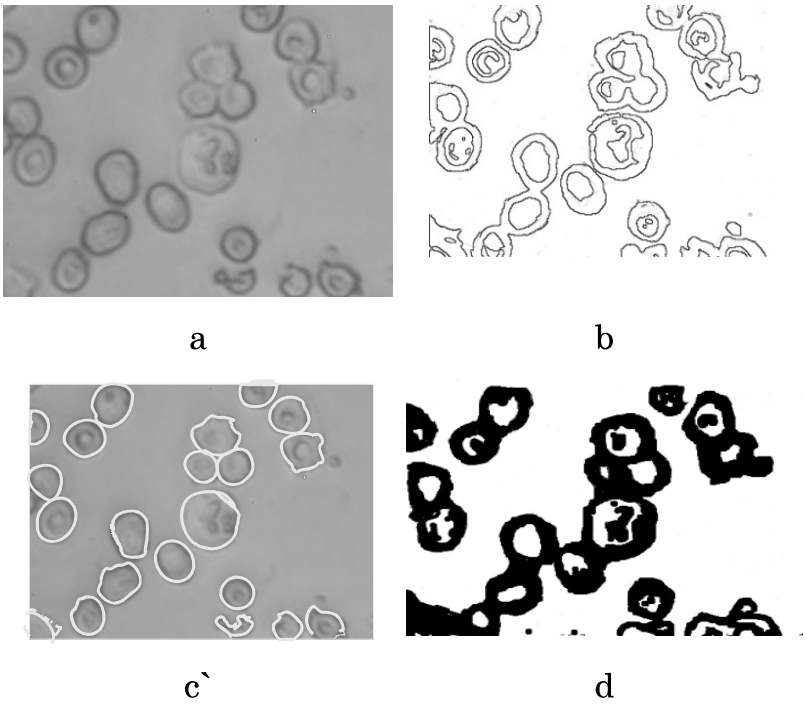
3. Approach to identify image points in the cell

3.1. Location of cells

Certainly, segmentation based on the boundary definition is divided into two stages: identifying of boundaries and merging of boundaries. In [20] the selection of boundary is considered as the several independent tasks, where each task has its own input information, the method of processing and output information. However, this approach can lead to incorrect results because of possible error propagation. In the present scheme at first are determined the dots of images that have a high probability of belonging the cell that is used to determine the approximate location of the cell. This process can be seen as the transformation of local information into global.

Based on the approximate location and model of the cell, we can re-evaluate the correctness of the assumptions about belonging of pixel to boundaries of the cell. For the detection of boundaries, we use the method of the boundaries determining of Cannes [21].

Essence of the method consists in finding of local areas with differences in brightness. Differences in brightness are sought by filtering on each axis one-dimensional filter lapsyan-haussyan. In the method of Cannes for the classification differences to "weak" and "strong" used two thresholds: “bottom” and “top”. "Weak" boundaries observed in the resulting image, if they are connected to the "strong" differences. Due to the influence of noise, there is a set of false directions in gray-scale image (see Fig. 1b). Many points of adjacent boundaries are connected. In the original image of blood cells (see Fig.1a), we can see a great variation in brightness levels at the boundaries of the cell. So, created boundaries, due to the influence of noise, contain a small amount of connected of image points. Thus, we can use the threshold value to determine whether the contour direction correctly indicates the border of the cell. Otherwise, we remove the contour appeared as a result of noise. Threshold values can be determined experimentally.



a b

c` d

Fig. 1. Results of image segmentation of blood cells: (a) the original image; (b) the image of boundary; (c) the proposed approach by us. (d) the method based on the histogram.

3.2. Detecting of image points in the cell by means of a dynamic clustering

After determining the approximate location of cells, use the method of dynamic clustering for finding points of the image possibly belonging to cells.

According to this method initially defined kernel k_j, which is a cluster. The kernel can be either a function or set of image points or other models.

To determine the belonging considered point to the cluster, enter a value $\Delta(y,k_j)$ that describes the similarity between the examined point y and cluster k_j. The method is implemented using the following actions:

1. Determine the initial core for each cluster.

2. For all considered points, perform the following rules for their classification: if $\Delta\left(y,k_j\right)\le t$, then $y\in\gamma_j$, where y indicates the considered point, t - the threshold value, and γ_j represents the j-th cluster.

3. Refresh core k_j. If thus to is no change kernels then stop execution, otherwise return to step 2.

Because the cell is ellipsoidal boundary, we use as kernel Gauss function that can be represented as:

$$K_j(y) = \frac{1}{2\pi\left\|\Omega_j\right\|}\exp -\frac{1}{2}\left(y-m_j\right)\Omega_j^{-1}\left(y-m_j\right)^T \quad (5)$$

where m_j - the average value of samples and Ω_j - covariance matrix.

The similarity between the examined point y and the cluster j is determined as follows:

$$\Delta\left(y,k_j\right) = \frac{1}{2}\left(y-m_j\right)\Omega_j^{-1}\left(y-m_j\right)^T + \frac{1}{2}log\left\|\Omega_j\right\| \quad (6)$$

Using the above method, we can find the points of the image, with a high probability of belonging to the cells. After detection of such points, we can search the ellipse in best possible way corresponding of contour cells in a relatively small area.

4. Using of clonal selection algorithm for the selection of the body boundary

4.1. Clonal selection algorithm

Clonal algorithm - is one form of evolutionary algorithms based on the use of the mechanisms of the immune system of vertebrates and humans. Studies have shown that clonal algorithm demonstrated to be very effective in solving optimization problems. Managing a population of individuals, it conducts the search for optimal solutions by means of mechanisms of reproduction, genetic variability and selection. Clonal selection principle used by the immune system to describe the basic characteristics of the immune response in antigenic stimulation. It is based on that fact that only the cells that able to recognize the foreign antigens can be selected and spread. The selected cells are subjected to mutation process to improve their affinity to selected antigens. Formally algorithm of clonal selection can be represented as [23,24,25]:

$$\text{CLONALG} = (P^l, G^k, l, k, m_{Ab}, \delta, f, I, \tau, AG, AB, S, C, M, n, d), \quad (7)$$

where P^l is space of search (space of forms); G^k is space representation; l is the length of vector of attributes (dimension of space of search); k is the length of antibody receptor; m_{Ab} is dimension of population of antibodies; δ is the expression function; f is the affinity function; I is the function of initialization of the initial

population of antibodies; τ is the condition of completion of algorithm work; AG is the subset of antigenes; AB is population of antibodies; S is the operator of selection; C is the operator of cloning; M is the mutation operator; n is the number of the best antibodies selected for cloning; d is the number of the worst antibodies subjected to substitution for new ones.

The process of converting a population of antibodies by clonal selection algorithm can be represented as a sequence of the following statements:

$$\mathrm{AB}_t \xrightarrow{\text{Selection (S)}} \mathrm{G}_S \xrightarrow{\text{Cloning (C)}} \mathrm{G}_C \xrightarrow{\text{Mutation (M)}}$$

$$\mathrm{G}_M \xrightarrow{\text{Repeat mutation (S)}} \mathrm{G}_S \xrightarrow{\text{Replacemen t(d)}} \mathrm{AB}_{t+1},$$

Whe.re t - is the number of generation, AB - is the population of antibodies (detectors), G_S - the subset of selected best antibodies, G_C - is the subset of clones, G_M - is the subset of clones after mutation.

Let us show the generalized stepwise description of the algorithm.

1. *Initialization.* Creation (usually by random generation) of the initial poulation of antibodies AB.

2. *Determination of affinity.* For every antibody AB_j, $AB_j \in AB$ determine its affinity relative to every antigen Ag_i, $Ag_i \in AG$. Write the result into the matrix of affinities $D: D = \left[\left|AG\right| \times m_{Ab}\right]$, and $d_{ij} = f\left(Ab_j, Ag_i\right), d_{ij} \in D$.

3. *Clonal selection and propagation.* Select from population n of each the best antibodies for every row of the matrix D and place them into separate population of clones $AB_C, \left|AB_C\right| = n \cdot \left|AG\right|$. It is necessary to generate clones of elements of the population AB_C proportionally to their affinity, i.e., the greater it is, the greater number of clones is generated and vice versa.

4. *Affinity maturation.* Subject to mutation all the clones of population AB_C with probability inversely proportional to affinities, i.e., probability of mutation is the greater, the lower is its affinity. Determine new affinity of every antibody AB_j, $AB_j \in AB_C$ similar to item 2 and obtain the matrix of affinities D_C. Select n antibodies from the population AB_C, for which the corresponding vector-column of the

 ISSN 1562-9945

matrix D_C gives the best generalized result of affinity, and transfer them into population of cells of memory M_R.

5. *Meta-dynamics.* Substitute the worst d antibodies of the population AB by new random individuals.

Substitute n antibodies of the population AB by cells of memory from M_R and pass to item 2 until the stoppage criterion is reached.

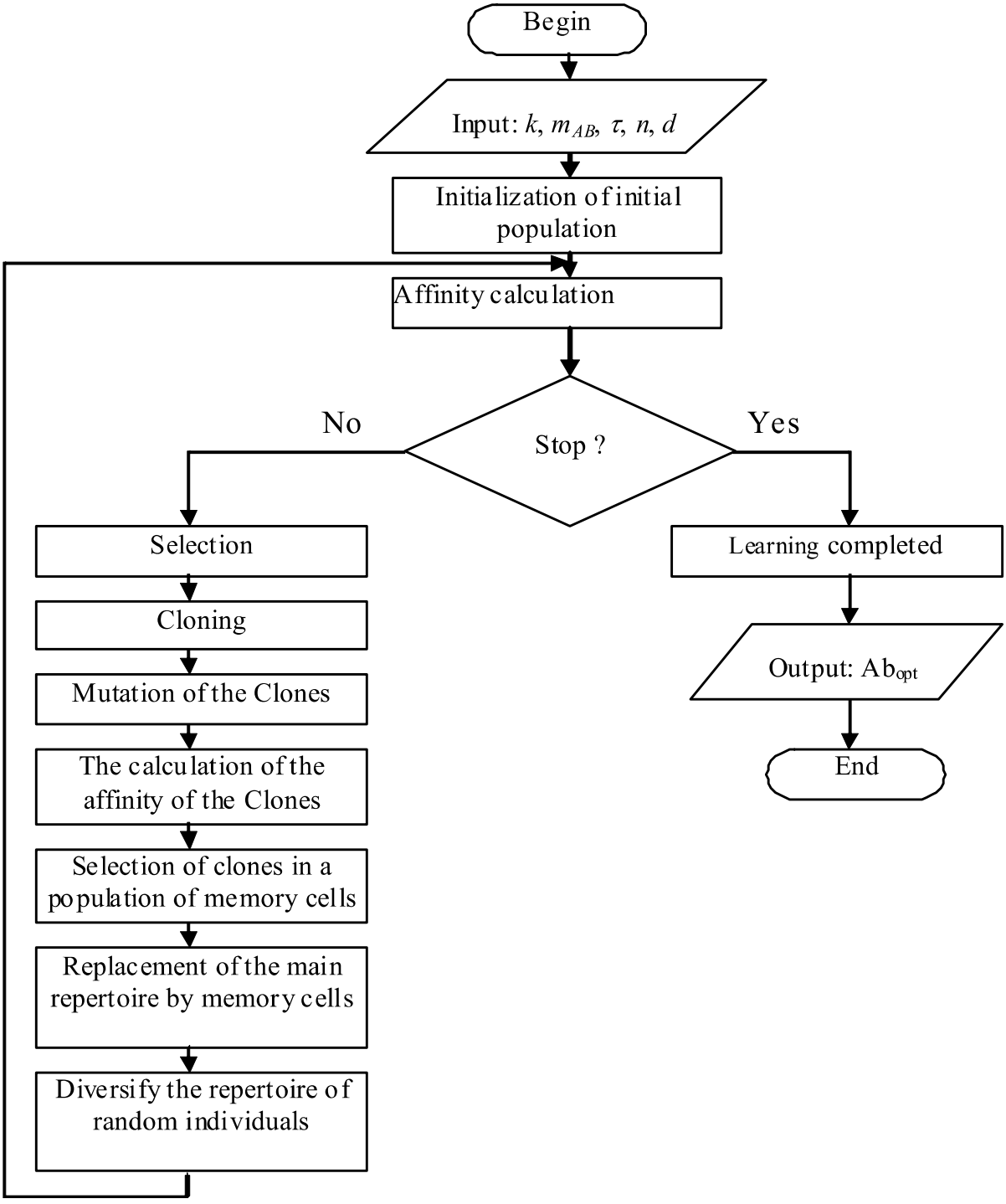


Fig. 2. Block-diagram of clonal selection algorithm.

4.2. Setting algorithm for task of allocation boundaries of cells

The input data for the solution of the problem are N possible points that are on the contour line of the cell. Each of these pixel has a unique number from 1 to N. To determine the ellipse we need to know only five of them. Let the vector $I = (I_1, I_2, ..., I_5)$ indices the five selected points. Assume that $I_i < I_j$ at $i < j, i, j = 1, ..., 5$. Affinity function in this case will be:

$$f(I) = \sum s(r_{I_i}^2) \quad (9)$$

where s - step function; s = 1 at r_j greater than or equal to the width of the template and s = 0 otherwise. This function counts the quantity of points that are within a certain distance of the ellipse. We denote by P the population of M antibodies $I_1, I_2, ..., I_M$. Let L - population size and GEN - number of generations. n_1 - the quantity of antibodies selected for cloning, and n_2 - the quantity of antibody clones elected from the population to restore the repertoire. In this case, the basic step by step description of the algorithm for the solution of the problem would look like this.

Step 1. **Initialization.** To initialize antibodies randomly to generate the five integers in the range from 1 to N inclusive. These numbers are the five points of index image. To initialize the entire population must repeat this process L times.

Step 2. **Selection and cloning.** When using affinity function, to assess of individuals of populations and to choose the best n_1 of them. Cloning involves the creation of copies selected individuals in proportion to their affinity.

Step 3. **Mutation and substitution.** For all newly created clones based on the likelihood of mutation and inversely proportional to their affinity, to conduct the operation being accidentally changed. For this task mutation operator is described as follows. Originally are generated two integers m and n (m ranging from 1 to 5, and n in the range from 1 to N). Next, the gene selected for mutation antibody with an index of m is replaced by the value of n. From population altered by mutation of clones are selected n_2 antibodies and replace the same number in the basic population to restore repertoire.

 ISSN 1562-9945

Step 4. **Replacing the worst antibodies.** To increase the diversity of the repertoire from the main population to reselect the antibody with lowest value of affinity and replace them with new randomly generated individuals.

Step 5. **Repeat** steps 2-4 until you reach stopping criterion.

5. Results of conducted experiments

In this section, we present results of experimental investigations of image segmentation with blood cells and give a comparison of methods based on histograms. The method is based on histogram is described as follows:

(1) obtain a histogram of the segmented image;

(2) according to the histogram to obtain the corresponding threshold of the segmented image;

(3) if the brightness level of a pixel is less than or equal to the threshold, the pixel belongs to the cluster object, otherwise it will be classified as background.

The threshold can be determined using the method of minimum error or the method of maximum entropy [18,19]. In our experiment, we set population size L= 200; meaning the number of antibodies selected, n_1= 70% of the population; hyper mutation probability P_m = 0.8; number of generations GEN = 100; and the number of the worst antibodies are replaced, n_2 = 30% of the population; cloning factor β = 0.8.

Fig. 1c shows the results of segmenting images experiments of blood cells. Image of contours obtained by the method of determining the boundaries by Cannes [22] method options below:

- upper threshold - 0.85;
- lower threshold - 0.79;
- covariance - 1.75.

Experimental results with boundary methods are shown in Fig. 1d. From Fig.1, are seen the benefits of our proposed method compared with the methods based on the histogram. The proposed approach is immune to noise. If the brightness near the boundaries of cells have insignificant scatter, the proposed method can allocates the border of cells and correctly to solve the problem of segmentation.

When two cells are very close to each other, simple methods based on histograms can not to distinguish them, while our method successfully identifies these cells. Thus, the proposed approach has the ability to work successfully with the imposition of boundaries cells.

The time spent on finding a pixel-owned cell, is about 20 sec to 1.6 GHz processor. The time spent on setup of contour model parameters is about 20 seconds.

6. Conclusion and further research

This paper proposes an approach to image segmentation of cells under significant noise effects, this method is based on a combination of dynamic clustering algorithm and clonal selection. In our algorithm, we use not only information about the circuit, but also the boundary cell model, assuming that the cell has an elliptical shape. Using prior knowledge about boundary cells, our method has a high resistance to noise.

Points of the image, which possibly belong to cells are determined by the method of dynamic clustering thus reducing the search space and time is spent on the optimization objective function by clonal algorithm. The obtained results suggest a possible direction for future research in the field of segmentation automation, which is especially important in the design of biomedical systems. Experimental studies have shown that the proposed approach is able to handle the partial overlapping of objects in images, but to improve of such ability the further studies are needed.

REFERENCES

1. Гонзалис Р., Вудс Р., Цифровая обработка изображений.М. Техносфера, 2005. – 1072 с.
2. Прэтт У. Цифровая обработка изображений: В 2 т. М.: Мирб 1982.
3. M.-P. Dubuisson, S. Lakshmanan, and A. K. Jain (1996) Vehicle segmentation and classification using deformable templates. IEEE Transactions on Pattern Analysis and Machine Intelligence, 18(3):293–308.
4. Ke Q, Jiang T, Ma S. A tabu search method for geometric primitive extraction. Pattern Recognit Lett 1997; 18:1443–51.
5. F. Yang and T. Jiang, “Cell image segmentation with the kernel-based dynamic clustering and an ellipsoidal cell shape model,” *J. Biomed. Inform.*, vol. 34, no. 2, pp. 67–73, 2001.

6. K.-M. Lee and W. N. Street (2000a) Automatic segmentation and classification using on-line shape learning. In: Proceedings of the 5th IEEE Workshop on the Application of Computer Vision, 64–70
7. K.-M. Lee and W. N. Street (2000b) Dynamic learning of shapes for automatic object recognition. In proceedings of the 17th ICML-2000 Workshop on Machine Learning of Spatial Knowledge, 44–49
8. K.-M. Lee and W. N. Street (2000c) A new approach of generalized Hough transform with flexible templates. In: Proceedings of the 2000 Inter- national Conference on Artificial Intelligence, volume III, 1133–1139
9. Спорыхин В. Я., Меркулова Е. В., Коков А. И. Автоматизированная подсистема распознавания и оконтуривания клеток // Вестник Херсонского государственного технического университета. – Херсон: ХГТУ, 2003 . – № 2(18). – С. 162–166.
10. Leandro Nunes de Castro, Fernando J. Von Zuben. Artificial immune systems: part 1 – basic theory and applications. TR - DCA 01/99 1999.
11. 11. Wu HS, Barba J, Gil J. A parametric fitting algorithm for segmentation of cell images. IEEE Trans Biomed Eng 1998; 45:400–7.
12. Garbay C. Image structure representation and processing discussion of some segmentation methods in cytology. IEEE Trans Pattern Anal Mach Intell 1986; 8:140–7.
13. Garrido A, Perez N. Applying deformable templates for cell image segmentation. Pattern Recognit 2000; 33:821–32.
14. Mouroutis T, Roberts SJ, Bharath AA. Robust cell nuclei segmentation using statistical modeling. BioImaging 1998; 6:79–91.
15. Simon I, Pound CR, Partin AW, Clemens JQ, Christensbarry WA. Automated image analysis system for detecting boundaries of liveprostate cancer cells. Cytometry 1998; 31:287–94.
16. Wu HS, Gil J. An iterative algorithm for cell segmentation using short-time Fourier transform. J Microsc 1996; 184:127–32.
17. Wu HS, Barba J, Gil J. Iterative thresholding for segmentation of cells from noisy images. J Microsc 2000; 197:296–304.

18. Kapur JN, Sahoo PK, Wong AKC. A new method for graylevel picture thresholding using the entropy of the histogram. Comput is Graph Image Process 1985; 29:273–85.
19. Kittler J, Illingworth J. Minimum error thresholding. Pattern Recognit 1986; 19:41–7.
20. Poggio T, Torre V. Ill-posed problems and regularization analysis in early vision. Proc AARPA Image Understanding Workshop 1984; 257–63.
21. Marr D. Vision: a computational investigation into the human representation and processing of visual information. San Francisco, CA: Freeman, 1982.
22. Canny J. A computational approach to edge detection. IEEE Trans Pattern Anal Mach Intell 1986; 8:679–98.
23. Литвиненко В.И., Фефелов А.П., Горавский С.П. Объектно-ориентированная реализация алгоритма клональной селекции//Радиоэлектроника. Информатика. Управление / Запорожье, 2003(9). С.81-88.
24. Литвиненко В.И., Горавский С.П., Фефелов А.П. Реализация модифицированного алгоритма клонального отбора для решения задач классификации //Міжнародна наукова конференція “Інтелектуальні системи прийняття рішень та прикладні аспекти інформаційних технологій” Т.1, Секція 1. с.127–131.
25. Грицик В.В., Литвиненко В.І., Опотяк Ю.В., Фефелов А.О.,Цмоць І.Г. Використання штучних імунних систем при розвязання задач мікроскопічної цитометрії // Інформаційні технології і системи Том 8 № 1, с.37–50

 ISSN 1562-9945