Proceedings of the 29th Annual International
Conference of the IEEE EMBS
Cité Internationale, Lyon, France
August 23-26, 2007.

Automatic Segmentation on Cell Image Fusing Gray and Gradient Information

Boqiang Liu, Cong Yin, Zhongguo Liu, and Yanyan Zhang
School of Control Science and Engineering, Shandong University, Jinan 250061, China

Abstract—**To develop an automatic classifying and diagnostic system for the hematopoietic cells from the blood and bone marrow smears stained with Wright-Giemsa, an automated segmentation algorithm fusing gray level, colorful information and mathematical morphological gradient is proposed for segmentation of the nucleated hematopoietic cells (including nucleus and cytoplasm). For the accurate segmentation of the nucleus, the conventional iterative threshold segmentation has been improved. Color information and prior knowledge are fully used by transaction of color spaces for the purpose of cytoplasm segmentation. In order to prevent over-segmentation, the morphological gradient information is used to mark the background, nucleus and cytoplasm. The edge detection is implemented in gray gradient image since the morphological gradient can detect the contour better than other conventional edge detection operators. The success rate is 95.5 % for nucleus and 92.6 % for cytoplasm. The results show that the method is valid and efficient to segment color images from blood and bone marrow smears.**

Key words—**cell image segmentation; gray space; iterative threshold segmentation; edge detection; morphological gradient**

I. Introduction

The clinical examination on the hematopoietic cells from the blood and bone marrow smears stained with Wright is essential for the homeopathy diagnosis and classification. The conventional approaches are blood cell examinations (i.e. number, morphology, etc) from peripheral blood and bone marrow smears by optical microscope. The main objective of this work is to recognize and classify different categories of blood cells. Generally in blood analysis, doctors firstly look for three different kinds of cells: red cells, white cells and blood platelets. The white cells often have nucleus, so it can be called the nucleated cells. These three categories are distinguished by their size and color. This purely visual observation has the problem of poor reproducibility between the observers and themselves [1]. Therefore, it is necessary to develop an automatic cell classifying and diagnostic system to complete a quantitative analysis and computer-aided detection of blood cell. Automatic image segmentation is the key step since the results of segmentation directly influences the subsequent feature extraction and recognition. In this paper, for our target, we want complete the segmentation of the nucleated cells. It is possible to get better segmentation results to adopt a fixed threshold value in color space when the stain is good, but it is difficult to satisfy this condition. So an improved iteration threshold segmentation algorithm combing gray-based image and color information is proposed to accomplish the two-step segmentation of the nucleus and cytoplasm effectively. The proposed segmentation system aims at detecting the nucleated cells automatically and segmenting the nucleus and cytoplasm precisely. Then based on the segmented images, the boundaries of cells should be enhanced and highlighted. Also, the good result of segmentation should help to lay the foundation for the following recognition and classification.

II. The Basic Principle For The Algorithm

The segmentation of the nucleated hematopoietic cells includes two aspects[2]: (1) For the purpose of calculating the morphological parameters of blood cell, the nucleated cells should be isolated from the complex background. Also the nucleus and cytoplasm should be divided into two different regions. (2) For the purpose of individual analysis, the single cell should be detected automatically from the image containing a number of cells.

A series of strategies contained in this algorithm are as follows. First, the optimal threshold value is obtained using the improved iteration threshold segmentation algorithm in the gray level image to detect the nucleus from the background effectively. Secondly, the saturation information in HSI color space is made full use of to segment the cytoplasm. After smoothing the segmented images, the nucleus and the cytoplasm are combined to an intact image and given the only mark for different regions. Then the morphological gradient is well suited to detect the contours of both nucleus and cytoplasm. Last, for the convenience analysis of the separated cell, every single cell should be detected automatically. Meanwhile, the performance table for judge the segmentation is displayed in the paper.

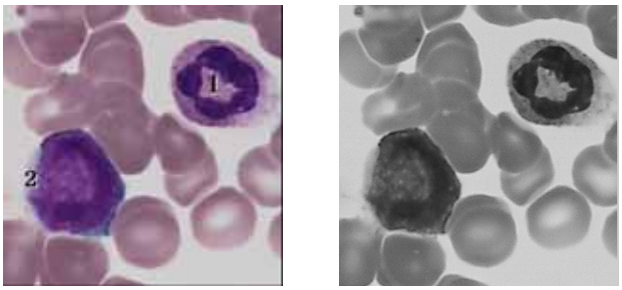
III. Using Improved Iteration Threshold Segmentation Approach To Segment Blood Cells

A. The segmentation of nucleus

The original image of blood cell is shown in Fig. 1(a) and its gray image is shown in Fig. 1(b). Observing the histogram in gray space (see Fig. 2), we can see the peak of the nucleus has a distance from the background. That is to say, there is a considerable distance between their gray values. Thresholding segmentation algorithm is just an approach to put the image into different grades and then to set up a threshold to determine the borders of a segmented object.

But there are many predominant particle sizes in the histogram, including the nucleus, the cytoplasm, the red blood cell and the background. The simple single threshold segmentation wouldn't be suited if we want to extract the nucleus precisely. So here the improved iteration threshold

segmentation algorithm is applied to segment the nucleus in which the improved strategy for the threshold is the key step.



(a) (b)

Fig. 1. Original image of the cell
(a) The color image (b) The gray image

B. The Improved Iteration Threshold Segmentation Algorithm

The main procedures of the iteration threshold segmentation algorithm are as follows[3]. First an approximate value is chosen as the estimated value for the initial threshold value. Then use the value to segment the initial image and to form the sub-images. The characteristics of the sub-images are utilized to select a new threshold which is used to segment the sub-image again. After several circles, make the wrong pixel points to reduce to a minimum. In this process the selection of the initial threshold is the key to determine the segmentation result. In the conventional iteration segmentation algorithm, the middle gray value for the entire image is chosen as the initial value. In fact, the gray values of the nucleus concentrate between 50 and 100 and there are little pixel elements (see Fig. 2). If the initial threshold is generated in the whole image, the final threshold will be focused on 100 and 150. Then the segmented nucleus will be embedded in the background. Therefore, for calculating the initial threshold, we first let:

$$T = \frac{\min(f(i,j)) + \max(f(i,j))}{2}. \quad (1)$$

Where, f stands for the image and f(i,j) stands for the gray value of a certain pixel in the image. Then T is the middle value of the entire image so that the whole image is compressed to 0-T-1 gray level. So the image has T grades after this translation. The aim of this compression is to make the nucleus outstand. The steps of the algorithm are as follows:

1. Calculate the initial threshold, let i=0, the calculating equation is

$$T^i = \frac{\sum_{k=0}^{T-1} h_k * k}{\sum_{k=0}^{T-1} h_k}. \quad (2)$$

Where, h_k is the numbers for the gray values equal to k.

2. Divide the image into two groups $R_1$and R_2 using the threshold T^i.

$$R_1 = \{f(i,j) | f(i,j) \geq T^i\}$$
$$R_2 = \{f(i,j) | 0 \leq f(i,j) \leq T^i\}$$

3. Calculate the mean gray values $\mu_1$and μ_2 for R_1 and R_2.

$$\mu_1 = \frac{\sum_{k=0}^{T^i} h_k * k}{\sum_{k=0}^{T^i} h_k} \quad (3)$$

$$\mu_2 = \frac{\sum_{k=T^i+1}^{T-1} h_k * k}{\sum_{k=T^i+1}^{T-1} h_k} \quad (4)$$

4. Calculate the new threshold

$$T^{i+1} = \frac{\mu_1 + \mu_2}{2}. \quad (5)$$

5. For an arbitrary decimal ξ, judge T^i-$T^{i+1} \leq \xi$?, if no, let $i=i+1, T^i=T^{i+1}$, then turn to 3, if yes, the iteration process ends.

6. Use T^i to do binary treatment for the initial image f(i,j).

$$f(i,j) = \begin{cases} 1(nucleus) & 0 < f(i,j) \leq T^i \\ 0(background) & T^i < f(i,j) \leq T \end{cases} \quad (6)$$

The image produced after segmentation based on our algorithm contains two kinds of objects: nucleus and background (see Fig. 3).

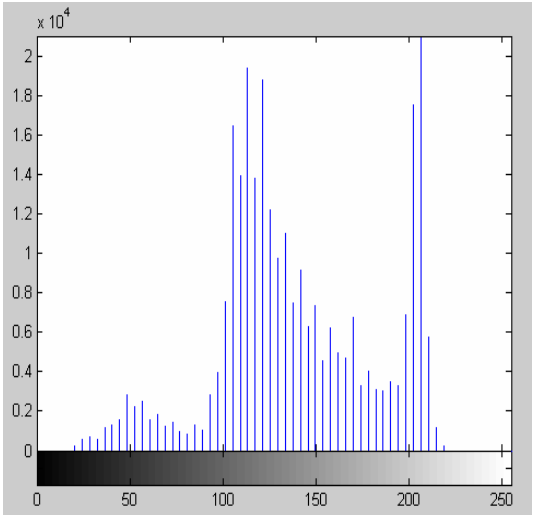


Fig. 2. The histogram of the cell based on gray image

C. The segmentation of cytoplasm

In order to extract the compact cell, the cytoplasm should also be removed from the background to segment. But different from the nucleus, there is no stable summit distribution with other objects for cytoplasm in gray-scale image. There are also certain types of nucleated cells which have little cytoplasm and even some of them have no cytoplasm such as small lymphocytic so that we can't identify

the cytoplasm just in gray space. But in HSI space, saturation of cytoplasm forms a great difference with other regions so the segmentation of cytoplasm is carried out in HSI space. The image is converted to its HSI equivalent using[4],[5],

$$H = \cos^{-1}\left\{\frac{(R-G)+(R-B)}{2\sqrt{(R-G)^2+(R-B)(G-B)}}\right\} \quad (7)$$

$$S = 1 - \frac{3}{R+B+G}[\min(R,G,B)] \quad (8)$$

$$I = \frac{R+G+B}{3}. \quad (9)$$

Each pixel in the image is represented by a vector of there components namely, H, S and I. The S- component, as it plays a more conspicuous, is given more weightage as compared to the other two. The iteration threshold segmentation algorithm is performed on this collection of vectors for the segmentation of cytoplasm. However, thresholding the image, the noise comes into the image. Then two post-processings are carried out based on the segmented images.

1. The noise is resolved performing a median filtering using a 3*3 window[6].

2. The logic and computation is carried out on the nucleus and cytoplasm which are segmented separately based on gray space and color space to integrate the whole cell. The two parts are also given different regional markers (gray values).

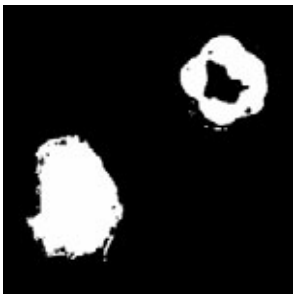


Fig. 3. The segmentation result of the nucleus

IV. Contour Detection Based On The Morphological Gradient

In this stage, the cell contours will be located and recognized and the different kinds of contours such as the nucleus' and the cytoplasm's will be marked. Edge detection is the basic step for the quantitative calculation and analysis of cell size, roundness, the number and other features. Its results directly impact on the analysis and diagnosis of the disease so that it will be impossible for cell morphology analysis if the edge detection is not ideal [7].

Mathematical morphology offers a powerful tool for extracting image components that are useful in the representation and description of region shape, such as boundaries, skeletons and texture. The morphological gradient is applied to enhance the boundaries of the cells. The morphological gradient of an image which is defined in terms of dilation $f \oplus g$ and erosion $f \Theta g$, is denoted *GRAD (f)* [8]:

$$GRAD(f) = (f \oplus g) - (f \Theta g). \quad (10)$$

Where, g stands for the structure element. Considering the shape of the cell is likely roundness, here 3*3 disk-shape structuring element is applied. As opposed to gradients obtained using such methods as Sobel operation, morphological gradient highlights sharp gray level transitions in the input image. The operation of the gradient calculation is shown in Fig. 4. Combined the technology of morphological gradient with thresholding segmentation, the contour is detected (see Fig. 5). Different gray values are given to the contours of nucleus and cytoplasm in order to distinguish between them.

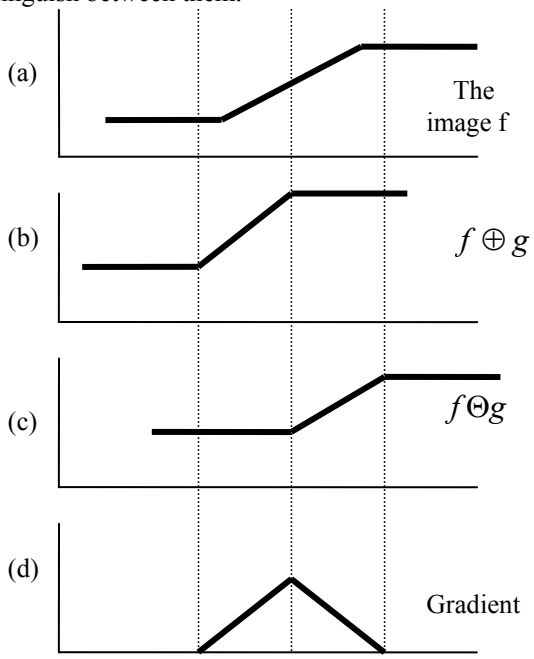


Fig. 4. The morphological gradient
(a) The image f (b) The dilation
(c) The erosion (d) The gradient

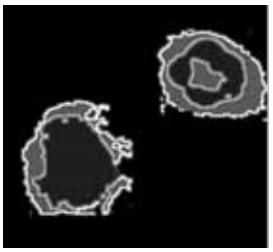


Fig. 5. The contour image detected by the gradient

V. Automatic Location For Single Nucleated Cell

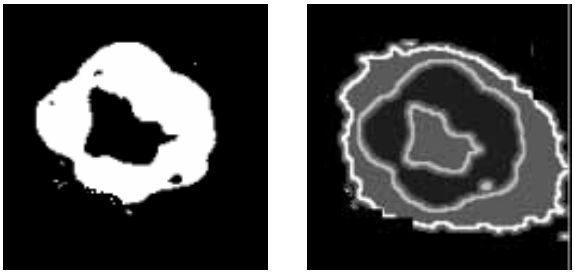
Since the original image involves a number of nucleated cells for identifying, a single nucleated cell should be separated and labeled for the purpose of individual analysis[2]. The main steps are as follows:

1. First scan the segmented image until all the

nuclear-mark points are found. Then we can search out their connected domains. Every domain will be given a new label until all the nuclear markers are updated.

2. Calculate the distance between the center (in general it means focus) of different connect domains. If the distance between them is less than the given range (usually given by experience), then they belong to the same cell (such as sub leaves nuclear) so they can be grouped in a category with the same label.

3. Calculate the center position of effective labeling regions. Make the position as the center to cut a certain sub-image of square (this square region should be able to include an intact blood cell). Thus the automatic detection of blood cells is completed. The detection result of the nucleated cell marked 1 in Fig. 1(a) is shown in Fig. 6.



(a) (b)

Fig. 6. The result of automatic detection of the single cell
(a) The detection result of the nucleus
(b) The detection result of the contour image

VI. The Criteria For Judging The Results Of Segmentation

For the evaluation of the segmentation algorithm, the visual judging method described in literature [9] is adapted in the paper. First outline the profile image of nucleus and cytoplasm manually and make off their respective regions, then calculate the number of pixels in each region. Since in front we have obtained the contour of the cell by employing the morphological gradient, so by comparing the pixels calculated in this way with the pixels calculated based on our algorithm, use the correct and error percentage to evaluate the segmentation effect. Table I shows the performance of the segmentation algorithm.

TABLE I
THE PERFORMANCE TABLE OF THE SEGMENTATION ALGORITHM

The number	Correct rate of nucleus (%)	Correct rate of cytoplasm (%)	Error rate of nucleus (%)	Error rate of cytoplasm (%)
1	95.5	92.6	4.5	7.4
2	94.5	93.4	5.5	6.6
average	95.0	92.5	5.0	7.0

VII. Conclusion

The segmentation system presented in this paper makes full use of the gray image, the gradient, the color information and the prior knowledge. The boundary of the cell is detected effectively and every different region is given a respective mark. The segmentation performance table verifies the validity of the algorithm in segmenting the cells. The result provides a useful data information for the next feature extraction.

Acknowledgment

This work was supported by Shandong province Natural Science Foundation. (Grant No. Y2005C68). The authors would like to thank Shandong Province-Established Hospital for providing the real cell data.

References

[1] Guclu Ongun, Ugur Halici, Kemal Leblebicioglu, Volkan Atalay, Meral Beksac, and Sinan Beksac, "An automated differential blood count system," Proceedings of the 23rd Annual EMBS International Conference, 2001, pp.2583-2586.

[2] Xiaomin YANG, Limin LUO, andYu WEI, "The study of automatic segmentation of the blood cell based on the computer," Chinese Journal of Biomedical Engineering, vol. 14, no. 4, pp.45-50,1995.

[3] Bin HE, Tianyu MA,Yunjian WANG, and Honglian ZHU, "Digital Image Processing Using Visual C++[M]," Publishing House of People Postation, Beijing, pp, 465-466, 2001.

[4] Rafacael C.Gonzalez and Richard E.Woods, "Digital Image Processing. Publishing House of Electronic Industry," Beijing, pp, 295-300, 2002.

[5] Sinha,N. and Ramakrishnan,A.G, "Automation of Differential Blood Count," Digital Object Identifier, 2(15-17), pp,547-551,2003.

[6] S.F.Bikhet, A.M.Darwish , H.A. Tolba and S.I.Shaheen, "Segmentation and classification of white blood cells," IEEE Intl. Conf. On Acoustics, Speech and Signal Processing, ICASSP, pp, 2259-2261,1999 .

[7] J. Theerapattanakul, J. Plodpai, and C. Pintavirooj, "An Efficient Method for Segmentation Step of Automated White Blood Cell Classifications," TENCON 2004, 2004 IEEE Region 10 Conference, 1(21-24):191-194, 2004.

[8] C. Di Rubeto, A. Dempster, S. Khan, and B.Jarra, "Segmentation of Blood Image using Morphological Operators," Proceeding 15th International Conference on Pattern Recognition, vol.3, pp, 397-400, 2000.

[9] Lezoray O, Elmoataz A, and Cardot H, "Segmentation of color images from serious cytology for automated cell classification [J]," Analytical and Quantitative Cytology and Histology, vol.2, No.4, pp, 311-322, 2000.