Automatic Detection of Red Blood Cells in Hematological Images Using Polar Transformation and Run-length Matrix

S. H. Rezatofighi*, A. Roodaki, R. A. Zoroofi
Control and Intelligent Processing Center of Excellence, Department of Electrical and Computer Engineering, Faculty of Engineering, University of Tehran, Tehran 14395-515, Iran
**h.tofighi@ece.ut.ac.ir*

R. Sharifian
Hematology-Oncology and BMT Research Center, Tehran University of Medical Sciences, Tehran, Iran

H. Soltanian-Zadeh
Control and Intelligent Processing Center of Excellence, Univ. of Tehran, Tehran, Iran & Henry Ford Hospital, Detroit, MI, USA.
hamids@rad.hfh.edu

Abstract

Many of diseases related to red blood cells can be diagnosed by analyzing the hematological images. The first step is to locate precisely the position of red blood cells. In this paper, a novel method based on polar transformation and run-length matrix is proposed for detecting red blood cells in hematological images. The multilayer perceptron was employed for classifying the feature vectors. This technique detected 3022 of the 3092 red blood cells existed in the images studied. The results illustrate capability of the proposed method for detecting red blood cells in hematological images.

1. Introduction

Circulation system is one of the most important organizations in the human body. In addition to the importance of cardiovascular system, some tiny red blood cells have the vital duty of carrying food and oxygen to different parts of the body and also collecting unwanted materials from them. Therefore, the whole body would suffer from the diseases that prevent red blood cells from doing their tasks.

Much diagnostic information can be extracted from the shape and size of the red blood cells, the number of the red blood cells in a particular sample of the blood, and the ratio between the area containing oxygen and the whole area of each cell [1]. Analyzing the hematological images manually is tiresome and time consuming. It also suffers from inter- and intra-observer variabilities. Hence, automation of this task will be helpful in identifying diseases related to the red blood cells accurately, besides saving the precious time of the hematologists.

Detection of the red blood cells is the first and most important step which presents several challenges. The images are typically damaged by artifacts. In addition, overlapping of the cells is another problem in detecting the red blood cells. Moreover, the hematological images suffer from variations in color and intensities [2]. Therefore, the techniques employed for this purpose should be robust to these problems.

Despite the importance of this procedure for the diagnosis of the red blood cells' diseases, only a few related researches are found in the literature. In [3], the image intensities were used as features and the Principal Component Analysis (PCA) was employed for feature reduction. Using local linear map neural network as classifier, they managed to detect the leukocytes in fluorescent images. On the other hand, in [4], Fisher Linear Discriminant (FLD) was used to improve the features extracted based on image's intensities, and Multi-Layer Perceptron (MLP) was employed for detecting the leukocytes in the fluorescent images. However, using intensities as features may not be very effective in detecting the overlapping cells. The second-order edge detection methodology for detecting the overall number of cells irrespective of their stain was proposed in [5]. Then, proliferating cells were located using PCA of the color image along with histogram thresholding.

Our search did not find any previous work on the automatic detection of red blood cells using hematological image analysis. As such, we seem to be

the first research group proposing an automated method for this purpose. The idea of our proposed method lies in the fact that the variations of characteristics of cells are more obvious in the Polar coordinates. Fig. 1 shows the different locations of a red blood cell with respect to the center of the image and their corresponding polar-transformed images. These situations occur when scanning the hematological images with an appropriate sweeping window.

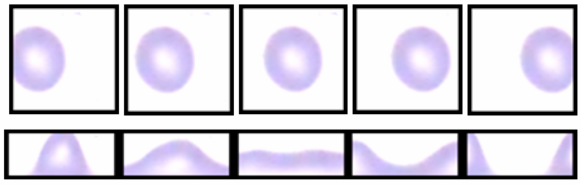

Fig. 1. Variations of the characteristic of shapes by moving the sweeping window. Second row is in the polar coordinates.

The location of the cell may be identified by studying the characteristics of the sub-images in the polar coordinate system. In particular, the lengths of the horizontal runs, for each specific intensity in each row of the sub-image in the polar coordinates, seem to provide valuable discriminative information. Our proposed method is based on this fact.

The rest of the paper is organized as follows. Section 2 describes the methods and the proposed algorithm employed for detecting the red blood cells in the hematological images. Section 3 provides the results obtained using the proposed method. Finally, Section 4 concludes the paper.

2. Methodology

In the feature extraction step, the image is swept by a window of a certain size and then, appropriate features are extracted from the windows and are assigned to the pixel located in the center of the sweeping window. Then, using Neural Networks, the feature vectors are classified into two classes, red blood cells and background,

2.1. The Pol-run feature extraction method

A number of features can be extracted from the polar-transformed images. As shown in Fig. 1, most of discriminative information lies in the differences of run lengths for a specific intensity in a specific row. The run-length method has the ability to code the length of runs related to each of the image intensities [6]. Hence, the proposed features are extracted from the run-length matrix of the polar image.

A gray level run is a set of repeated pixels having the same gray level value. The length of the run is the number of pixels in the run. The run-length matrices encode textural information related to the number of times each gray level appears in the image by itself, the number of times it appears in pairs, and so on.

The run-length matrix can be defined in four directions (0°, 45°, 90°, and 135°). Its (i,j)th element gives the number of times a gray level i - 1, i = 1, . . . , Ng, appears in the image with run length j , j = 1,2, ..., Nr. This is an Ng × Nr array, where Nr is the largest possible run length in the image. Features such as short run emphasis, long run emphasis, gray-level nonuniformity, run nonuniformity, and run percentage have previously been extracted from the run length matrix and represent different textural properties [6].

However, these five features only encode the general information of intensity variations. Thus, we lose the specificity provided by the run-length matrix needed for discriminating the exact places of the red blood cells from their neighbors. Therefore, some new features need to be extracted from the run-length matrix with the ability to describe the details of differences between the image windows that contain the cell in their center from the other ones.

We studied each of the rows of the run-length matrix related to a specific intensity separately. We observed that most of the discriminative information is found in the maximum number of occurrences multiplied by its run length in each row of the run-length matrix. This defines the first proposed feature. The drawback of this feature is its lack of uniqueness. For improving the reliability of our method another feature is extracted from each row of the run-length matrix which is the sum of every occurrence multiplied by its corresponding run length. Assuming $a_{i,j}$ as the (i,j)th element of the run-length matrix, two features are extracted from the I^{th} row of the run-length matrix by:

$$f_1 = J_{\max(a_{I,j})} \times \max(a_{I,j})$$
$$f_2 = \sum_{j=1}^{N_r} j \times a_{I,j} \quad (1)$$

In a nutshell, the image is first swept by an appropriate window, and then each window is transformed to the Polar coordinates assuming that the reference point of the polar transformation lies in the center of the window. Next, a run-length matrix for each row of the polar-transformed image in the

horizontal direction is constructed. Finally, two aforementioned features are extracted from each row of the run-length matrix. So, the dimension of the feature vectors depends on the size of window and the number of gray levels existing in the image. For example, if the size of the sweeping window is N*N and the number of gray levels is Ng, then the dimension of the feature vectors would be $2 \times N \times Ng$.

2.2. Multilayer Perceptron

Multilayer Perceptrons (MLPs) are a kind of Artificial Neural Network (ANN) classifiers consisting of hidden layers in addition to the input and output layers [7]. These classifiers have been extensively used in the field of pattern recognition especially when the number of feature vectors is high [4]. Here, we employ a two layer perceptron for classifying the extracted feature vectors.

3. Results and Discussion

Our proposed method was tested on 22 Blood smear slide images acquired by light microscope from stained peripheral blood using a Digital Camera-Sony-Model No.SSC-DC50AP with magnification of 100. The resolution of images is 720*576 pixels.

Since the dimensionality of the pol-run features depends on the number of gray levels, the image intensities were first decimated to 16 gray levels. The size of the sweeping window was empirically chosen to be 15*15 so that all the red blood cells fit in it.

The whole database was divided into two parts: the train set consisting of six images and the test set consisting of the rest of the images.

First, other components such as white blood cells, platelets and some staining trashes were removed by a fast segmentation algorithm based on Gram-Schmidt orthogonalization [8]. The resulting images were swept by a 15*15 window. Then, each sub-image was transformed to the polar coordinates. As the number of intensities must remain fixed after transformation, the nearest neighbor method was used for interpolation.

Then, a run-length matrix was constructed for each row of the polar-transformed image. Next, two aforementioned features were extracted from each row of the run-length matrices.

In the classification section, the feature vectors were classified using a two layer perceptron with 40 neurons in the hidden layer. If the output after global thresholding is 1, it shows that the decision of the network is that the pixel belongs to the red blood cell class and the thresholded output -1 implies that the pixel is belongs to the background.

Since the image generated by the neural network was noisy, a post processing step was employed for improving the results. First, the image was eroded using a structural element of a certain size. Next, the image was dilated using the same structural element used for the erosion. Then, connected components smaller than a pre-specified size were removed. Finally, for each connected component, a point was selected as the center point (see Fig. 2.).

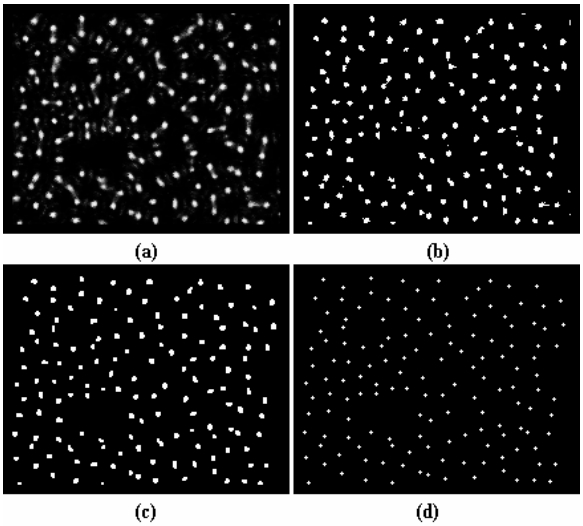

Fig. 2. (a) Image generated by the neural network, (b) Image after global thresholding, (c) Image after post processing, (d) Image showin center points of the connected components.

Totally there were 3092 cells in the test set. 3022 cells were detected in the right places yielding the true positive rate of 97.73%. The false positive rate of the algorithm was quite low as a total of 4 wrong cells were identified. A sample result generated by the proposed method is illustrated in Fig. 3.

4. Conclusions

In this paper, a novel method was introduced for detecting the red blood cells in hematological images. This technique is based on length of runs extracted from polarized sweeping windows, which we call "Pol-Run." It can be inferred from the results that the Pol-Run features are able to discriminate objects with different shapes and almost equal sizes from the background. These features even are capable of detecting overlapping objects.

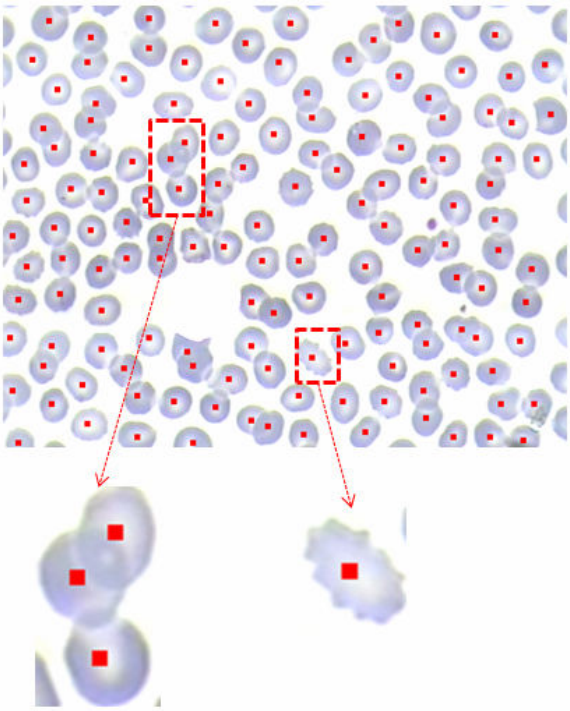

Fig. 3. Result of our proposed method on a typical hematological image. Zoomed images show the result for three overlapping cells and a jagged cell.

However, one of the limitations of this proposed method is in detecting objects with large size variations. Also, due to dependency of dimension of the feature vectors on the size of the objects, classifying them is complicated when large objects are the main goal. Therefore, in this situation, one can decimate the size of the images to an appropriate size by a fix proportion as long as the details of the objects are preserved. Then, the decimated images can be probed for the desired objects. Finally, the obtained points as locations of objects can be mapped into the original image.

Much diagnostic information can be extracted from the shape of the red blood cells and their number. In general, there are three specific shapes for the red blood cells. These shapes are shown in Fig. 4. As a future work, we are planning to detect and count different red blood cells based on their shapes using the Pol-Run features.

Fig. 4. Three different forms for red blood cells.

5. References

[1] Guyton, C. and E. Hall, *Text Book of Medical Physiology*, Elsevier Saunders, Pennsylvania USA, 2006.

[2] N. Guo, L. Zeng, and Q. Wu, "A Method Based on Multi-spectral Imaging Technique for White Blood Cell Segmentation", *Computers in Biology and Medicine,* vol. 37, 2006, pp. 70-76.

[3] T.W. Nattkemper, H.J. Ritter, and W.Schubert, "A Neural Classifier Enabling High-throughput Topological Analysis of Lymphocytes in Tissue Sections", *IEEE Trans Inf. Technol. Biomed.*, vol. 5, no. 2, 2001, pp. 138–149.

[4] X. Long, W.L. Cleveland, and Y.L. Yao, "A New Preprocessing Approach for Cell Recognition", *IEEE Trans Inf. Technol. Biomed.*, vol. 9, no. 3, 2005, pp. 407–412.

[5] C.G. Loukas, G.D. Wilson, B. Vojnovic, and A. Linney, "An Image Analysis-based Approach for Automated Counting of Cancer Cell Nuclei in Tissue Sections", *Cytometry,* Part A, 55A, 2003, pp. 30-42.

[6] Theodoridis, S. and K. Koutroumbas, *Pattern Recognition*, Elsevier academic press, USA, 2003.

[7] Bishop, C.M. , *Neural Networks for Pattern Recognition*, Clarendon Press, Oxford, UK, 1997.

[8] S.H. Rezatofighi, H. Soltanian-Zadeh, R. Sharifian, and R.A. Zoroofi, "Segmentation of Nucleus of White Blood Cells Using Gram-Schmidt Orthogonalization", Submitted *to International Symposium on Telecommunication 2008,* Tehran, Iran.