A New Approach to White Blood Cell Nucleus Segmentation Based on Gram-Schmidt Orthogonalization

S. H. Rezatofighi
Control and Intelligent Processing Center of Excellence, University of Tehran, Tehran, Iran
e-mail: h.tofighi@ece.ut.ac.ir

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

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

R.A. Zoroofi
Control and Intelligent Processing Center of Excellence, University of Tehran, Tehran, Iran

***Abstract*—The differential counting of white blood cells provides invaluable information to hematologist for diagnosis and treatment of many diseases. Manually counting of white blood cells is a tiresome, time-consuming and susceptible to error procedure. Due to the tedious nature of this process, an automatic system is preferable. In this automatic process, Segmentation of white blood cells is one of the most important stages. The nucleus of white blood cells has the most information about the type of white blood cells, thus an exact segmentation seems to be helpful for other stages of automatic recognition of white blood cells. In this paper, we introduced a novel method based on orthogonality theory and Gram-Schmidt process for segmenting the nuclei of white blood cells. For evaluation of results, we compared our proposed method with a hematologist manual segmentation. Results show robustness of this technique for segmentation of nuclei, while this method is very simple to implement.**

Keywords- gram-schmidt process, hematological image, nucleus segmentation, orthogonality, white blood cell

I. INTRODUCTION

White blood cell composition reveals important diagnostic information about the patients. Substituting automatic detection of white blood cells for manually locating, identifying, and counting different classes of cells is an important topic in the domain of cancer diagnosis. Microscopic differential white blood cell count is still performed by hematologists, being indispensable in diagnostics with malignance suspicious. While its value as a reference method for blood samples containing abnormal cells remains indisputable, it is slow and subjective and its reproducibility is poor [1]. Therefore, automation of this task is very helpful for improving the hematological procedure and accelerating diagnosis of many diseases.

Automatic recognition of white blood cells in hematological images usually consists of four major steps, including: preprocessing, image segmentation, feature extraction and classification (see Fig. 1).

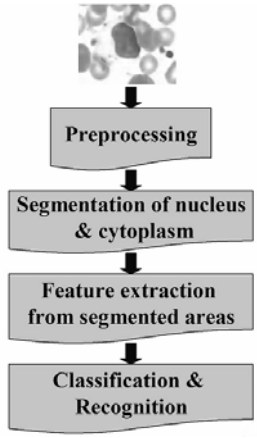


Figure 1. Common block diagram for white blood cell recognition

The segmentation step is very crucial because the accuracy of the subsequent feature extraction and classification depends on the correct segmentation of white blood cells. It is also a difficult and challenging problem due to the complex nature of the cells and uncertainty in the microscopic images [2]. Therefore, this step is the most important challenge in many literatures and improvement of nucleus and cytoplasm segmentation has been the most common effort in many researches.

The most common segmentation method is using edge detection [3] method which is based on HSI model of a color image. In [4]-[7], they try to segment the nucleus and cytoplasm based on selecting color features and thresholding on the histogram. Besides histogram thresholding and edge detection, clustering [8], and region growing [9] methods are often used for this purpose. Fuzzy clustering method and bayes classifier is applied for segmenting the nucleus of white blood cells [10]. For segmenting the white blood cells, multi-spectral imaging technique is a new method which is introduced in [11]. Results from this paper show that this method is very efficient for this purpose. But, execution of such system is very cost-consuming and it is not easy for making prevalent in all laboratories. In [12], [13] and [14], the watershed method, the deformable models and the

combination of these two methods are applied, respectively. Finally, other methods which are applied on blood images in previous works are based on energy [15], morphology [16], neural network [17], fuzzy [18], [21], model based systems [19].

Nuclei have the most invaluable information for determining the type of white blood cells. Then, quick and accurate segmentation of nucleus is necessary for automatic recognition of white blood cells. In most of hematological images analyzed for recognition of white blood cell's type, color of nuclei is often violet with different intensities and saturation levels. Also, by changing the blood's samples, variation of amount of intensities and saturations is increased; especially, when these images are acquired by different cameras and microscopes. Therefore, methods based on thresholding, HSI model and other methods which use color features are not capable of segmenting images accurately. On the other hand, some methods which are not based on color features are generally complicated and time-consuming.

In this paper, a method is introduced based on Gram-Schmidt orthogonalization for segmentation of nuclei of white blood cells. For validation, results of our proposed method are compared with manually segmentation by expert. Achieved results demonstrate that this method is very efficient for segmentation of nuclei of white blood cells; although it is very simple.

The rest of the paper is organized as follows. In section 2 we will introduce Gram-Schmidt orthogonalization briefly. In sections 3, segmentation of images by this method is discussed. Experimental Results are outlined in Section 4. Finally, Paper is concluded in section 5.

II. THE GRAM-SCHMIDT ORTHOGONALIZATION

In mathematics, particularly linear algebra and numerical analysis, the Gram–Schmidt process is a method for orthogonalzing a set of vectors in an inner product space, most commonly the Euclidean space R^n. The Gram–Schmidt process takes a finite, linearly independent set $S = \{v_1, \ldots, v_n\}$ and generates an orthogonal set $S' = \{u_1, \ldots, u_n\}$ that spans the same subspace as S.

We define the projection operator by

$$proj_u v = \frac{< u, v >}{< u, u >} u = < u, v > \frac{u}{< u, u >} \qquad (1)$$

where <u, v> denotes the inner product of the vectors u and v. This operator projects the vector v orthogonally onto the vector u.

The Gram–Schmidt process then works as follows:

$$\begin{aligned} u_1 &= v_1, & e_1 &= \frac{u_1}{\|u_1\|} \\ u_2 &= v_2 - proj_{u_1} v_2, & e_2 &= \frac{u_2}{\|u_2\|} \\ u_3 &= v_3 - proj_{u_1} v_3 - proj_{u_2} v_3, & e_3 &= \frac{u_3}{\|u_3\|} \\ &\vdots & &\vdots \\ u_k &= v_k - \sum_{j=1}^{k-1} proj_{u_j} v_k, & e_k &= \frac{u_k}{\|u_k\|} \end{aligned} \qquad (2)$$

The sequence $u_1, \ldots, u_k$ is the required system of orthogonal vectors, and the normalized vectors $e_1, \ldots, e_k$ form an orthonormal set [20].

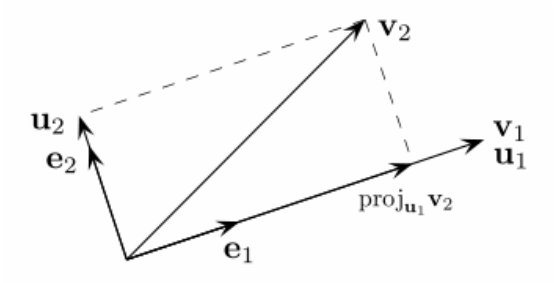


Figure 2. The first two steps of Gram-Schmidt process

Based on this method, for a linearly independent set $S = \{v_1, \ldots, v_n\}$, we can find a vector which has maximum orthogonality with one desired vector v_k and also minimum orthogonality with other vectors in N dimensional spaces. This vector w_k can be calculated according to following formula:

$$w_k = v_k - \sum_{j=1}^{k-1} proj_{v_j} v_k, \qquad (3)$$

Therefore, result of inner product of set S by w_k is:

$$\begin{cases} < v_j, w_k >= 0 & j \in 1,..,n \ \ and \ \ j \neq k \\ < v_k, w_k >= K & K \neq 0 \end{cases} \qquad (4)$$

In Fig. 3, relation between w_2 and v_1, v_2 and v_3 is shown graphically in three dimensional spaces.

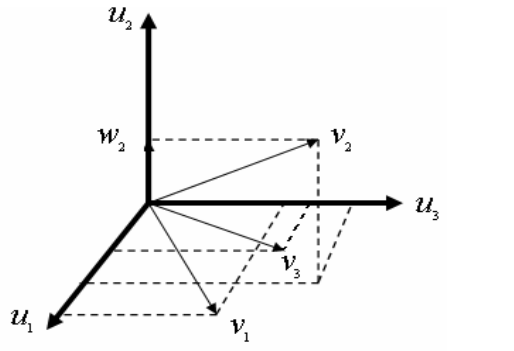
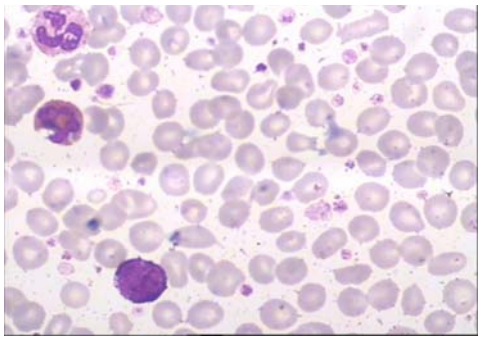


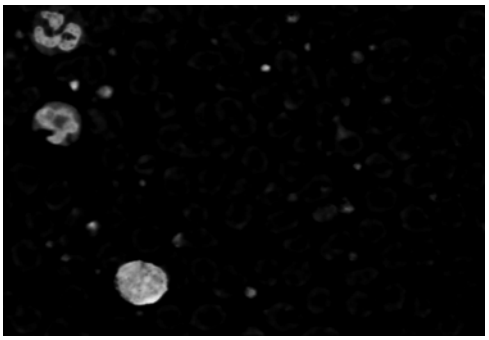
Figure 3. relation between w and v1, v2, and v3 in 3D spaces

III. SEGMENTATION OF NUCLEUS OF WHITE BLOOD CELLS

To apply Gram-Schmidt orthogonalization [20] for the segmentation of nucleus of white blood cells, a 3-D feature vector is defined for each pixel using the RGB components of the images. Then, according to the Gram-Schmidt method, a weighting vector *w* is calculated for amplifying the desired color vectors and weakening the undesired color vectors. As shown in Fig. 4, the inner product of the weighting vector and the pixel vectors defined from the original image produce a composite image which has maximum intensity in regions with violet color and minimum intensity in other regions.



(a)



(b)

Figure 4. (a) Original image, (b) Resulting image after multiplying pixels of original image by vector calculated by Gram-Schmidt method

Next, by choosing an appropriate threshold based on the histogram information, we segment the image. The resulting segmentation is noisy and contains both of nucleus areas and platelets. Because these platelets are smaller than the nuclei of the white blood cells, by removing small components, the nuclei are segmented.

Figure 5. final result after thresholding and removing fine areas

As mentioned in the introduction, the color of nuclei is violet with different intensities and saturation levels. Also, by changing the blood samples, the variations of intensity and saturation may increase. Due to this large variation of color, three different weighting vectors are calculated for the Gram-Schmidt orthogonalization. For each vector, an image which is similar to Fig. 4(b) is obtained and an appropriate threshold is calculated for each image. These thresholds are computed based on weighted mean and maximum of their histograms. Finally, after thresholding and removing small (fine) components, we apply the logical "AND" operation to the results. Fig. 6 shows a flow chart for our proposed scheme.

IV. RESULTS AND DISCUSSION

The proposed method was tested on approximately 251 blood smear slide images acquired by light microscope from stained peripheral blood using the Digital Camera-Sony-Model No. SSC-DC50AP with magnification of 100. The resolution of the images is 720*576 pixels. The digital images were classified by a hematologist into the normal leukocytes: basophil, eosinophil, lymphocyte, monocyte, and neutrophil. Also, areas related to the nuclei were manually distinguished by an expert. The automatic segmentation results were quantitatively evaluated using the manual segmentation results using the similarity measure given in Eq. (5).

$$Ts = \frac{(A_{program} \,\&\, A_{expert})}{\max(A_{program}, A_{expert})} \quad (5)$$

Here, $A_{program}$ is the segmented area by the proposed method and A_{expert} is the area determined by the expert hematologist.

Table 1 presents the similarity measures for the segmentation of nucleus of each type of white blood cells.

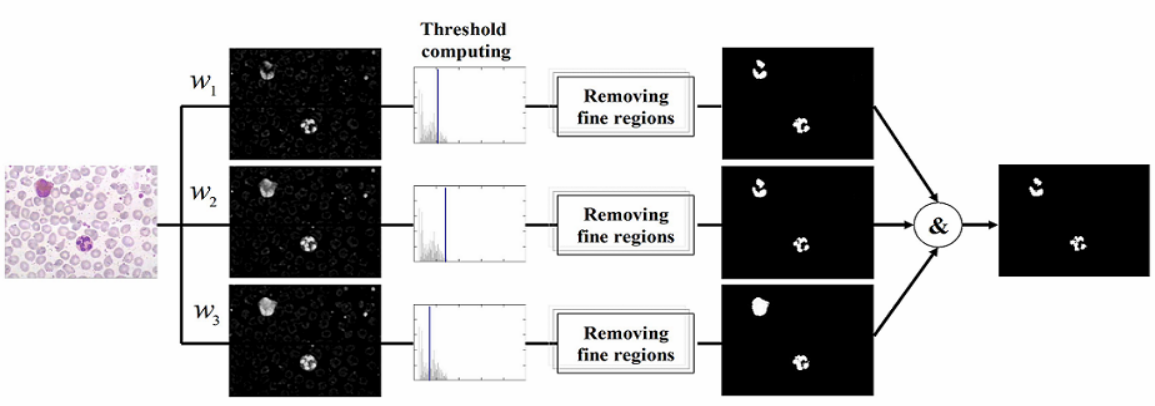


Figure 6. Block diagram for proposed scheme

TABLE I. RESULTS OF SEGMENTATION FOR EACH CLASS OF WHITE BLOOD CELLS

	Basophil	*Eosinophil*	*lymphocyte*	*monocyte*	*neutrophil*	**Total**
Nucleus segmentation	94.7%	90.81%	88.86%	96.7%	94.05%	93.02%

According to Table 1, the result of nucleus segmentation for the lymphocyte class is lower than the other classes. The main reason is that the color of the cytoplasm is very similar to the color of the nucleus in many of the lymphocytes samples. Therefore, the segmentation error for this type of the white blood cell is lager than the others.

V. CONCLUSION

In this research, we propose a new technique to segment nuclei of white blood cells in hematological images. This technique use orthogonality theory based on Gram Schmidt process for amplifying the desired color vectors and weakening the undesired color vectors. Due to the large variation of color in our data, we were forced to find three vectors instead of one. Finally, after choosing an appropriate threshold and removing the small components, "AND" operation is applied for final segmentation results. This method is very simple to implement; whenas it is very quick and efficient. But for establishment of this method on new data, we should calculate new vectors which are suitable for new data.

ACKNOWLEDGMENT

The authors would like to thank Dr. Reza Safaee and Ms. Hajar Mirmongereh for helping us gathering the hematological images and dataset.

REFERENCES

[1] D. M. U. Sabino, L. d. F. Costa, E. G. Rizzatti, and M. A. Zago, "A texture approach to leukocyte recognition," *Real-Time Imaging*, vol. 10, pp. 205-216, 2004.

[2] J. WU, P. Zeng, Y. Zhou, and C. Oliver, "A novel color image segmentation method and its application to white blood cell image analysis," *ICSP*, 2006.

[3] H. Hengen, S.L. Spoor, and M.C. Pandit, "Analysis of blood and bone marrow smears using digital image processing techniques, " in: *M. Sonka, J.M. Fitzpatrick, (Eds.), Proceedings of the SPIE, Medical Imaging 2002: Image Processing*, vol. 4684, pp. 624–635, 2002.

[4] I. Cseke, "A fast segmentation scheme for WBC images," *IEEE Conf. Image, Speech and Signal Analysis*, pp.530-533, 1992.

[5] S. F. Bikhet, A. M. Darwish, H. A. Tolba, and S. I. Shaheen, "Segmentation and clustering of WBC," *IEEE Conf. Acoustics, Speech, and Signal Processing*, vol. 4, pp. 2259-2261, 2000.

[6] Q. Liano and Y. Deng, "An accurate segmentation method for white blood cell images," *IEEE Conf. Biomedical Imaging*, pp. 245-248, 2002.

[7] J. Angulo and G. Flandrin, "Automated detection of working area of peripheral blood smears using mathematical morphology," *Analytical Cellular Pathology*, vol. 25, pp. 37-49, 2003.

[8] M. Trivedi and J. C. Bezedek, "Low-level segmentation of acrial images with fuzzy clustering," *IEEE Trans. On System, Man. And Cybernetics*, vol. 16, no 4, pp. 589-598, 1986.

[9] J. M. chassery and C. Garbay, "An iterative segmentation method based on contextual color and shape criterion," *IEEE Trans. On Pattern Analysis and Machine Intelligence*, vol. 6, no. 6, pp. 795-800, 1984.

[10] N. T. Umpon," Patch-based white blood cell nucleus segmentation using fuzzy clustering, "*ECTI Trans. Electrical. Electronic. Communications*, vol. 3, no. 1, pp. 5–10, 2005.

[11] 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, pp. 70-76, 2006.

[12] K. Jiang, Q. M. Liao, and S. Y. Dai, "A novel white blood cell segmentation scheme using scale-space filtering and watershed clustering," *Proceedings of the Second International Conference on Machine Learning and Cybernetics*, pp. 2820-2825, 2003.

[13] F. Zamani and R. Safabakhsh, "An unsupervised GVF snake approach for white blood cell segmentation based on nucleus," *ICSP*, 2006.

[14] J. Park and J. M. Keller, “Snakes on the watershed,” *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 23, No. 10, 2001.

[15] B. Ravi Kumar, D. K Joseph and T. V. Sreenivas, “Teager energy based blood cell segmentation,” *IEEE Conf. Digital Signal Processing*, vol. 2, pp. 619-622, 2002.

[16] C. Ruberto, A. Dempster, S. Khan, and B. Jarra, “Analysis of infected blood cell images using morphological operators,” *Image and Vision Computing*, vol. 20, pp. 133-146, 2002.

[17] H. Sheikh, B. Zhu, and E. Micheli, “Blood cell identification using neural networks,” *IEEE Conf. Biomedical Imaging*, pp. 245-248, 1996.

[18] J. Park and J. Keller, “Fuzzy patch label relaxation in bone marrow cell segmentation,” *Proc. IEEE Int'l Conf. Systems, Man. and Cybernetics*, pp. 1133-1138, Oct. 1997.

[19] B. Nilsson, A. Heyden, A. B. Cellavision and S. Lund, “Model-based segmentation of leukocytes clusters,” *IEEE Conf. of Pattern Recognition*, vol. 2, pp. 727-730, 2002.

[20] H. Soltanian-Zadeh, D. J. Peck, D. O. Hearshen, and R. R. Lajiness-O’Neill, “Model independent method for fMRI analysis,” *IEEE Trans. Med. Imag.*, vol. 23,no. 3, 2004.

[21] E. Montseny, P. Sobrevilla, and S. Romani, “A fuzzy approach to white blood cells segmentation in color bone marrow images,” *IEEE Int. Conf. on Fuzzy Systems*, pp. 173-178, 2004.