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

S. H. Rezatofighi H. Soltanian-Zadeh R. Sharifian R.A. Zoroofi
Control and Intelligent Processing Center of Excellence, University of Tehran, Tehran, Iran
e-mail: h.tofighi@ece.ut.ac.ir
Hematology-Oncology and BMT Research Center, Tehran University of Medical Sciences,
Tehran, Iran
e-mail: hamids@rad.hfh.edu
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).

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 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 orthogonalizing a set of vectors in an inner product space, most commonly the Euclidean space $\mathbb{R}^n$. The Gram-Schmidt process takes a finite, linearly independent set $S = \{v_1, ..., v_n\}$ and generates an orthogonal set $S' = \{u_1, ..., u_n\}$ that spans the same subspace as $S$.

We define the projection operator by

$$proj_u v = \frac{\langle u, v \rangle}{\langle u, u \rangle} u = \langle u, v \rangle \frac{u}{\langle u, u \rangle}$$

(1)

where $\langle u, v \rangle$ 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:

$$u_1 = v_1, \quad e_1 = \frac{u_1}{\|u_1\|}$$

$$u_2 = v_2 - proj_{u_1} v_2, \quad e_2 = \frac{u_2}{\|u_2\|}$$

$$u_3 = v_3 - proj_{u_1} v_3 - proj_{u_2} v_3, \quad e_3 = \frac{u_3}{\|u_3\|}$$

$$\vdots$$

$$u_k = v_k - \sum_{j=1}^{k-1} proj_{u_j} v_k, \quad e_k = \frac{u_k}{\|u_k\|}$$

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

Based on this method, for a linearly independent set $S = \{v_1, ..., v_n\}$, 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:

$$w_k = v_k - \sum_{j=1}^{k-1} proj_{v_j} v_k$$

(3)

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

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

(4)

In Fig. 3, relation between $w_2$ and $v_1$, $v_2$ and $v_3$ is shown graphically in three dimensional spaces.
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 3. relation between w and v1, v2, and v3 in 3D spaces](image)

**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.

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

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).

\[
T_S = \frac{(A_{\text{program}} \& A_{\text{expert}})}{\max(A_{\text{program}}, A_{\text{expert}})} \quad (5)
\]

Here, \( A_{\text{program}} \) is the segmented area by the proposed method and \( A_{\text{expert}} \) 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 5. final result after thresholding and removing fine areas](image)
TABLE I. RESULTS OF SEGMENTATION FOR EACH CLASS OF WHITE BLOOD CELLS

<table>
<thead>
<tr>
<th>Class</th>
<th>Basophil</th>
<th>Eosinophil</th>
<th>Lymphocyte</th>
<th>Monocyte</th>
<th>Neutrophil</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nucleus segmentation</td>
<td>94.7%</td>
<td>90.81%</td>
<td>88.86%</td>
<td>96.7%</td>
<td>94.05%</td>
<td>93.02%</td>
</tr>
</tbody>
</table>

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 larger 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 uses 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


