Journal of Innovative Optical Health Sciences
Vol. 7, No. 1 (2014) 1450007 (8 pages)

DOI: 10.1142/S1793545814500072

World Scientific
www.worldscientific.com

A METHOD OF LEUKOCYTE SEGMENTATION BASED ON S COMPONENT AND B COMPONENT IMAGES

YIPING YANG, YIPING CAO* and WENXIAN SHI
Department of Optical Electronics
Sichuan University, Chengdu
Sichuan 610064, P. R. China
**caoyping@mail.sc.cninfo.net*

Received 3 July 2013
Accepted 29 August 2013
Published 19 December 2013

A leukocyte segmentation method based on S component and B component images is proposed. Threshold segmentation operation is applied to get two binary images in S component and B component images. The samples used in this study are peripheral blood smears. It is easy to find from the two binary images that gray values are the same at every corresponding pixels in the leukocyte cytoplasm region, but opposite in the other regions. The feature shows that "IMAGE AND" operation can be employed on the two binary images to segment the cytoplasm region of leukocyte. By doing "IMAGE XOR" operation between cytoplasm region and nucleus region, the leukocyte segmentation can be retrieved effectively. The segmentation accuracy is evaluated by comparing the segmentation result of the proposed method with the manual segmentation by a hematologist. Experiment results show that the proposed method is of a higher segmentation accuracy and it also performs well when leukocytes overlap with erythrocytes. The average segmentation accuracy of the proposed method reaches 97.7% for segmenting five types of leukocyte. Good segmentation results provide an important foundation for leukocytes automatic recognition.

Keywords: Image segmentation; leukocyte; component image; B component image.

1. Introduction

The leukocyte is a crucial kind of blood cell in human blood and plays an important role in people's immune system. In general, leukocytes are divided into five types[1] of neutrophil, eosinophil, basophil, monocyte and lymphocyte. Different types of leukocytes have different functions in people's immune system, so various diseases will cause different leukocyte changes in quantity and shape. Leukocyte detection has very important significance for clinical medical examination.

The traditional leukocytes detection mainly relies on manual operation, however, leukocytes automatic identification technology has made a

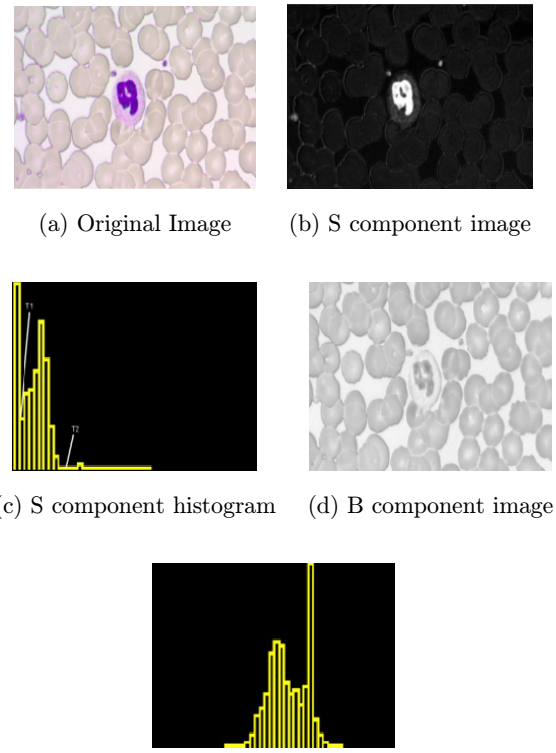
huge progress with the development of computer science and image processing technology. Image segmentation plays an important role in the leukocyte automatic identification technology. A number of image segmentation methods have been proposed in the recent decades, for example, the region growing method.[2,3] However, segmentation accuracy of this method relied greatly on proper seed extraction. By using the watershed algorithm,[4,5] Jiang and colleagues[4] introduced a method based on watershed clustering, this scheme successfully avoids the variety and complexity in image space, but it cannot perform well in the case of high density of cells, where lots of erythrocytes are closed to the leukocyte. Neural network algorithm[6–8] is proposed by some researchers, they have not considered different intensity of same objects in different images in Giemsa staining method, and calculations are more complicated. In the segmentation method based on fuzzy theory,[9,10] Nipon Theera-Umpon *et al.*,[10] divide the color areas and distribute each pixel into each area, and segment single cell images of white blood cells in bone marrow into two regions, nucleus and non-nucleus. This method is based on the fuzzy C-means clustering and mathematical morphology. Segmentation method based on deformable model,[11–13] and other methods.[14,15] Cseke[16] presented an algorithm based on the B component image, this method can realize leukocyte segmentation quickly and simply, but it is not able to implement segmentation accurately when leukocyte is overlapped with erythrocyte. Ravi Kumar *et al.*[17] propose a method based on color model, they define a new edge operator on the G component of the processed color image to highlight the nucleus boundary which is effective to segment nucleus. They segment cytoplasm from background and red blood cell to make use of the morphological method, which require some assumptions that may be untenable in many cases, such as color of the background may be very similar to leukocyte.

A large number of experiments show that the leukocyte images can be segmented by thresholds in S component and B component images, respectively to obtain two binary images. Gray values are the same at every corresponding pixels in cytoplasm region, but opposite in other regions for the two binary images. This feature shows that "IMAGE AND" operation can be applied on the two binary images to segment the cytoplasm region of leukocyte, and leukocyte nucleus region can be obtained by the method proposed in Ref. 18. The "IMAGE XOR" operation is conducted between the cytoplasm region binary image and the nucleus region binary image to get a binary image. Then the whole leukocyte region can be accurately segmented after doing some relative operation, such as "IMAGE CLOSE" operation, "IMAGE OPEN" operation, and so on. A method of leukocyte segmentation based on S component and B component images is proposed. In this work, the method is able to realize leukocyte segmentation when leukocyte overlaps with erythrocyte and excluded interference of the stains whose color is similar to leukocyte nuclei. It can obtain better result in different density of cells than Jiang's method.[4] Comparing with Ravi Kumar's approach,[17] the method proposed in this paper is able to perform better for various colored backgrounds. The proposed method possesses less calculation than region growing method and fuzzy theory.

2. Space Vectors Selection and Extraction Methods

2.1. *Selection and analysis of space vectors*

RGB and HSI are commonly used to describe color space. A variety of colors are obtained by superposed red (R), green (G) and blue (B) color channels for RGB color space. And the RGB color space is usually used in devices for color display. H (Hue), S (Saturation), I (Intensity) compose the HSI color space, which is suitable for human visual characteristic, so HSI is considered for image processing. Computing the histogram of H component, I component and S component images, respectively, it is easy to find that there are two troughs existing in S component histogram, as is shown in Fig. 1. Figure 1 (a) is the original microscopic image of peripheral blood smear. Figure 1(b) is the S component image corresponding to Fig. 1(a), 1(c) is the histogram corresponding to Fig. 1(b). The gray values at the two troughs are assumed to be T_1 and T_2. Gray values less than T_1 belong to regions of background, gray values between T_1 and T_2 belong to regions of erythrocytes or leukocytes cytoplasm, and gray values above T_2 belong to the other regions. T_1 can be used to remove background. In order to separate the region of leukocyte alone from the regions of erythrocyte, other space vector must be considered.



(a) Original Image (b) S component image

(c) S component histogram (d) B component image

(e) B component histogram

Fig. 1. S component and B component images and corresponding histogram.

A great many experiments show that gray values of background and leukocyte regions are bigger than erythrocyte in the B component image, as shown in Fig. 1(d), where the B component image is considered to remove erythrocytes. Fig. 1(e) is the histogram corresponding to Fig. 1(d), it is easy to find that there are two obvious peaks in the histogram of B component image. We use the single threshold adaptive method to calculate the minimum gray value between the two peak values, where the value will be used to segment leukocyte images in the B component image.

2.2. *Threshold selection method of space vector*

2.2.1. *Threshold selection method of the S component image*

According to the analysis mentioned above, T_1 and T_2 can be found in the S component image. Otsu method[19] is applied to accurately find the value of T_1 which is used to remove background.

Assuming that the total pixels of the image are N, and range of the gray values is $\{0, 1, 2, \ldots, 255\}$, the total pixels with gray value i are n_i, and $N = \Sigma_{i=0}^{255} n_i$, so the probability of gray value i is:

$$p_i = \frac{n_i}{N} \quad \text{and} \quad \sum_{i=0}^{255} p_i = 1 \tag{1}$$

T_1 and T_2 are regarded as the critical points, range of T_1 and T_2 are $\{0, \ldots, 255\}$, value of T_2 is bigger than T_1. Gray values of S component image can be divided into three types: $C_1\{0, \ldots, T_1\}, C_2\{T_1 + 1, \ldots, T_2\}$ and $C_3\{T_2 + 1, \ldots, 255\}$, so probabilities of each type are as follows:

$$\begin{cases} \omega_1 = \sum_{k=0}^{T_1} p_k \\ \omega_2 = \sum_{k=T_1+1}^{T_2} p_k \\ \omega_3 = \sum_{k=T_2+1}^{255} p_k \end{cases} \tag{2}$$

Means of each type are as follows:

$$\begin{cases} \mu_1 = \sum_{j=0}^{T_1} \frac{jp_j}{\omega_1} \\ \mu_2 = \sum_{j=T_1+1}^{T_2} \frac{jp_j}{\omega_2} \\ \mu_3 = \sum_{j=T_2+1}^{255} \frac{jp_j}{\omega_3} \end{cases}. \tag{3}$$

The overall mean of the S component image is as follows:

$$\mu_0 = \sum_{i=0}^{255} ip_i. \tag{4}$$

Assuming that the interclass variance is σ^2, the function of T_1 and T_2 is as follows:

$$\sigma^2(T_1, T_2) = \omega_1(\mu_1 - \mu_0)^2 + \omega_2(\mu_2 - \mu_0)^2 + \omega_3(\mu_3 - \mu_0)^2. \tag{5}$$

When threshold values of T_1 and T_2 meet the following formula:

$$\sigma^2(T_1, T_2) = \max\{\sigma^2(T_1^*, T_2^*)\}, \quad 0 < T_1^* < T_2^* < 256. \tag{6}$$

The T_1^* and T_2^* are two arbitrary values which are compatible with the condition: $0 < T_1^* < T_1^* < 256$ and an optimal set of threshold T_1 and T_2 is selected by maximizing σ^2. T_1 can be used to separate erythrocytes and leukocytes from background.

2.2.2. *Threshold selection method of the B component image*

The optimal segmentation threshold of the B component image is obtained by the iteration method,[20] and the concrete implementation is as follows:

(1) $R_{\min}$ and $R_{\max}$, the minimum and maximum gray values of the B component image, are found to set the initial threshold T:

$$T = \frac{R_{\min} + R_{\max}}{2}. \tag{7}$$

(2) Gray values of the B component image are divided into two groups by T: C_0 $\{0, 1, \ldots, T\}$ and C_1: $\{T, \ldots, 255\}$, then μ_1 and μ_2, the average gray values of the two groups, are calculated respectively.

(3) Recalculating a new threshold:

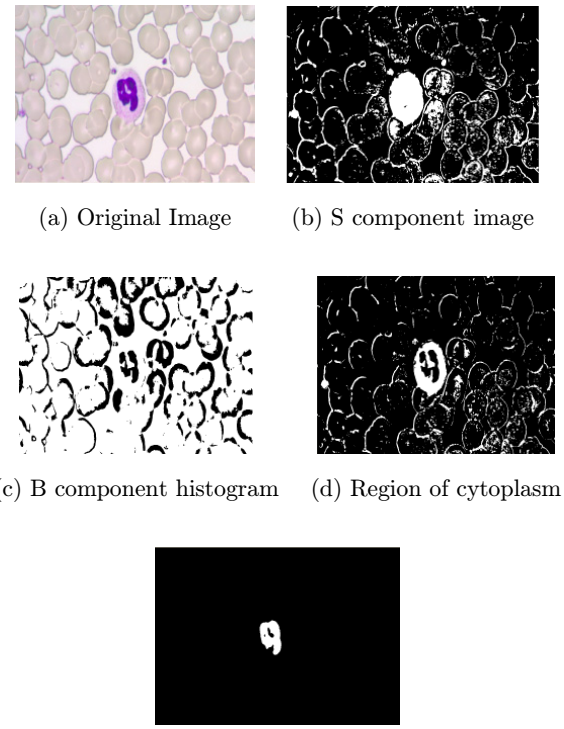
$$T' = \frac{(\mu_1 + \mu_2)}{2}. \tag{8}$$

(4) Condition of ending the circulation: $|T' - T| < \varepsilon$ (ε is preset, and $\varepsilon = e^{-6}$). If the result can not meet the condition, then we make $T = T'$, and steps 2 and 3 were repeated. If the condition has been met, T' is the best segmentation threshold.

(5) Gray values which are greater than or equal to T' are assigned to 1, less than T' are assigned to 0.

3. Segmentation of Leukocytes by Fusing the S Component and B Component Image

We segment the leukocyte image in S component and B component images through the method mentioned above. Figure 2(a) is the original microscopic image with 1024*768 sizes at 100 times. In this image, leukocyte is overlapped with erythrocytes, and there are some stains in it. Figure 2(b) is the binary image in the S component image after segmentation. Figure 2(c) is the binary image in the B component image after segmentation. Comparing Fig. 2(b) with 2(c), we obtain a regularity, as is shown in Table 1. Table 1 demonstrates that gray



(a) Original Image (b) S component image

(c) B component histogram (d) Region of cytoplasm

(e) Region of nuclei

Fig. 2. Binary image of cell nuclei B component and S component.

values are the same at every corresponding pixels in the leukocyte cytoplasm region, the gray values are 1, but opposite in the other regions for the two binary images. With this feature, "IMAGE AND" operation is conducted among Figs. 2(d) and 2(c) to obtain the region of leukocyte cytoplasm, as is shown in Fig. 2(d). Leukocyte nucleus region can be obtained by the method proposed in Ref. 18, as is shown in Fig. 2(e), the centroids belonging to the leukocyte nucleus region should be calculated. "IMAGE XOR" operation is conducted between Figures 2(d) and 2(e), result is shown in Fig. 3(a). It is easy to find the leukocyte cytoplasm region in Fig. 3(a), but some small punctiform and linear white regions can also be found, they are boundaries of erythrocyte or stains. We draw a square whose center is the centroid of nucleus region which the length of a side is R (empirical value, R is set 216 in experiments). This square contains the whole leukocyte region, and gray values of regions outside the square are assigned to 0 to remove some small punctiform and linear white regions, the result is shown in Fig. 3(b). However, some small punctiform and linear white regions cannot be excluded,

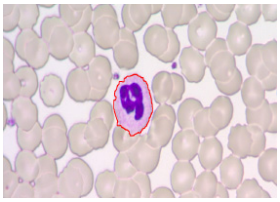
Table 1. Comparing the binary picture of B component of with S component.

	1	0
S component	Leukocyte nucleus, erythrocyte, cytoplasm of leukocyte	Background, center of erythrocyte, leukocyte nucleus, some leukocyte nucleus
B component	Background, cytoplasm of leukocyte, center of erythrocyte	Erythrocyte, leukocyte nucleus

(a) IMAGE XOR operation

(b) Removing area of out of square

(c) Segmentation result



(d) Drawing contour

Fig. 3. Process effect.

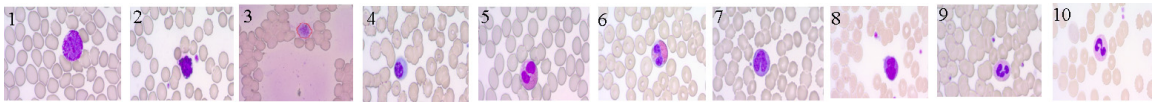
circularity and acreage are considered to exclude them. The punctiform white regions can be removed by acreage, and the linear white regions can be removed by circularity. Formula of circularities is shown in Eq. (9), where S is acreage of white regions and L is the perimeter. The smaller value the circularity is, the more away from shape the circular would be.

$$C = 4\pi S/L^2 \tag{9}$$

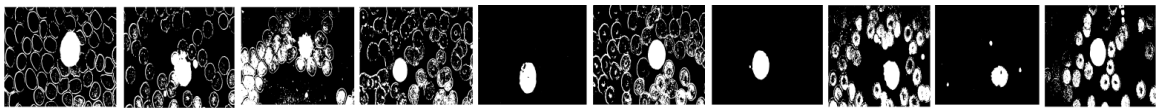
"IMAGE OPEN" operation is used in Fig. 3(b), After that the connected components are labeled by making use of eight-neighboring theory to work out the circularities and acreages of white regions. Gray values of regions whose circularities the less than a value (empirical value) are set to 0, acreages less than the value (empirical value) are also set to 0, the rest of small punctiform and linear white regions can be excluded completely. "IMAGE CLOSE" operation should be applied to smoothen the leukocyte contours. The result of segmentation is in Fig. 3(c), which shows that the method is able to perform well when leukocyte overlaps with erythrocytes, and excludes interference of the stains whose color is similar to leukocyte nuclei. Sobel operation will be used to extract the pixel position of leukocyte cytoplasm contours to describe a line corresponding to the leukocytes cytoplasm contours, as is shown in Fig. 3(d).

4. Experiments and Results Analysis

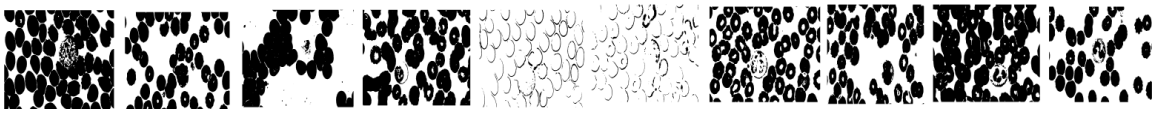
The method is implemented using VC6.0 based on OpenCV. Peripheral blood smears are prepared with Wright-Giemsa stain. Degrees of dye absorption are different for every blood smear and are dependent on some influencing factors, such as environment, features of leukocytes, and so on. Lots of microscopic images with 1024*768 sizes at 100 times are captured. Ten representative samples are chosen to analyze the experiment results. Figure 4(a) are 10 original microscopic images named Image1–10: Image1 and Image2 show basophils, erythrocyte distribution of the Image1 is concentrated, and leukocyte is overlapped with erythrocytes in image1. Image2's erythrocyte concentration is lower than the first and shape is irregular. The Image3 and the Image4 show lymphocyte. For Image3, background color is similar to leukocyte cytoplasm, and the leukocyte is broken a little bit. Background color and leukocyte cytoplasm color have a great difference in Image4. Image5 and Image6 show eosinophils. The leukocyte is overlapped with erythrocytes in the Image5, degree of dye absorption of the Image5 is high. For Image6, degree of dye absorption is low and background color is light. Image7 and Image8 show monocytes. Image7's degree of dye absorption is high. For Image8, the degree of dye absorption is lower, and contains some stains whose color is similar to leukocyte. Image9 and Image10 show neutrophils. Image9's degree of dye absorption is high, and the leukocyte is overlapped with erythrocytes. Image10's degree of dye absorption is low, background color and erythrocyte



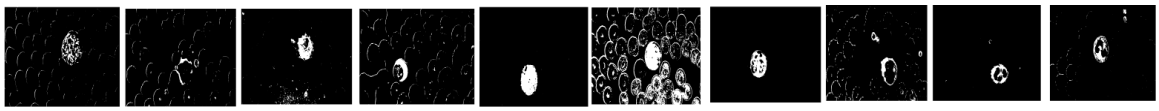
(a) Original images



(b) S component images

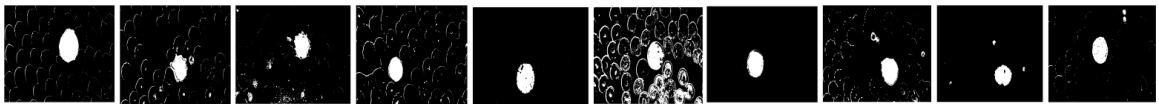


(c) B component images



(d) IMAGE AND operation

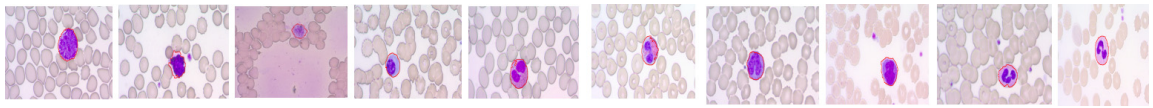
(e) Cell nucleus



(f) IMAGE XOR operation

(g) Removing region of out of square

(h) Segmentation result



(i) Describing Contour

Fig. 4. Experiment process.

Table 2. Segmentation results of five types of leukocyte.

Types of leukocyte		Neutrophil	Eosinophil	Basophil	Monocyte	Lymphocyte
Numbers of samples		141	45	14	104	87
Result of segment	Good	139	44	13	100	82
	Medium	1	1	1	2	3
	Poor	1	0	0	2	2
Detectable proportion		98.6%	97.8%	93%	96.5%	96.5%

color is light, erythrocyte distribution is less concentrated than the Image9.

The experimental processes are shown from Figs. 4(b) to 4(i): Fig. 4(b) shows S component binary images corresponding to Fig. 4(a). Figure 4(c) shows B component binary image corresponding to Fig. 4(a). Figure 4(d) are the results of "IMAGE AND" operation among Figs. 4(b) and 4(c). Figure 4(e) are nucleus regions. Figure 4(f) are the results of "IMAGE XOR" operation between Figs. 4(d) and 4(e). Figure 4(g) are the results after removing the regions out of squares. Figure 4(h) are leukocyte regions after dealing with by making use of circularity. Figure 4(i) are the last results of segmentation. The experiments show that the method is suitable for five types of leukocytes. Experiments in different cases obtain good results whether leukocyte cytoplasm color is similar to background or not, whether leukocytes connect with erythrocytes or not, whether the shape of the leukocytes is regular or not and whether the erythrocyte distribution is concentrated or not. The contours of leukocyte are retrieved with lines, as is shown in Fig. 4(i).

The segmentation results of the proposed method are evaluated using the similarity measure between the experiments result and the manual segmentation results. The formula of similarity is given in Eq. (6).

$$T_s = \frac{A_{\text{program}} + A_{\text{expert}}}{\max(A_{\text{program}}, A_{\text{expert}})}. \tag{10}$$

Here, A_{program} is the segmented area by the proposed method and A_{expert} is the area obtained by the expert hematologist. The experiment results are shown in Table 2. 385 of the 394 total samples perform well, and the average accuracy of segmentation reaches to 97.7%. Accuracies of segmentation are above 96% for segmenting five types of the leukocytes except basophils for they are only appeared in the blood with serious disease. Although the basophil samples are smaller, the accuracy can also reach to 93%. Experimental results have shown the universal applicability and stability of the proposed method.

5. Conclusions

A leukocyte segmentation method based on S component and B component images has been proposed. By making use of S component and B component images from the captured microscopic images at 100 times, an effective mathematical model is established. After some image processing procedures, leukocytes can be segmented with higher accuracy. The experiment results show that the method is suitable for segmenting five types of leukocytes, it also achieves good performance when leukocyte is overlapped with erythrocyte. The proposed leukocyte scheme reduces calculation and improves operating efficiency. Good segmentation results provide an important foundation for automatic recognition from the five types of leukocytes.

References

1. D. Huang, K. Hung, Y. Chan, "A computer assisted method for leukocyte nucleus segmentation and recognition in blood smear images," *J. Syst. Softw.* **85**, 2104–2118 (2012).
2. J. M. Chassery, C. Garbay, "An iterative segmentation method based on contextual color and shape criterion," *IEEE Trans. Pattern Anal. Mach. Intell.* **6**(6) 795–900 (1994).
3. O. Lezoray, A. Elmoataz, H. Cardot, G. Gougeon, M. Lecluse, H. Elie, M. Revenu, "Segmentation of cytological images using color and mathematical morphology," *Eur. Congress Stereology*, Vol. 18, pp. 1–14 (1999).
4. K. Jiang, Q.-M. Liao, S.-Y. Dai, "A novel white blood cell segmentation scheme using scale-space filtering and watershed clusting," *International Conf. Machine Learning and Cybernetics*, Vol. 10, pp. 12–19 (2003).

5. J. Ao, S. Mitra, R. Long, B. Nutter, S. Antani, "A hybrid watershed method for cell image segmentation," *2012 IEEE Southwest Symp. Image Analysis and Interpretation*, Vol. 4, pp. 29–32 (2012).
6. Y. Fang, C. X. Zheng, C. Pan *et al.*, "White blood cell image segmentation using on-line trained neural network," *IEEE Engineering in Medicine and Biology 27th Annual Conf. Shanghai, China*, pp. 6476–6479 (2005).
7. S. T. Wang, M. Wang, "A new detection algorithm based on fuzzy cellular neural networks for white blood cell detection," *IEEE Trans. Inf. Technol. Biomed.* **10**(1), 5–10 (2006).
8. A. Khashman, "Blood cell identification using a simple neural network," *Int. J. Neural Syst.* **18**(5), 453–458 (2008).
9. P. Sobrevilla, E. Montseny, J. Keller, "White blood cell detection in bone marrow images," *Proc. 18th Int. Conf. North Amer. Fuzzy Inf. Process. Soc.*, New York, NY, pp. 403–407 (1999).
10. N. Theera-Umpon, "Patch-Based white blood cell nucleus segmentation using fuzzy clustering," *ECTI Trans. Electrical Eng, Electronics, and Communications*, Vol. 3, pp. 15–19 (2005).
11. F. Sadeghian, Z. Seman, "A framework for white blood cell segmentation in microscopic blood images using digital image processing," *Biological Procedures Online* **11**(1), 196–206 (2009).
12. B. C. Ko, J.-W. Gim, J.-Y. Nam, "Automatic white blood cell segmentation using stepwise merging rules and gradient vector flow snake," *Micron.* **42**, 695–705 (2011).
13. K. Li, Z. Lu, W. Liu, J. Yin, "Cytoplasm and nucleus segmentation in cervical smear images using radiating GVF snake," *Pattern Recognit.* **45**, 1255–1264 (2012).
14. C. Pan, D. Sun Park, S. Yoon, J. C. Yang, "Leukocyte image segmentation using simulated visual attention," *Expert Syst. Appl.* **39**, 7479–7494 (2012).
15. M. Saraswat, K. V. Arya, H. Sharma, "Leukocyte segmentation in tissue images using differential evolution algorithm," *Swarm Evol. Comput.* **11**, 46–54 (2013).
16. Cseke, "A fast segmentation scheme for white blood cell images," *Pattern Recognition, 11th IAPR Int. Conf., C: Image, Speech and Signal Analysis, Proc.* Vol. 3, 530–533 (1992).
17. B. R. Kumar, D. K. Joseph, T. V. Sreenivas, "Teager energy based blood cell segmentation," *Int. Conf. Digit. Sig. Process.* Vol. 2, pp. 619–622 (2002).
18. J. Wu, P. Zeng, Y. Zhou, C. Olivier, "A novel color image segmentation method and its application to white blood cell image analysis," *Proc. 8th IEEE ICSP* **2**, 245–248 (2006).
19. N. Otsu, "A threshold selection method from gray level histograms," *IEEE Trans. Syst. Man Cybem.* **9**(1), 62–66 (1979).
20. A. Perez, R. C. Gonzalez, "An iterative thresholding algorithm for image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.* **9**(6), 742–751 (1987).