Precise segmentation of White Blood Cells by using multispectral imaging analysis techniques

Qiongshui Wu*, Libo Zeng, Hong Zheng, Ningning Guo
Electronic Information School, Wuhan University, Wuhan, 430079, P.R.China
**E-mail: qswu@whu.edu.cn Tel: +86-27-68775472*

Abstract

Counting of different classes of white blood cells in bone marrow smears can give pathologists valuable information regarding various hematological disorders. For automation imaging analysis techniques, precise segmentation of White Blood Cells is quite challenging due to the complex contents in bone marrow smears. Far more different from traditional color imaging analysis methods, we introduced multispectral imaging techniques. After a high quality image was acquired, the spectrum of each pixel was directly fed into a trained Support Vector Machine (SVM) for classification, and then morphological binary operations were performed to correct the small error-classified regions. Mass of experiments showed that the segmentation results are highly satisfactory and inspiring. It shows that the introduction of multispectral imaging analysis techniques into White Blood Cells detection is a success. Multispectral imaging analysis is a promising technique in biomedicine.

Keywords: *White blood cells, bone marrow image, multispectral imaging analysis, segmentation, SVM*

1. Introduction

Counting of different classes of white blood cells in bone marrow smears can give pathologists valuable information regarding various hematological disorders. Currently, the automatic bone marrow microscopic images analysis presents significant difficulties due to: (1) The contents of this kind of images are so complex, there are more than 20 different classes cells present[1]. (2) Image quality is significantly affected by the staining and illumination inconsistences. Staining and illumination inconsistences can lead to gray value variety, which is a big trouble for image segmentation. (3) The big variability of the white blood cells characteristics. The cells are frequently clustered, and there is no clear boundary between the nucleus and cytoplasm in many cases. These issues make precise segmentation a very challenging problem.

In the past decades, a lot of meaningful work based on color images was done in bone marrow images segmentation[2-7]. But for standard color image analysis, there are many problems[8]: Firstly, in image acquisition process, the images are illuminant-dependent and the image quality is influenced by spectral characteristics of the imaging system, such as spectral responses of lens, optic ununiformity and throughput properties, which makes image reproduction very difficult. For example, the same sample may produce quite different image quality even if it is imaged under another same type of instrument, so, the robust of segmentation algorithms is very exigent. And the second, the standard digital RGB (Red, Green, Blue) values are the integrated results over the whole spectrum range, thus a lot of spectral information is unavoidably lost, which means that a mass of chemical and physical information is buried under these spectral characteristics. While in many applications, such as cell detection, this amount of spectral information is very valuable for differentiating slightly color difference.

In order to remedy these problems and get better segmentation results, we introduced multispectral imaging[9] analysis techniques for automatic detection of White Blood Cells in bone marrow microscopic images. In this paper, a pixel-level classification approach based on spectral information instead of traditional grey intensity or spatial information was presented for precise segmentation.

2. Apparatus

A schematic diagram of the multispectral imaging microscope apparatus is shown in Figure 1. This apparatus consists of a BX41 microscope (Olympus Corporation), a *Liquid Crystal Tunable Filter (LCTF)* device and its controller (VariSpecTM, Cambridge

978-0-7695-3391-9/08 $25.00 © 2008 IEEE
DOI 10.1109/ICINIS.2008.105

Research &Instrumentation, Inc.), a three-dimensional automation stage, a Pentium IV computer (PC), a cooled monochrome *Charge Coupled Device (CCD)* camera (Penguin 600L, Pixera Corporation) and etc.

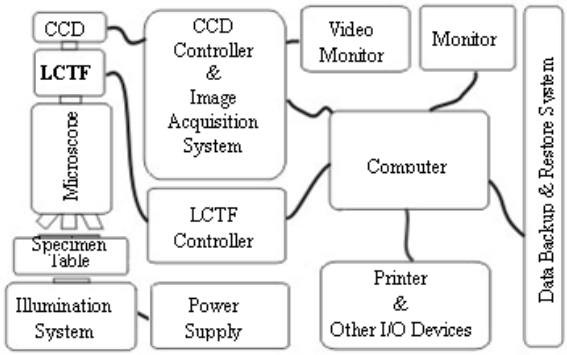

Figure 1: Schematic diagram of a multispectral imaging microscope apparatus based on LCTF

Bone marrow smears were prepared by Wright method, which were provided by the Department of Hematological, Zhongnan Hospital of Wuhan University.

The detailed description of this apparatus and image acquisition method can refer to our previous work [10].

3. Method

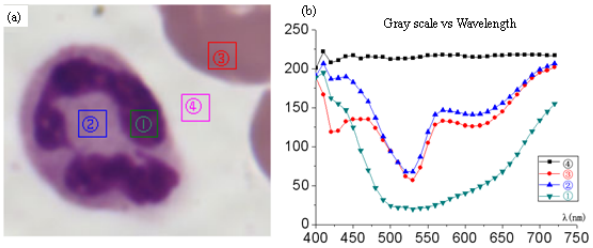

Figure 2: (a) A typical marrow bone image which consists of a polymorphonuclear leucocyte. •nucleus • cytoplasm •erythrocytes •background. (b)Transmitted spectrum of different parts marked in (a). Each rectangle region in (a) is regarded as a point, whose gray scale is the average in the rectangle.

Figure 2 (a) is a typical bone marrow image. From Figure 2 (b), we can find that, for bone marrow images, different regions have quite different spectra, though their colors may be very similar. So, we use spectral information other than traditional RGB color to differentiate different regions.

For each pixel P(i,j) in a multispectral image, the spectrum (As shown in Figure 2 (b)) can be expressed as a vector x:

$$x = (w_{\lambda_1}, \cdots, w_{\lambda_n})^T \quad (1)$$

w_{λ_k} is grey intensity at wavelength λ_k, n is the count of wavelength band, here, it is 33.

Our segmentation method is a pixel level classification based on *Support Vector Machines (SVM)* [11-12]. The spectral vector of each pixel was directly fed into a *SVM* classifier, and all the pixels were classified into four classes: nucleus, cytoplasm, erythrocytes and background.

The segmentation steps include two processes: training and classification.

The training process can be outlined below:

(1) Thirty-five multispectral images(100x Objective) were manually segmented under the direction of a pathologist.

(2) Many interested regions of nucleus, cytoplasm, erythrocytes and background (we label the different region classes with 1,2,3,4 individually) were manually marked with different color mask according to their region classes. And then the spectrum of each pixel in the marked regions was extracted as feature vector, and its region class is used as class type. All the spectral features and class types were saved as standard training dataset.

(3) When all the thirty images were completed, all the saved training dataset were fed into a *SVM* classifier for training. There are totally 3,530 nuclear pixels, 6,399 cytoplasmic pixels, 1,696 erythrocytic pixels and 1,275 background pixels were saved as training sample sets. After training, the model results were obtained.

For tested image, the spectral vector of a pixel was fed into the trained classifier for classification, and the output of the classifier is the class type of the pixel.

For the SVM classifier, we used *C-support Vector Classification(C-SVC)[13,14]*. Since the classification is a multi-class classification, we used the one-against-one approach[15, 16] in which k(k − 1)/2 classifiers are constructed and each one trains data from two different classes. In classification we use a voting strategy: each binary classification is considered to be a voting where votes can be cast for all data points, and finally, each point is designated to be in a class with maximum number of votes. A polynomial was selected as the kernel function, and the degree of the polynomial is set to the value of 3. The program code and more detailed discussion about *C-SVC* algorithm can be found at Dr. Chih-Jen Lin's Home Page[17].

Since *C-SVC* is not very fast, if every pixel in a multispectral image is classified with *C-SVC*, it will be very time-consuming. In order to enhance the classification speed, background regions were firstly

removed by applying a constant threshold value of 200 to the grey image at wavelength 530nm. From Figure 2 (b), we can easily find that the difference between background and other regions is the most significant, so a constant threshold is enough for the removal of background.

After segmented by the *C-SVC*, a series binary operations were applied to correct the small error-classified regions and make the edge of a cell more smooth: Firstly, hole filling was performed to fill the small holes in the cell areas; And then, a watershed algorithm was performed to split the connected areas; And then, small areas were deleted; Finally, binary open operation was performed to make the edge more smooth.

4. Results and discussion

One hundred and eighteen multispectral images other than the training sets acquired with 100x objective were used to evaluate the segmentation algorithms. There are 245 cells in these 118 images. Part of segmentation results are shown in Figure 3.

The segmentation results were evaluated with a pathological expert according to his satisfaction. For all the 118 images, the segmentation of the background and erythrocytes were extremely well, there were very few errors. For the 245 cells, they can be segmented out 100%. For nuclear segmentation, 235 cells were satisfactory. For cytoplasmic segmentation, 227 cells were satisfactory.

In all the cells, the correct ratio of acidophilic granulocytes and basophiles is not very high. One reason is that the color of the nucleus has a big variance, and another reason is that they occupy no more than 1 percent of all the cells, which makes enough training samples not available. For the cytoplasmic segmentation, the correct ratio of metarubricytes is not very high, this is because their cytoplasmic color is quite similar to erythroblast, which were wrong segmented as erythroblast.

5. Conclusions and future work

A mass of experiments showed that our segmentation results are highly satisfactory. From Figure 3, we can conclude that the segmentation results are pretty good. Using spectral information for the detection of *White Blood Cells* is a successful exploration and SVM-based classifier is also a success in spectrum classification.

Compared to traditional imaging analysis method, our method has following advantages: (1) Introduction of spectral information into segmentation makes the segmentation algorithm more popular and extensible. For traditional method, an effective algorithm for object detection may cause wrong results if the analyzed target changes. For example, if there is an algorithm very effective in red object segmentation, when used in blue object detection, the algorithm and parameters have to be modified or it will not work correctly. While for our method, algorithm modification is not needed, just changing the training samples is OK. (2) Multispectral image contains more information than traditional color image, which make the analysis method more flexible. For the later feature extraction or classification, it can provide more meaningful features, which is helpful for the improvement of classification accuracy.

Though preliminary results show that our multispectral imaging analysis method in bone marrow smears application is promising and inspiring, yet there is still a lot of room for improvements: (1) The wavelengths we used are from range 400 nm to 720 nm with an increment of 10 nm, if a small increment is used, better results might be obtained. But it is really a time-consuming work to do this, and it will occupy a tremendous disk space of our computer. (2) Boosting image acquisition and segmentation speed. Currently, its speed is not high enough for practical applications. But it is a new research in bone marrow image segmentation, and it is promising in biomedical image processing. (3) Concentrating on lower magnification images (40x objective) analysis. One advantage is that it can increase the scan speed than with 100x objective. Another advantage is that no immersion oil is needed for 40x objective. A preliminary experiment shows that it is feasible and practical.

6. Acknowledgments

We expresses appreciation to Ms. Kejian Zhang of Hematological Department of Zhongnan Hospital, Wuhan University for providing us with bone marrow smears and giving us directions on medical issues. Our work is funded by the Chinese National Ministry of Science and Technology under the grant 2001BA210A02.

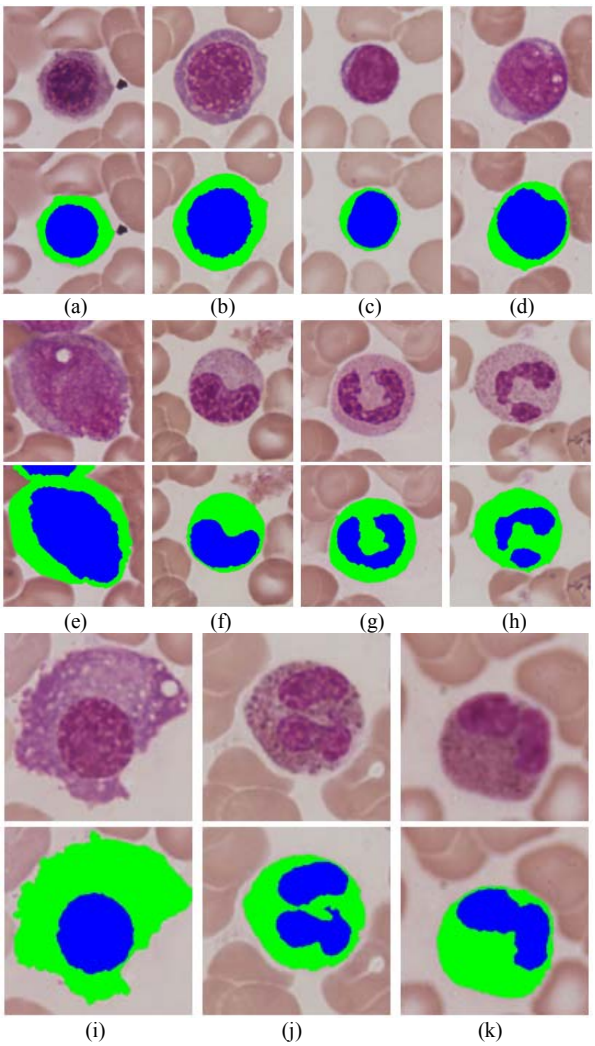

Figure 3: Segmentation results. The upper row is the original image and the next row is the segmented result. Blue color represents nuclear area and green color represents cytoplasmic area. (a) metarubricyte (b) polychromatic normoblast (c) lymphocyte (d) monocyte(e) neutrophilic myelocyte (f) neutrophilic metamyelocyte (g) band-cell (h) polymorphonuclear leucocyte(i) plasmocyte (j) acidophil leukocyte (k) basicyte

References

[1] Bennett, J.M., Catovsky, D., and Daniel, M.T., *Proposals for the classification of the acute leukaemias.* British Journal of Haematology, 1976. **33**(4): p. 451-458.

[2] Theera-Umpon, N., *White blood cell segmentation and classification in microscopic bone marrow images*, in *Fuzzy Systems And Knowledge Discovery, Pt 2, Proceedings*. 2005, Springer-Verlag Berlin: Berlin. p. 787-796.

[3] Nilsson, B. and Heyden, A., *Segmentation of complex cell clusters in microscopic images: Application to bone marrow samples.* Cytometry Part A, 2005. **66A**(1): p. 24-31.

[4] Zhang, X.W., Song, J.Q., et al., *Extraction of karyocytes and their components from microscopic bone marrow images based on regional color features.* Pattern Recognition, 2004. **37**(2): p. 351-361.

[5] Montseny, E., Sobrevilla, P., and Romani, S., *A fuzzy approach to white blood cells segmentation in color bone marrow images*, in *2004 Ieee International Conference on Fuzzy Systems, Vols 1-3, Proceedings*. 2004, Ieee: New York. p. 173-178.

[6] Meschino, G.J. and Moler, E., *Semiautomated image segmentation of bone marrow biopsies by texture features and mathematical morphology.* Analytical And Quantitative Cytology And Histology, 2004. **26**(1): p. 31-38.

[7] Hengen, H., Spoor, S., and Pandit, M. *Analysis of blood and bone marrow smears using digital image processing techniques*. 2002. San Diego, CA, United States: The International Society for Optical Engineering.

[8] Hardeberg, J.Y., *Acquisition and Reproduction of Colour Images: Colorimetric and Multispectral approaches*. 1999, Paris: Ecole Nationale Supérieure des Télécommunications.

[9] Levenson, R.M. and Hoyt, C.C., *Spectral imaging and microscopy.* American Laboratory, 2000. **32**(22): p. 26-33.

[10] Wu Q.S.,Zeng, LB,et al., *Analysis of blood and bone marrow smears using multispectral imaging analysis Techniques*, in Medical Imaging 2005: Image Processing, Pt 1-3, edited by J. M. Fitzpatrick, J. M. Reinhardt, Proceedings of SPIE Vol. 5747 (SPIE, Bellingham, WA, 2005): 1872-1882.

[11] Gunn, S.R., *Support Vector Machines for Classification and Regression*. 1998, Faculty of Engineering and Applied Science, Department of Electronics and Computer Science, University of Southampton.

[12] Cortes, C. and Vapnik, V., *Support-vector network.* Machine Learning, 1995. **20**: p. 273-297.

[13] Cristianini, N. and Shawe-Taylor, J., *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*. 2000, Cambridge: Cambridge University Press.

[14] Vapnik, V., *Statistical Learning Theory*. 1998, New York: Wiley.

[15] Friedman, J., *Another approach to polychotomous classification*. 1996, Department of Statistics, Stanford University.

[16] Hsu, C.W. and Lin, C.J., *A simple decomposition method for support vector machines.* Machine Learning, 2002. **46**(1-3): p. 291-314.

[17] Lin, C.J., *A Library for Support Vector Machines*. 2003, http://www.csie.ntu.edu.tw/~cjlin/libsvm/index.html.