Automatic Leukocyte Classification Based on Microscope Images

Piyamas Suapang[1, a], Methinee Thongyoun[2,b] and Sorawat Chivapreecha[3,c]

[1] Biomedical Engineering Program, Department of Physics, Rangsit University,

Pathumthani, 12000, Thailand

[2, 3] Department of Telecommunications Engineering, King Mongkut's Institute of Technology Ladkrabang, Bangkok, 10520, Thailand

[a] piyamas_suapang@yahoo.com, [b] methinee_thy@yahoo.com, [c] sorawat@telecom.kmitl.ac.th

Keywords: Leukocyte Segmentation; Leukocyte Classification; White Blood Cell Count

Abstract. Numbers of white blood cells in different classes help doctors to diagnose patients. A technique for automating the differential count of white blood cell is presented. The proposed system takes an input, color image of stained peripheral blood smears. The process involves segmentation, feature extraction and classification. The segmentation procedure, a novel simple algorithm, is proposed for localization of white blood cells and the different cell components are separated with automatic thresholding. Features extracted from the segmented nucleus are motivated by the visual cues of shape, color and texture. This research uses the Artificial Neural Network for implemention and uses the different combinations of feature sets. The results presented here are based on trials conducted with normal cells. For training the classifiers, a library set of 233 patterns is used. The tested data consists of 134 samples and produced correct classification rate close to 88.10 %.

Introduction

A typical blood microscope image has been digitalized by a CCD and acquired by a frame-grabber system [1]. The microscope inspection of blood slides provides important qualitative and quantitative information concerning the presence of hematic pathologies. Principal cells present in the blood are red blood cells, and the white cells (leucocytes). Leucocyte cells containing granules are called granulocytes (composed by neutrophil, basophil, eosinophil). Cells without granules are called agranulocytes (lymphocyte and monocyte). The percentage of leucocytes in human blood typically ranges between the following values: neutrophils 50-70%. eosinophils 1-5%. basophils 0-1%. monocytes 2-10%, lymphocytes 20.45% [2]. These cells provide the major defense against infections in the organism and their specific concentrations can help specialists to discriminate the presence or not of very important families of pathologies (i.e. the presence of mononucleosis, hepatitis diabetes, allergy, arthritis, anaemia, and many others).

From decades this operation is performed by experienced operators, which basically perform two main analyses. The first is the qualitative study of the morphology of the cells and it gives information of degenerative and tumoral pathologies such as leukemia. The second approach is quantitative and it consists of differential counting the white blood's cells. Unfortunately, the accuracy of cell classification and counting is strongly affected by individual operator's capabilities. In particular, the identification and the differential count of blood's cells is a time-consuming and repetitive task that can be influenced by the operator's accuracy and tiredness. The automated classification of the peripheral white blood cells (leukocytes) has been the subject of this study. The peripheral blood leukocytes provide a very interesting and challenging medium for the study for biological image processing and classification. The taxonomy of the various blood cells has had a classic history in histology [3], and a qualitative, verbally descriptive taxonomy has been established in the hematological literature [4]. One of the rapidly emerging areas of pattern recognition has been medical picture processing and image classification [5], [6]. Radioisotope scanning [7], [8], breast

cancer diagnosis [9], the classification of Papanicolaou smears [10], chromosome analysis [11], [12], and the classification of white blood cells [13-15], have all proven to some extent to be amenable to pattern recognition techniques.

This paper focuses on the problem of automatic leukocytes classification using microscope images i.e. important extracting leukocytes from others blood's components, morphological and shape features for identification and features are used as input for neural network. The proposed system firstly indicates individual leukocytes from others blood cells, secondly it extracts morphological and shape features and finally it classifies the leukocytes by neural networks. Therefore, the automatic leukocytes classification using microscope image is proposed to improve speed and accuracy of the performance.

Methodology

Segmentation. First, the captured image file was split into its three component bands (red green and blue see Fig. 1). The result was three grayscale files one for each of the red green and blue components of the image captured by the camera. Histogram analysis was used to examine three grayscale components (corresponding to the red, green and blue bands) of 30 images covering all five basic white blood cell types. It was found that the green component was consistently a better discriminator between the purple nuclear material and the rest of the image.

We did utilize thresholding to produce a binary bitmap image from the green band bitmap for each image. An automatic threshold value was selected by minimum value to discriminate between nuclear and nonnuclear pixels, as every white blood cell has a nucleus. We used only the green band for this nuclear thresholding, as, in this band, the nuclear material is much darker than either the cytoplasm (with the exception of basophilic cytoplasm) or the background. Experimentation showed that a automatic threshold value of 95 (on a scale of 0-256) gave an acceptable discrimination between nuclear and non-nuclear pixels, leaving nuclear pixels black on a white background (Fig. 2 (a)).

Due to edge effects on the captured images, there were many nonnuclear pixels around the perimeter of the image. An edge erosion filter, which simply set all of the pixels within 3 pixels of the edge of the image to white, was developed and used (Fig. 2 (b)). This removed the dark bands around the edge of the image.

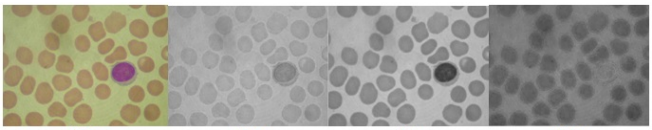

Fig. 1 *Raw captured image and 3 single color components.*

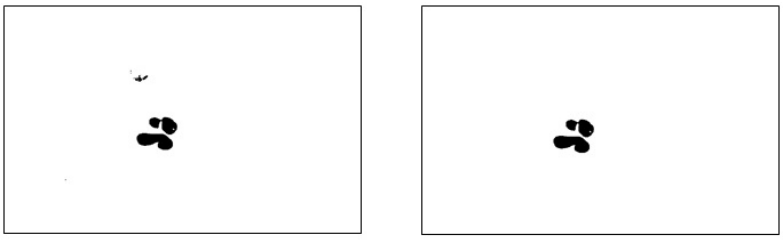

***Fig.* 2** *The results of (a) thresholding and (b) erosion.*

Feature Extraction. Most white blood cells are roughly circular in shape, though some (monocytes, in particular) may deviate significantly from the circular). Each region of interest was thus centered on the center of the nuclear matter within a cell. Thus, it was known that the center of each region of interest was within the cell. However, for any other pixel, it was not known whether

that pixel fell within the cell or outside it. The number of pixels within the circle was also determined. It was immediately noticed that some non-nuclear material (particularly platelets which are also stained purple by the May-Grunwald Giemsa staining protocol) also came out as black on these images. Platelets were found to be very small in comparison to blobs of nuclear matter. The color bands were the red, green and blue captured by the hardware (camera + composite video capture card) together with the color ratios green/red and green/blue. What's more, texture features are computed by the color variance (σ^2) of each pixel on the green band for the nuclear matter.

Neural Network Architecture for Pattern Classification. Van derHeijden states that neural networks are processing structures "consisting of many interconnecting processing elements (neurons)." These artificial neurons are connected together to form neural networks. In this example, a number of inputs are each connected to each of a number of neurons in an intermediate layer. The neurons in the hidden layer are each connected to all the output neurons (one in this case). This is an example of a fully connected feed forward neural network. Feed forward networks of this type can be trained by back propagation. This is a procedure that trains the network by making small adjustments to the weights of each neuron in the direction that reduces the error at that neuron's output. The input layer used white blood cell feature 15 features. The hidden layer was designed by 5 nodes. Finally, the results of output layer were equal the number of white blood cells in different classes and determined from the probability of class membership.

Results and Discussion

Fig. 3 shows that the discrimination between nuclear and nonnuclear pixels is selected after a automatic threshold value. These leukocytes features is determined form the basis of the feature set. These parameters can be used as efficient features for inputs of classifiers. The test applied these leukocytes features are carried out and the results are shown in the Table I. The results presented here are based on trials conducted with normal cells. For training the classifiers, a library set of 233 patterns is used. The tested data consists of 134 samples and produced correct classification rate close to 88.10 %.

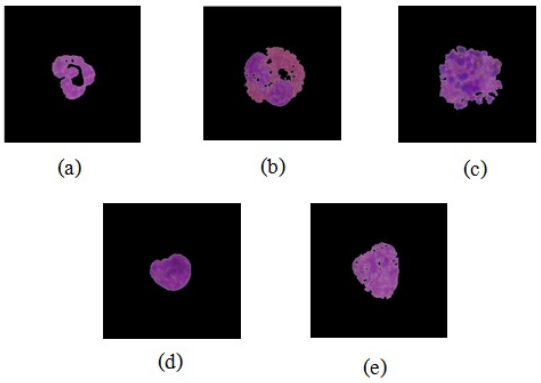

Fig. 3 *The nucleus segmentation.*

Table 1 The results of testing by leukocytes features set

Leukocytes	Samples	Correct		Incorrect	
		Numbers	%	Numbers	%
Neutrophils	39	35	89.74	4	10.26
Eosinophils	30	23	76.67	7	23.33
Basophils	5	4	80.00	1	20.00
Lymphocytes	30	30	100.00	0	0.00
Monocytes	30	26	90.00	4	10.00
Total	134	118	88.10	16	11.90

Conclusion

This paper presented a methodology to achieve a fully automated detection and classification of leucocytes by microscope color images identifying the following classes: Basophil, Eosinophil, Lymphocyte, Monocyte and Neutrophil. Experiments show that the final classification module implemented by means of a parallel classifier composed by back propagation neural classifiers achieves an accurate solution with minor computational complexity than traditional nearest neighbor classifier. Results indicate that the morphological analysis of blood‘s white cells is achievable and it offers remarkable classification accuracy. Further studies will be focused on identification of tumor deformations in the cell morphology for a fully-automated diagnostic system.

References

[1] F. Cillesen, W. Der Meer, “Atlas of Blood Cell Differentiation”, Elsevier Science B.V., Amsterdam The Netherlands. 1998.

[2] “Blood”. Keith Breden Taylor and Julian B. Schorr, Colliers Encyclopaedia. Vol 4, 1978.

[3] A. Hughes, A History of Cytology. New York: Aberlard-Schuman, 1959.

[4] G. A. Daland and T. H. Ham, A Color Atlas of Morphologic Hematology. Cambridge, Mass.: Harvard University Press, 1967.

[5] R. S. Ledley, "High-speed automatic analysis of biomedical pictures," Science, pp. 216-223, Oct. 1964.

[6] P. G. Stein, L. E. Lipkin, and H. M. Shapiro, "Spectra II: general-purpose microscope input for a computer," Science, vol. 166, pp. 328- 333, 1969.

[7] J. W. Butler, "Automatic analysis of bone autoradiographs," in Pictorial Pattern Recognition, G. C. Cheng et. al., Eds. Washington, D.C.: Thompson Book Co., pp. 75-85, 1968.

[8] M. R. Evans and J. W. Sweeney, "A computer technique for investigating and rationalization of scintillation scan reading," Methods Inform. Med., vol. 6, no. 1, pp. 24-27, 1967.

[9] F. Winsberg et al., "Detection of radiographic abnormalities in mammograms by means of optical scanning and computer analysis," Radiology, vol. 89, pp. 211-215, 1967.

[10] G. L. Wied et al., "Taxonomic intracellular analytic system (TICAS) for cell identification”, Acta. Cytol., vol. 12, pp. 180-204, 1968.

[11] J. W. Butler, M. K. Butler, and A. Stroud, "Automatic classifications of chromosomes-II," in 1964 Proc. Rochester Conf. Data Acquisition and Processing in Biology and Medicine. New York: Pergamon, 1965.

[12] J. Hilditch and D. Rutovitz, "Chromosome recognition," Ann. N.Y. Acad. Sci., vol. 157, pp. 339-364, 1969.

[13] Pramit Ghosh, Debotosh Bhattacharjee, Mita Nasipuri and Dipak Kumar Basu, “Automatic White Blood Cell Measuring Aid for Medical Diagnosis”. Process Automation, Control and Computing (PACC), pp. 1-6, 20-22 July, 2011.

[14] Sawsan F. Bikher, Ahmed M. Darwish, Hany A. Tolba, Samir I. Shaheen. “Segmentation and Classification of White Blood Cells”. IEEE International Conference on Acoustics, Speech, and Signal Processing, vol 6, pp. 2259 - 2261, 5-9 June, 2000.

[15] Neelam Sinha, A.G.Ramakrishnan. “Automation of Differential Blood Count”. TENCON 2003: Conference on Convergent Technologies for the Asia-Pacific Region, vol 2, pp. 547 - 551, 15-17 Oct, 2003.