International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 2 Issue 6, June - 2013

Feature Extraction And Segmentation Of White Blood Cells

Vaishali Phapale[1], Prof. V. D. Hajare[2] , And Prof. H. T. Patil[3]
Department of Instrumentation Engineering,
Cummins College Of Engineering For Women
Karvenagar, Pune

Abstract— Color image segmentation has becoming more popular for computer vision due to its important process in most medical analysis tasks. The paper presents a set of preprocessing and segmentation algorithms along with a set of features that are able to recognize and classify different categories of normal white blood cells. One of the main tasks is the segmentation of white blood cell (WBC) where the WBC composition reveals important diagnostic information of a patient. In this paper the unsupervised segmentation technique namely k-means clustering algorithm is used to the segmentation process. By implementing the proposed segmentation technique, the segmented WBC which consists of nucleus regions can be achieved by using the CIE-Lab technique and k-means clustering.

***Index Terms*—CIE-Lab, Feature Extraction, k-means clustering , Segmentation, White Blood Cells (WBC)**

I. Introduction

The main objective of this work is to recognize and classify different categories of normal white blood cells. In blood analysis , doctors look first for three different kinds of cells: red blood cells, white blood cells and blood platelets. These three types of cells are distinguished by their size and color. In this system used gray level images because white blood cells appear darker than the red cells and blood platelets. In terms of size the blood platelets are the smallest and the white cells the largest. White blood cells categorized into five normal subclasses depending on the morphology of their outer contour and their nuclei. These classes are Eisonophil, lymphocyte, Monocyte, Basophil and Neutrophil. Leukocytes play a significant role in the diagnosis of different diseases such as leukemia and different types of infections. The hematologist requires two types of blood count for diagnosis and screening. The first one is called the Complete Blood Count (CBC) and the second one is called the Differential Blood Count (DBC). CBC could be done by instruments called cytometer and could successfully be performed automatically. On the other hand, DBC is more reliable but currently it is a manual procedure to be done by hematology experts using microscope. In DBC, an expert counts 100 white blood cells on the smear at hand and computes the percentage of occurrence of each type of cell counted. The results reveal important information about patient's health status. The main part of each white blood cell (WBC) is its nucleus which contains chromosomes. There are few complications in extracting information from WBCs due to wide variation of there cells in shape, size, edge, and position.

The three types of WBCs granulocytes are neutrophils, eosinophils, basophils and granulocytes are monocytes, lymphocytes. The most common WBC, the Neutrophils 54 to 62% of the leucocytes. Neutrophil nuclei may have two to five lobes. Comprising only about 3% of the leucocytes , the eisinophils have bilobed nuclei and uniform sized cytoplasmic granules. Basophils comprising less than 1% of the circulating WBC , the basophils are smaller than other granulocytes. They are easy to identify because of the numerous granules in the cytoplasm. Monocytes comprising 2 to 8% of circulating WBC monocytes are twice the size of red blood cells. They have oval or kidney bean shaped nuclei. Lymphocytes are easy to recognize due to the large nucleus

II LITERATURE SURVEY

Numerous segmentation methods for cell images from peripheral blood smears or bone marrow smears have been proposed. Ongun [1] Segmented the object in a microscopic image based on morphological preprocessing followed by iteration algorithms like snake edge detection. Xiong [2] generic features to quantify the goodness of good working areas without requiring the extraction of individual cells. Jiang[7] proposed WBCs segmentation by different methods like snake and watershed in color space. Dorini [8]enhanced the segmentation accuracy by using scale- space operators. The automatic morphological method of Scotti [9] was based on the morphological analysis of WBCs and of normal cells. Their proposed system extracts the morphological indexes. Kumar [10] used Teager energy operator for nucleus segmentation with edges, which are detected effectively by Teager energy operator, but it was restricted because of low contrast between the gray level of cells and background.

III METHODOLOGY

A. Image Acquisition

The first step is to acquire the image of blood samples, were obtained from D Y Patil Hospital, Pathology Department, Pimpri, Pune. Images captured at 100X magnification and saved in jpg format. Database contains five types of normal White Blood Cells and some of them abnormal peripheral blood cells.

B. RGB to CIE Lab

The CIE (Commission International de I'Eclairage- the International Commission on Illumination) designated in1931 the following specific wavelength values to the three primary colors: blue=435.8 nm, green = 546.1 nm and red = 700 nm[5].

CIE Lab color space system is based on a sequential or continuous Cartesian representation of three orthogonal axes L^* a^* b^* .Coordinate L^* represents clarity (L^* =0 black and L^* = 100 colorless) a^* green/red colour component (a^* > 0 red, a^* < 0 green) an b^* blue/yellow colour component (b^*>0 yellow, b^*<0 blue)

L^* a^* b^* defined [6].

$$L^* = 116(Y/Y_n)^{1/3} - 16$$
$$a^* = 500\,[(X/X_n) - (Y/Y_n)] \quad (1)$$
$$b^* = 200 - (Y/Y_n)^{1/3} - (Z/Z_n)^{1/3}$$

where X_n, Y_n, are the CIE XY stimulus values of the reference white point(the subscript n suggests 'normalized').

C. Nucleus Segmentation

The aim of nucleus segmentation is separating the nucleus from the other parts of a cell and a microscopic blood smear image. A typical peripheral blood smear image consist of four components : red cells (unnucleated cells), white blood cells nucleus , cytoplasm, and background which contains platelets and spot noise. WBCs appear in a different color from red cells and the other parts of a peripheral blood smear images. In this system using color information , after applying several preprocessing steps, nuclei are extracted by means of K- means method. Figure 1 shows the WBCs nucleus segmentation procedure.

At first the microscopic image is separated into RGB components. The median filter is applied only on R, G to decrease spot noise and maintain the edge quality as much as possible. After that, the RGB image is reconstructed by filtered R and G component and the main B component, the nucleus keeps it's own color but the background of the image and RBC's become brighter. Need some color system which can predominant the color of nucleus from the color of background and RBC's. This color system is Luv color which has three independent channels. It is shown that this color system specify the nucleus from the other object of the image. The dependency of color channels is decreased in LUV system rather than RGB because of more Euclidean distances.

Classify the colors of Luv system by applying k-means clustering method on Luv color system and mapping this clusters again to RGB color system. These three classes are shown in Figure 3. To specify which image indicates the nucleus, the mean value of R channel is calculated from each image and image with minimum value would be the candidate of nucleus. Then, the mask of nucleus is made by thresholding in grayscale image and after hole filling, a mask with extra white objects which must be omitted from the nucleus mask.

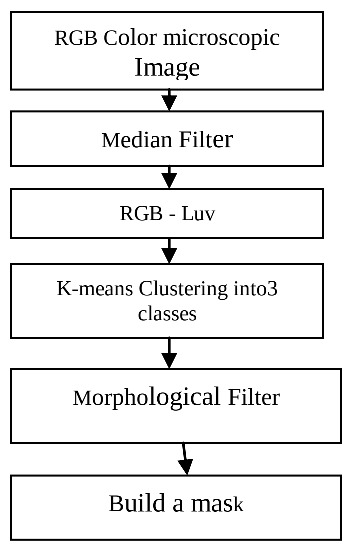

Figure 1 Nucleus segmentation procedure

D Segmentation of White Blood Cell using K-means Clustering

In this paper , at first, nuclei are extracted by clustering the microscopic images into three color clusters in Luv color system using *K*-means method [3].

The K-means is a clustering method which is one of the most popular unsupervised learning algorithm due to it's simplicity. The K-means clustering has been used for image segmentation. K-means clustering is based on minimizing the objective function, shown in equation 2

$$J = \sum_{i=1}^{n} \sum_{j=1}^{k} \left\| x_i - c_j \right\|^2$$

(2)

Where n is the number of data, K is the number of cluster shown in figure1, x_i is the i-th sample, c_j is the j-th centre of the cluster.

Each data will be clustered into 3 groups for all analyses [4].

Classify the colors of Luv system into three classes which can be done by applying simple K-means clustering method on Luv color system and mapping this cluster again to RGB color system. These three classes are shown in figure2.

The algorithm is composed of the following steps:

1. Select randomly k points into the space represented by the pixels that are being clustered. These points represent the initial cluster centre, c_j.
2. Assign each data to the nearest centre.
3. When all data have been assigned , recalculate the new centre position.
4. Repeat step 2 and 3 until the centers are no longer move . This will produce a separation of the object into group from which the metric to be minimized can be calculated.

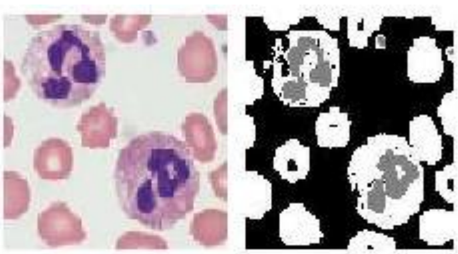

Figure 1(a) Original Image (b) Cluster Image

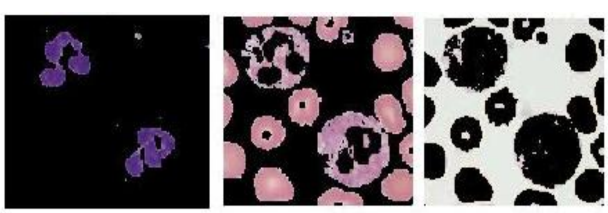

Figure 2 Luv color classes

Extract candidate zone for nucleoli, the curve let transform is applied on extracted nucleus. Curvelet transform is an appropriate transform for detecting detailed information in images due to it's optimality for extraction of 2D singularity based features shown in equation 2

Number of Nuclei Lobes	Ratio of Nuclei to Cell	Ratio of perimeter to Nuclei	Entropy of Nuclei
1	*0.6874*	*0.1281*	*0.8334*
2	*0.3799*	*0.0724*	*0.2344*
1	*0.8845*	*0.0505*	*0.2509*
1	*0.5513*	*0.0047*	*0.3991*
4	*0.3087*	*0.1024*	*0.2523*

The main image is decomposed to different sub bands with different resolution. To apply curvelet transform shown in equation 2, the proposed algorithm in is used , decomposes an N×N Image as follows:

$$f(x,y) = c_j(x,y) + \sum_{j=1}^{J} w_j(x,y) \qquad (3)$$

$c_j(x,y)$ - shows the low pass coefficient of image

w_j - Detail coefficient of the image in scale of 2^{-j}

The output of this algorithm is composed of J+1 sub band by the size of N×N. Wrapping curvelet transform in 4 levels and 8 angles is applied on the nucleus. When all data have been assigned , recalculate the new centre position.

E . Feature Extraction

The most important task in pattern recognition is selecting the proper diagnostic features, describing the image by the numerical values, and enabling the automatic system to perform the recognition. Usually many of the features using texture, geometrical and statistical analyses of the image. The textural features reflect the statistical arrangement of the pixels in the image. The sum and difference histograms of the gray levels of the neighboring pixels are analyzed at difference direction and based on such analysis, the mean value and entropy are treated as the features.

A feature vector is an n-dimensional vector that contains a set of values where each value represents a certain

feature, this vector use to classify an object. Use the area, centre area and perimeter of the object.

The area of the ith objects is defined as follows

Table 1 : WBC feature extraction

$$Area = \sum_{i=1}^{n} \sum_{j=1}^{m} b_{ij} \qquad (4)$$

Where b_{ij} - is the value of binary image (0,1)at the pixel coordinate(i, j) within an m × n image. Area of nuclei and cell taken to calculate ratio of nuclei to cell. Nuclei area is depends on number of lobes present in white blood cells. Ratio of nuclei to cell is minimum when nuclei area is more, it is maximum when nuclei area is less shown in table 1.

Perimeter is defined as the total pixels that constitute the edge of the object, provide information about shape of the object. Entropy converts any class other than logical to uint 8 for the histogram counts calculation so that pixel values are discrete and directly corresponds to a bin value. Ratio of Perimeter to nuclei is maximum when nuclei perimeter is small and minimum for nuclei perimeter is large.

Entropy calculates for extracted nuclei from image shown in equation 5. Entropy of nuclei varies for each type of white blood cell.

$$Entropy = -\sum_{g=0}^{L-1} p(g)\, log_2[p(g)] \qquad (5)$$

Where p(g) probability of gray level g in image.

F . Experimental Results

The proposed method was tested on approximately 30 blood smear slide images acquired by light microscope from stained peripheral blood using the Digital Camera with magnification 100. White Blood Cells features such as number of nuclei lobes, ratio of nuclei to cell, ratio of perimeter to nuclei and entropy shown in Table 1 for five types of White Blood Cells. Experimental results show that method works successfully in the segmentation of white blood cell images, the cells and their nuclei are detected out effectively.

References

[1] Guclu Ongun,Ugur Halici,Kemal Leblebicioglu,Volkan Atalay, Feature extraction and classification of blood cells for an automated differential count system, Turkey.

[2] Xiong, S.H.Ong, J.H.Lim,N.N.Tung,J.Liu,D.Racoceanu,K.Tan,A.Chong, Automatic working area classification in peripheral blood smears using spatial distribution features across scales, Singapore.

[3] Ramin Soltanzadeh, Hossein, Rabbani, and Ardeshir Talebi ,Extraction of Nucleolus Cadidate Zone in White Blood Cells of Peripheral Blood Smear Images Using Curvelet Transform, Isfahan University of Medical Sciences , Isfahan,Iran.

[4] A.S.Abdul Nasir , M.Y.Mashor, Unsupervised Colour Segmentation of White Blood Cell for Acute Leukaemia Images, University Science Malaysia.

[5] Rafael C. Gonzalez, Richard E. Woods, Digital Image Processing ,Third Edition.

[6] Method OIV- MA-AS2-11, Determination of chromatic characteristics according to CIELab,R2006.

[7] K. Jiang, Q. M .Liao ,and S. Y. Dai, A novel white blood cell segmentation scheme using scale- space filtering and watershed clustering, in proceedins of the International Conference on Machine Learning and Cybernetics, vol.5, November 2003.

[8] L. B. Dorini , R. Minetto, and N. J. Leite, White blood cell segmentation using morphological operators and scale – space analysis , in proceedings of the 20th Brazilian Symposium on Computer Graphics and Image Processing , October 2007.

[9] E.Scotti , " Automatic morphological analysis for acute leukemia identification in peripheral blood microscope images", in Proceedings of the 14th International Conference On Digital Signal Processing , 2002

[10] B. R. Kumar , D. K. Joseph, and T.V. Sreenivasc, " Teager energy based blood cell segmentation ," in Proceedings of the 14th International Conference On Digital Signal Processing , vol. 2, pp.619-622, 2002