Red Blood Cell Recognition using Geometrical Features

Jameela Ali[1],AbdulRahim Ahmad[2], Loay E. George[3],Chen Soong Der[4],Sherna Aziz[5]
[1]College of Graduate Studies
[2,4]College of Information Technology
University Tenaga National -Malaysia
[3,5]Baghdad University-Iraq

Abstract

This paper presents the research on analysis and extraction of features of red blood cells for anemia recognition. Images from the blood samples collected at a hospital were used. Three geometrical features were used to distinguish between normal and anemic cells; Fourier descriptors, aspect ratio and moments. The City block distance measure was used as a criterion to determine the similarity degree between the tested samples and the established templates. Test results indicate that combination geometrical features gave high discriminative power approaching 98%.

Keywords: Red blood cells, pattern recognition, Fourier descriptors, aspect ratio, moments.

1. Introduction

Pattern recognition is a field of research that examines the process and the design of systems to identify patterns in the data. Recognition system has emerged as a great challenge for computer vision. The longer term aim is to enable it to achieve near human level recognition for large number of categories under wide variety of conditions [1].

The red blood cell (RBC) recognition system can be used for educational purposes in medical schools and assist in the development of workers in the field of Hematology. RBCs come in a variety of shapes and textures, depending on the types of blood disease suffered by the patient. The variation is especially so in anaemia [2]. For example, the shape of a normal RBC is a biconcave disk, with 6 to 9 µm in diameter and 1.5 to 2.5 µm thick. In the peripheral smear of the sample slide, RBCs are a nucleate and contain predominantly haemoglobin that is distributed to form a dense outer rim with a paler centre that occupies approximately one third of the diameter of the cell[3]**,** The red color of blood is the result of a pigment called hemoglobin, which consists of iron and protein. Increase or decrease in the concentration of hemoglobin can result in different shapes, colors and sizes of the RBC and thus also can affect the textures [4][5].

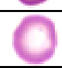
Shape is one of the most important image features due to the fact that shape can effect human perception. Shape features has been extensively applied in RBC recognition to distinguish between normal cells and infected cells [6]. For infected cells, there are also many different shapes, four of which are; Sickle, Echnocyte, Teardrop and Ellipse which relate to four different types of anemia. Many shape representations and retrieval methods exists. However, most of those methods either do not well represent shape or are difficult to be normalized (making matching hard to do). One of the best methods is the one based on Fourier descriptors (FD). It achieved both representation and normalization well [10].

Textures another most widely used feature and has been an active topic in machine intelligence and pattern analysis since the 1950s. Texture features to discriminate different patterns of images by extracting the dependency of intensity between pixels and their neighboring pixels or by obtaining the variance of intensity across pixels [8]. Another three types of anemia are; Stomatocyte, Target and Hypochromic, also differentiated by different shapes of the cell. Moments-based texture analysis method has been introduced in medical images. Texture features are extracted by calculating moments in the texture pixels neighborhoods. Their capability to discriminate different textures has been verified by Wu [9]. In this research, we analyze and compare between eight different RBC shapes as show in table 1.

IJCSI International Journal of Computer Science Issues, Vol. 10, Issue 1, No 1, January 2013
ISSN (Print): 1694-0784 | ISSN (Online): 1694-0814
www.IJCSI.org

In this research, sample images of anemic and normal blood samples were collected and processed to obtain shape features (Fourier descriptors, moment and aspect ratio) to be used for training a recognition system. This paper is structured as follows; section II, discusses the proposed method. In the following section, the types of features are discussed followed by the discussion on experiments and the results in section IV. Finally, section V concludes.

Table1. Classification according to their shapes &textures

No.	Scientific cell name	Image cell
1.	Normal RBC	
2.	Teardrop RBC	
3.	Echnocyte RBC	
4.	Elliptocytes RBC	
5.	Sickle RBC	
6.	Target RBC	
7.	Stomatocytes RBC	
8.	Hypochromic RBC	

2. Methodology

The suggested RBC recognition scheme consists of three stages; (a) isolating target area, (b) determining geometrical features and finally to (c) recognizes the RBC as falling into one of the different types of anemia.
(a) The first stage involves image processing steps required to determine the background and target colors and to isolate the cell area (target) from the surrounding. Then, the external boundary pixels of the cell cut-out are traced.
(b) In the second stage the trace points are used to determine some adopted geometrical features like Fourier descriptors, aspect ratio and moments which have been used to describe the shapes of the blood cells.
(c) In the last stage, 100 different image samples for the normal and abnormal blood cells are used to train the recognizer to recognize 8 kinds of RBC, where 7are infected blood cells and one is of normal type. They are used as test materials to establish the template feature vector for each kind of blood cells. The City Block Distance (CBD) measure was used as a criterion to determine the similarity degree between the tested samples and the established templates that is found. This will permit the ordering of characteristic features and selection of the most useful features.

2.1 System layout

The proposed RBC recognition system is mainly designed to recognize anemia infected cells from normal cells, and recognize the types of anemia by using the extracted cell's features. The system is developed with the help and supervision of an expert (physician). It includes two main phases; training and recognition phase. Each phase involves three subs-stages; preprocessing, feature extraction and recognition. Fig. 1 shows the block diagrams of the training phase.

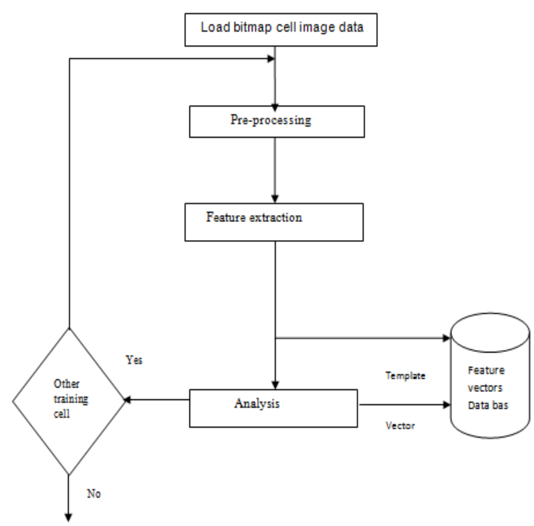


Fig (1) Block diagram of training phase

3. Feature Extraction

We focus on shape description because. Shape description cans be broadly categorized into two types, boundary based and region based. Boundary based methods use only the contour or border of the object shape and completely ignore its interior. Hence, these methods are also called external methods. The region based techniques take into account into account internal details (like holes etc) besides the boundary details. Recognition of a shape by its boundary is the process of comparing and recognizing shapes by analyzing the shapes 'boundaries [10];

3.1 Fourier Descriptors (FD):

Fourier Descriptors is most widely used in boundary based method [11]. The first set of features is Fourier

Descriptors(FD). In this technique, after retrieving RBCs binary shape images using Fourier Descriptors. Due to Fourier descriptors are used to describe the objects shape in terms of its spatial frequency content [12]

Fourier Descriptors based on the following Equation

$$C(n) = \frac{1}{m}\sum_{i=1}^{m-1}\left(\frac{\Delta xi}{\Delta yi}\cos\frac{2\pi ni}{L}\right)$$

$$S(n) = \frac{1}{m}\sum_{n=1}^{m-1}\left(a_n\frac{\Delta yi}{\Delta xi}\sin\frac{2\pi ni}{L}\right)$$

Where

m is the number of contour points

$\Delta xi = xi - xi+1$

$\Delta yi = yi - yi+1$

$\Delta \gamma i = \sqrt{(\Delta xi)^2 + (\Delta yi)^2}$

$L = \sum \Delta \gamma i$

(xi,yi) is the colume and row number of i thcountour point

Table (2) show the Fourier Descriptors of RBCs results.

3.1 MomentsInvariants(M):

Moments is one of complet of geometric in spatial domain in Region Based methods[13].moments is the second automated method for RBC image feature extraction .According to [14], Which was derived equations moments. A moments based on the following Equation

$$aMom(n) = \frac{1}{k}\left|\sum_{i\in c}\left|(x_i' + jy')^n\right|\right| \quad , \ n = 1,2,3,4$$

Where,

$$x_i' = \frac{1}{L}\left[(x_i - \bar{x})\cos\theta - (y_i - \bar{y})\sin\theta\right]$$

$$y_i' = \frac{1}{L}\left[(x_i - \bar{x})\sin\theta + (y_i - \bar{y})\cos\theta\right]$$

$$\bar{x} = \frac{1}{m}\sum_{i\in c} x_i, \quad \bar{y} = \frac{1}{m}\sum_{i\in c} y_i$$

$$L = \min_{i\in c}\left(|x_i - \bar{x}|, |y_i - \bar{y}|\right)$$

Table (2) show a moment of RBC results

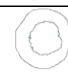
3.2 Aspect Ratio (AR).

The third and final method of automated feature extraction represents Aspect Ratio. Aspect Ratio is one of the most common examples of a Shape factor that represents quantities in shapes that have no dimensions used in image analysis[15].

$$ExtAspRat = \frac{(No\ of\ External\ Boundary\ Pixels)^2}{Total\ No\ of\ Cell\ Pixels}$$

Table (2) show Aspect Ratio of RBC results

Table (2) show the result of geometrics features

NO	Scientific cell name	Image cell	FD	AM	ER
1	Normal RBC		0.4971	0.01172	10.26
2	Teardrop RBC		0.3946	0.00268	19.11
3	Echnocyte RBC		0.3131	0.00802	41.19
4	Elliptocytes RBC		0.4111	0.00398	16.54
5	Sickle RBC		0.0540	0.00236	24.78
6	Target RBC		0.4955	0.01178	10.18
7	Stomatocytes RBC		0.5063	0.01101	11.86
8	Hypochromic RBC		0.4691	0.00687	13.12

4. City-Block Distance

The RBCs shape features extracted from the three methods above are presented to the City-Block Distance measurement for testing to make matching with the feature values in a reference database. In particular, City-Block Distance is a classifier that matches values of input features with values from features in a reference database. The use of City-Block Distance relies on four assumptions of distance function between points. For all points x, y, and z, a distance function ***D(x,y or z)*** satisfies the following properties: (a) Non-negativity: ***D(x, y) = 0***. (b) Reflexivity: ***D(x, y) = 0*** if and only if ***x = y***. (c) Symmetry: ***D(x, y) = D(y, x)*** and (d) Triangle inequality: ***D(x, y) + D(y, z) = D(x, z)* [16]**.

City-Block Distance achieves this classification and matching of image features by measuring the distance in between two connected pixels. Given two pixels at position (x_1,y_1) and (x_2,y_2), measurement by City-Block Distance can be expressed using function[16]

$$D_{12} = |x_1 - x_2| + |y_1 - y_2|$$

To assess the discrimination power of each adopted feature the following criteria was adopted

$$P_k = \frac{(n_c - 1)\sum_{j=1}^{n_c} \sigma(j,k)}{2\sum_{i=1}^{n_c - 1}\sum_{j=i+1}^{n_c}\left|\overline{F}(i,k) - \overline{F}(j,k)\right|}$$

Where

$$\overline{F}(j,k) = \frac{1}{n_j}\sum_{i=0}^{n_j - 1} f(j,i,k)$$

$$\sigma(j,k) = \sqrt{\frac{1}{n_j}\sum_{i=0}^{n_j - 1}(f(j,i,k) - \overline{F}(j,k))^2}$$

Where, nj is the number of training cell images belong to j-class, F (i,j, k) is the kth feature extracted from ith sample that belong to jth class, k=0,1,…….number of features. $\overline{F}(j,k)$ is the mean of kth feature for jth class. $\sigma(j,k)$ is the standard deviation of kth feature for jth class. Table (3) Discrimination power of adopted geometrical features.

Table (3) show discrimination poer of geometric features

NO	Featurs	discrimination power
1	MO	60.6%
2	AR	62%
3	FD	64%
4	FD,MO	86%
5	FD,AR	88%
6	MO,AR	86.5%
7	FD,MO,AR	98%

5. Experimental Results

The test results indicate that, for table (2) the cells divided into two groups to spherical and non-spherical according to (EAR) values. Where Target, Stomatocytes, and Hypochromic are spherical groups. Where value close to the normal cells value. Whilst Sickle, Echnocyte and Teardrop are non –spherical cells.

For table (3) first, discrimination power for FD better than AM and EAR respectively. Secondly collect together three features given high accuracy equal to98%.

6. Conclusion

Automated image-recognition systems provide significant benefits for medical test analysis. Since medical images are highly variants the development of reliable recognition processes is difficult. Blood recognition system is a difficult application in medical diagnoses, because the cells have several shapes, color and size. In this paper, we focus on three geometrical features (Fourier Descriptors, Moments and Aspect ratio) to extract features to 8 types of RBCs and used City-Block Distance method to distinguish between 8 different shapes. The results indicate that discrimination power of FD better than AM and EAR respectively and grouping the three features given high accuracy in discrimination power equal to98%.

References

[1] Woo Chaw Seng and Seyed Hadi Mirisaee, “A New Method for Fruits Recognition System”, *MNCC Transactions on ICT*, 2009 Vol. 1, No. 1, June.
[2] Ferat Sahin, “A Radial Basis Function Approach to a Color Image Classification Problem in a Real Time Industrial Application”, Master's thesis, Virginia polytechnic institute, Blacksburg, 1997.
[3] M. H. DE KEIJZER" Automated counting nucleated red blood cells in blood samples of newborns" Clin. Lab. Haem. 2002, 24, 343–345.
[4] F. Domino, Robert R. Sharp, Steven Lipper" NMR Chemistry Analysis of Red Blood Cell Constituents in Normal Subjects and Lithium-Treated Psychiatric Patients" BIOL PSYCHIATRY 1985; 20:1277-1283.
[5] Sveta Kabanova, Petra Klinbongard"Gene expression analysis of human red blood cells"International Journal of Medical Sciences, 2009,6(4):156-159.
[6] Vinay Saxena"Fourier descriptors under rotation, scaling, translation and various distortion forhand drew planar curves"Journal of Experimental Sciences, 20123(1): 05-07ISSN: 2218-1768.
[7] A.Karid, L.E.Nugroho"Acomprative experiment of several shape methods in recognizing blants"international journal of computer science & information technology, June 2011, (IJCSIT),Vol3, No3.
[8] Fabian Timm, Thomas Martinetz"Statistical Fourier Descriptors for Defect Image Classification" international conferanc of pattren recognition Ratzeburger Allee, 2011, 160, 23538 L¨ubeck, Germany.
[9] Ke Wu, Carole Garnier" A preliminary study of moment-based texture analysis for medical images"Conf Proc IEEE Eng Med Biol Soc. 2010: 5581–4.doi: 10.1109/IEMBS.2010.5626789.
[10] Zhang, D., & Guojun, L. A Comparative Study on Shape Representation Using Fourier Descriptors with Different Shape Signatures. Churchil: Australia 2009.
[11] Persoon, E., & Fu, K. Shape Discrimination Using Fourier Descriptors. IEEE Trans. On Sytems, Man andCybernetics, 7(3) 1995.
[12] Loay E. George,Shatha M. Noor"Handwritten Arabic (Indian) Numerals Recognition Using Fourier Descriptorand Structure Base Classifier"Journal of Al-Nahrain University Vol.14 (2), pp.215-224, June, 2011
[13] Eakins and M.Graham " contour-based image retrieval " JTAP, 1999, Generic,89.

[14] JAMEELA ALKRIMI," Medical Diagnosis by Analyses of Blood Cell Image" Master theses 2006.
[15] Tieng, M., Q., & Boles, W., W. Recognition of 2D Object Contours using the Wavelength Transform Zero-Crossing Representation IEEE Trans,1997, 19(8).
[16] International Association for Pattern Recognition. (2003), what is Pattern Recognition. IAPR Newsletter, 25(1), 1.