Journal of Network and Innovative Computing
ISSN 2160-2174 Volume 1 (2013) pp. 74 - 86

Texture-Based Leukocyte Image Retrieval Using Color Normalization And Quaternion Fourier Transform Based Segmentation

Prabir Sarkar[1], Madhumala Ghosh[2] and Chandan Chakraborty[3]

School of Medical Science and Technology
Indian Institute of Technology, Kharagpur, India
[1]*prabir.jis@gmail.com*
[2]*madhumala.ghosh@gmail.com*
[3]*chandanc@smst.iitkgp.ernet.in*

***Abstract*: This work aims to develop texture-based microscopic image retrieval methodology for leukocyte recognition. This approach includes four consecutive steps viz. color normalization technique, noise reduction, segmentation method and textural feature extraction. The quadratic approach provides best normalization which is followed by noise removal by median filter. Comparative studies among five filters have been experimented to search the best suited filter for this particular type of microscopic image of peripheral blood smears. In segmentation step, the segmentation is done by quaternion Fourier transform (QFT) and this method is compared with literature. The comparative study reveals that amongst three segmentation techniques, quaternion Fourier transform gives higher segmentation accuracy (92.13%) to identify the leukocyte nucleus. Comparative analysis in feature extraction step shows that Gabor-wavelet features provides higher precision i.e., 82.5% for (normal leukocyte) query image and histogram based features gives the better precision (76.53%) for recognizing abnormal leukocyte.**

***Keywords*:** CBIR, Quaternion Fourier transform, Gabor wavelet, CML Leukocytes, blood smear image**.**

I. Introduction

Content-based image retrieval (CBIR) approach uses visual contents to search the image from large scale databases according to user's interests. The term 'visual content' refers to colors, shapes, textural and any other information of the image. The 'content-based' means a search which will analyze the visual content of the image rather than the metadata such as keywords, tags, or descriptions associated with the image. CBIR involves retrieval of images similar to a query image in terms of visual features or visual information extracted from the images. It has been used in various applications such as digital forensic investigation, fingerprint matching, face detection, DNA matching and also used in medical science for finding critical information towards various disease diagnostics.

Light microscopic imaging of peripheral blood smears generates significant clinical information about blood cells leading to various blood related disorders. In view of this, large collections of digital blood smear images are developed in the form of microscopic image repository to keep track of patients' hematological records. Amongst various types of hematological disorders, leukemia is one of the most well known types of blood cell disorder which affects the quality of life of a human being a lot and hence its early prediction is required for proper therapeutic intervention. There are mainly four types of leukemia - chronic myeloid leukemia (CML), acute myeloid leukemia (AML), chronic lymphoid leukemia (CLL) and acute lymphoid leukemia (ALL). In addition, leukocytes itself are also of five types viz., neutrophil, eosinophil, basophil, monocyte and lymphocyte; those of which are visually differentiated by mainly their nucleus shape-size and color as well as textural information of cytoplasm.

As early screening leads to reduce the mortality rate due to leukemia, there is an indeed need of developing semi or fully automatic diagnostic scheme for blood microscopic images. Various attempts by the researchers have been made towards this, but most of these are based on either classifiers or knowledge-based approaches. In effect, the detection process becomes time-consuming. Under such circumstances, CBIR plays a crucial role by retrieving the most similar image to classify the query image through pattern matching without using all the images together and hence this takes comparatively less time to provide the diagnostic result. In view of this, this work addresses the development of a content based leukemia image retrieval methodology from the large microscopic image database.

CBIR is the major research topic in medical image processing domain due to the large microscopic blood smear image database. It has become one on the active and fast research area in the field of computer vision over past decades [1, 2]. Early techniques were not generally based on visual features

but on the textual annotation of images. In other words, images were first annotated with text and then searched using a text-based approach from traditional database management system [3]. Annotating images with text is a cumbersome and expensive task for large image databases and is often subjective, context-sensitive and incomplete. In the beginning of the 1980's text based retrieval system for image is present, but the start of CBIR such as Query by Image and video Content (QBIC) system by IBM researcher [4] in the 1995. Since then, research on CBIR has developed rapidly. So many commercial systems were designed for image retrieval such as Blobworld system [5], PicHunter [6], Virage [7], Photobook [8], etc. Most of the retrieval systems have similar architectures for image storage and access methods, visual feature extraction and similarity distance measurements. In medical applications, approaches have been proposed on CBIR system which is specially designed to support medical tasks [9-11]. Designing CBIR for medical images in the different imaging modalities requires very specific and particular diagnostic information. Many CBIR systems exist in different medical imaging modalities such as computer tomography (CT), magnetic resonance images (MRI), positron emission tomographic (PET), X-ray images, ultrasound images, light microscopic images [12-15] etc. In the context of light microscopic blood smear image retrieval, in [16] proposed automatic method of a CBIR system for blood smear images using global feature color histogram and wavelet transform. But only using global feature over medical images to retrieve similar image, may fail to find similar type of images [17]. Therefore region based feature or local feature is required for efficient medical image retrieval system [18-19]. In [20] worked in image retrieval for leukemia affected blood cells in which they have reviewed on the technical aspects of image acquisition, analysis, search and retrieval. They summarized that the importance of segmentation is the key step for retrieve similar leukemia blood smear image. For identification of leukocyte region, there are varieties of automatic or semi-automatic methods have been proposed to segment leukocytes [21-23]. From those literatures we find different segmentation method. Here automatic, unsupervised and robust segmentation technique is required to develop CBIR system. In [42] the paper describes a preliminary study of developing different class of leukemia using microscopic blood smear images. In [43] Bimodal system for retrieval of pathology images where region of interest in the image segmented and invariant shape descriptors used for similarity measure. Segmentation and classification methods are used for performing similarity measure [44] in an automated expert system for lymph node Hemopathology. In [45] acute leukemia identification is done from peripheral blood smear image. They extract feature from lymphocyte nucleus and cytoplasm for classification. Comprehensive literature survey showed that no significant attempt has been made towards texture based pattern matching for leukemia detection in addition to morphometric similarity.

In order to overcome the problem of retrieval of leukocyte blood smear image, we propose quaternion Fourier transforms based segmentation technique to retrieve similar images. Our CBIR system can make a strong impact in diagnostic, research and education for blood smear images. Rest of the paper is organized as follows: In section II we have described the framework of the proposed method. In section III we have described color correction and different preprocessing technique of blood smear image. In section IV we have proposed quaternion Fourier transform based segmentation technique of leukocyte nucleus. In section V, we have explained different textural features and distance measurement techniques for retrieval of normal or CML leukocyte images. In section VI we have shown the experimental results with further discussion.

II. Framework of Proposed Methods

In CBIR, the user may search a large collection of medical images that are similar to a query image. The search is based on perceptual similarities of the attributes color, texture and shape. Here the proposed method is based on color image segmentation where our objective is to separate leukocytes from the background. Before segmentation to make a standardization of microscopic blood smear image we applied color correction method and preprocessing technique for noise removes. Using Quaternion Fourier transform first we obtain reconstructed image of blood smear image. From the reconstructed image we extract leukocyte regions. For identification of healthy and CML leukocyte, Gabor wavelet features are extracted from region of interest and similarity is computes between a query image and database image feature vector. Overall work flow diagram of leukocyte blood smear image retrieval is shown in figure 1.

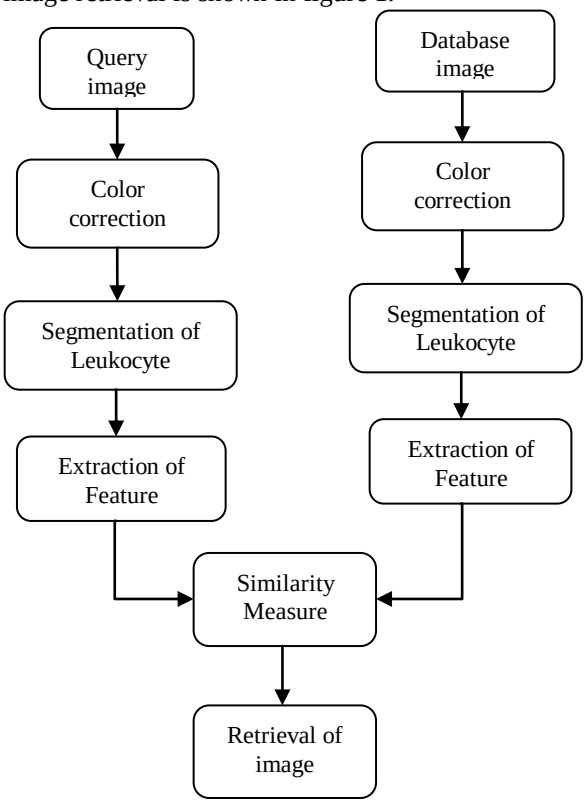

Figure1. Work flow diagram for leukocyte blood smear image retrieval

III. Color Correction and Preprocessing

A. Microscopic image data collection

A set of 100 healthy and 100 CML light microscopic image samples was collected from Dept. of Pathology, Midnapur Medical College & Hospital, West Bengal. All these images were grabbed by expert pathologists using Leica Observer (Leica DM750, Switzerland) compound microscope under 100X oil objective (NA 1.5150) in JPEG format with image resolution of 640×480.

B. Color Correction Technique

The color of microscopic blood smear images may vary from sample to sample and this color variation is mainly generated from the camera calibration, light intensity as well as staining variability. Therefore it is very much important and beneficial approach to make a standardization of the microscopic image color to a particular scale. Sometimes, it has been observed that, in case of microscopic blood smear image, any variation from the color value intensity level can be contributed to a change in illumination. Gray world assumption, retinex theory and quadratic mapping are three common method used for illumination correction [34]. The gray world and retinex theory lead to methods that are very efficient and produce a linear correction to the R and B channels. Quadratic mapping is an effective technique which is combination of the two gray world and retinex theory method together. Basically, the gray world assumption is defined as a technique which balances the different color channel of an RGB image with respect to gray value in order to maintain an average of gray across the entire image of microscopic blood smear images.

1) Gray world assumption

Let us consider a microscopic blood smear image as $I(p, q)$, where p and q describe the coordinates of the image having size M×N. $I_r(p, q)$, $I_g(\mathrm{p}, q)$ and $I_b(p, q)$ are respective red, green and blue channels of the of microscopic blood smear images. The average value of red, green and blue channel is defined as,

$$R_{avg} = \frac{1}{MN}\sum_{p=1}^{M}\sum_{q=1}^{N} I_r\ (p, \mathrm{q}) \tag{1}$$

$$G_{avg} = \frac{1}{MN}\sum_{p=1}^{M}\sum_{q=1}^{N} I_g\,(p, \mathrm{q}) \tag{2}$$

$$B_{avg} = \frac{1}{MN}\sum_{p=1}^{M}\sum_{q=1}^{N} I_b\,(p, \mathrm{q}) \tag{3}$$

Therefore by keeping the green channel as constant and the gain for red and blue channels are calculated as follows,

$$G_\alpha = \frac{G_{avg}}{R_{avg}} \quad \text{And} \quad G_\beta = \frac{G_{avg}}{B_{avg}} \tag{4}$$

Finally, the adjusted value of red and blue channel are termed as,

$$\grave{I}_r\,(p,q) = G_\alpha \times I_r\,(p,\ \mathrm{q}) \tag{5}$$

$$\grave{I}_b\,(p,q) = G_\beta \times I_b\,(p,\ \mathrm{q}) \tag{6}$$

Therefore the final output image after illumination correction is represented by

$$I_{Gimg} = \begin{bmatrix} \grave{I}_r\,(p,q) \\ I_g\,(p,q) \\ \grave{I}_b\,(p,q) \end{bmatrix} \tag{7}$$

2) Retinex theory

It is argued that the perceived white associated with the maximum cone signals of human visual system. As such, the method for white balance should be to equalize the maximum values of the red

$$R_{\max} = \max_{p,q}\left\{I_r\,(p,q)\right\} \tag{8}$$

$$G_{\max} = \max_{p,q}\left\{I_g\,(p,q)\right\} \tag{9}$$

$$B_{\max} = \max_{p,q}\left\{I_b\,(p,q)\right\} \tag{10}$$

The gain for red and blue channels are calculated as follows,

$$R_\alpha = \frac{G_{\max}}{R_{\max}} \ \text{And}\ R_\beta = \frac{G_{\max}}{B_{\max}} \tag{11}$$

Then adjust the red and blue pixels by,

$$\tilde{I}_r\,(p,q) = R_\alpha \times I_r\,(p,q) \tag{12}$$

$$\tilde{I}_b\,(p,q) = R_\beta \times I_b\,(p,q) \tag{13}$$

Therefore the final output image after retinex theory illumination correction is represented by

$$I_{Rimg} = \begin{bmatrix} \tilde{I}_r\,(p,q) \\ I_g\,(p,q) \\ \tilde{I}_b\,(p,q) \end{bmatrix} \tag{14}$$

3) Quadratic mapping

To incorporate the both gray world and retinex methods quadratic approach has come. As with the quadratic mapping methods above we keep the green channel unchanged. Let the change to the red and blue channels be represented as

$$\tilde{I}_r\,(p,q) = \mu I_r^2\,(p,q) + \nu I_r\,(p,q) \tag{15}$$

$$\tilde{I}_b\,(p,q) = \mu I_b^2\,(p,q) + \nu I_b\,(p,q) \tag{16}$$

In the equation (15) and (16) where (μ,ν) are the parameters for automatic illumination correction using quadratic mapping. To satisfy the gray world assumption, we require that (17) and (18),

$$\sum_{p=1}^{M}\sum_{q=1}^{N}\tilde{I}_r(p,q)=\sum_{p=1}^{M}\sum_{q=1}^{N}I_g(p,q) \tag{17}$$

i.e.

$$\mu\sum_{p=1}^{M}\sum_{q=1}^{N}I_r^2(p,q)+\nu\sum_{p=1}^{M}\sum_{q=1}^{N}I_r(p,q)=\sum_{p=1}^{M}\sum_{q=1}^{N}I_g(p,q) \tag{18}$$

Similarly to satisfy the Retinex theory, we need (19),

$$\mu\max_{p,q}\left\{I_r^2(p,q)\right\}+\nu\max_{p,q}\left\{I_r(p,q)\right\}=\max_{p,q}\left\{I_g(p,q)\right\} \tag{19}$$

From the (18) and (19) both equations can be representing in a matrix. In equation (20) as follows,

$$\begin{pmatrix}\sum\sum I_r^2 & \sum\sum I_r\\ \max I_r^2 & \max I_r\end{pmatrix}\begin{pmatrix}\mu\\ \nu\end{pmatrix}=\begin{pmatrix}\sum\sum I_g\\ \max I_g\end{pmatrix} \tag{20}$$

The illumination correction for blue channels can be computed in the same manner. From this equation (20) we can find the value of (μ,ν) using Cramer's rule. Using this (μ,ν) value red and blue color channel adjusted and green channel remain constant to get color corrected image.

C. Noise reduction of microscopic blood smear image

Several filtering techniques are incorporated to filter out the noises present in the images. Here, we assume that the original image corrupted by noise as $g(a,b)$ and S_{pq} represents the set of coordinates in a rectangular sub image window (neighborhood) of size m × n, which is centered at a point (p, q). The filtered image is represented by $f(p,q)$. The filters used in the comparative study are described as follows,

1) Median filter

Median filter [32] is very popular type of filter which provides an excellent amount of random noise reduction capability with considerably less blurring than the linear type of smoothing filters of the same size.

$$f(p,q)=\underset{(a,b)\in S_{pq}}{median}\left\{g(a,b)\right\} \tag{21}$$

2) Laplacian filter

Hence this filter sharpens the image while the Laplacian image [32] is superimposed over the original one. Here, Laplacian filter is defined as,

$$g(p,q)=f(p,q)+c\left[\nabla^2 f(p,q)\right] \tag{22}$$

Where, ∇ = represents the second derivative.

3) Alpha trimmed mean filter

Alpha trimmed mean filter [32] is used for removing the mixed type of noises like salt and pepper noise, Gaussian noise. Suppose $g_r(a, b)$ represents the remaining ($mn - d$) pixels in the sub image S_{pq}. A filter which is formed by the average value of the remaining pixels is defined as the alpha trimmed mean filter. Here, it is defined as,

$$f(p,q)=\frac{1}{mn-d}\sum_{(a,b)\in S_{pq}}\left\{g_r(a,b)\right\} \tag{23}$$

4) First order statistics filter

This filter uses the first order statistics [32] of the neighborhood of the centre pixel in the sub image S_{pq}. The filter function is defined as,

$$f(p,q)=\bar{g}(p,q)+k(p,q)(g(p,q)-\bar{g}(p,q)) \tag{24}$$

Where, $f(p, q)$ is the estimated impulse noise reduced value of pixels at location (p, q) on the image with an original value g(p, q). $\bar{g}(p,q)$ is the local mean of the $m \times n$ neighborhood around and including g(p, q). $k(p,q)$ is the weight factor with $k\in[0,1]$. The factor $k(p,q)$ is a function of local statistics and it is computed as,

$$k(p,q)=\frac{1-\bar{g}^2(p,q)\sigma^2(p,q)}{\sigma^2(p,q)(1+\sigma^2)} \tag{25}$$

Where, $\sigma^2(p,q)$ is the local noise variance in the moving window (W = m ×n) and σ^2 is the variance of noise in the whole image respectively. Hence the noise variance of the whole image is computed as,

$$\sigma^2=\sum_{x,y}\frac{\sigma^2(p,q)}{\bar{g}^2(p,q)} \tag{26}$$

Hence in our problem, the window size is taken as 5×5.

5) Midpoint filter

Midpoint filter computes the mid-point [32] between minimum and maximum values within the area encompasses by S_{pq}. Therefore, the value of restored image f at point (p, q) is calculated by,

$$f(p,q)=\frac{1}{2}\left[\max_{(a,b)\in S_{pq}}\left\{g(a,b)\right\}+\min_{(a,b)\in S_{pq}}\left\{g(a,b)\right\}\right] \tag{27}$$

This filter actually is the combination of order statistics as well as averaging filters and displays better result for randomly distributed noise structures (e.g. Gaussian or uniform noise).

D. Quantitative performance measures for filter

The preprocessing steps are considered according to some quantitative performance measures of the considered filters as well as enhancement procedures. We have taken a common representation of original microscopic blood smear image as $g(p,q)$ and the filtered image as $f(p,q)$.

1) Mean square error (MSE)

Mean square error [32] is a typical metric which is used to calculate the measure of change in quality between the original image and preprocessed image.

$$MSE = \frac{1}{MN}\sum_{p=1}^{M}\sum_{q=1}^{N}(g(p,q)-f(p,q))^2 \quad (28)$$

2) Root mean square error (RMSE)

Root mean square error (RMSE) [32] is defined as the square root of MSE value. Actually RMSE helps to find out the standard deviation of the mean square error rate between the original image and preprocessed image.

$$RMSE = \sqrt{\frac{1}{MN}\sum_{p=1}^{M}\sum_{q=1}^{N}(g(p,q)-f(p,q))^2} \quad (29)$$

3) Signal to noise ratio (SNR)

Signal to noise ratio [32] defines the comparative ratio value of average power of signal to the estimated noise component present within the signal. Here this ratio is calculated in between the original image and estimated noise component present in that image. This ratio is defined as,

$$SNR = 10\log_{10}\left(\frac{\sum_{p=1}^{M}\sum_{q=1}^{N}(g^2(p,q)+f^2(p,q))}{\sum_{p=1}^{M}\sum_{q=1}^{N}(g(p,q)-f(p,q))^2}\right) \quad (30)$$

4) Peak signal to noise ratio (PSNR)

Peak signal to noise ratio [39] is used to calculate the ratio of the maximum possible power of an image signal and the noise which affects the fidelity of the representation. The calculation of PSNR is advantageous because of its easier computational approach but this ratio should not represent the perceptual quality of the image. PSNR is formed as,

$$PSNR = 10\log_{10}\left(\frac{(\max I)^2}{MSE}\right) \quad (31)$$

Here 'max I' represents the maximum possible pixel value of that particular image (in between 0-255 range of gray values).

5) Average difference (AD)

Average difference (AD) [39] calculates the difference between the original image and filtered image. A lower value of AD proves that more noise is removed from the image and image is more apparent in look. AD is computed as,

$$AD = \frac{1}{MN}\sum_{p=1}^{M}\sum_{q=1}^{N}(g(p,q)-f(p,q)) \quad (32)$$

6) Normalized absolute error (NAE)

Normalized absolute error (NAE) [39] helps to find out how much the filtered image is close to the original one i.e. how much noise is removed from the image. So, larger value of NAE indicates poor quality of filtered image. Here, NAE is computed as,

$$NAE = \frac{\sum_{p=1}^{M}\sum_{q=1}^{N}\left|(g(p,q)-f(p,q))\right|}{\sum_{p=1}^{M}\sum_{q=1}^{N}\left|(g(p,q))\right|} \quad (33)$$

7) Universal image quality index (UIQI)

The universal image quality index (Q) [38] relates the three basic factors of the original and enhanced images i.e. loss of correlation, luminance distortion and contrast distortion.

$$Q = \frac{1}{K}\sum_{p=1}^{K}Q_p \quad (34)$$

Where, $Q_p = \frac{\sigma_{k,fg}}{\sigma_{k,f}\,\sigma_{k,g}} \cdot \frac{2\overline{f}_k\,\overline{g_k}}{\overline{f}_k^{\,2}+\overline{g_k}^{2}} \cdot \frac{2\sigma_{k,f}\,\sigma_{k,g}}{\sigma_{k,f}^2+\sigma_{k,g}^2}$

$\overline{g_k}$ and $\overline{f_k}$ are the mean of sub image window for original and enhanced image, $\sigma_{k,f}$ and $\sigma_{k,g}$ represent the standard deviation of k[th] sub image window **W** for original and enhanced image, $\sigma_{k,fg}$ is the covariance between them. Here window size is 5×5.

IV. Segmentation Technique

A. Quaternion based segmentation

The light microscopic blood smear image is a color image can be represented using quaternion as follows:

$$f(p,q) = r(p,q)i + g(p,q)j + b(p,q)k \quad (35)$$

Where (p,q) pixel indicates coordinates. $r(p,q)$, $g(p,q)$ and $b(p,q)$ denote R, G and B components of the pixel respectively. Quaternion approach introduced by Hamilton [24] has been used in the many color image processing applications. The application of quaternion Fourier transforms on color images was reported by Ell and Sangwine to represent color image pixels using quaternion and then implement the quaternion Fourier transform [25]. A quaternion has four components, one real and three imaginary. A pure quaternion *q* in its form

$$q = w + xi + yj + zk \quad (36)$$

Where *w, x, y* and *z* are real numbers; *i, j* and *k* are the imaginary parts. The hyper-complex or quaternion color image is represented as

$$f(p,q) = r(p,q)\mu_1 + g(p,q)\mu_2 + b(p,q)\mu_3 \quad (37)$$

Where μ_1, μ_2, μ_3 are unit pure quaternion, $\mu_1^2 = \mu_2^2 = \mu_3^2 = -1$, $\mu_1 \perp \mu_2$, $\mu_2 \perp \mu_3$, $\mu_1 \perp \mu_3$ and $\mu_3 = \mu_1\mu_2$. Ell and Sangwine defined Quaternion Fourier transform (QFT) of $f(p,q)$ can be described as follows

$$QFT[u,v]=\frac{1}{\sqrt{MN}}\sum_{p=0}^{M-1}\sum_{q=0}^{N-1}e^{-\mu_i 2\pi\left(\left(qv/M\right)+\left(pu/N\right)\right)}f(p,q) \quad (38)$$

Where $i\in\{1,2\}$ In the frequency domain (u,v) is the co-ordinate location of an image pixel and in the special domain (p,q) is the co-ordinate location of an image pixel. After computing QFT we set amplitude spectrum as a constant (A=1) and $QFT[u,v]$ contain only phase spectrum in frequency domain which carries location information [26]. After that Inverse quaternion Fourier transform (IQFT) of $QFT[u,v]$ can be represented as

$$f(p,q)=\frac{1}{\sqrt{MN}}\sum_{p=0}^{M-1}\sum_{q=0}^{N-1}e^{\mu_i 2\pi\left(\left(qv/M\right)+\left(pu/N\right)\right)}QFT[u,v] \quad (39)$$

And whose result is presented as $t(p,q)$:

$$t(p,q)=i_0(p,q)+i_1(p,q)\mu_1+i_2(p,q)\mu_2+i_3(p,q)\mu_3 \quad (40)$$

After taking absolute value of $t(p,q)$ we reconstruct image which gives accurate location of leukocyte based on as follows:

$$R(p,q)=\square\, t(p,q)\,\square^2 \quad (41)$$

Leukocytes are detected according to the reconstructed image obtained. Then morphological operations have been used for identification of leukocytes from the reconstructed image. It removes platelets or artifact from the reconstructed image. Afterwards, the leukocyte nuclei are segmented from reconstructed image.

Pseudo code for Finding Segmented leukocyte nuclei from a blood smear image

Step 1: Input the leukocyte color image.
Step 2: Perform color correction and preprocessing using quadratic mapping and median filter respectively.
Step 3: Construct Quaternion image q (n, m) using RGB color channel.
Step 4: Perform quaternion Fourier transform on quaternion image.
Step 5: Set the amplitude spectrum as a constant value (A=1) and get original phase spectrum.
Step 6: Then perform inverse Quaternion Fourier transform on original phase spectrum.
Step 7: Take absolute value of reconstructed image.
Step 8: Convert into gray image.
Step 9: Convert into binary image to perform morphological operation.
Step 10: Perform dilation and erosion operation using disk-shaped structuring element with radius 1 and fill the nuclei.
Step 11: Get segmented leukocyte nucleus image.

V. Feature Extraction and matching

A. Feature extraction

The main purpose of texture-based retrieval is to find images or regions with similar texture. It is assumed that we are interested in images or regions that have homogenous texture.

1) Gabor wavelets

A two dimensional Gabor function [27] [37] is explained as,

$$G_{l,k}(p,q)=\frac{1}{2\pi\sigma_p\sigma_q}\exp\left(-\frac{1}{2}\left(\frac{p^2}{\sigma_p^2}+\frac{q^2}{\sigma_q^2}\right)+2\pi jwp\right) \quad (42)$$

At this point, *w* denotes modulation frequency; σ_p and σ_q indicates the standard deviations of the Gaussian envelope function. The self-similar Gabor wavelets are obtained through the generating function:

$$G_{l,k}(p,q)=a^{-l}G\begin{bmatrix}a^{-l}(p\cos\theta+q\sin\theta),\\ a^{-l}(-p\sin\theta+q\cos\theta)\end{bmatrix}\quad a>1 \quad (43)$$

Where p and q specify the scale and orientation of wavelet respectively and a^{-l} is a scale factor, l and k are integer, the orientation θ is given by θ = kπ/K, and K is the number of orientations. Therefore, for a given input of image $I(p,q)$, its Gabor wavelet transform is estimated as

$$X_{l,k}(p,q)=I(p,q)*G_{l,k}(p,q) \quad (44)$$

For *l* = 1, 2. . . S and *k* = 1, 2. . . K. Here, the parameters K and S are number of orientation and number of scales respectively. The mean and the standard deviation are used as features of this wavelet function and are represented by the following equations,

$$\mu_{l,k}(p,q)=\frac{1}{M\times N}\sum_{p=1}^{M}\sum_{q=1}^{N}\left|X_{l,k}(p,q)\right| \quad (45)$$

$$\sigma_{l,k}(p,q)=\frac{1}{M\times N}\sum_{p=1}^{M}\sum_{q=1}^{N}\left(\left(\left|X_{l,k}(p,q)\right|-\mu_{l,k}\right)^2\right)^{\frac{1}{2}} \quad (46)$$

Hence, $\mu_{l,k}$ and $\sigma_{l,k}$ construct the feature components of the feature vector. Putting K = 6 (as no of orientations) and S = 4 (no of scales), a feature vector of length 48 is obtained.

$$F=\{\mu_{11},\sigma_{11},.....,\mu_{46},\sigma_{46}\} \quad (47)$$

2) Wavelet

Nine wavelet features are extracted from the normal and leukemia leukocytes image after decomposing. Here, 'Daubechies' wavelet [28] is used here to measure the wavelet features. The details' coefficients are extracted at level 4. Then the percentage of energy corresponding to the approximation is considered.

3) Entropy measure

Entropy is the measure of uncertainty associated with randomness. Here we have consider five different types of

entropy measures including Shannon, Havarda and Charvat, Kapur's entropy, Renyi's [29], and yeager's measure. Here considering the gray level histogram H_i , $i=\{0,1,2.....,N_{l-1}\}$, where N_l is the number of distinct gray levels of the leukocyte.

a) Shannon Entropy

$$S=-\sum_{i=0}^{N_l-1}H_i\log_2(H_i) \tag{48}$$

b) Havarda and Charvat's entropy

$$HC=\frac{1}{1-\alpha}\left(\sum_{i=0}^{N_l-1}H_i^{\alpha}-1\right), where \alpha\neq 1,\alpha>0 \tag{49}$$

c) Kapur's entropy

$$K_{\alpha,\beta}=\frac{1}{\beta-\alpha}\log_2\frac{\sum_{i=0}^{N_l-1}H_i^{\alpha}}{\sum_{i=0}^{N_l-1}H_i^{\beta}},\alpha\neq\beta,\alpha>0,\beta>0 \tag{50}$$

d) Renyi's entropy

$$R=\frac{1}{1-\alpha}\log_2\sum_{i=0}^{N_l-1}H_i^{\alpha}\text{ , where }\alpha\neq 1,\alpha>0 \tag{51}$$

e) Yager's measure

$$Y=1-\frac{\sum_{i=0}^{N_l-1}|2H_i-1|}{|M\times N|} \tag{52}$$

4) Histogram based first order statistical textural features

From the leukocytes we have computed five types of first order statistical feature viz. mean, variance, skewness, kurtosis and energy [30].

5) GLCM based textural features

Firstly we have calculate Gray level Co-occurrence matrix (GLCM) from segmented leukocyte nucleus which is square matrix of $N\times N$, where N is the number of gray levels. Haralick described statistical features using GLCM that can describe the texture of the leukocyte. Some features like Correlation, Contrast, Homogeneity, and Energy are calculated [31].

B. Distance measure

1) Weighted-Mean-Variance (WMV)

Consider two images patterns u and v and let $F^{(u)}$ and $F^{(v)}$ represent the corresponding feature vectors. Then the distance between the two patterns in the feature space is defined to be,

$$d(u,v)=\sum_l\sum_k d_{lk}(u,v) \tag{53}$$

Where, $d_{lk}(u,v)=\left|\frac{\mu_{lk}^{(u)}-\mu_{lk}^{(v)}}{\alpha(\mu_{lk})}\right|+\left|\frac{\sigma_{lk}^{(u)}-\sigma_{lk}^{(v)}}{\alpha(\sigma_{lk})}\right|$

$\alpha(\mu_{lk})$ and $\alpha(\sigma_{lk})$ are the standard deviations of the respective features over the entire database [36].

2) Euclidean distance

All features combination is compared with Euclidean Distance measurement [33]. It is also called L2 distance. If $u=(x_1,x_2,....,x_n)$ and $v=(y_1,y_2,....,y_n)$ are n dimensions feature vector for query image and database image respectively. Euclidean Distance between u and v is given by

$$E(u,v)=\sqrt{\sum_{i=1}^{n}(x_i-y_i)^2} \tag{54}$$

VI. Results and discussion

In this paper we have described the color correction method for microscopic blood smear image and explained the need for automatic illumination correction technique. Then we discuss about three color correction method namely gray world, retinex and quadratic mapping which is widely used for this task. In figure 2 we have shown color corrected microscopic blood smear image. During the preprocessing study three different highly experienced medical experts of Midnapur Medical College & Hospital, are assigned and simultaneously appointed for the evaluation of the color corrected images. On the basis of visual assessment of the experts we have chosen quadratic mapping technique for color correction as a best method.

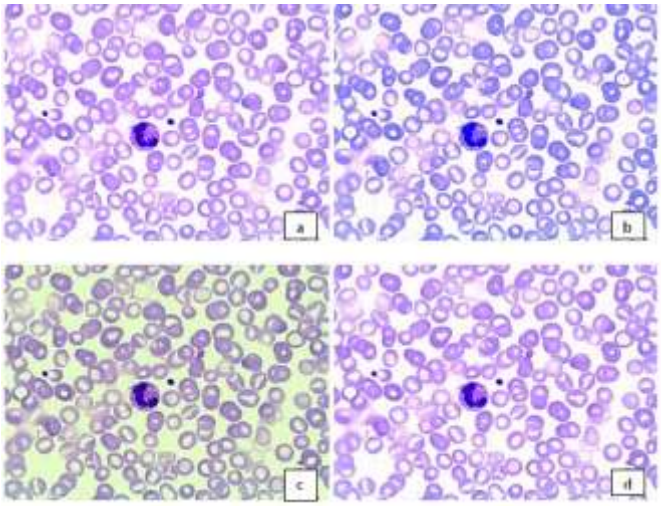

Figure 2 a) original image b) quadratic mapping image c) gray world image d) retinex image

After color correction we have applied five filtering technique for noise removable. Lower value of MSE, RMSE, NAE, AD indicates the fact that there is very small change between

original and filtered image. In other case higher value of SNR, PSNR reports that there is very small change. The value of UIQI closer to 1 indicates the best possible image quality as well as intact structural similarity. In figure 3 we have shown enhanced microscopic blood smear image. On the basis of visual assessment of the experts and quantitative performance measure we have chosen median filter for noise removal as a best method.

Table 1: Quantitative performance measure of **Local first order statistics filter**

	MSE	RMSE	NAE	AD	PSNR	SNR	UIQI
R	3.273	1.793	0.004	0.444	43.132	8.938	0.999
G	2519.03	48.610	0.195	-38.541	14.731	-0.016	0.943
B	6.744	2.586	0.006	0.643	39.913	0	0.999

Table 2: Quantitative performance measure of **Mid-point filter**

	MSE	RMSE	NAE	AD	PSNR	SNR	UIQI
R	189.335	13.484	0.029	0.502	25.718	-0.0004	0.999
G	260.479	16.067	0.046	0.137	24.051	-0.002	0.997
B	82.874	9.088	0.017	0.645	28.976	0	0.999

Table 3: Quantitative performance measure of **Laplacian filter**

	MSE	RMSE	NAE	AD	PSNR	SNR	UIQI
R	363.838	18.708	0.038	5.013	22.869	0.012	0.999
G	473.243	21.609	0.067	1.923	21.496	0.023	0.999
B	221.086	14.754	0.025	5.602	24.820	0.0004	0.999

Table4: Quantitative performance measure of **Alpha trimmed mean filter**

	MSE	RMSE	NAE	AD	PSNR	SNR	UIQI
R	6872.024	82.853	0.342	80.863	9.769	0.001	0.845
G	5204.548	72.139	0.347	68.433	10.967	0.002	0.839
B	7417.822	86.126	0.339	85.148	9.428	7.633	0.848

Table 5: Quantitative performance measure of **Median Filter**

	MSE	RMSE	NAE	AD	PSNR	SNR	UIQI
R	26.165	5.096	0.011	-0.572	34.019	-0.000	0.999
G	24.801	4.953	0.014	-0.325	34.278	-0.0001	0.999
B	26.102	5.083	0.009	-0.936	34.051	3.674E-05	0.999

From the table 1 it has been observed that local first order statistics filter has low MSE in R and B color channel (3.273 and 6.744) but in case G color channel the MSE value is high (2519.031). RMSE value is low in R and B channel (1.793 2.586 respectively) but in case G channel is high (48.610). In case of all color channels NAE and AD value is low for that filter. But PSNR value is high in case of all channels. SNR value is low for G and B channels but R channel SNR value is high. UIQI value is closer to 1. So this filter is not so good for this kind of medical application of image.

From the table 2 it has been observed that Midpoint filter has high MSE value for RGB color channel (189.335 260.479 and 82.874 respectively). RMSE, NAE and AD value is low for RGB color channel. PSNR value is higher for RGB color channel but SNR value is not high. UIQI value is closer to 1.

From the table 3 it has been observed that Laplacian filter has high MSE value for RGB color channel (363.838 473.243 and 221.086 respectively). RMSE and AD value is high for RGB color channel. But NAE value is low. PSNR value is high for RGB color channel but SNR value is low. UIQI value is closer to 1.

From the table 4 it has been observed that Alpha trimmed mean filter has higher MSE, RMSE and AD value for RGB color channel separately. PSNR and SNR value not so high. UIQI value is not closer to 1. So this filter is not good for medical application image.

From the table 5 if has been observed that median filter has lowest MSE value for RGB channels (26.165, 24.801, 26.102 respectively) which reflects the fact that the filtered image consist a very little amount of information loss from the input image. This filter has lowest RMSE, NAE and AD value. In this filter PSNR value is high but SNR value is low. UIQI value is closer to 1 for RGB color channel. Finally based on expert assessments and visual interpretations, median filter is found to be the best choice among all filters. During this assessment of choosing the proper filter, the filtered and original images are shown to medical experts. They separately assess each filter on the basis of their visual perspective and experience and provide their valuable reports. From the reports, it is observed that median filter and midpoint filter are mostly preferred.

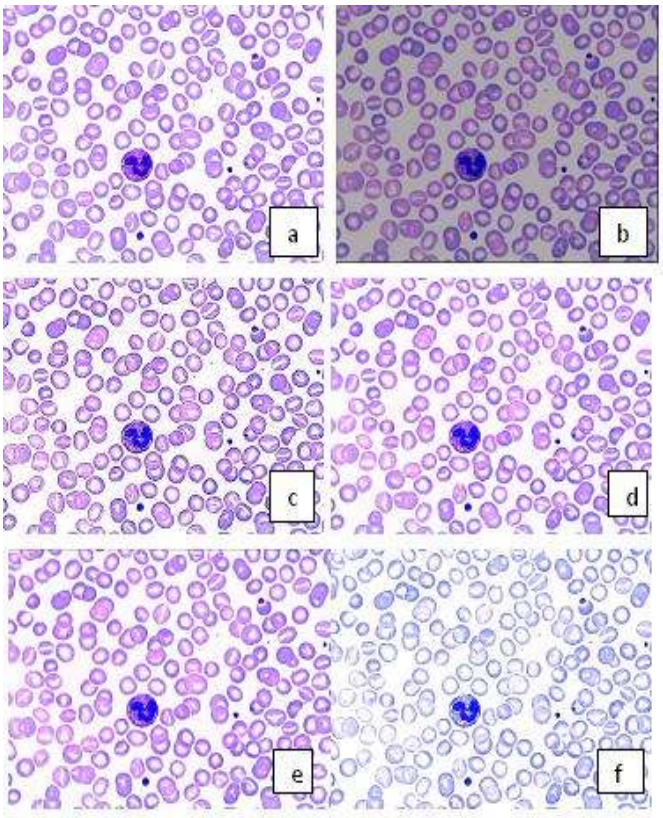

Figure 3. a) Original image b) Alpha trimmed mean filter c) Laplacian filter d) Median filter e) Midpoint filter f) Local first order statistics

Here, characterization of leukocytes images has been performed with 50 normal and 50 CML leukocytes. Leukocytes are segmented using QFT and Gabor wavelet based textual features are extracted from those segmented leukocyte nucleus. Ground truth image were prepared by expert editing using Adobe Photoshop 7.0 version software. We can see from figure 4 (b) and (e) the reconstructed image after quaternion Fourier transform where we can detect the leukocyte nuclei as well as abnormal nuclei (CML).

As shown in the reconstructed images, the high frequency content in the blood smear images which is normal leukocyte or abnormal leukocyte in the blood smear image normal or abnormal leukocyte areas are the high frequency content in the image. Then these reconstructed images converted into binary image and remove all small area connected component. To segment leukocyte accurately we use disk-shaped structuring element with radius 1 to perform dilation and erosion operation. In figure 4 (c) and (f) have shown the final segmented images.

In the figure 5 shows that the original image, ground truth image and other four segmented images. The segmentation performance of our propose technique is evaluated by determine the percentage of accuracy [16], as described by

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \times 100\% \qquad (55)$$

Where, TN, TN, FP and FN are true positive, true negative, false positive and false negative. The percentage of accuracy as in (55) is calculated based on comparison of the pixels that represent resultant segmented image and the pixels represent ground truth image.

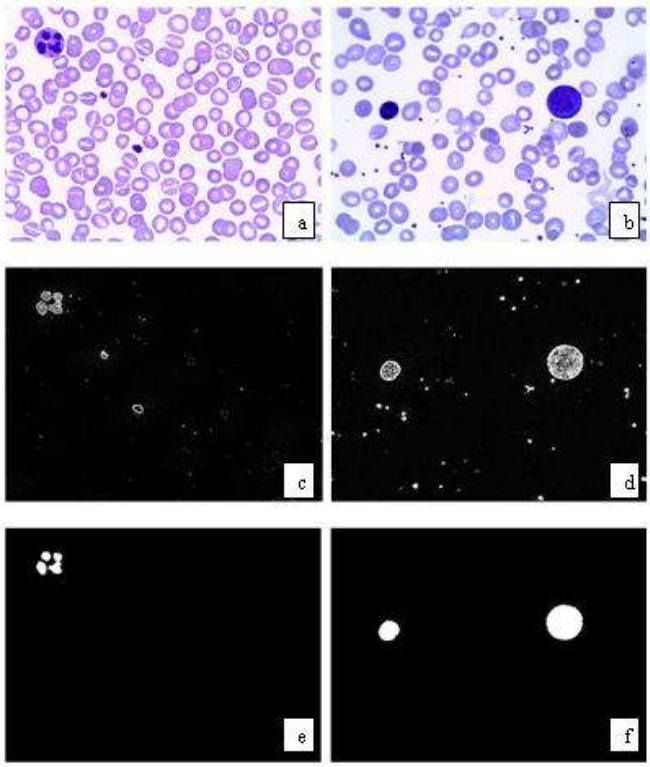

Figure 4. a)-b) Original Normal image and Original CML image c)-d) corresponding QFT based reconstructed images e)-f) After morphological operation segmented image

Table 6. Comparison based on average Accuracy based on 20 normal ground truth images

Methods	**Accuracy**
Quaternion based Segmentation	92.13%
Entropy based Segmentation	86.84%
Histogram based Segmentation	90.55%
Otsu's thresholding	44.33%

In table 6 shows the segmentation performance using Quaternion Fourier transform based, Entropy based, Histogram based and Otsu's thresholding [21]. The results indicate that, segmentation using Quaternion Fourier Transform is produced slightly better performance with the average accuracy of 92.13% based on 30 microscopic blood smear images. Comparing with other method Otsu thresholding [21], histogram based segmentation [32], and entropy based segmentation [40] produced the accuracy are 44.33%, 90.55% and 86.84% respectively. In the automatic threshold selection technique we observed many draw back. Here in QFT based segmentation technique no need to depend on threshold selection. In this QFT based segmentation technique is automatic and unsupervised.

After segmentation of leukocyte nucleus, we extract textural feature from gray-scale leukocyte nucleus image. In case of the Gabor-wavelet feature 48 component feature vectors are

used. The comparisons are made with Daubechies wavelet [28], Entropy measures including Shannon, Havarda and Charvat, Kapur's entropy, Renyi's and yeager's measure [29], histogram based first order statistical textural features viz. mean, variance, skewness, kurtosis and energy [30] and GLCM based textural feature viz. Correlation, Contrast, Homogeneity, and Energy [31] are used here. The Weighted-Mean-Variance distance measure is used for Gabor-wavelet textural feature similarity. For other textural feature Euclidean distance is measured. The performance of a retrieval system can be measured in terms of its precision [41]. It is defined as,

$$precision = \frac{A}{A+B} \tag{56}$$

Where A represents the number of relevant records retrieved, B represents the no. of irrelevant records retrieved. Here average precision is computed for query image.

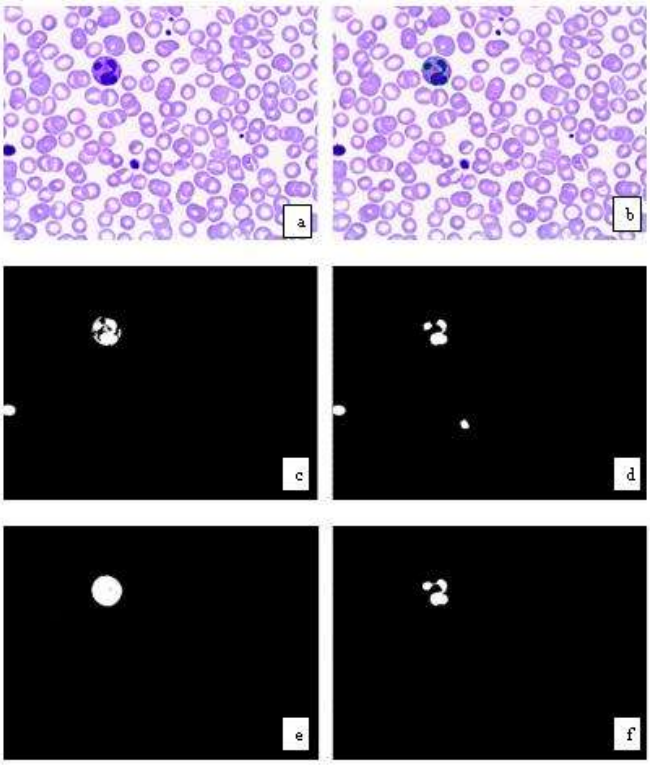

Figure 5. a) Original Normal image b) Ground truth image c) Entropy based image d) Histogram based image e) Otsu's thresholding image f) Quaternion based image

Distance measure is calculated between the query image's features and the same of the database's image features vector. Hence we get an array of distance values. In the database there are two type of images; healthy leukocytes and CML. Then distances are sorted in increasing order and top 10 images are retrieved from the whole database. It is observed from table 7 that Gabor-wavelet, wavelet, Entropy, GLCM and Histogram based textural descriptors [35] provides 82.5 %, 71.418%, 71.267%, 74.296%, and 78.526% average precision for normal query image respectively and from table 8 that textual descriptor provides 73.423 %, 64.988%, 69.825%, 67.135%, and 76.532% average precision for CML query image respectively. In the table 7 Gabor wavelet feature gives the best precision average for nor query image but in table 8 CML query image histogram based feature perform better than Gabor feature. The experimental results indicate that Gabor-wavelet feature set are quite robust for whole microscopic blood smear image database.

Table 7. Computation of precision (%) for corresponding Normal query image

Query image	*Gabor*-wavelet	Wavelet	Entropy	GLCM	Histogram
Q1	83.333	83.333	71.428	76.923	100
Q2	83.333	66.666	66.666	71.428	76.923
Q3	76.923	58.823	71.428	71.428	76.923
Q4	71.428	62.5	55.555	66.666	58.823
Q5	58.823	66.666	71.428	62.5	66.666
Q6	76.923	83.333	62.5	71.428	76.923
Q7	83.333	71.428	62.5	71.428	83.333
Q8	90.909	71.428	83.333	76.923	83.333
Q9	100	83.333	76.923	83.333	90.909
Q10	100	66.666	90.909	90.909	71.428
Avg.	**82.5**	**71.418**	**71.267**	**74.296**	**78.526**

Table 8. Computation of precision (%) for corresponding CML query image

Query image	*Gabor*-wavelet	Wavelet	Entropy	GLCM	Histogram
Q1	90.909	71.428	62.5	71.428	66.666
Q2	76.923	62.5	58.823	76.923	76.923
Q3	55.555	62.5	76.923	62.5	55.555
Q4	76.923	76.923	58.823	58.823	83.333
Q5	83.333	71.428	90.909	66.666	90.909
Q6	71.428	58.823	62.5	71.428	83.333
Q7	62.5	55.555	71.428	58.823	83.333
Q8	83.333	52.631	62.5	66.666	76.923
Q9	66.666	66.666	76.923	71.428	71.428
Q10	66.666	71.428	76.923	66.666	76.923
Avg.	**73.423**	**64.988**	**69.825**	**67.135**	**76.532**

VII. Conclusion

CBIR system can be used in disease diagnosis, education and research purpose. In case of medical diagnosis scenario the usefulness of the CBIR system lies in the accuracy of proper image segmentation technique. If the precision is high then it is obvious that the CBIR system is power full of diagnosis of a disease. Our proposed methodology is automatic, unsupervised, and quite robust which provided better retrieval sensitivity. In future study we are willing to extend this study for other leukemia disease.

Acknowledgment

Authors acknowledge Dept. of Information Technology, Govt. of India. (Ref. No. - IIT/SRIC/SMST/DPR/2009-2010/15) for financial support to carry out this work. Authors also acknowledge BMI lab of School of Medical Science and Technology and Dr. A K Maity, Dr. M Pal of Midnapur Medical College & Hospital, West Bengal for their guidance, clinical support and pathological knowledge.

References

[1] Muller H., Michoux N., Bandon D., and Geissbuhler A., "A review of Content-based Image Retrieval Systems in

Medical Application- Clinical Bennefits and Future Directions," *International Journal of Medical Informatics*, 2004, 73(1), pp. 1-23.

[2] Blaser A., Database Techniques for Pictorial Applications, *Lecture Notes in Computer Science*, Vol.81, Springer Verlag GmbH, 1979.

[3] P.G.B. Enser, Pictorial information retrieval, *J. Document*. 51 (2) (1995) 126—170.

[4] M. Flickner, H. Sawhney, W. Niblack, J. Ashley, Q. Huang, B. Dom, M. Gorkani, J. Hafner, D. Lee, D. Petkovic, D. Steele, P. Yanker, *Query by image and video content: the QBIC system*, IEEE Computer. 28 (9) (1995) 23—32.

[5] C. Carson, M. Thomas, S. Belongie, J.M. Hellerstein, J. Malik, Blobworld: a system for region-based image indexing and retrieval, in: D.P. Huijsmans, A.W.M. Smeulders (Eds.), *Proceedings of the Third International Conference On Visual Information Systems (VISUAL'99),* no. 1614 in Lecture Notes in Computer Science, Springer-Verlag, Amsterdam, The Netherlands, 1999, pp. 509—516.

[6] I.J. Cox, M.L. Miller, S.M. Omohundro, P.N. Yianilos, Target testing and the PicHunter Bayesian multimedia retrieval system, in: *Advances in Digital Libraries (ADL'96)*, Library of Congress, Washington, D.C., 1996, pp. 66—75.

[7] J.R. Bach, C. Fuller, A. Gupta, A. Hampapur, B. Horowitz, R. Humphrey, R. Jain, C.-F. Shu, The Virage image search engine: an open framework for image management, in: I.K. Sethi, R.C. Jain (Eds.), *Proceedings of the SPIE Conference on Storage & Retrieval for Image and Video Databases* vol. 2670, San Jose, CA, USA, 1996, pp. 76—87.

[8] Pentland A., Picard W. R., Sclaroff S., Photobook: tools for content-based manipulation of image databases, *Int. J. Computer. Vision*, 18(3) (1996) 233-254.

[9] Korn P, Sidiropoulos N, Faloutsos C, Siegel E, Protopapas Z: "Fast and effective retrieval of medical tumor shapes". *IEEE Trans KDE*, 10(6): 889–904, 1998.

[10] Shyu CR, Brodley CE, Kak AC, Kosaka A, Aisen AM, Broderick LS: "ASSERT – A physician-in-the-loop content based retrieval system for HRCT image databases". *Computer Vision and Image Understanding*; 75(1/2): 111– 132, 1999.

[11] Lehmann T. M., Guld M. O., Thies C., Plodowski B., Keysers D., Ott B., Schubert H., " IRMA - Content-Based Image Retrieval in Medical Applications", *Medinfo2004;* 2004*:* 842*–846.*

[12] Nazari R. M., Fatemizadeh E., "A CBIR System for Human Brain Magnetic Resonance Image Indexing", *International Journal of Computer Applications* (0975 – 8887) Volume 7– No.14, October 2010.

[13] Chen K., Lin J., Zou Y., Yin G., "Content-based Medical Ultrasound Image Retrieval Using a Hierarchical Method", *Proceedings of IEEE 2nd International Congress on Image and Signal Processing*, pp. 1–4 (2009)

[14] Dimitrovski I., Loskovska S., "Content-Based Retrieval System for X-Ray Images", *2nd International Congress on Image and Signal Processing*, 2009, CISP '09, DOI: 10.1109/CISP.2009.5304151.

[15] Hao Wi, Jinmall Kim, Weidong Cai, David Dagan Feng, "Volume of Interest (VOl) Feature Representation and Retrieval of Multi-Dimensional Dynamic Positron Emission Tomography Images", *Proceedings of 2004 International Symposium on Intelligent Multimedia, Video and Speech Processing* October 20-22, 2004 Hong Kong

[16] Woo Chaw Seng, Seyed Hadi Mirisaee, "Evaluation of a Content-Based Retrieval System for Blood Cell Images with Automated Methods", *Journal Medical System* 35:571-578, DOI 10.1007/s10916-009-9393-3

[17] Ruey-Feng Chang, Chii-Jen Chen, Chen- hao liao, "Region-based Image Retrieval Using Edgeflow Segmentation and Region Adjacency Graph" *Proceedings of the IEEE Internation Conference on Multimedia and Expo (ICME*), 2004.

[18] Yinan lu, yong Quan, Zhenhua Zhang, Gang Wang, "MST Segmentation for Content-Based Medical Image Retrieval", *Proceedings of the International Conference on Computational Intelligence and Software Engineering*, CISE 2009.

[19] Cheng-Chieh Chiang, Yi-Ping Hung, Hsuan Yang, Greg C. Lee, "Region-based image retrieval using color-size features of watershed regions", *Journal of Visual Commun. Image* R.20 (2009) 167-177, DOI: 10.1016/j.jvcir.2009.01.001

[20] Rajendran S., Arof H., Ibrahim F., Yegappan S., "Review on Technical Aspects of Image Acquisition, Analysis and Retrieval for Leukemia Cells", *Proceedings of the IFMBE*, 2008, Volume 21, Part 3, Part 4, 230-233, DOI: 10.1007/978-3-540-69139-6_60

[21] Hamghalam M., Ayatollahi A., "Automatic Counting of Leukocytes in Giemsa-Stained Images of Peripheral Blood Smear", *Proceedings of the International Conference on Digital Image Processing* IEEE 2009, DOI: 10.1109/ICDIP.2009.9

[22] Rezatofighi H. S., Soltanian-Zadeh H., Sharifian R., Zoroofi A. R., "A New Approach to white Blood Cell Nucleus Segmentation Based on Gram-Schmidt Orthogonalization", *Proceedings of the International Conference on Digital Image Processing* IEEE 2009, DOI: 10.1109/ICDIP.2009.19

[23] Ghosh M., Das D., Chakraborty C., Ray K. A., "Automated leukocyte recognition using fuzzy divergence" *Micron* 41 (2010) 840-846, DOI: 10.1016/j.micron.2010.04.017

[24] Hamilton R. W., Elements of quaternions. London, U.K.: *Longmans Green*, 1866.

[25] Todd A. Ell, Stephen J. Sangwine, "Hypercomplex Fourier Transforms of color Images", *IEEE Trans. on Image Processing*, VOL. 16, NO. 1, Jan 2007, DOI: 10.1109/TIP.2006.884955

[26] Fang Y., Xiong W., Lin W., Chen Z., "Unsupervised Malaria Parasite Detection Based on Phase Spectrum", *Proceedings of the 33rd Annual Internation Conference of IEEE EMBS*, Boston, USA, 2011

[27] Manjunath S. B. and Ma Y. W., "Texture features for browsing and retrieval of image data", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 18 (8), August 1996, pp. 837-842

[28] Krishnan R. M. M., Chakraborty C., and Ray K. A., "Wavelet based texture classification of oral histopathological sections", *Microscopy: Science, Technology, Applications and Education*, FORMATEX 2010

[29] Pharwaha S. P. A., Singh B., "Shannon and Shannon Measures of Entropy for Statistical Texture Feature Extraction in Digitized Mammograms", *Proceedings of the World Congress on Engineering and Computer Science* 2009 Vol II, WCECS 2009, ISBN: 978-988-18210-2-7.

[30] A. Materka, M. Strzelecki, Texture Analysis Methods A Review, Technical University of Lodz, Institute of Electronics, *COST B11 report*, Brussels 1998. 350

[31] Das K. D., Ghosh M., Chakraborty C., Maity K. A., Pal M., "Probalistic Prediction of Malaria using Morphological and Textual Information", *Proceedings of the 2011 International Conference on Image Information Processing* (ICIIP 2011), 978-1-61284-861-7/11, IEEE-2011.

[32] Gonzalez C. R., Woods E. R., Digital image processing, 3rd Edn. *Prentice Hall*, New York, 2008.

[33] Vadivel A., Majumdar K. A., Sural S., "Performance comparison of distance metrics in content-based image retrieval applications", *Proc. of International Conf. on Information Technology*, Bhubaneswar, India, pp. 159-164.

[34] Lam Y. E., "Combining Gray World and Retinex Theory for Automatic White Balance in Digital Photography" *Proceedings of the Ninth International Symposium on Consumer Electronics*, 2005 (ISCE 2005). DOI: 10.1109/ISCE.2005.1502356.

[35] Sarkar P., Chakraborty C., Ghosh M., "Content-Based Leukocyte Image Retrieval Ensembling Quaternion Fourier Transform and Gabor-wavelet Features", *Proc. of 12th International Conference on Intelligent Systems Design and Applications* (ISDA), 2012. DOI: 10.1109/ISDA.2012.6416562.

[36] Rubner Y., Puzicha J., Carlo T., Buhmann M. J., "Empirical Evaluation of Dissimilarity Measures for Color and Texture" *Computer Vision and Image Understanding* 84, 25-43 (2001)DOI: 10.1006/cviu.2001.0934.

[37] Zhang D., Wong A., Indrawan M., Lu G., "Content –based Image Retrieval Using Gabor Texture Features", *Proc. Pacific-Rim Conference on Multimedia*, PP. 392-399, Sydney, Australia, Dec, 2000.

[38] Sheet, D., Part, S., Chakraborty, A., Chatterjee, J., & Ray, A. K. (2010). Image Quality Assessment for Performance Evaluation of Despeckle Filters in Optical Coherence Tomography of Human Skin, *IEEE EMBS Conference on Biomedical Engineering & Sciences* (IECBES 2010), Kuala Lumpur, Malaysia, 30th November, 499–504. 2010.

[39] Poobal S., Ravindran G., "The Performance of Fractal Image Compression on Different Imaging Modalities using Objective Quality Measures*", International Journal of Engineering Science and Technology (IJEST),* ISSN : 0975-5462, Vol. 3 No. 1 Jan 2011.

[40] Ghosh M., Das D., Chakraborty c., " Entropy based divergence for leukocyte image Segmentation", *Proceedings of 2010 International Conference on Systems in Medicine and Biology* 16-18 December 2010, IIT Kharagpur, India

[41] Singh S. M., Hemachandran K., "Content-Based Image Retrieval using Color Moment and Gabor Texture Feature" *IJCSI International Journal of Computer Science* Issues, Vol. 9, Issue 5, No 1, September 2012, ISSN (Online): 1694-0814

[42] Kasmin F., Prabuwono S. A., Abdullah A., "Detection of leukemia in human blood sample based on microscopic images: A study" *Journal of Theoretical and Applied Information Technology*, Vol. 46, No 2, December 2012, ISSN (Online): 1992-8645

[43] Comaniciu D., Meer P., Foran D., Medl A., "Bimodal system for interactive indexing and retrieval of pathology images", *Proceeding of Fourth IEEE work shop on Applications of computer vision (WACV'98)*, DOI: 10.1109/ACV.1998.732861

[44] Olivieri N. D., Vega F., "Image prototype similarity matching for lymph node hemopathology", *Proceeding of 15th international conference on pattern recognition*, Vol *2, September* 2000, DOI: 10.1109/ICPR.2000.906067.

[45] Scotti F., "Automatic morphological analysis for acute leukemia identification in peripheral blood microscope images", *Proceeding of IEEE International conference on Computational Intelligence for measure systems and applications CIMSA 2005*, VOI: 10.1109/CIMSA.2005.1522835

Author Biographies

Prabir Sarkar received his B-Tech in Information Technology from WBUT, West Bengal, India, in 2009. He is currently an MS student at the School of Medical Science and Technology, Indian institute of Technology, Kharagpur, India. His research interests include pathological image analysis and computer vision.

Madhumala Ghosh received her B-Tech in Electronics and Telecommunication Engineering from WBUT, West Bengal, India in the year 2005. Then she received her M-Tech in Intelligent Automation and Robotics from Jadavpur University, West Bengal, India in 2008. She completed her PhD from School of Medical Science and Technology, IIT Kharagpur in the field of pathological image processing.

Chandan Chakraborty is currently an Assistant Professor of School of Medical Science and Technology, IIT Kharagpur. He completed his BSc and MSc and PhD in Applied Statistics. He received young scientist award by President of India for his work in 2007. He is an IEEE member. He has authored over 80 research papers in refereed journals, books, and conference proceedings. His current research interests including pattern recognition for medical imaging, biostatistics, and medical image analytics and informatics.