Improved Technique for Detection of Malaria Parasites within the Blood Cell Images

Ms. Snehal Suryawanshi, Prof. V. V. Dixit

Abstract—Though there is considerable progress still there is need to improve accuracy, speed, automation level, adaptablilty towards new applications. Hence this paper proposes a new technique of image segmentation by Poisson distribution using minimum error thresholding. Then the seed points are detected by a novel method combining multiscale laplacian of Gaussian with gabor filtering. Finally the features extracted are compared with database to check whether the blood cell images are infected by malaria parasites.

Index Terms—Image thresholding, Poisson distribution, multiscale LoG, Gabor filtering, Malaria diagnosis, Artificial Neural Network, Back propagation.

1 INTRODUCTION

Malaria is avoidable and curable, but it is still one of the most harmful diseases in developing countries. Although countries across Africa, where the incidence of malaria is the highest, have scaled up malaria control strategies, effective control and treatment present enormous logistical difficulties, as many at-risk populations live in extreme poverty in remote rural areas [1].

A number of methods have been proposed for automatic parasite detection in Giemsa stained blood films based on different approaches. These approaches include pixel-based parasite detection [6], detection based on morphological processing of segmented parasites [2, 3], or detection by extracting image features from the segmented cells [4].

2 RELATED WORK

Ross al [4], in his work proposed an image processing technique is described that is used to identify erythrocytes and possible parasites present on microscopic slides. The algorithm consists of pre-processing of the image, image analysis, segmentation, features generation and classification of erythrocytes as infected with malaria or not.

Di Ruberto et al [3], presents a system, where objects have been detected by means of an automatic thresholding on single components of the RGB and HSV histogram based on a morphological approach. The paper describes morphological methods for both cell image segmentation and parasites detection. The planned method uses Granulometries to evaluate the size of the red cells and the nuclei of parasites and regional maxima to detect the nuclei of parasites. Morphological techniques, like thinning, gradient, reconstruction by dilation, and morphological filters are also utilized for pre or post-processing of the images.

Diaz et al [6], evaluates a color segmentation techniques for separation of pixels into three different classes: parasites, red blood cell and background, based on standard supervised classification techniques. In this sytem four different supervised classification algorithms– KNN, Naive Bayes, SVM and Neural network are evaluated on different color spaces – RGB, normalized RGB, HSV and YC_bC_r.

3 OBJECTVE

To develop an automated tool to analyze malaria parasite within the blood cell images.

4 SYSTEM ARCHITECTURE

The system flow chart is as shown below,

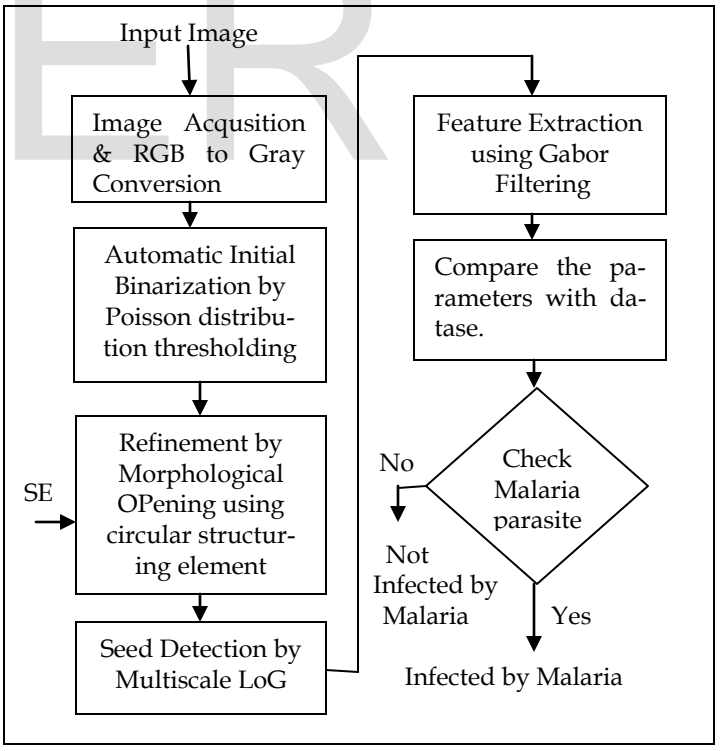


Fig. 1 Flow chart, outlining the main steps of the proposed system.

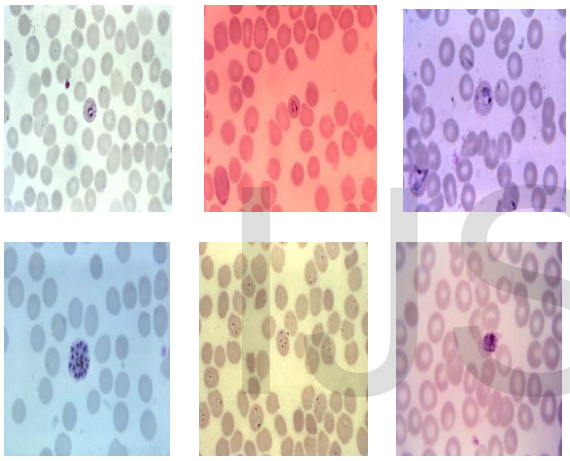
4.1 Image Acquisition

Images of Giemsa stained blood smears were selected from the Public Health Image Library [5]. Images are available in different magnifications and sizes. Giemsa stain is used to differ-

International Journal of Scientific & Engineering Research Volume 4, Issue 7, May-2013
ISSN 2229-5518

entiate nuclear and cytoplasmatic morphology of platelets, RBCs (red blood cells), WBCs (white blood cells) and parasites. Giemsa staining solution stains up nucleic acids and, therefore, parasites, white blood cells, and platelets, which contain DNA, are highlighted in a dark purple color. RBC are usually colored in slight pink colors.

- The images are available in TIFF format with the resolution of 2 to 3 megapixels.
- By scanning, we obtained the digital images and, these images contain a part of the noise and artifact from the sample and from the microscope light also noise from the chemical development process or from the scanner.
- Images exhibit high variability in color tone, intensity, contrast, and illumination. The overall color tone varies significantly from grayish, blue, purple, and pink to yellowish and it may even change from the center of the image to its borders (Fig. 2.).



Fig, 2. Samples of available stained blood smear images showing differences in color tone and illumination. Image courtesy of CDC/ Dr. Mae Melvin, Steven Glenn [5]

4.2 Image pre-processing

The pre-processing stage is designed to remove unwanted effects from the image and to adjust the image as necessary for further processing

4.2.1 RGB to Gray Conversion

The microscopic image which is being acquired is converted from RGB to gray scale in order to reduce the processing time thrice. RGB to gray conversion is done by averaging all the three components i.e. R (red), G (green) and B (blue) which results in gray scale,

$$\text{Gray Scale (G S)} = (R+G+B)/3 \quad (1)$$

4.3 Cell Segmentation

4.3.1 Poisson Distribution thersholding

To segment foreground from background Poisson distribution based on minimum error thresholding algorithm is used on gray scale enhanced image [11], [12]. The Poisson distribution is a discrete probability distribution for the counts of events that occur randomly in a given interval of time (or space). Whereas Minimum error thresholding method finds the optimum threshold by optimizing the average pixel classification error rate directly, using either exhaustive search or an iterative algorithm [10].

Here by using Poisson distribution minimum error thresholding we get a bimodal histogram, which gives us threshold value and thus this separates foreground and background.

$$t^* = \arg\min\{\mu - P_0(t)(\ln P_0(t) + \mu_0(t)\ln \mu_0(t)) - P_1(t)(\ln P_1(t) + \mu_1(t)\ln \mu_1(t))\} \quad (2)$$

4.3.2 Opening

The results from above step are then morphologically operated to get the clear view of the parasite i.e. for the purpose of refinement, we have used erosion followed by dilation which is basically called as Opening process [9]. This smoothen the image with respect to foreground. The structuring element used is circular element, since region of interest is circular in shape.

4.3.3 Foreground Extraction

The output of morphological operation is subtracted from gray scall and then again divided from same. This forms a mask which is multiplied byoriginal image.

$$\text{Foreground Extraction} = ((255\text{-Output of Opening})/255) \times \text{Original image} \quad (3)$$

4.4 Feature Extraction

4.4.1 Multiscale LoG

In order to localize the seed point multiscale LoG filter is used, since it extracts connected cells from the cluster by assigning individual marker per cell in advance [7]. In this method, the Gaussian filtering is combined with Laplacian to break the image where intensitie varies to detect the edges effectively. The LoG filter is given by,

$$\text{LoG}(x,y,\sigma) = (\partial^2 G(x,y,\sigma))/\partial x^2 + (\partial^2 G(x,y,\sigma))/\partial y^2 \quad (4)$$

where σ is scale value, $G(x,y,\sigma)$ is a Gaussian with 0 mean and scale σ. Gaussian suppresses high frequency noise. Whereas Laplacian achieves maximum peak at $\sigma = r/\sqrt{2}$. The location of peak obtained by this technique are robust, which is one of the advantageous point. Also it provides information about boundaries of touching nuclei [7].

4.4.2 Gabor Filtering

Now, a set of Gabor filter with different frequencies and orientations may be helpful for extracting useful features from an image. Gabor filter continuously keeps monitoring intensity variations of localized portion. Hence it marks high intensity features. Gabor function in space domain is given as [15],

International Journal of Scientific & Engineering Research Volume 4, Issue 7, May-2013
ISSN 2229-5518

$$g(x, y) = s(x, y)\, w_r(x, y) \quad (5)$$

where s(x, y) is complex sinusoid, known as carrier, $w_r(x, y)$ is 2D Gaussian shaped function known as the envelop.

A Gabor filter is constructed by modulating a sine/cosine wave with a Gaussian. Decomposition of a signal is accomplished using a quadrature pair of Gabor filters, with a real part specified by a cosine modulated by a Gaussian, and an imaginary part specified by a sine modulated by a Gaussian.

The features can be obtained by using OR as well as AND logic. The first one will give better accuracy but more computations while the later one provides with reduced accuracy but lesser computations. Finally, we get intensity feature points

4.5 Comparison and Classification

4.5.1 Euclidean Distance Classifier

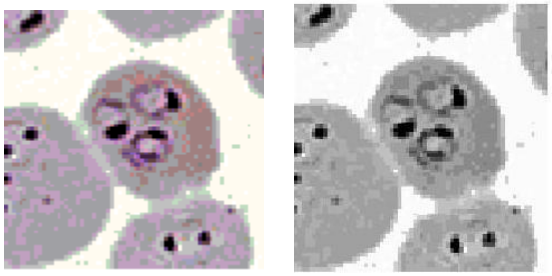
Linear Euclidean classifier classifies the intensity feature points linearly. Basically it is partition of space using a decision boundary into two separate decision regions. Features under the same label or region are contigious in nature. Here we use decision tree. It is important to note that tree boundaries are always linear and axis parallel. For separation we measure Euclidean distance between the two feature points, equation for the same in Cartesian co-ordinate system is given as,

$$D(x, y) = \sqrt{(y_1 - x_1)^2 + (y_2 - x_2)^2 + \cdots + (y_n - x_n)^2} \quad (6)$$

where, $x = (x_1, x_2, \ldots, x_n)$ and $y = (y_1, y_2, \ldots, y_n)$ are the two points in Euclidean space. It gives distance between points x and y.

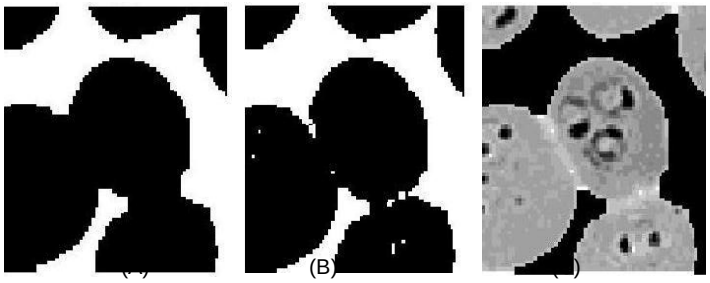
5 RESULTS AND DISCUSSION

After following the above mentioned algorithm each an every step gives output in the form of images. For this system human blood cell image is given as input, it is processed by applying various techniques, finally it is checked whether the human being is affected by malaria. The results are as shown below,



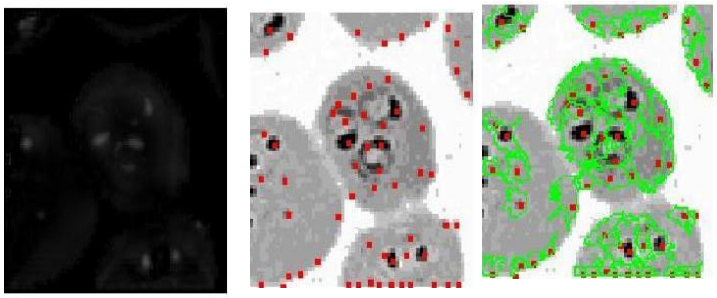
(A) (B)

Fig. 3 Steps of system design algorithm (A) Original image, (B) RGB to gray conversion.



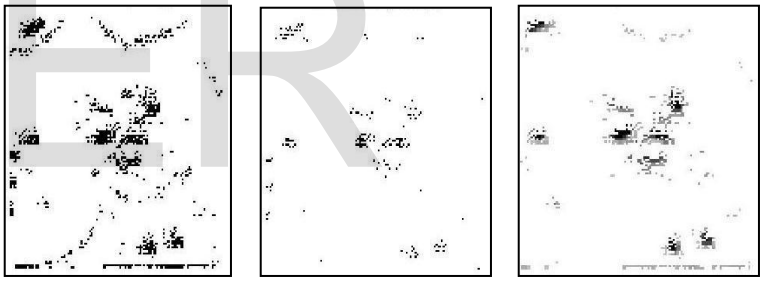
(A) (B) (C)

Fig. 4 Steps for Segmentation (A) Poisson distribution thresholded image, (B) morphological opening, (C) Extracted foreground.



(A) (B) (C)

Fig. 5 Steps in features localization by multiscale LoG (A) Refined LoG, (B) Seed Point Distribution, (C) Clustered Sections.



(A) (B) (C)

Fig. 6 Feature Extraction using Gabor Filter (A) Gabor Features (OR logic) (B) Gabor Features (AND Logic) (C) Intensity Feature Points.

6 CONCLUSION

The work presented in this paper is based on some recently extended techniques, which are robust, accurate and easy to implement. Poisson distribution based minimum error thresholding algorithm automatically binarizes image, which is refined by morphological opening and hence foreground is being extracted. Further work consists of localization of seed pointby firstly refinement, then seed point distribution and finally clustering the sections. Thus, these clusters are used to generate the intensity feature using Gabor filtering which basically uses AND Logic & OR Logic. These features create a decision boundary in order to differentiate whether affected by malaria disease. The system is applicable only for diagnosis treatment needs to done only as per the physician.

International Journal of Scientific & Engineering Research Volume 4, Issue 7, May-2013
ISSN 2229-5518

REFERENCES

[1] WHO Malaria Fact Sheet no. 94 Dec 2011. Available at: http://www.who.int/mediacentre/factsheets/fs094/en/

[2] W.-K. Chen, *Linear Networks and Systems.* Belmont, Calif.: Wadsworth, pp. 123-135, 1993. (Book style)

[3] H. Poor, "*A Hypertext History of Multiuser Dimensions,*" *MUD History,* http://www.ccs.neu.edu/home/pb/mud-history.html. 1986. (URL link *include year)

[4] K. Elissa, "*An Overview of Decision Theory,*" unpublished. (Unplublished manuscript)

[5] R. Nicole, "*The Last Word on Decision Theory,*" *J. Computer Vision,* submitted for publication. (Pending publication)

[6] C. J. Kaufman, Rocky Mountain Research Laboratories, Boulder, Colo., personal communication, 1992. (Personal communication)

[7] D.S. Coming and O.G. Staadt, "*Velocity-Aligned Discrete Oriented Polytopes for Dynamic Collision Detection,*" *IEEE Trans. Visualization and Computer Graphics*, vol. 14, no. 1, pp. 1-12, Jan/Feb 2008, doi:10.1109/TVCG.2007.70405. (IEEE Transactions)

[8] S.P. Bingulac, "*On the Compatibility of Adaptive Controllers,*" *Proc. Fourth Ann. Allerton Conf. Circuits and Systems Theory*, pp. 8-16, 1994. (Conference proceedings)

[9] H. Goto, Y. Hasegawa, and M. Tanaka, "*Efficient Scheduling Focusing on the Duality of MPL Representation,*" *Proc. IEEE Symp. Computational Intelligence in Scheduling (SCIS '07)*, pp. 57-64, Apr. 2007, doi:10.1109/SCIS.2007.367670. (Conference proceedings)

[10] J. Williams, "*Narrow-Band Analyzer,*" PhD dissertation, Dept. of Electrical Eng., Harvard Univ., Cambridge, Mass., 1993. (Thesis or dissertation)

[11] E.E. Reber, R.L. Michell, and C.J. Carter, "*Oxygen Absorption in the Earth's Atmosphere,*" Technical Report TR-0200 (420-46)-3, Aerospace Corp., Los Angeles, Calif., Nov. 1988. (Technical report with report number)

[12] L. Hubert and P. Arabie, "*Comparing Partitions,*" *J. Classification,* vol. 2, no. 4, pp. 193-218, Apr. 1985. (Journal or magazine citation)

[13] R.J. Vidmar, "*On the Use of Atmospheric Plasmas as Electromagnetic Reflectors,*" *IEEE Trans. Plasma Science*, vol. 21, no. 3, pp. 876-880, available at http://www.halcyon.com/pub/journals/21ps03-vidmar, Aug. 1992. (URL for Transaction, journal, or magzine)

[14] J.M.P. Martinez, R.B. Llavori, M.J.A. Cabo, and T.B. Pedersen, "*Integrating Data Warehouses with Web Data: A Survey,*" *IEEE Trans. Knowledge and Data Eng.*, preprint, 21 Dec. 2007, doi: 10.1109/TKDE .2007.190746.(PrePrint)

[15] Tutorial on Gabor Filters, Copyright c 1996, 2002 Javier R. Movellan.

- *Ms. S. R. Suryawanshi is currently pursuing masters degree program in signal processing engineering, E&TC dept. of Sinhgad College of Engineering in Pune University, India, Mob. -9665013518. E-mail: snehal.09suryawanshi@gmail.com*
- *Prof. V. V. Dixit is currently working as Associate Professor, E&TC dept. of Sinhgad College of Engineering in Pune University, India, Mob. -9822777265. E-mail: vvdixit73@gmail.com*