

# Computer Aided Medical Diagnosis for the Identification of Malaria Parasites

S.F. Toha<sup>1</sup> and U.K. Ngah<sup>2</sup>

**Abstract:** This paper presents one of the applications of digital image processing in artificial intelligence particularly in the field of medical diagnosis system. Currently in Malaysia the traditional method for the identification of Malaria parasites requires a trained technologist to manually examine and detect the number of the parasites subsequently by reading the slides. This is a very time consuming process, causes operator fatigue and is prone to human errors and inconsistency. An automated system is therefore needed to complete as much work as possible for the identification of Malaria parasites. The integration both soft computing tools has been successfully designed with the capability to improve the quality of the image, analyze and classify the image as well as calculating the number of Malaria parasites.

## I. INTRODUCTION

Nowadays, as the computational power increases, the role of automatic visual inspection becomes more important. There are four identified species of this parasite causing human Malaria, where there are normally three main parasites usually found in Malaysia namely, *P. Falciparum*, *P. Vivax*, and *P. Malariae* [1]. It is a disease that can be treated in just 48 hours, yet it can cause fatal complications if the diagnosis and treatment are delayed. Image processing and artificial intelligence techniques are introduced that may provide a valuable tool for improving the manual screening of specimens. There are two types of blood available for the detection of Malaria parasites [1] that are thin blood smear, to find the species of Malaria parasites while thick blood smear is done to find the density of Malaria parasites per red blood cells (used in this study). There have been a number of relevant approaches in the literature using computational intelligence and microscopic images. Toha and Ngah [2] presented the use of fuzzy logic and two soft computing tools to identify the type and development stage of Malaria parasites using thin blood smear image. Ngah and Ho [3] developed a pi fuzzification algorithm to analyse the intensity value of micro calcification. They use a fuzzy c-mean clustering in order to classify the mammography micro calcification image into few image clusters

## II. RESEARCH APPROACH

The Thick Blood Smear Image analysis is focusing on how to count the number of Malaria parasites exist in each digitized red blood specimen image. The image analysis can be done in both *manual* and *automatic*. Figure 1 shows the overall process of the system.

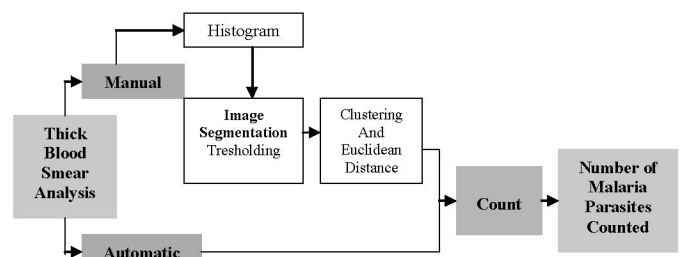


Fig. 1. Program Use-Case Diagram

### A. Histogram

The information about an image is known directly from the result of image histogram. In order to calculate histogram of an image, an array with the same size as total of grey level is developed with a starting value of 0. The image is then scanned pixel by pixel and the pixel's grey level is used as the array index. Refer to Equation (1) for histogram  $h_i$ .

$$h(i) = \frac{n(i)}{n} \quad (1)$$

$n(i)$  = Summation of the number of pixels within the same grayscale value,  $i$ .

$n$  = Summation of all the number of pixels in an image.

$i$  = 0 to N which represent image intensity.

### B. Image segmentation - threshold

Thresholding is a non-linear operation that converts a grayscale image into a binary image where the two levels are assigned to pixels that are below or above the specified threshold value [4]. The threshold value within 1 to 255 gained is used as an input for image segmentation. All the intensity value which lies below the threshold value will be set as 0 while for the intensity value above the input value will be set as 255. This will result to a segmented image where black color will represent the background of the digitized red blood specimens while the white color will characterize the Malaria parasites existed in the image. Equation (2) is used to calculate threshold value:

<sup>1</sup> Department of Mechatronics, Faculty of Engineering, International Islamic University Malaysia, Malaysia

<sup>2</sup> School of Electrical and Electronic Engineering, Universiti Sains Malaysia, Malaysia  
Email: tsfauziah@iium.edu.my, umik@eng.usm.my

$$t2 = (tb/nb + to/no/2.0) \quad (2)$$

- t2 = threshold value when the threshold value is not the same as the average value of grey level image.
- tb = sum of grey level value for overall pixel under the average value of grey level image.
- nb = the number of pixel under the average value of grey level image.
- to = sum of grey level value for overall pixel over the average value of grey level image.
- no = the number of pixel over the average value of grey level image.

C. The algorithm of counting the number of malaria parasites - image clustering and euclidean distance

Cluster analysis is used to identified the parasites exist since each and every parasite is different in term of size and locality. It will focus on detecting the white pixel with a grey level value of 255 with four basic cluster analysis steps:

1. Data collection and selection of the variables
2. Generation of a similarity matrix
3. Decision about number of clusters
4. Validation of cluster solution

The Euclidean Distance algorithm is used to measure the mathematical distance for the average distance value lies between pixels and thus the pixels are classified to their nearest cluster.

1. Calculate the mean value for each data set
2. Define the points of area with the same distance with the average value of data set.
3. All of the points defined determine the line of discrimination that will separate between clusters.
4. The distance between two points will be defined as Equation (3):

$$d2 = (x2 - x1)^2 + (y2 - y1)^2 \quad (3)$$

where d is the Euclidean distance, point p1 will be at (x1,y1) and point p2 will be at (x2,y2).

D. Counting the number of malaria parasites

The cluster algorithm used in detecting the Malaria parasites in C++ program is as Equation (4):

$$\text{If (USERDATA [ ] [ ] = 255)} \quad (4)$$

The algorithm is used to check each and every pixel until the pixel with 255 grey level value is found. The second pixel found will be compared to the first one using Euclidean distance. The Euclidean value calculated will determine whether the second pixel will be grouped with the first pixel or it will be classified as a member of another cluster. This process will continue until the entire pixel with 255 grey level value are scanned.

III. RESULT

The analysed image must be called up first and it will be appeared in the child window of the software as shown in Figure 2. The result from image thresholding will appear immediately as shown in Figure 3.

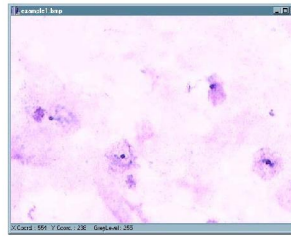


Fig. 2. Original Thick Blood Smear Image

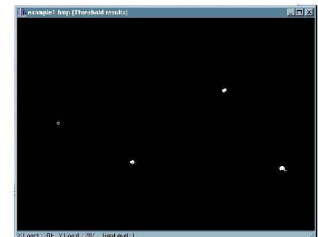


Fig. 3. Threshold Image Result

Figure 4 shows the result from thick blood smear image analysis.

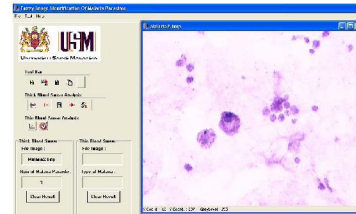


Fig. 4. Result of Counting the Number of Malaria Parasites Found in the Thick Blood Smear Image

IV. CONCLUSION

The integration of soft computing tools, in this project which are digital image processing and C++ programming language, has lead to the generation of more powerful, intelligent and efficient system. The system is designed with the capability to first improve the quality of the image, analyzed the image and thus classified the image whereby the end result is calculating the number of Malaria parasites inside the digitized thick blood smear image.

REFERENCES

- [1] Microbiology Syllabus, December 2005. Citing Internet sources URL
- [2] [http://www.ratsteachmicro.com/2003\\_HCOE\\_PDFs/2004\\_Malaria\\_syllabus.pdf](http://www.ratsteachmicro.com/2003_HCOE_PDFs/2004_Malaria_syllabus.pdf)
- [3] Toha S.F. and Ngah U.K. "Fuzzy Logic Based System for the Identification of Malaria Parasites from Human Blood Specimens", Proceedings for The Third International Conference on Artificial Engineering and Technology, Karambunai Sabah, 2006.
- [4] Ngah U.K. and Ho E.L. "Detection and Analysis of Mammography Moco Calcification", Technical Journal, School of Electrical and Electronics Engineering, USM, Volume 9, August 2003.
- [5] Haralick R.M., Shanmugam K. and Dinstein I., "Textural Features for Image Classification", IEEE Transactions on Systems, Man and Cybernetics. SMC vol. 3 no. 6 page 610-620. 1973.
- [6] Editors Chen CH and Wang PSP, Handbook of Pattern Recognition and Computer Vision, Third Edition, 2005.