Pertanika J. Sci. & Technol. 25 (S): 275 - 282 (2017)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Tuberculosis Bacteria Counting Using Watershed Segmentation Technique

M. K. Hassan[1,2]*, N. F. H. Eko[1] and S. Shafie[1,2]

[1]*Department of Electrical and Electronics, Faculty of Engineering, Universiti Putra Malaysia, 43400 UPM, Serdang, Selangor, Malaysia*
[2]*Institute of Advanced Technology, Universiti Putra Malaysia, 43400 UPM, Serdang, Selangor, Malaysia*

ABSTRACT

Tuberculosis (TB) is the second biggest killer disease after HIV. Therefore, early detection is vital to prevent its outbreak. This paper looked at an automated TB bacteria counting using Image Processing technique and Matlab Graphical User Interface (GUI) for analysing the results. The image processing algorithms used in this project involved Image Acquisition, Image Pre-processing and Image Segmentation. In order to separate any overlap between the TB bacteria, Watershed Segmentation techniques was proposed and implemented. There are two techniques in Watershed Segmentation which is Watershed Distance Transform Segmentation and Marker Based Watershed Segmentation. Marker Based Watershed Segmentation had 81.08 % accuracy compared with Distance Transform with an accuracy of 59.06%. These accuracies were benchmarked with manual inspection. It was observed that Distance Transform Watershed Segmentation has disadvantages over segmentation and produce inaccurate results. Automatic counting of TB bacteria algorithms have also been proven to be less time consuming, contains less human error and consumes less man-power.

Keywords: Automated bacteria counting, Image Processing, Watershed Segmentation, Graphical User Interface

INTRODUCTION

Tuberculosis (TB) is a disease caused by bacteria. There is evidence of its existence in the Americas since Pre-Columbian times. Signs of this disease was discovered in the preserved spines of Egyptian mummies as described by Allaoui (2012). The TB is now the second biggest killer disease after HIV/AIDS with an estimation of 9.4 million new cases and 1.68 million deaths recorded by the World Health Organization in 2010 (Sandy,

ARTICLE INFO

Article history:
Received: 24 August 2016
Accepted: 03 Jun 2017

E-mail addresses:
khair@upm.edu.my (M. K. Hassan),
farahhidayah3272@gmail.com (N. F. H. Eko),
suhaidi@upm.edu.my (S. Shafie)
*Corresponding Author

ISSN: 0128-7680

2015). The TB disease is caused by the mycobacterium tuberculosa bacteria, which is a member of mycobacteriaceae family and actinomycetales ordo (Rachna& Mallikarjuna, 2013). The bacteria can infect and attack any part of the organs such as kidney, heart, liver, bones, brain and other but it predominantly affects the lungs especially those who have a weak body. Fortunately, with the development of modern technologies and proper treatments, TB disease is curable. The number of TB patients who have been successfully treated especially in United States has been increasing since 1993 (Ibnu et al., 2012; Daniel, 2006).

Researchers have come up with various techniques to examine TB bacteria. The latest technology uses microscope imaging to count TB bacteria. Acid-fast bacillus stain (AFB) or sputum stain test is performed on the sputum sample taken from the patients' lung (Daniel, 2006; Kauleshwar et al., 2013). This test determines whether the patients have positive or negative TB infection. These staining dye process provides good contrast between the bacilli bacteria and the background, thus, improving detection. The image of the bacteria will be captured using camera to be viewed through computer screen (Rachna & Mallikarjuna, 2013). With the help of computer-aided system such as MATLAB software, the image of TB disease can be further processed and diagnosed. In order to perform an automated counting of TB bacilli in the image, some processing algorithms have been introduced such as Image Acquisition techniques, Image Pre-processing techniques, Image Segmentation techniques and others. In Image Processing, segmentation technique is the most important stage. There are two approaches that can be applied in segmentation, which are the frontier approaches and region approaches. Watershed segmentation combines both the approaches (Uppal & Raman, 2012). Thus, it results in rapid detection of both edges and regions. Matrix Laboratory, also known as MATLAB, is a high-level language, which is widely used by engineers and scientist all over the world (Allaoui, 2012). Generally, it has been used extensively in mathematics, engineering and science. In addition, it covers all varieties of fields such as signal and image processing, control system, communication and many more. Another component in MATLAB is Graphical User Interface (GUI) in which the user is allowed to interact with the system using graphical icon and visual indicators (Raof et al. 2011). It is also user-friendly.

In this research, images of the stained sputum were processed and diagnosed. The sample images of TB bacteria were obtained from the Pathology Department of Hospital Serdang, Selangor where they were detected using microscope and captured using digital camera. The counting was done using the proposed algorithm of Watershed Segmentation techniques via Image Processing Toolbox MATLAB in software. The image is displayed through GUI for the ease of the user. The advantages of using watershed segmentation are: it reduces human error, consumes less man-power and saves time compared with manual TB diagnosis (Ibnu et al., 2012).

METHODOLOGY

General Algorithm of Image Counting Using MATLAB

Figure 1 shows block diagram of algorithm used for Image Counting in this research. These algorithms consist of Image Acquisition, Image Pre-Processing and Image Segmentation.

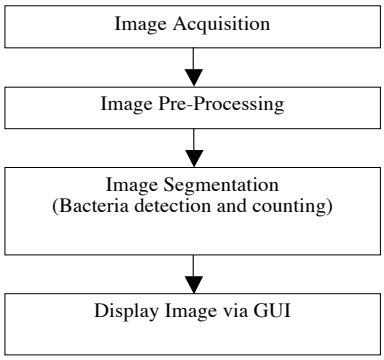

Figure 1. Block Diagram of the Algorithm used for Bacteria Counting using MATLAB

Image Acquisition. In the Image Acquisition, the stained-sputum sample image that consists of TB bacteria was obtained from Pathology Department of Hospital Serdang, Selangor. The stained-sputum images were captured through the microscope and read as input to the MATLAB software for analysis. The image must be in specific format such as JPG, JPEG, PNG, BMT (Steve, 2010).

Image Pre-Processing. There are several steps in the image pre-processing stage. In order to eliminate any unwanted objects, Image Pre-Processing is performed on the input image before the segmentation stage. The technique consists of reading an image, applying the colour thresholding segmentation before converting it into greyscale colour, enhancing the contrast of the greyscale image, and binarisation. The processed image is now ready for Image Segmentation. (Rachna & Mallikarjuna, 2013).

Image Segmentation. Image Segmentation is aimed at identifying and differentiating between the bacteria with the background image so that the former can be detected and counted (Allaoui, 2012). The segmentation process is the process of decomposing an image into sub-images of individual character. The segmentation will only stop when unwanted objects have been eliminated. In this project, two techniques of Watershed Segmentation had been applied and compared in order to separate and count overall amount of bacteria including any overlap between the bacteria that existed in the sample of images. To avoid any over segmentation, morphology segmentation was implemented before Watershed Segmentation.

Watershed Segmentation. In the acquired image, overlapping always happen. These overlapping bacteria need to be separated in order to ensure a high degree of accuracy in counting. Therefore, Watershed Segmentation was implemented to solve the problem. It can define the border between two regions which has been set up by the morphology segmentation (Kauleshwar et al., 2013). In this research, two methods involving Watershed Segmentation was used and the results verified. The first technique directly uses Watershed Segmentation on the image, also known as Watershed Distance Transform Segmentation method. Distance transform of a binary image is the distance from every pixel to the nearest non-zero valued

pixel (Pedro & Daniel, 2012). The second method, Marker Based Watershed Segmentation, limits the number of regions by using the internal markers to specify the object of interest. On the other hand, the external marker is those pixels, which belongs to the background. Each region that is being segmented consists of external markers and single internal marker All of these markers result in the separation between the overlapped bacteria and avoid over segmentation from happening (Parvati et al., 2008).

Process in detecting and counting of bacteria. The shape of the TB bacteria is almost similar to the ellipse. By applying the parametric form of the ellipse equation, the outline of the ellipse is plotted over each of the segmented edge of TB bacteria object in the sample of images. In addition, by using 'region props' function in Image Processing Toolbox, the regions such as major axis length, minor axis length and centroids for each of the detected bacteria can be measured and finally detected and counted.

RESULTS AND DISCUSSION

Figure 2(a) shows TB bacteria sample images obtained from the Pathology Department of Hospital Serdang, which are processed using the Image Processing algorithm using Watershed Segmentation, the Distance Transform and Marker Based method. This image was read as an input and displayed via the GUI on computer screen. The TB bacteria are shown in red colour pixels while the non-TB bacteria are shown other colour pixels such as blue, green and so on. Meanwhile, Figure 2(b) shows the implementation of colour thresholding segmentation on the sample image in Figure 2(a). The red colour pixel which represents the TB bacteria had increased rapidly overtaking the other colour pixels. Hence, this is a good method to detect TB bacteria. Next, the red colour was extracted from the background while the other colours were rejected to background.

Figure 2(c) shows the conversion of the Figure 2(b) into the greyscale colour so that the hue and saturation information of the RGB image can be eliminated while retaining the luminance of the image. The greyscale colour shown in Figure 2(c) is the extracted red pixels which is the TB bacteria together with some noises (unwanted objects). All of the noises will be further eliminated in the segmentation process. Figure 2(d) shows the binary image of TB bacteria sample image. The previous image need to be converted into the binary pixel image first before proceeding into segmentation. This is because some segmentation techniques only processed the image in binary pixel format. As illustrated in Figure 2(d), the white colour or also known as pixel 1 is the TB bacteria while the black colour which is pixel 0 is the unwanted object that had been rejected into the background. Figure 2(d) needs more filtration process in order to eliminate the existing noises. Figure 2(e) illustrates the edges of all the TB bacteria that had been detected by using Sobel Edge Detection. All complete connecting of pixel 1 is counted as 1 bacteria. Figure 2(f) displays the image after applying Gaussian filter and morphological segmentation.

From Figure 2(a) - (e), it can be noticed that the image consists of the TB bacteria and non-TB bacteria (noises). Gaussian filter was used to remove some of the unwanted noises. In order to totally remove the unwanted noises, morphology segmentation techniques had been

added together with the Gaussian filter. Morphology Segmentation techniques that has been used in this research is the erosion followed by dilation techniques. As a result, the image consists only TB bacteria and free from all noises. Furthermore, Watershed Segmentation is needed in order to recognise and separate any overlap that occurs between the TB bacteria. Figure 2(g) shows the image after applying the Watershed Distance Transform Segmentation Method. Figure 2 (h) is the image after it has undergone Marker Based Watershed Segmentation Method. It can be observed that the total amount of TB bacteria in the images is based on the total colour present in the images. Any overlap between the bacteria had been separated and filled by colour. Figure 2(i) and (j) view the same sample images that had been processed by using two different methods. Figure 2(i) shows the detected TB bacteria by using Watershed Distance Transform Segmentation Method. On the other hand, Figure 2(j) shows the TB bacteria detection by using Marker Based Watershed Segmentation Method. The red colour ellipse line in Figure 2(i) and (j) show the eclipse shape of detected TB bacteria that had been separated (if any overlap existed). By comparing both of the images, Figure 2(i) and (j), different amount of TB had been detected. Figure 2(i) shows 46 TB bacteria had been detected and for Figure 2(j), only 45 TB bacteria had been detected in the sample of images. The blue inside the red eclipse line is the centroid of detected TB bacteria. The manual inspection for this sample of images show that total amount of 42 bacteria. Moreover, the exact length of minor axis and major axis of the eclipse shape of TB bacteria can be known and calculated. Additionally, the exact location of X-centroid and Y centroid for each of the bacteria in the image can be located.

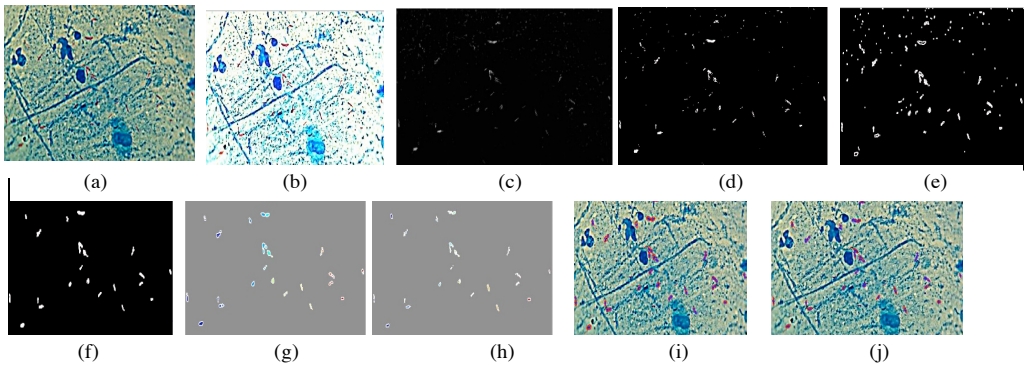

Figure 2. TB bacteria sample images

Many sample images were obtained from Pathology Department of the Hospital Serdang. However,, most of the sample images have different light intensity and result in different pixel values. Hence, the value of RGB adjustment must be different for each of the sample images. To gain images with similar light intensity, the camera distance and light intensity must be controlled. In addition, auto gain control function must be disabled. Table 1 shows seven different sample images of TB bacteria with almost similar light intensity and were used in this research. By analysing Table 1, total amount of TB bacteria that is counted by the Marker Based Segmentation method is more accurate with the average percentage of 81.08

%. Meanwhile, the accuracy of Distance Transform Watershed Segmentation only achieved an average percentage of 59.06 %. All of the percentage accuracies were benchmarked with manual inspection that had been done in Hospital Serdang.

Table 1
Comparison between automated counting and manual counting

Image	Manual Count	Automated Count			
		Distance Transform Segmentation Method	Percentage Accuracy (%)	Marker Based Segmentation Method	Percentage Accuracy (%)
A	1	4	25.0	1	100.0
B	2	4	50.0	2	100.0
C	42	46	91.3	45	93.3
D	1	2	50.0	1	100.0
E	9	10	90.0	10	90.0
F	4	8	50.0	1	25.0
G	14	8	57.1	9	64.3
Average Percentage (%)		59.06		81.80	

Figure 3 shows a comparison between manual counting (blue lines) and automatic counting using Distance Transform Watershed Segmentation (red line) and Marker Based Segmentation method (green line). The graph shows that the counting value that used Marker Based Segmentation method is closer to the manual counting than the counting value that applied Distance Transform Watershed Segmentation method.

The reason behind the lack of accuracy in Distance Transform Watershed Segmentation method is because of over segmentation . This method is very sensitive to any local minimum in the image. It tends to detect the lines of the watershed transform where each local minimum gives rise to a region. Hence, false objects can be detected. The worst can happen when it separates the connected pixel and leads to the wrong counting process. To overcome this

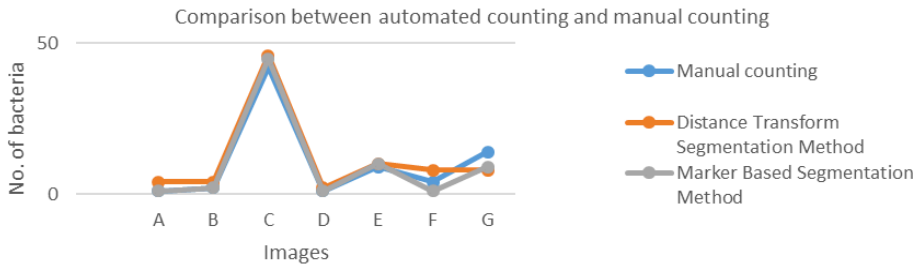

Figure 3. Graph in comparing both, manual counting and automatic counting using Watershed Segmentation

problem, the number of minima should be reduced and calculation of too many regions should be avoided. Marker Based Watershed Segmentation can be applied in order to reduce the number of regional minima together in solving the over segmentation problem. To increase the accuracy of the segmentation, the contrast of greyscale image needs to be enhanced before proceeding to the Image Segmentation technique. There are some advantages of automatic counting compared with the manual counting as the latter has several weaknesses which can be tackled through automatic counting. The list of weaknesses can be described as follow:

a. Counting Time per slide of TB Bacteria
   The manual counting time is around 10-15 minutes per slide. On the other hand, automatic counting is done between one and three 1-3 minutes.
b. Human error
   Human error is common in manual counting. It has to be counted three times in order to validate the count. The first count usually always suffers from over count or under count. On the contrary, automatic counting of bacteria that had been validated can reduce any errors.
c. Man power
   In manual counting, more manpower is needed to count the bacteria. However, with the existence of automatic counting, less man power is needed because the counting is done within the algorithm in the software.
d. GUI
   The front panel of GUI to ease the counting process had also been developed as shown in Figure 4. It offers the user two different methods for bacteria counting. It was developed using MATLAB environment.

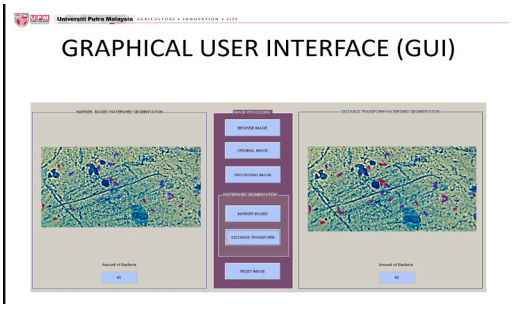

Figure 4. GUI front panel

CONCLUSION

In conclusion, the most suitable algorithm techniques in Image Processing, such as Image Acquisition, Image Pre-Processing and Image Segmentation, for counting the total amount of TB bacteria in the stained-sputum image was successfully implemented, applied and validated. Marker Based Segmentation method had been proven to be the best method in counting the TB

bacteria with an average percentage accuracy of 81.08%. On the contrary, Distance Transform Watershed Segmentation only achieved an average percentage of 59.06%. The GUI had also been successfully developed for displaying the image and user interaction. Automatic and manual counting of TB bacteria had been discussed and it was proven from this research that the former is the best technique in order to reduce human error and man power, with minimum time consumed.

ACKNOWLEDGEMENTS

A big Thank You to Pathology Department of Hospital Serdang in providing all of the stained-sputum sample images that contain Tuberculosis (TB) bacteria. This project would not have been completed without the Hospital's cooperation. We also record our gratitude to Universiti Putra Malaysia under Putra Grant GP-IPB/2013/9412702 which funded this research.

REFERENCES

Allaoui, A. E. (2012). Medical Image Segmentation by Marker Controlled Watershed and Mathematical Morphology. *The International Journal of Multimedia and Its Applications, 4*(3), 1-9.

Daniel, T. (2006) The history of tuberculosis. *Respiratory Medicine, 100*(11),1862-1870.

Ibnu, S. A. Kusworo, G. R., & Nelly, M. (2012). Development of Algorithm Tuberculosis Bacteria Identification Using Color Segmentation and Neural Network. *International Journal of Video and Image Processing and Network Security, 12*(1), 9-13.

Kauleshwar, P. N., Devvrat, L. A., & Dheeran, U. (2013). Character Recognition Using Matlab's Neural Network Toolbox. *International Journal of u- and e- Service, Science and Technology, 6*(1), 13-20.

Parvati, K., Prakasa, B. R., & Mariya, D. M. (2008). Image Segmentation Using Gray-Scale Morphology and Marker-Controlled Watershed Transformation. *Discrete Dynamics in Nature and Society, 2008*, 1-8.

Pedro, F. F., & Daniel, P. H. (2012). Distance Transforms of Sampled Functions. *Theory of Computing, 8,* 415–428.

Rachna, H., & Mallikarjuna, M. S. (2013). Detection of Tuberculosis Bacilli using Image Processing Techniques. *International Journal of Soft Computing and Engineering, 3*(1), 2231-2307.

Raof, R. M., Mashor, R. A., & Noor, S. (2011). Image Segmentation of Ziehl-Neelsen Sputum Slide Image for Tubercle Bacilli Detection. *Journal of Engineering and Medical Science, USM, 1*(1), 365-378.

Sandy, C. R. (2015). *Sputum Stain for Mycobacteria*. Retrieved May 16, 2015, from: http://www.healthline.com/.healthline.com

Steve, E. (2010). Visualizing region props ellipse measurements. Retrieved May 20, 2015, from http://blogs.mathworks.com/

Uppal, N., & Raman, G. (2012). Computational Approach to Count Bacteria Colonies. *International Journal of Advances in Engineering and Technology, 4*(1), 364-372.