

MYCOBACTERIUM TUBERCULOSIS BACILLI CELLS IDENTIFICATION USING MOMENT INVARIANT

JADHAV M.E.^{1*}, KALE K.V.², BAHETI M.J.³

¹Institute of Management Studies & Information Technology, Aurangabad,

²Department of Computer Science & IT, Dr. Babasaheb Ambedkar Marathwada University, Aurangabad

³Department of Computer Engg, SNJB'S College of Engg, Chandwad, Nashik

* Corresponding Author: Email- muktijadhav@gmail.com

Received: September 29, 2011; Accepted: November 03, 2011

Abstract- Mycobacterium Tuberculosis(M.TB) bacilli is causative agent of Tuberculosis. Manual detection method for M.TB is time consuming, tedious & sometime it may confuse with some stain residue & non tuberculosis bacilli. For this reason the need of atomization is required for exact identification of tuberculosis. We present a new method for M.TB bacilli cells recognition method using Moment invariant and Gaussian member function. The object i.e. M.TB bacilli cells are extracted using color segmentation method, from Ziehl-Neelsen (ZN) stained sputum smears images. These images have blue color background on which red color M.TB bacilli cells. The objects are extracted using color segmentation (thresholding) method and boundary detection. The various extracted shapes contain M. TB bacilli cells, non TB cells, stain residue etc, these extracted objects are preprocess using morphological dilation and then moment invariant feature extracted. Our main objective is to recognizes M.TB bacilli cells with the help of color segmentation & shape base feature moment invariants using Gaussian member function. The recognition rate of M.TB bacilli cells is obtaining 98.17%.

Key words - Mycobacterium Tuberculosis (M.TB), ZN-Stained (Ziehl -Neelsen), Moment invariant, Gaussian membership function.

Introduction

The Mycobacterium Tuberculosis bacilli is the etiology of Tuberculosis(TB). Tuberculosis is air borne disease and its diagnosis in early stage is critical. The pulmonary (lung) tuberculosis is the origin of extra-pulmonary tuberculosis. Pulmonary TB diagnosis by sputum smear microscopy. Sputum smear microscopy is relatively simple, least expensive, easy to perform[1,2]. In low and middle-income countries smear microscopy with the Ziehl-Neelsen staining technique, due to its simplicity, rapidity, reproducibility, low cost, and effectiveness in detecting infectious cases[3].

However manual screening technique is time consuming, labor-intensive, tedious[4]. It may require skilled person for diagnosis of M.TB bacilli. So the need of atomization for improve test sensitivity, better diagnostic accuracy of results. The accuracy can improve by increasing number of images and can be analyzed by the computer. However color of the M.TB bacilli is the main discernment property using ZN-stain. ZN staining is widely used and globally relevant to national TB program. It makes Mycobacterium TB bacilli cell has red in color and sputum cells has blue color background. But some time it may confuse with stain residue, non TB bacilli and artifacts so the need of another shape base parameter required improving accuracy of results. The segmentation of foreground ie. Red color MTB cells and background ie. Blue color sputum smear is achieved by

color segmentation. Moment invariant is the shape based feature used for proper recognition of M.TB bacilli using Gaussian membership function. The live images are acquired from 100X (high power) microscopic field of ZN-stain sputum smear acid fast bacilli slide, is the source data for the proposed work. The images are captured by CCD camera. In this work 75 images of sputum smear are preprocessed for image enhancement for reduction of noise. From the enhanced sputum smear images 219 objects are extracted after segmentation is the source data for feature extraction. Various authors addressed identification of M.TB bacilli they are using image data from the image library we are using live data images of ZN stain sputum smear from the TB hospital Aurangabad. We are inventing new technique on live images using moment invariant & Gaussian membership function it is easy to implement and gives better accurate result in fewer amounts of data.

Literature Survey

Several authors have addressed segmentation of bacteria particle. Verpoulos et al.[4][5]. Worked on gray level images used an identification method based on shape descriptors and Neural network classifiers showing a sensitivity(ratio of true positive decision against the total number of positive cases) of 94.1%. Wilkinson [6] proposed a rapid multi resolution segmentation technique based on computing different

thresholds for different areas of a grey level image. Other authors use the color information as the key discriminate factor either for bacteria segmentation and identification[7],[8].

Manuel G. Forero et al.[9][10][11] proposed a technique of bacilli segmentation based on chromatic information, a multi thresholding image segmentation techniques and simple color filtering the second approach based on the use of gray-level morphological operators only to the green channel, the third approach is bacilli detection based on heuristic knowledge extracted from the bacilli shape contour. It uses the color information for image segmentation and finally a classification tree is used to categorize if a sample is positive or negative. M.G.Forero et.al.[12] used a Gaussian mixture model was used to characterize the bacillus class and a Bayesian decision rule was employed for the identification of M.TB cells. R.A.A.Raof et al.[13] proposed the thresholding procedures involved setting of boundaries based on grey values or intensities of image pixels. Hue color component based approach is proposed to segment the bacilli by adaptive choice of the hue range[14]. Rethabile Khutlang et al. proposed, different classifiers are compared; the sensitivity of all tested classifiers is above 90% for the identification of a single bacillus object using all extracted features[15]. All the above works are carried on the microscopic images of sputum smear for identification of M.TB bacilli.

Flow of Processing

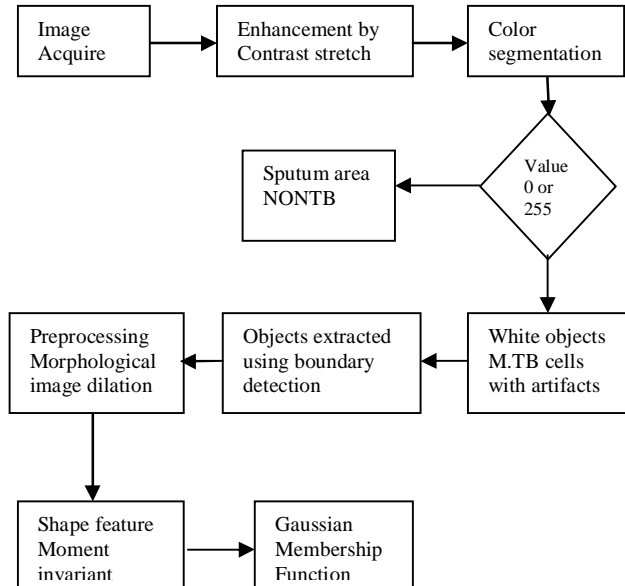


Fig.1 Flowchart of M.TB identification process.

Color Segmentation

Segmentation means the partitioning of entire image into small segment. The goal of segmentation is to simplify and/or change the representation of an image into something that is more meaningful and easier to analyze. Segmentation is generally used to locate objects and boundaries in images.

Thresholding is a method used for image segmentation. It is the way to extract the objects from background by selecting threshold. This method is based on a clip-level (or a threshold value) to turn a RGB scale image into a binary image. The image $f(x,y)$ composed of light object on dark background. The threshold T is require to select for extract object from background. Then any point (x,y) in the image at which $f(x,y) > T$ is called object point otherwise point is called background point[16].

The segmented image $g(x,y)$ is given below.

$$g(x, y) = \begin{cases} 1 & \text{if } f(x, y) > T \\ 0 & \text{if } f(x, y) \leq T \end{cases} \quad (1)$$

There should be more than one thresholding value at a time where T_1 is lower threshold and T_2 is the upper threshold value[16].

$$g(x, y) = \begin{cases} 0 & \text{if } f(x, y) < T_1 \\ 1 & \text{if } T_1 < f(x, y) \leq T_2 \\ 0 & \text{if } f(x, y) > T_2 \end{cases} \quad (2)$$

In RGB imaging, pixel is characterized by three RGB values. However, with multi spectral images such as RGB images, it can be difficult to specify the selection criteria for segmentation. The main objective to filter out blue pixels and retain the reddish pixels which are the mycobacterium TB cells. The properties of the RGB pixels are being studied to extract the important features from the image. This is based on the color information, the color thresholding algorithm should be able to extract the pixels of mycobacterium and reject pixels of other objects. Using this thresholding RGB image get segmented.

Following is the Microscopic digital image of Sputum sample capture by CCD camera.

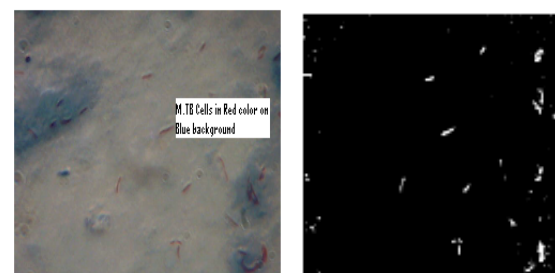


Fig. 2 Original image Fig.3 Segmented image
Fig. (2),: Original image of Microscopic field of ZN-Stain sputum smear on Acid Fast Bacilli slide Fig. (3), : Image after Segmentation.

The segmented image shows two values either 255 (white pixel) which indicate the area of M.TB pixel or 0 (black pixel) sputum area. We get final value which is either 255 (white pixel) or 0 (black pixel) i.e M.TB cells are white in color on black color background. The obtained image contains M.TB cells with some stained

residues, non decolorized epithelial cells and non TB bacillus. For proper recognition of M.TB bacilli the objects boundaries is extracted using boundary detection technique. The extracted objects processed from morphological operation such as dilation. Then the shape based features of these objects are extracted.



Fig. 4, Various shapes object are obtained after color segmentation algorithm and boundary detection.

Moment invariant

For exact recognition of M.TB. bacilli we need to extract some features of the organisms. Shape is a fundamental property of an organism i.e. object, The shape descriptor moment invariants are one of the most popular, widely accepted & used contour-based shape descriptors which is a set derived by Hu (1962).

The recognition of objects regardless of their position, size and orientation The idea of using moments in shape recognition gained prominence when Hu (1962), derived a set of invariants using algebraic invariants[17].

$$f(x, y), \quad (x, y) = 0, \dots, M-1$$

$$m_{pq} = \sum_{x=0}^{M-1} \sum_{y=0}^{M-1} (x)^p \cdot (y)^q f(x, y) \quad pq = 0, 1, 2, 3 \dots$$

(3)

The moments $f(x, y)$ translated by an amount (a, b) , are defined as,

$$\mu_{pq} = \sum_x \sum_y (x+a)^p \cdot (y+b)^q f(x, y) \quad (4)$$

Thus the central moments μ'_{pq} or μ_{pq} can be computed from (2) on substituting $a = \bar{x}$ and

$$b = \bar{y} \text{ as, } \bar{x} = \frac{m_{10}}{m_{00}} \text{ and } \bar{y} = \frac{m_{01}}{m_{00}},$$

$$\mu_{pq} = \sum_x \sum_y (x - \bar{x})^p \cdot (y - \bar{y})^q f(x, y) \quad (5)$$

When a scaling normalization is applied the central moments change as,

$$\eta_{pq} = \mu_{pq} / \mu_{00}^{\gamma}, \quad \gamma = [(p+q)/2] + 1 \quad (6)$$

In particular, Hu (1962), defines seven values, computed by normalizing central moments through order three, that are invariant to object scale, position, and orientation. In terms of the central moments, the seven moments are given as,

$$M_1 = (\eta_{20} + \eta_{02})^2$$

$$M_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2$$

$$M_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} + \eta_{03})^2$$

$$M_4 = (\eta_{30} - \eta_{12})^2 + (\eta_{21} + \eta_{03})^2$$

$$M_5 = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]$$

$$M_6 = (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03})$$

$$M_7 = (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] - (\eta_{30} + 3\eta_{12})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]$$

(7)

The seven moments values obtained from each object like these way we obtain values of 219 objects means 1533 features. Then applied Gaussian membership function method on the moment invariant features

Gaussian Membership function

A membership function provides a measure of the degree of similarity of an element to a fuzzy set. Membership functions can take any form, but there are some common examples that appear in real applications. Membership functions can either be chosen by the user arbitrarily, based on the user's experience (MF chosen by two users could be different depending upon their experiences, perspectives, etc.).

This can also be designed using machine learning methods (e.g., artificial neural networks, genetic algorithms, etc.). There are different shapes of membership functions; triangular, trapezoidal, piecewise-linear, Gaussian, bell-shaped, etc.

For an unknown input numeral x , the features are extracted using the invariant moments model. The membership function is chosen as,

$$\mu_i = \exp(x_i - M_i)^2 / 2\sigma_i^2 \quad (8)$$

where x_i is the i^{th} feature of the unknown numeral.

If all x_i 's are close to μ_i which represent the known statistics of a reference character, then the unknown numeral is identified with this known numeral because all membership function values are close to 1 and hence the average membership function is almost 1 [18]

Let, $M_i(r)$ and $\sigma_i^2(r)$ belong to the r^{th} reference numeral with $r = 0, 1, \dots, 9$, we then calculate the average membership as,

$$\mu_{av}(r) = 1/c \sum_{i=1}^c \exp(x_i - M_i)^2 / 2\sigma_i^2 \quad (9)$$

where c denotes for the number of fuzzy sets. Then $x \in r$ if $\mu_{av}(r)$ is the maximum for $r=0, 1, \dots, 9$.

The concluding results were commenting either as recognized or misrecognized objects.

Table 1- Result of Gaussian Membership function.

Application	Training data	Testing data	Features moment invariant	Classes	Accuracy
M.TB bacilli cells recognition	219	219	7	2	98.17%

Result and Discussion

The first parameter for segmentation of M.TB bacilli from sputum smear is color based segmentation using thresholding objects are separated but sometime it may confused with non TB bacilli and stain residues. The boundary detection technique applied to extract objects from segmented image. The second parameter is shape based moment invariant for extracting features of obtained shape, here 1533 features obtain from 219 objects. Results are expressed per view field and not by bacillus recognition performance and therefore a valid comparison are not possible. The recognition rate using this technique is 98.17%. These results are low as compared to [12] but for only invariant moment based approach it has shown promising results for our database. We will try to improve the recognition rate by adding some more features to the moment invariants in future scope.

Conclusion

We implemented the technique of color segmentation allows suppression of unwanted data. The only objects those color close to the prototypical bacilli are retained. The second key technique is that it is based on the analysis and screening of bacillus shapes using moment invariant. By applying the Gaussian membership function i.e. TB and non-TB bacilli cells are classified with overall recognition rate of 98.17%.

References

- [1] Albert H., Heydenrych A., Brookes R., Mold R.J., Harley B., Subotsky E. (2002) *Int. J. Tuber Lung Dis*, 6,529–37.
- [2] Muzaffar R., Batool S., Aziz F., Naqvi A., Rizvi R. (2002) *Int. J. Tuber Lung Dis*, 6,635–40.
- [3] Luna J.A.C. (2004) *International Union against Tuberculosis and Lung Disease*, Paris.
- [4] Veropoulos K., Learmonth G., Campbell C., Knight B. and Simpson J.(1999) *Analytical and quantitative cytology and histology* 21(4),277.
- [5] Veropoulos K., Campbell C., Learmonth G., Knight B., and Simpson J.(1998) *Proceedings of the 8th*

International Conference on Artificial Neural Networks, 2, 797-802.

- [6] Wilkinson M. (1996) *Fluorescence Microscopy and Fluorescent Probes*. Slavik, J.,261-266.
- [7] Alvarez-Borrego J., Mourino R., Cristobal G. and Pech J. (2000) *Int. conf. on pattern recognition*, 2847, 283.
- [8] Demantova P., Sakamoto D., Ioshii S. and Gamba H. (2001) *Proceedings of the II latin american congress on biomedical engineering*.
- [9] Forero M., Sierra E., Alvarez-Borrego J., Pech J., Cristobal G., Alcalá L. and Desco M. (2001) *SPIE Proc. Algorithms and systems for optical information processing*, V, 4471, 293-304.
- [10] Forero M., Sroubek F., Alvarez-Borrego J., Malpica N., Cristobal G., Santos A., Alcalá L., Desco M. and Cohen L. (2002) *SPIE Proc. Photonic Devices and Algorithms for Computing IV*, 4788.
- [11] Manuel G. Forero, Gabriel Cristobal and Josue Alvarez-Borrego (2003) *The Applications Of Digital Image Processing*, *SPIE Proc*, XXVI, 5203,71-81.
- [12] M. G. Forero, G Cristobal M. Desco (2006), *Journal of Microscopy*, Vol. 223,120–132.
- [13] Raof R.A.A., ZalehaSalleh, Sahidan S.I., Mashor M.Y.(2008) *5th International Conference on Electrical Engineering, Computing Science and Automatic Control IEEE,CFP08827-CDR*, 978-1-4244-2499-3/08, 212-217.
- [14] Vishnu Makkapati, Ravindra Agrawal and Raviraja Acharya (2009) *5th Annual IEEE Conference on Automation Science and Engineering Bangalore, India*, 978-1-4244-4579-0/09.
- [15] Rethabile Khutlang, Sriram Krishnan, Andrew Whitelaw, Tania S Douglas (2009) *IEEE* 978-1-4244-3932-4/09.
- [16] Gonzalez R. C. and Woods R. E. (2008) *Pearson Education*, 3rd Ed.
- [17] Muharrem Mercimek, Kayhan Gulez and Tarik Velimumcu (2005) *Sadhana* 30(6),765–775.
- [18] Hanmandlu M. and Murthy O.V. R. (2005) *Fuzzy Model Based Recognition of Handwritten Hindi Numerals*. *Proc ICCR*, 490-496.