Separation and Counting of Blood cells using Geometrical Features and Distance Transformed Watershed

A.Sai Prasad, K.Surya Latha, S. Koteswara Rao

Abstract—In the field of biomedicine, blood cells are complex in nature; the microscopic images of a blood stream contain RBCs, WBCs and Platelets. Pathological inspection of an infected cell based on the disease, is solely dependent on subjective assessment which usually leads to significant inter-observer variation in grading and subsequently resulting in late diagnosis. However automatic assessment of required cell count still remains a challenging task as many of the cells are clumped in an image and proper segmentation is the primary aspect. This paper aims at segmentation of blood cells for counting. Auto threshold, Chessboard distance measure and watershed are used for segmentation of blood cells.

Keywords- Canny Edge detection, Clumped cells, Distance transform, Geometrical Features, Watershed.

I. INTRODUCTION

Blood is a highly specialized tissue composed of more than 4,000 different kinds of components. Four of the most important ones are RBC cells, WBC cells, Platelets, and plasma. In one liter human blood 3 to 5 trillion red blood cells, 4.5 to 12 billion white blood cells and 120 to 350 billion Platelets circulate [5]. We can easily differentiate RBCs from WBCs, on the basis that the former ones have no nucleus. The plasma content is seldom measured. We can differentiate the Platelets from the other blood components by the size factor. In fact blood cells identification and counting is carried out in laboratories through human observation. However, the automatic cell recognition in bright field microscopy is a difficult task due to immense variability of cell appearance and overlapping of cells. Digital image processing techniques can help the hematologist in their analysis and diagnosis. Image segmentation techniques can be used for separating the RBC, WBC and Platelets in an image. The prime reason for segmenting the image is to define the boundaries of the blood cells enabling features to be extracted without the inclusion of extraneous material. Harms and Aus [4] state that "segmentation is the first, important step in image analysis". Russell and Norvig [7] similarly stated that segmentation is a key step towards organizing the array of image pixels into regions that would correspond to semantically meaningful entities in the scene. Rosenfeld discusses two forms of segmentation. Pixel based image segmentation, in which pixels are classified independently, is seen as simpler, but has a number of drawbacks particularly in the area of local consistency. Region based segmentation, where the goal is to split the image into distinct connected regions is seen as a better alternative.

Both forms of segmentation require some experimentation to develop a good semantic model that can be used to split or merge regions [12]. Sonka, Hlavac and Boyle [8] nominate thresholding as the simplest segmentation process, as it is computationally cheap and fast. It is only suitable for objects that do not touch each other and where their gray levels are clearly distinct from background levels. Harms and Aus[4] use a model of the object of interest in their cell segmentation algorithm. In this method, they use color difference to identify the nuclear material in one focal plane, they then shift the focus and use absolute color to identify myelin sheath material in a different focal plane and use a luminosity (gray scale) image in a third focal plane to identify cytoplasm. The three images are then overlaid, nuclear sub regions are combined if they are close enough. This method has the advantage of being able to work with the sharpest images possible of both the nucleus and the cytoplasm boundary. A limitation of this technique is that a microscope digital focus control, with a 0.1 micron resolution is mandatory for reproducibly measuring tissue sections of different focus settings. In this paper we proposed an improved segmentation method based on geometrical feature of the blood cells. It is known from the observations that various cells differ greatly with the size. For example the Eosinophils have the size of 8 to10 micrometers, while mega Caricocyte have up to 100 micrometers. These features are used to segment the WBCs, RBCs and Platelets. The results obtained from these features are used to count the bloods cells in the image. The remainder of this paper is organized as follows. Section II describes the proposed model; Section III describes separation of clumped cells. The implementation and results are given in Section IV; Conclusion is given in Section V.

II. PROPOSED MODEL

In this section, we describe the steps that are performed in the improved method for segmenting the blood cells. The deficiencies of the other methods are highlighted, and the modifications that we made in order to remove those deficiencies are described. The method operates on binary images obtained from initial segmentation, and consists of several well-defined stages, presented in Fig. 1. Initially the acquired image is converted into the gray scale image for edge detection. We used canny edge detection for finding the boundaries of the blood cells. In the next step the edged image is processed by morphological operations which preserve the essential shapes of the objects and removal of noise. After these preprocessing steps we separate the white

blood cells, red blood cells and Platelets from the entire image plane by using the geometrical features of the blood cells. Finally the separated images of individual cells are watershed for separating the clumped cells for better counting.
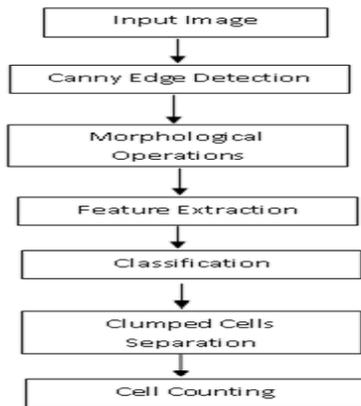

Fig. 1. General Structure of Proposed Method

A. Canny edge detection

The Canny edge detection operator is the most popular edge detection, with three objectives such as optimal detection, good localization and single response. Canny was the first to demonstrate that Gaussian filtering is optimal for edge detection [9]. The Gaussian operator is optimal for image smoothing. The Gaussian operator

$g(x, y, \sigma)$ is given by

$$g \sigma$$

By differentiation , for unit vectors $U_x = [1,0]$ and $U_y = [0,1]$ along the cordinate axes, the gaussian operator becomes as follows

$$+$$

$$= - \qquad - \qquad (1)$$

The coefficients of a derivative of Gaussian template are calculated by using equation (1) which combines with the first order differentiation with Gaussian smoothing. This is a smoothed image and so the edge will be a ridge of data. To mark an edge at the correct point, we can convolve an image with an operator which gives the first derivative in a direction normal to the edge. , the first derivative of a Gaussian functioning 'g' in the direction of the normal

$$(2)$$

Where can be estimated from the first order derivative of the guassing function g convolved with the image I and scaled appropriately as:

$$(3)$$

The location of the true edge point is then at the maximum point of G_n, convolved with the image. This maximum is evolved when the differential (along) is zero.

$$(4)$$

By substituting equation (2) in equation (4)

$$= 0 \qquad\qquad (5)$$

This equation provides the basis for an operator which meets one of the canny's criterian that edges should be detected in the correct place. This is a non-maximum suppression, which is equivalent to retaining peaks, which thins the response of the edge detection operator to give edge points that are in the right place, without multiple responses and with minimal response to noise. Fig.2. shows the comparison of the canny edge detection with the other methods such as Prewitt, Sobel, Roberts, Laplacian of Gaussian (LOG). It can be observed that the detected image using canny preserves the finite details of the edges of blood cells.
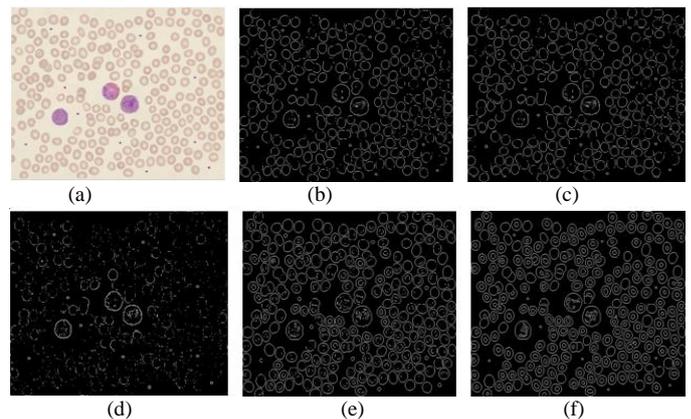

(a)   (b)   (c)

(d)   (e)   (f)

Fig2. Comparison among the various edge detectors(a) Original image, (b) Prewitt edge detector, (c) Sobel edge detector, (d) Roberts edge detector (e) LOG edge detector, (f) Canny edge detector.

B. Morphological Operations

The term morphology is applied on the binary images by considering the connectedness of the pixels in the image [6]. The effect of any morphological processing depends on the structuring element being used and is a key element in any morphological operation. The various morphological operations used are dilation, erosion, opening and closing. Dilation joins the broken lines forming contour delineating region of interest. Erosion has the effect of removing small isolated features, of breaking apart thin, joining regions in a feature and of reducing the size of solid objects by 'eroding' them at the boundaries. In the classification of blood cells, the half filled objects or cells are removed, that is, the blood cells at the boundaries are removed by using the morphological operation opening. Closing enlarges the boundaries of foreground regions in an image. Thus the basic processing techniques help in the accurate results for the classification of blood cells through which the desired blood cells are retrieved.

C. Feature Extraction

It is very complex to classify the blood cells from raw image under human observation. A raw image can't be used directly as features due to high variations in morphology which are coupled with rotation and scaling factors. A feature to be used in classification must provide rotation,

scale invariance and capable of capturing the morphological characteristics [3]. In this paper we considered geometrical features to classify the blood cells. A normal blood cell is primarily one of two major particles: a RBC with a normal Probability Distribution Function (PDF) around 6.0 to 8.5 μm [2] or a WBC (10 to 18 μm) [11] which includes a nucleus and cytoplasm is about 1 to 3 times bigger than normal and mature RBCs. Moreover, WBCs are classified into five main shape groups with varying degrees of non-convexity [9]. We use size characteristics as an effective factor to distinguish between the two main types of RBC and WBC cells. In hematology, geometrical features are widely used, as various blood cells differ greatly by their size or nucleus shape. Geometrical features are computed on the basis of Region of Interest ROI (R) , which has well defined closed boundary composed of a set of geometrical features such as *area, perimeter, centroid, tortuosity (Area/ Perimeter), Compactness C-given by the formula: perimeter²/area and radius* [3]. The blood cells area can be found by the number of pixels enclosed by the cell boundaries: A = {Centroid(R)}. The centroid is the average of all points of R.

= , = . The radius is measured by the average length of the radial segment defined by the border points R= .

D. Separating RBC, WBC and Platelets:

Thus far, an image is formed with solid objects; before separating the Clumped cells and counting, RBCs, WBCs and Platelets should be separated into three sub images. This task could be done by a step- by-step iterative method: Find the *area(A)* of the objects in the blood stream image (Image after morphological (operations). Obtain minimum and maximum area of the objects. Set the value of maximum area as the area of WBC's. Estimate the size of the RBC and Platelets using the size of WBC. Let the A_1, A_2 and A_3 represents the *average area* of the *WBC's, RBC's* and *Platelets*. Based on average areas, as given in Table. I, of the blood streams we can separate WBCs, RBC's and Platelets as shown in Fig 3. The objects which have area 20% of WBC area are considered as Platelets. Iteratively find the entire objects area and save them in a separate image plane. The objects which have area that are between 40% of A_1 and 25% of A_1 are considered as RBC's.

Iteratively find the entire objects area and save them in a separate image plane.

Table I: Average Area of blood cell in the image plane.

Image	Average Area			% of A2 w.r.to A1	% of A2 With A3
	WBC	RBC	Platelets		
	A1	A2	A3		
1	2667	643.86	192.5	24.14	12.84
2	2541.5	863.13	199.5	33.96	13.75
3	1834.5	459.23	194.5	25.03	18.78

III. CLUMPED CELLS SEPARATION

Clumped cells appear in blood smear images with various degree of overlapping. Splitting these clumped cells effectively can be done by spatial information of image plane. Among the various algorithms for segmentation of overlapped objects, Watershed segmentation is an attractive method and tends to favor in the attempts to separate touched or overlapped objects which is one of the most difficult image processing operations. The major idea of Watershed Segmentation is based on the concept of *topographical representation* of image intensity. The basic concept is based on visualizing a gray level image into its topographic representation, which includes three basic notations: *minima, catchment basins* and *Watershed lines*. If we imagine the bright areas have "high" altitudes and dark areas have "low" altitudes, then it might look like the topographic surface illustrated by Fig 4. In this surface there are three types of points: (1) Points belonging to the different minima; (2) Points at which water would fall with certainty to a single minimum; and (3) points at which water would be equally likely to fall to more than one such minimum. The first type of points forms different minima of the topographic surface. The second type points which construct a gradient interior region is called catchment basin or watershed of that minimum. The third type of points form crest lines dividing different catchment basins, which is termed as watershed lines.[6]
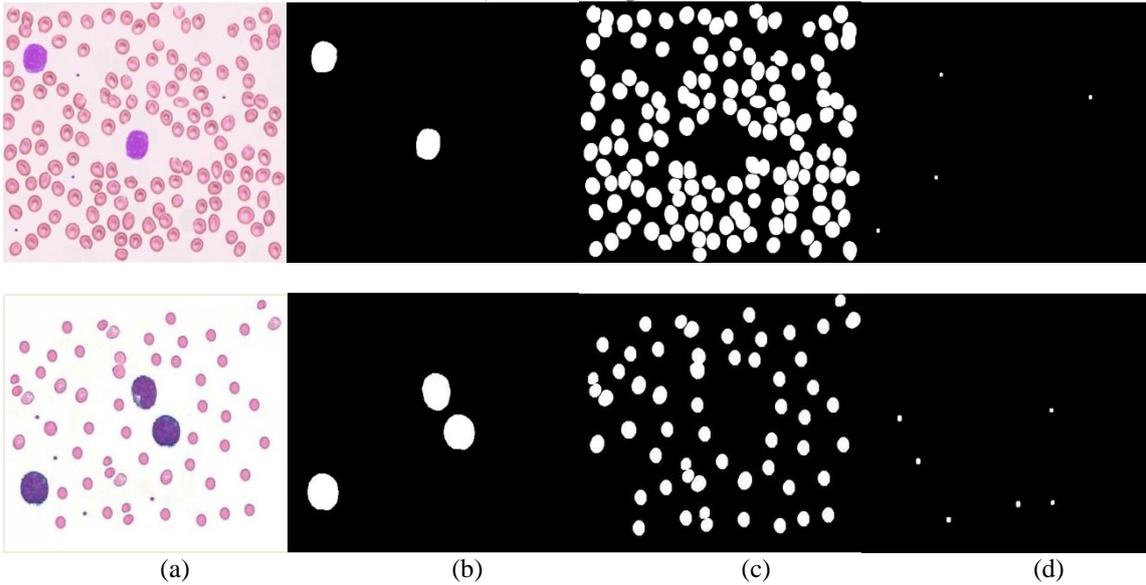
(a) (b) (c) (d)
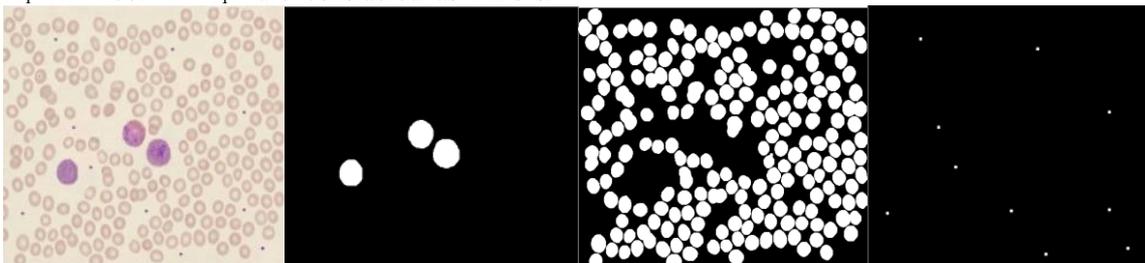
Fig3. Sub image column (a) shows an original blood stream images (1,2,3) with RBCs, WBCs, Platelets, (b) shows the separated WBCs, (c) shows the separated RBCs and (d) shows the separated Platelets

A. Distance Transforms

Watershed segmentation produces good results for gray level images with different minima and catchment basins. For binary images, however, there are only two gray levels 0 and 1 standing for black and white. If two cells are connected together in a binary image like Fig. 6, only one minimum and catchment basin will be formed in the topographic surface. To use watershed to segment the connected cells, we need to use distance transforms (DTs) to preprocess the image to make it suitable for watershed segmentation.
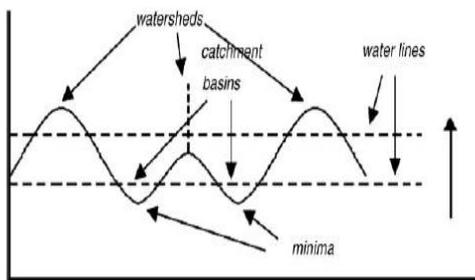

Fig. 4. Watersheds and Catchment basins. Basins are defined as connected regions with local minimum to which any incident rainfall will flow. The watershed can be visualized as those ridges from which water would be equally likely to fall towards either of two or more catchment basins.

In this article, we define the DT of a binary image as the distance from every pixel of the object component (black pixels) to the nearest white pixel. There are many different metrics to define the distance between two pixels $[x_1, y_1]$ and $[x_2, y_2]$ in a digital image. Several commonly used DT functions [9] for image processing are:

Euclidean:
$$D_{Euclidean}((x_1, y_1),(x_2, y_2)) = \sqrt{(x_1} \quad . \tag{8}$$

Quasi-Euclidean:
$$D_{Quasi-Euclidean}((x_1,y_1),(x_2,y_2)) = (\quad -1) * (\quad + \tag{9}$$

City Block:
$$D_{Cityblock}((x_1,y_1),(x_2,y_2)) = \quad + \tag{10}$$

Chessboard:
$$D_{Chessboard}((x_1,y_1),(x_2,y_2)) = \max(\quad , \tag{11}$$

Different distance transform discs of radius $k = 3$. A set of pixels at distance $\leq k$ according to a distance metric is called a *disc* of radius k [6]. Fig. 6 illustrates the effects of applying different DTs on synthetic image with overlapping of two objects. The synthetic image is shown in Fig6 and the results of applying various distance transforms on blood cells is shown in Fig.7.
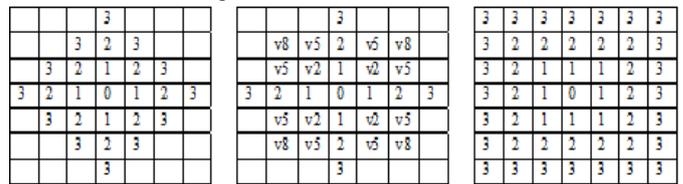

(a) (b) (c)

Fig. 5 shows discs of radius k=3 according to the distance functions of equations (8), (10) and (11), (a) Euclidean (b) City Block and (c) Chessboard
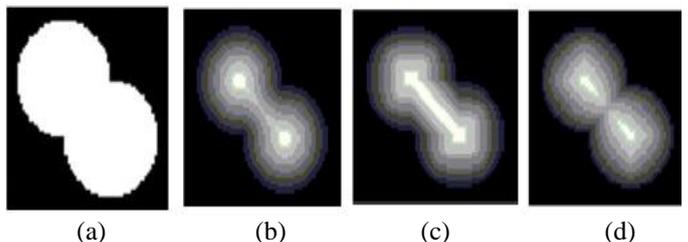

(a) (b) (c) (d)

Fig6. Illustration of DTs. (a) Synthetic image of two overlapped blobs (b)-(d) Euclidean, City block, Chessboard distance transforms.

IV. RESULTS AND EXPERIMENTAL COMPARISION

The experimntal results for the clumped red blood cell segmentation is shown in Figure 6. The results so obtained are by applying morphological watersheding with different distance transforms given in the equations (2)-(6). It can be observed that by using euclidean, city block and quasi euclidean distance tranforms, the watershed images are oversegmented i.e it is segmenting the blood cells which are not overlapped and gives more count than the actual red blood cells. The watersheded image (e) with chessboard distance transform fuction has segmented almost all the clumped cells. The red blood cells count is approximate to that of the manual count. The actual count of red blood cells are 200, the count obtained from chessboard DT is 201, euclidean DT is 215, city block DT is 202, and from the quasi-euclidean is 211.
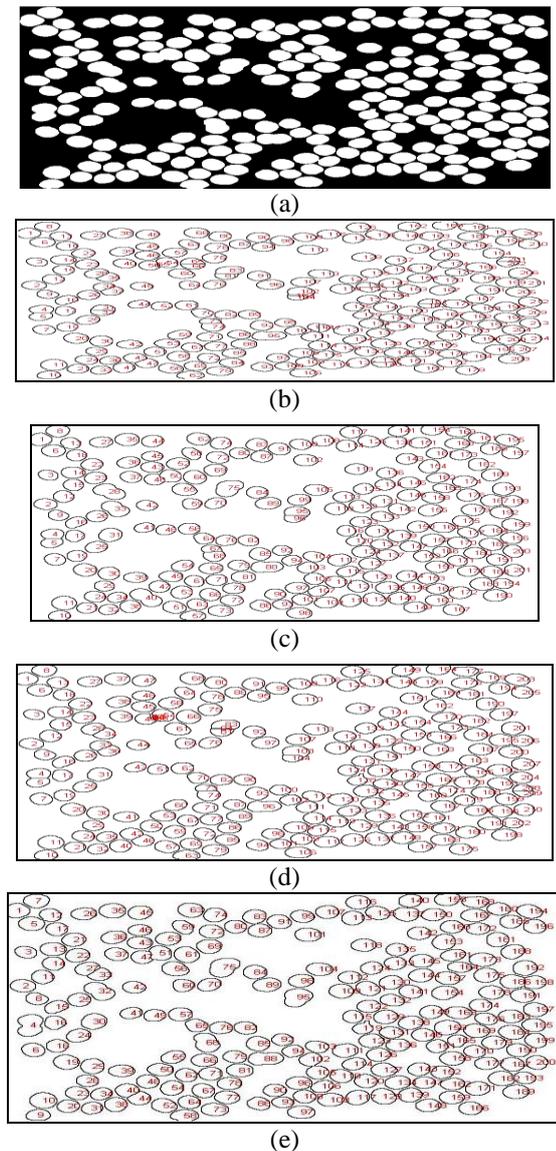

(a)

(b)

(c)

(d)

(e)

Fig 7. (a) shows the RBC image before watershed, (b)-(e) are the watershed images obtained using Euclidean distance, City block, Quasi-Euclidean and Chessboard distance transformations

V. CONCLUSION

In this paper, a simple step-by-step procedure for detection and segmentation of blood smear particles has been presented using the geometrical features together with a new approach for segmenting the clumped red blood cells. Experimental results indicate that the current analysis of blood cells is accurate and it offers remarkable segmentation accuracy. The performance of the proposed method is evaluated by comparing the automatically extracted blood particles with manual segmentations and other traditional techniques. Furthermore, the introduced method being simple and easy to implement is best suited for medical applications in clinical settings.

REFERENCES

[1] Watershed Segmentation for Binary Images with Different Distance Transforms, Qing Chen, Xiaoli Yang, Emil M. Petriu, IEEE Transactions- 2006.

[2] M.Habibzadeh et.al.,(2011) 'Counting of RBCs and WBCs in noisy normal blood smear microscopic images', Medical Imaging 2011: Computer-Aided Diagnosis, Proc. of SPIE Vol. 7963, 79633I

[3] Magudeeswaran Veluchamy et.al., (2012) ,' Feature Extraction and Classification of Blood Cells Using Artificial Neural Network American Journal of Applied Sciences 9 (5): 615-619, 2012 , ISSN-1546-9239

[4] Harry Harms and Hans Magnus Aus, "Tissue Image Segmentation with Multicolor, Multifocal Algorithms", Pattern Recognition Theory and Applications, Edited by P.A. Devjver and J. Kittler, Springer-Verlag Berlin, Heidelberg 1987 Thesis A.R.J. Katz 56

[5] Wim Van der Meer, "Blood Cell Morphology: Controversies and Alternatives", Thesis Radboud University Nijmegen, The Netherlands, ISBN 90-9020526-8.

[6] Digital Image Processing, 2nd Edition, Rafael C. Gonzalez, Richard E. Woods.

[7] Stuart J. Russell and Peter Norvig, "Artificial Intelligence A Modern Approach", 1995, Prentice-Hall, Inc. Upper Saddle River, New York.

[8] Milan Sonka, Vaclaw Hlavac and Roger Boyle, "Image processing, Analysis and Machine Vision", Chapman and Hall, London, 1993.

[9] Feature Extraction and Image Processing-2nd Edition, Mark.S.Nixon, Albertos.Aguado.

[10] S. M. Lewis, B. J. Bain, and I. Bates, Dacie and Lewis, chapter 5:Blood cell morphology in health and disease practical hematology, 10th ed. Philadelphia: Churchill Livingstone Elsevier, pp. 107-110, 2006.

[11] Newland J. The peripheral blood smear. Goldman L,Ausiello D, eds. Cecil Medicine. 23rd ed. Philadelphia, PA: Saunders Elsevier, chap 161, 2007.

[12] Rosenberg, Azriel, "Image Analysis, Problems, Progress and Prospects", Pattern Recognition, Vol 17, Number 1, 1984 pp3-12. Thesis A.R.J. Katz 57.

AUTHOR'S PROFILE

A. Sai Prasad received the M.Sc degree in Physics from Andhra University, Andhra Pradesh, India, in 2002, and the M.Tech degree in Computer Science with specialization in AI & Robotics from Andhra University, Andhra Pradesh, India, in 2009. Presently, he is an Assistant Professor and HoD of IT-Department in Vignan's Institute of Engineering for Women, Visakhapatnam, Andhra Pradesh, India from. He had 11 years of teaching experience both in Physics and Computer Science. He is a member of Computer Society of India.

K. Suraya latha received the M.Sc degree in Computer Science from Andhra University, Andhra Pradesh, India, in 2004 and pursuing the M. Tech (Computer Science Engineering) degree from JNTU Kakinada. She worked as Asst. Professor in Prakasam Engineering College, Andhra Pradesh, India, from 2005 to 2010.

Dr. S. Koteswar Rao. Received his B.Tech from Govt. College of Engineering, Anantapur, JNTU (AP) in Electrical Engineering, the Ph.D (Engg) degree from college of Engineering, Andhra University, Visakhapatnam in 2010. He wrote a Text book on 'Passive Underwater Target Tracking using Bearings-only Measurements'. He has presented 85 papers in National and International Conferences. He has got 29 paper publications in International Journals. His areas of research interests includes Image Processing, Digital signal processing, biomedical signal processing, Statistical Signal Processing, Adaptive signal processing. He retired as Scientist- 'G' from DRDO, NSTL in 2012. Presently, he is Principal and Senior Professor in ECE Dept, Vignan's Institute of Engineering for Women (VIEW), Visakhapatnam.