A Novel Method of White Blood Cell Segmentation and Counting

J.B. Nemane¹, V. A. Chakkarwar²
Dept. of CSE, Govt. College of Engg. Aurangabad (MS), India
¹jyotinemane@yahoo.co.in,²vrush.a143@gmail.com

Abstract: The quantity of WBC Cell is very important for the doctor for doing diagnosis for various diseases. So, exact counting of cell plays very important role. Manual process of counting is a very time consuming task, and has less precision. So an automatic counting process is very important to get high precision results. For the counting of cell requires segmentation of the cell, for classifying purpose. In this paper, we propose a new algorithm to segment the image. Firstly, the proposed algorithm is used to remove noise. Second algorithm used for segmentation of image is by using watershed algorithm. The watershed is just a method to determine separate regions can be counted efficiently.

Keywords: Gradient, Segmentation, White Blood Cell

1. INTRODUCTION

Cell segmentation is a challenging problem due to both the complex nature of the cells and the uncertainty present in video microscopy. Manual methods for this purpose are onerous, imprecise and highly subjective, thus requiring automated methods that perform this task in an objective and efficient way.

Automated detection and classification of white blood cells is a major step in diagnosis of several diseases like Acute Lymphoblastic Leukaemia. The traditional procedure requires a haematologist to manually count and classify the cells with the help of a microscope.

An automated diagnosis system will alleviate the workload and the influence of subjective factors. Automated detection involves removal of red blood cells and platelets from the background. The main drawback of the existing methods is their inefficiency in handling cell images originating from different sources and environment. There are three types of cells in normal human blood: red blood cells (RBCs), white blood cells (WBCs) and blood platelets. Generally, RBCs are simple and similar. While WBCs contain nucleus and cytoplasm and can be categorized into five classes:

1. Neutrophil
2. Eosinophil
3. Basophil
4. Monocyte
5. Lymphocyte

The nucleus of each of the above types has a unique shape, and this is the most important feature used in cell classification. In addition to the shape of the nucleus, the “phils” category has granules with in the blood cell where as “cytes” category does not have granules.

2. REVIEW WORKS

White blood cell segmentation from the cell image background involves subtraction of red blood cells, platelets and other objects mixing in the microscopic images. Illumination inconsistencies and cell occlusion are the main reasons that make cell segmentation challenging.

Cell classification has widespread interest especially for clinics and laboratories. For example, patient’s blood cells counting is used to extract information about other cells that are not normally present in peripheral blood but may be released in certain disease processes by the haematologist.

Patient’s blood cells counting were performed manually by medical technologists by viewing slide prepared with blood sample of the patient under microscope. A manual count will also give information about other cells that are not normally present in peripheral blood but might be released in certain disease.

Unfortunately, the accuracy of cell classification and counting is strongly affected by individual operator’s capabilities. In particular, the identification and differential count of blood’s cell is a time consuming and repetitive task that can be influenced by operator’s accuracy and tiredness. In an effort to overcome the tedious and time consuming task of human experts in counting white blood cells in bone marrow or peripheral blood, many automated techniques have been proposed.

Thresholding was one of the earliest methods implemented for image segmentation. The simplicity of implementation and its intuitive properties gave image thresholding a central
A Novel Method of White Blood Cell Segmentation and Counting

Position in applications of image processing. In the case of cell segmentation, thresholding was followed by morphological operations in most of the cases. Thresholding is computationally cheap and fast.

Region-based approaches work on a certain criterion of homogeneity, usually of same or similar brightness or colour. All the pixels that satisfy the homogeneity criterion were grouped together into a region. These approaches include region splitting and merging approaches, seeded region growing approaches.

Model-based approaches: In this approach, the nucleus and the cytoplasm are extracted by building models. The cell model must include features that discriminate leukocyte regions from others and, intuitively encourage the formation of round, homogeneous regions of certain size.

Fuzzy methods have gained sufficient significance in recent times and are now used in major image segmentation techniques. In fuzzy technique was used with the aim to allow a good processing of both vagueness and in determination characteristics of images, and the analysis of monochrome instead of colour images.

R. Sukesh Kumar et al. discussed about an approach for colour image segmentation using higher order entropy as a textural feature for determination of thresholds over a two dimensional image histogram. Two basic models for colour images are the RGB (Red, Green, Blue) colour model and the HIS (Hue, intensity, saturation) colour model. Two methods of colour image segmentation used RGB space as the standard processing space. These techniques might be used in blood cell image segmentation. Colour images are very rich source of information, because they provide a better description of a scene as compared to gray scale images. Hence, colour segmentation becomes a very important issue.

3. PROPOSED METHOD

Let $X = \{x_i\}_{i=1}^n$ be a set of n data points in d-dimensional space, $\mathbb{R}^d$. The multivariate kernel density estimator with Gaussian kernel and a symmetric positive definite d & d bandwidth matrix $H$, computed at the point $x$ is given by

$$f(x) = \frac{1}{n(2\pi H)^{d/2}} \sum_{i=1}^{n} \exp \left( -\frac{1}{2} d^2(x, x_i, H) \right)$$

Where

$$d^2(x, x_i, H) = (x - x_i)^T H^{-1} (x - x_i),$$

is the Mahalanobis distance from $x$ to $x_i$. By computing the gradient of $f^h(x)$

$$\nabla f^h(x) = \frac{n^{-1}}{n(2\pi H)^{d/2}} \sum_{i=1}^{n} (x - x_i) \exp \left( -\frac{1}{2} d^2(x, x_i, H) \right)$$

After some algebra we have

$$m(x) = H \frac{\nabla f^h(x)}{f^h(x)}$$

Where

$$m(x) = \frac{\sum_{i=1}^{n} x_i \exp \left( -\frac{1}{2} d^2(x, x_i, H) \right)}{\sum_{i=1}^{n} \exp \left( -\frac{1}{2} d^2(x, x_i, H) \right)} - x.$$ 

Assume now that the data points $x_i$ are extracted from an input image. Then the vector components of $x_i$ contain both the spatial lattice information $x_{si} = (x_i; y_i)^T$ and range information $x_{ci} = (c_{i1}; c_{i2}; c_{i3})^T$ where $c_{i1}$; $c_{i2}$ and $c_{i3}$ are three color components at position $(x_i; y_i)$. Then $x_i = (x^T_{si}; x^T_{ci})^T$ is a point in joint spatial-range domain. We assume that the bandwidth matrix $H$ is diagonal having the diagonal terms equal to $\sigma^2$s for the spatial part and $\sigma^2$R for the range part. Then the Mahalanobis distance in Eq.(2) can be rewritten as:

$$d^2(x, x_i, H) = \frac{||x_{si} - x_{si}||^2}{2 \sigma^2 S} + \frac{||x_{ci} - x_{ci}||^2}{2 \sigma^2 R}$$

Where

$$||x_{si} - x_{si}||^2 = (x - x_i)^T + (y - y_i)$$

The color distance in the RGB space is defined as,

$$||x_{ri} - x_{ri}||^2 = (r - r_i)^2 + (g - g_i)^2 + (b - b_i)^2$$
Let \( Y = \{ y_i \}_{i=1}^n \) and \( Z = \{ z_i \}_{i=1}^n \) be a set of data points in joint spatial-range domain of the processed data and the result filtered image respectively. The MS filter is formulated as follows.

**Algorithm 1:**

1. Set \( t = 0 \), and initial the error \( \varepsilon \).
2. Set \( Y_0 = x_i \) for \( i = 1 \ldots n \).
3. Calculate the convergent value \( y_i \) for \( i = 1, \ldots, n \) by

\[
Y_{i}^{t+1} = \frac{\sum_{j=1}^{n} z_i \exp \left(-\frac{1}{2}d^2(y^T z_{i}, H) \right)}{\sum_{j=1}^{n} \exp \left(-\frac{1}{2}d^2(y^T z_{i}, H) \right)}
\]

3.1 Update

3.2 If \( d^2(y_i^t, y_i^{t+1}, H) > \varepsilon \) go to step 4.

4. Set \( Z_i = (Xs_i^T, Y_{r_i}^T ) \), where \( Y_{r_i} \) is the range part of the convergent value \( Y_i \).

Each data point \( y_i \) is initialized at \( x_i \) and during calculating the convergent in step 3.1 and 3.2, \( y_i \) moves iteratively along the gradient direction in both spatial and range domain. Finally, it converges to a local mode in the joint spatial-range domain. The advantage of the MS filter is that the image structure does not change during iterations. Then the MS filter will achieve better image structure preservation. The MS filter can remove noise while preserving edges or boundaries of the local structure by choosing the suitable \( \sigma^2 \) and \( \gamma_r \).

**3.1 Region Merging Algorithm**

Given a data point of the filtered image \( z = ((x, y)^T, (y)^T)^r \), its four neighbors are defined as follows:

\[
Z_0 = ((x-1, y)^T, (y)^T)^r_0,
\]
\[
Z_1 = ((x+1, y)^T, (y)^T)^r_1,
\]
\[
Z_2 = ((x-1, y)^T, (y)^T)^r_2,
\]
\[
Z_3 = ((x+1, y)^T, (y)^T)^r_3
\]

On the other hand, \( z \) is called the center of its neighbors.

**Region Merging Algorithm**

1. If \( ||Y^T, (y)^T r_i|| > \sigma^2 R \) then return.
2. If \( z_i \) is not labeled as 0, then return
3. Label \( z_i \) as \( \text{Idx} \)
4. Recursively call
   a. Region Grow(\( z_i \), \( z_i0 \), \( \text{Idx} \)),
   b. Region Grow(\( z_i \), \( z_i1 \), \( \text{Idx} \)),
   c. Region Grow(\( z_i \), \( z_i2 \), \( \text{Idx} \)),
   d. Region Grow(\( z_i \), \( z_i3 \), \( \text{Idx} \)),
5. return

Let \( L = \{ 1, 2, \ldots \} \) be a set of labels of regions, and the unlabeled region is labeled as 0. According to the first assumption, the Region growing algorithm is designed as follows. Given a data point \( z \) which is labeled as \( \text{Idx} \) (\( \text{Idx} \in L \)) its neighbors are also labeled as, \( \text{Idx} \) if \( ||Y^T, (y)^T r_i|| > \sigma^2 R \).

**4. WATERSHED ALGORITHM**

For a segmentation purpose, the length of the gradient is interpreted as elevation information.

Gradient Image Relief of the gradient

![Figure 1 (a) Gradient Image (b) Relief of the gradient](image1)

Figure 1 (a) Gradient Image (b) Relief of the gradient

**Figure 2 Watershed of the gradient**
A Novel Method of White Blood Cell Segmentation and Counting

**Figure 3 Watershed of the gradient (relief)**

**Algorithm for Watershed Algorithm**

1. Use the red channel for a better clearance of particles.
2. Apply quality enhancement functions such as deblurring, denoising.
3. Use Granulometry as RBC size estimator.
4. Estimate WBC size upon medical references and RBC size.
5. Proper binary conversion (Otsu method).
6. Use Proper edge detection method (Canny solution).
7. Merging binary and edge detection outputs (Otsu and Canny outputs).
8. Filling current image and make a solid objects.
9. Separate WBCs from RBCs (Mask function).
10. Improving the quality of the results of Watershed Algorithms. The quality of the best can be further improved in watershed algorithm by using colormap technique, some image processing filters and the Granulometric Analysis.

**4.1 Colour Map Technique**

For performing the segmentation of image in watershed algorithm the basic criteria is to get border of the cell distinctly so that we can detect the cell more accurately. To obtain the accuracy we can perform the channel split of the image i.e. R, B and G channel. The detect the border in R channel and then further respective steps can be perform on other channels of the image.

**A. Image Processing Filters**

The quality of the picture contents can be further improved by applying image processing filters like Deblurring (Enhancing Contrast) Denoising (eliminating Blood platelets, and other unwanted noise) Smoothing (reducing the number of connected components).

**B. Granulatity Analysis**

In Images RBCs are found a swell as WBCs. Hence we need two separated images for RBCs and WBCs. In this paper, we used size characteristics as an effective factor to distinguish two main objects. Granulometry algorithm is divided into following steps:

- e. morphological openings
- f. Structure Elements (SE model referring to object shapes)
- g. Maximum Size of SE (which is calculated according to primary estimation based on haematology science, and consideration of a noticeable point to enhance algorithm accuracy and speed)
- h. Increasing size SE up to max (Loop Function)
- i. \( Ps (k) = \) Sum of all the pixels in the image, opened with k-th SE.
- j. Normalized function \( N (k) = \frac{PS (k)}{PS (0)} \), where \( PS (0) \) is the direct of the original image, (SE Size=1 pixel).
- k. The Granulometric density function can be defined as \( G (k) = N (k+1) - N (k) \).

**Figure 4: WBC Types in Blood Smear image**
A Novel Method of White Blood Cell Segmentation and Counting

Separation of WBC and RBC

So far, a solid binary image is formed; everything is acceptable to apply the Watershed algorithm. Different objects including WBC, RBC are separated into labelled regions for each object. It is possible to enhance Watershed algorithm by employing watersnake. But in separated labelled regions there is not any discrimination between WBCs and RBCs, they cannot be distinguished from another. Thus before applying the Watershed method on the whole image we’d better form a new image only with WBCs based on size characteristics.

4.2 Proper Binary Image

Watershed procedure needs binary image as input. Because of intrinsic characteristics of staining blood smear, particles color and their intensity cannot be easily distinguished from the background. All binary conversions are not 100% practical for blood smear images as shown in Fig. 6.

![Figure 5: Watershed algorithm for the single WBC cell.](image)

![Figure 6: (Left to Right): average binary, Otsu binary method](image)

4.3 Proper Edge Detection

Otsu algorithm has better efficiency in comparison with average algorithm, but it is not satisfactory. To achieve a desirable image, membrane detection should be followed. Experimental results proved that canny filters demonstrate borders of WBC.

4.4 Circle Mask

We notice that morphological operators such as Erosion or Dilation are not the best choices and always need some restrictions which are not available in all blood smear image such as agglutination between RBCs which make problems in using these simple morphological operations. A reliable method by our proposed method can consist of:

- Applying Granulometry over blood smear image and saving density value pixels referred to RBC size (that’s called intensity_area_prime_max)
- Estimating WBC size from RBC size calculated from Granulometry, medical references and an acceptable marginal range: \( C_1 = \text{RBC size} \times 2 - 5 \) (margin)
- Moving surrounded mask over blood smear image and detecting the exact matching objects of the same size and then changing all its intensity into 0.
- Applying circular mask function and Granulometry in a closed loop by an initial radius value (\( C_1 = \text{RBC size} \times 2 - 5 \text{ (margin)} \)) and then increasing integer radius(C1) value until we get similar density value pixels output as we had in first running Granulometry over main image. Saving WBC indicator in a new image variable. Omitting possible noises by applying median filter 5*5.

5. CONCLUSION

In cases of the RGB color space, the white blood cell images can be segmented by using the proposed algorithm, but the segmented images are highly over segmented in the cytoplasm region. This paper has presented a new literature based on Granulometry, mask function and pre-processing considerations to achieve a fully automated detection and segmentation of blood cells in order to separate labeled regions for Watershed algorithm.

REFERENCES
A Novel Method of White Blood Cell Segmentation and Counting

[2]. JumaAl-Muhairy, Youssef Al-Assaf," Automatic white blood cell segmentation based on image processing”, IFAC, 2005