

Methods to reduce over-segmentation in red blood cell images.

Sheetal Jadhav

Department of Electronics and

Communication Engineering

National Institute of Technology

Surathkal, Karnataka

Email: 14ec222.sheelat@nitk.edu.in

Abstract—The study or analysis of the morphology blood cells and also their count can provide us with great information, which can be useful to detect foreign parasites, which may at times be harmful in the cells and hence able to detect any diseases. However, the microscopic study and analysis of blood cells is extremely a difficult and tedious job, but naturally due to minute sizes of the cells, their peculiar morphology, or simply because of the simple fact that they all tend to form blobs, which makes it very difficult to manually perceive them. However, image segmentation has made this study a lot easier. The task of image segmentation is to partition an image into non-overlapping regions based on intensity or textural information of the pixels. This article presents a better implementation of well-known region-based image segmentation method, the watershed transform. This implementation was intended mainly for segmenting images specifically of red blood cells and get an exact count of them in the end. The count is important because RBCs contain haemoglobin, which carries oxygen to our body's tissues. The number of RBCs we have can affect how much oxygen our tissues receive. Our tissues need oxygen to function effectively. This article follows two methods of placing the seed points in the image to perform watershed segmentation and avoid the over-segmentation caused.

I. INTRODUCTION

Segmentation, which means partitioning of an image into background and objects of interest such as cell nuclei, is central to biological studies. Automated image segmentation for cell analysis is generally a difficult task due large variability and complexity of data with across cell types and application contexts, with weak contrast, touching nuclei, diffused background and varying size and shape of cell nuclei all posing challenges to existing methods. Nevertheless, screening the literature published on this subject, we find that vast majority of cell segmentation methods are based on a basic few approaches as follows:

- 1) Intensity Thresholding
- 2) Feature Detection
- 3) Morphological Filtering
- 4) Region Accumulation
- 5) Deformable model fitting

There exist innumerable approaches for single cell segmentation based on the outline of the above methods. Some of these include global/adaptive thresholding which is a straightforward method for situations where cell nuclei are well separated and have relatively evenly distributed background intensities.

A review of adaptive thresholding is found in watershed methods, which are parameter free and often used for more complicated problems, this is based on a recursive immersion simulation process. Another popular approach uses active contours, such as snakes and various set level approaches, in which objects are represented by smooth contours or level sets. Other algorithms address specific problems by exploiting the unique features of given images. For example, Wu et al. (1998) used an elliptical model for cell detection and segmentation; Berlemont et al. (2007) used a feature adapted beamlet model to detect DNA filaments in fluorescent microscopy; Wang et al. (2007) proposed a method that combines intensity and shape information for cell segmentation; Li et al. (2008) developed a gradient flow tracking method to address the segmentation problem of touching cell nuclei. As mentioned earlier this article puts forth two different methods for seed generation for watershed transform in order to avoid over-segmentation that usually happens due to the watershed transform method.

II. WATERSHED TRANSFORM

The watershed transform is the method of choice for image segmentation in the field of mathematical morphology. The watershed transform can be classied as a region-based segmentation approach. The intuitive idea underlying this method comes from geography: it is that of a landscape or topographic relief which is ooded by water, watersheds being the divide lines of the domains of attraction of rain falling over the region. An alternative approach is to imagine the landscape being immersed in a lake, with holes pierced in local minima. Basins (also called catchment basins) will fill up with water starting at these local minima, and, at points where water coming from different basins would meet, dams are built. When the water level has reached the highest peak in the landscape, the process is stopped. As a result, the landscape is partitioned into regions or basins separated by dams, called watershed lines or simply watersheds. When simulating this process for image segmentation, two approaches may be used: either one first finds basins, then watersheds by taking a set complement; or one computes a complete partition of the image into basins, and subsequently finds the watersheds by boundary detection. To be more explicit, we will use the expression watershed transform to denote a labelling of the

image, such that all points of a given catchment basin have the same unique label, and a special label, distinct from all the labels of the catchment basins, is assigned to all points of the watersheds. We note in passing that in practice one often does not apply the watershed transform to the original image, but to its (morphological) gradient. This produces watersheds at the points of grey value discontinuity, as is commonly desired in image segmentation. However, watershed methods, which usually yield mosaic image representations of original images, often suffer from serious over-segmentation problem. Marker-controlled and hierarchical watershed methods are then often employed as remedies to the over-segmentation problem by either providing strong prior information (marker) to pre-process the image before applying watershed method (Beuchar, 1992), or by applying merging rules to post process over-segmented regions from the watershed transform (Beuchar, 1994; Najman & Schmitt, 1996; Umesh Adiga & Chauduri, 2001; Mao et al, 2006).

III. INITIAL SEGMENTATION

Prior to applying watershed transform, the usual adaptive thresholding approach was taken and performed on the infected red blood cell images with us. With this approach it became easier to segregate the cells which were not in a cluster and relatively easy to segment from the ones which formed blobs and cluster and segmenting them was a tedious and not a straightforward task. As shown in the figure, there is the original image with the red blood cells and three other images which indicate the segmentation of the cells which were easiest to segment, the relatively easy but still clustered cells and the cells which had formed huge blobs and were not at all straightforward to segment, respectively. Whereas the conventional thresholding operator uses a global threshold for all pixels, adaptive thresholding changes the threshold dynamically over the image. This more sophisticated version of thresholding can accommodate changing lighting conditions in the image, e.g. those occurring as a result of a strong illumination gradient or shadows.

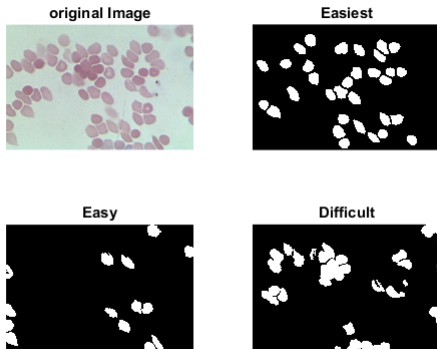


Fig. 1. Initial Segmentation of cells

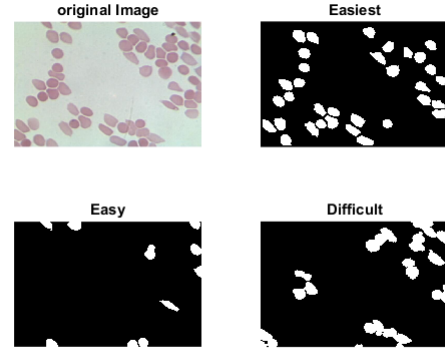


Fig. 2. Initial Segmentation of cells

IV. SEED GENERATION BASED ON CENTROIDS

The use of the watershed transform in complex medical cell images ends up in over-segmentation. In order to obtain regions that belong to the objects of interest, generation of seeds in such a manner so as to create watersheds almost exactly at the boundaries of the cells, so that we can get an exact count of them is extremely essential. This method is more concentrated to be applied on images in which cells are clustered and in the form of blobs. In this method firstly all the blobs are separated and the number of cells in each blob are counted and displayed. Then we calculate or plot the centroid of each cluster (blob) of cells. Once that is done, we go on the divide this blob in four quadrants with its centroid at the centre. The centroid of the four quadrants is again calculated. The centroids calculated in the step are used as seed points to apply the watershed transform algorithm. If the blob is to huge, further division of the blob is done, till the point it seems that there is no cluster of cells. This approach of making centroids as the seed points was adopted mainly because,

- 1) Centroid is always symmetric with respect to an object in this case a cell, hence it would perfectly act as a seed point to generate watersheds on the cell images.
- 2) ***ask sir for more points if any ***

This image with seed points was then given to the watershed transform algorithm for final segmentation. The image below shows the final segmentation of the red blood cells images and also the over-segmented image as well. An image of the seed points (seed points are black) is shown below.

V. SEED GENERATION BASED ON DISTANCE FROM THE BOUNDARY OF A CELL

This is the second method which was adopted to generate seed points in order to perform watershed transform for segmentation. In this approach in a region of approximately equal to one cell size we calculate the distance of each point form the boundary of cell, and select that point which is farthest from the boundary. This selected point is taken as a seed point. Thus the image with the seed points is then provided to the watershed algorithm for final cell segmentation. This approach



Fig. 3. Seed points based on centroids

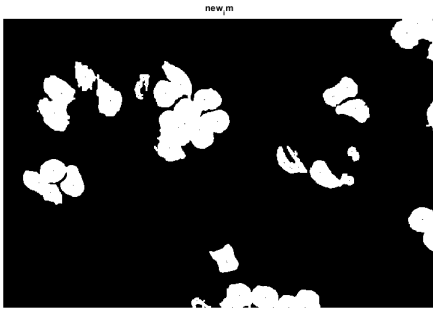


Fig. 4. Seed points based on distance from the boundary

proves a little better result than the previous method mentioned because,

1) In spite of clustering of cells guaranteed at least one seed point is generated for every cell, which is very essential thing to form watersheds and this was not necessarily happening with the previous method due to the clustering of cells.

2) More and appropriate segments are generated even in highly clustered patch of the image.

The figure shows a binary image with the seed points (seed points are black) is shown, also the final segmented image with this method and the over-segmented image is as well shown.

VI. RESULTS & SUMMARY

The results of both the methods have been summarized here. The first figure shows the over segmented image and the second figure shows the better segmented image for each of the two methods of two different. This simulation was done on MATLAB.

VII. CONCLUSION

The watershed transform is in original form a segmentation tool, that produces over-segmentation, but this drawback can be extinguished by several methods, especially with aid of other segmentation technique. As can be deduced from our results, the methods incorporated are quite novel and promising approaches. The approaches put forth in this article were intended to simplify the segmentation process of mainly the red blood cell images in order to get their count. The

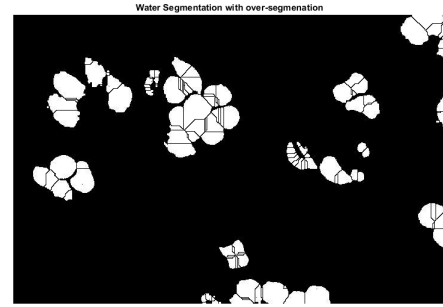


Fig. 5. Over-Segmented RBC image

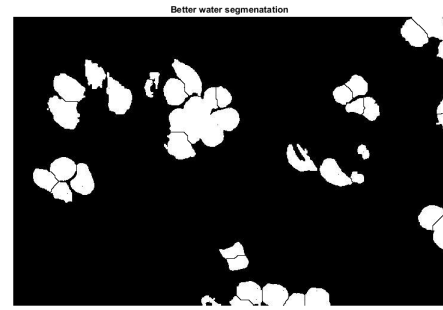


Fig. 6. Improved Segmentation with first method

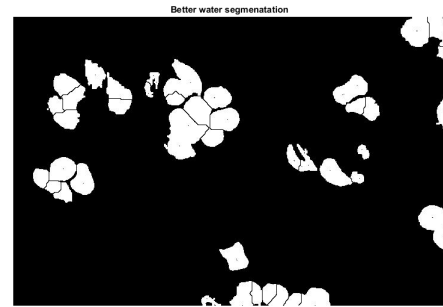


Fig. 7. Improved Segmentation with second method

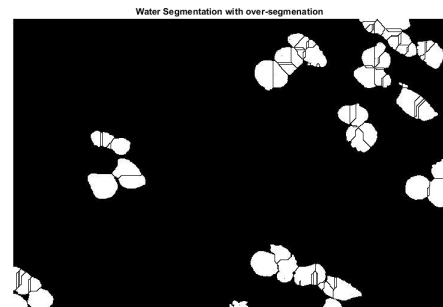


Fig. 8. Over-Segmented Image

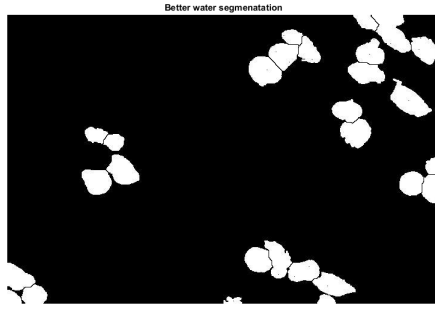


Fig. 9. Improved Segmentation with first method

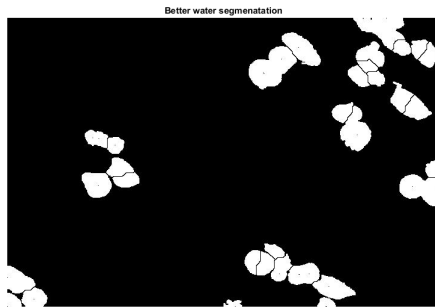


Fig. 10. Improved Segmentation with second method

morphology of red blood cells suggest they do not have a nucleus, which makes the intensity levels uniform inside the cell. If there were a nucleus in the cell, there would be an intensity gradient which would make the task of placing seed points much easier, as one could place it on the nucleus itself to get a perfectly segmented image. However, the absence of the nucleus makes it difficult for us to place the seed point, and thus we have come up with the idea to place the seed points so as to get nearly perfected cell images in order to get the cell count. Out of the two methods mentioned above, the second method seems to give better results than the first one, as the first method does not guarantee a seed point within each cell. When each cell has a seed point, it becomes easy to generate the watersheds and hence give segments matching the actual cell boundaries. In the future we would like to continue our research over watershed-based classification by using edge linking and merging and other features.

REFERENCES

- [1] Mariela Azul Gonzalez, Gustavo Javier Meschino, Virginia Laura Bal-larin, *Solving the over segmentation problem in applications of Water-shed Transform* \LaTeX , 0.5em minus 0.4em Grupo de Procesamiento de Imagenes. Departamento de Electronica. Facultad de Ingenieria. Universidad Nacional de Mar del Plata, Argentina
- [2] LAMIA JAAFAR BELAID AND WALID MOUROU, *IMAGE SEGMENTATION: A WATERSHED TRANSFORMATION ALGORITHM* \LaTeX , 3rd ed.
- [3] Wojciech Bielecki *Oversegmentation avoidance in watershed-based algorithms for color images* \LaTeX , 3rd ed.