



**HAL**  
open science

## Automated Identification of Plasma Cell in Bone Marrow Images

Mourtada Benazzouz, Ismahan Baghli, Mohammed Lamine Benomar, Youcef Benmouna, Mohammed Ammar

► **To cite this version:**

Mourtada Benazzouz, Ismahan Baghli, Mohammed Lamine Benomar, Youcef Benmouna, Mohammed Ammar. Automated Identification of Plasma Cell in Bone Marrow Images . 2015. hal-01466770

**HAL Id: hal-01466770**

**<https://hal.science/hal-01466770>**

Preprint submitted on 13 Feb 2017

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Automated Identification of Plasma Cell in Bone Marrow Images

Mourtada BENAZZOUZ  
Computer science department  
Genie Biomedical Laboratory  
Abou Bekr Belkaid University  
Tlemcen, Algeria

Email: m\_benazzouz@mail.univ-tlemcen.dz

Ismahan BAGHLI  
Amine BENOMAR  
Genie Biomedical Laboratory  
Email: i\_baghli@mail.univ-tlemcen.dz

Youcef BENMOUNA  
and Mohammed AMMAR  
Genie Biomedical Laboratory  
Abou Bekr Belkaid University

**Abstract**—The classification and the count of different White Blood Cells (WBC) in microscopy images allow the assessment of a wide range of important hematology pathologies such Myeloma. This article aims particularly myeloma disease; this pathology is manifested by a proliferation of a type of cell called plasma cells. This paper presents a robust and accurate novel method for processing WBC(leukocyte) using a combination of ideas. The segmentation of cells is achieved in two phases, the first is to extract the nucleus from the green channel of the image using the otsu's thresholding method, and the second operates through a region growing on the nucleus in order to delimitate a cytoplasm. After that, a cell's characterization step is performed where desired features can be extracted in order to classify types of leukocytes and thus allow the counting of plasma cells. The proposed process have been successfully tested to a large number of images, showing promising results in terms of segmentation and classification of plasma cells, and also applied to varying cell appearance and image quality, encouraging future works.

## I. INTRODUCTION

There are three types of cells in the bone marrow: 1) red cells or erythrocytes are simple, similar and delivering oxygen to the body tissues. 2) platelets, also called thrombocytes, are blood cells whose function is to stop bleeding. 3) leukocytes contain nucleus and cytoplasm are the cells of the immune system that are involved in protecting the body against both infectious disease and foreign invaders. The proportion of counts of distinctive types of white blood cells in the bone marrow, called differential counts, provides precious information to doctors for diagnosis. Enumerate leukocyte is still performed by hematologists, being indispensable in diagnostics in several diseases. In general, leukocytes are usually manually classified in laboratories by using microscopes; this traditional method is very tedious and time consuming. Therefore, an automatic differential counting system is preferred.

In previous works, there are variety methods for segmentation and wbc's classification. We only mention the work of the last ten years. Haralicks techniques are performed by [1] in order to enhance the characterization of bone marrow images regarding texture information about the leukocyte nucleus and cytoplasm.

In [2], the idea is to investigate information about the nucleus alone, is much easier than the segmentation of the entire cell followed by a classification of WBC with a neural network.

Combining pixel-wise classification with template matching to locate erythrocytes and use a level-set approach in order to get the exact cell contours of leukocyte nucleus and plasma regions as well as erythrocyte regions was presented in [3]. [4] after a segmentation of the nucleus, focuses on recognition steps, based on reduction features by principle component analysis (PCA) to obtain suitable features. The genetic algorithm based k-means clustering approach is used to classify the five kinds of leukocyte in the reduced dimensions. In a another paper [5], the authors begin with the nucleus segmentation based the image pre-processing with SMMT (selfdual multiscale morphological toggle has used to ensure the accuracy of two well-known image segmentation techniques, namely, watershed transform and Level-Set methods, then the cytoplasm region is identified with two different schemes, based on granulometric analysis and on morphological transformations. Recently [6] composed an algorithm of markov random filed with k-means and enhancement methods that provides us the segmentation stage truly without luminance and unsuitable stained smear.

Finally, we can also mention the recent work of our research team. The first propose a segmentation scheme using pixel classification based on the fusion of information. The adopted model is guided by the two strategies offered by information fusion, i.e. classifying separately the data from different sources(color spaces) then merging decisions, or combining these data to classify them [7]. The second [8] deals with new segmentation framework based on evidence theory, called ESA (Evidential Segmentation Algorithm).

The rest of paper is organized as follows. The next Section precisely describe the proposed WBC segmentation approaches, in Section 3 presents a detailed of characterization of cells. The effectiveness of our framework is verified by a set of experiments in Section 4. Finally, we conclude in Section 5.

## II. IMAGE SEGMENTATION

We start an image pre-processing method to separate cells which is an efficient way for detection and classification of leukocytes. We performed multiple tests over color histograms and according number of authors [9], we arrive at the following conclusion: a threshold value of 100 from  $[0 - 255]$  give an acceptable discrimination between nuclear and non-nuclear pixels.

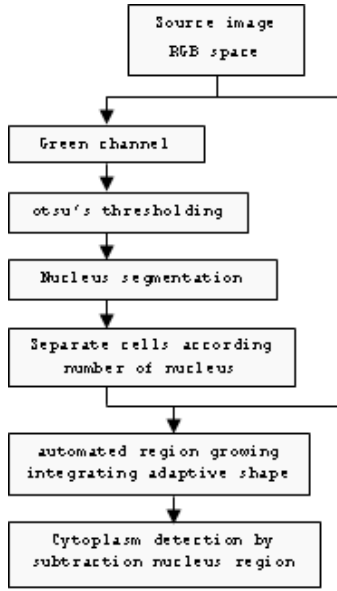


Fig. 1. Segmentation stages

A non-supervised of nucleus region detection is processed before nucleus color segmentation for image  $I$  dimension reduction, using the G channel from RGB color coordinates.

This method leads to image reduction from  $1024 \times 768$  pixels to  $200 \times 200$  pixels; the calculation takes into account the nuclear center, so a square smaller than  $I$  can be specified containing the leukocyte.

#### A. Otsu's thresholding method

Otsu is a nonparametric and unsupervised method of automatic threshold selection for picture segmentation. Otsu is based on a very simple idea: Find the threshold that minimizes the weighted within-class variance. the weighted within-class variance is :

$$\sigma_w^2(t) = q_1(t)\sigma_1^2(t) + q_2(t)\sigma_2^2(t) \quad (1)$$

where the class probabilities are given by:

$$q(t) = \sum_{i=1}^t P(i) \quad (2)$$

and the class means is calculated :

$$\mu(t) = \sum_{i=t+1}^l \frac{iP(i)}{q_2(t)} \quad (3)$$

which  $t$  is maximum range of green channel in histogram and  $p(i)$  is a probability green value  $i$ . All kind of nucleus can be extracted with Otsu method. Figure 2 shows the result.

#### B. Region Growing

Region Growing can be presented as an algorithm for segmentation of intensity images which is robust, rapid, and free of tuning parameters. The method, however, requires the input of a number of seeds, either individual pixels or regions, which will control the formation of regions into which the

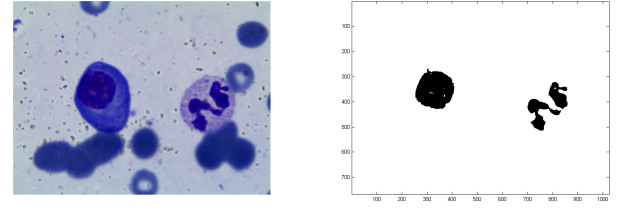


Fig. 2. Otsu's segmentation

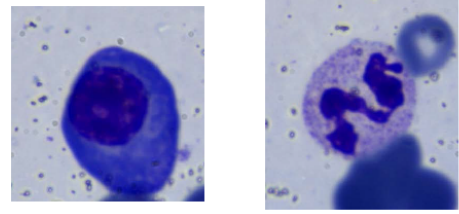


Fig. 3. Separate cells

image will be segmented[10]. The concept like that of other methods by pixel aggregation, is to start with a point that meets a detection criterion and to grow the point in all directions to extend the region.

#### 1) Shape Prior for cytological image segmentation :

For segmentation purposes we define a region of interest as a blob(nucleus region obtained in previous step by otsu), exhibiting a high contrast relative to its local background. The best boundary for the region is a set of connected pixels exhibiting predefined contrast properties.

Since the shape of WBC is always round or oval, we added the circularity index as a second criterion in the region growing algorithm after aggregation of neighboring pixels; i.e. we will test the form obtained at each iteration by comparing it with a of circularity threshold. The major goal of this new criterion is to avoid the algorithm to grow to an area belonging to a neighboring cell that touches or overlaps with our Region of Interest.

$$Circularity = \frac{1}{4\pi} \times \frac{P^2}{A} \quad (4)$$

where  $P$  is Perimeter and  $A$  is Area [11]. In the case of a circle, the roundness parameter is minimal and is equal to 1.

In order to do a segmentation assessment, we compare the results obtained by our segmentation scheme with ground-truth

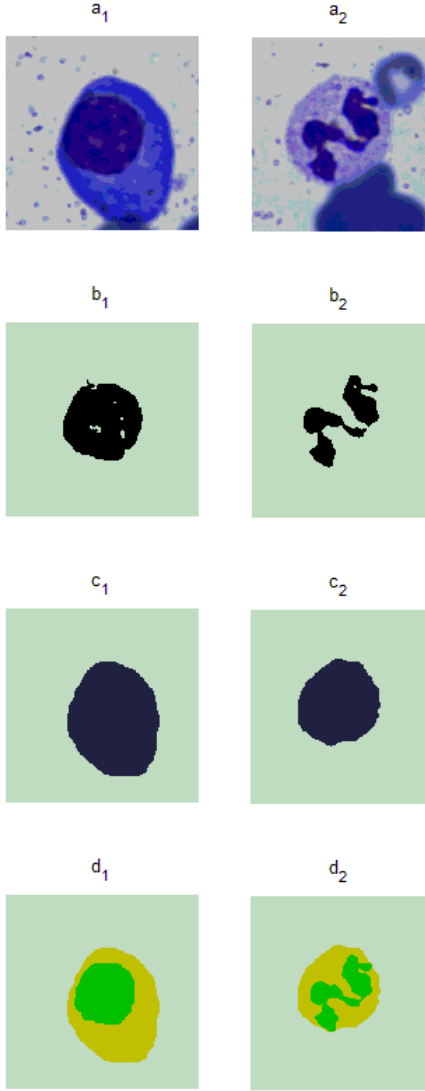


Fig. 4. Segmentation results: ( $a_1, a_2$ : original cells), ( $b_1, b_2$ : nucleus by OTSU), ( $c_1, c_2$ : region growing), ( $d_1, d_2$ : yellow cytoplasm obtained by subtracting from the two last stages)

images(obtained manually); The Table I shows the segmentation results of the two surfaces (nucleus and cytoplasm), which are the major components of our ROI.

|                  | Nucleus | Cytoplasm |
|------------------|---------|-----------|
| <b>Precision</b> | 96.67%  | 91.44%    |
| <b>Rappel</b>    | 98.73%  | 94.04%    |

TABLE I. SEGMENTATION ASSESSMENT.

The first observation is that the nucleus region is highly satisfactory using the OTSU method : Fig 4 ( $b_1, b_2$ ).

Region growing is shown in the next line in Fig 4 ( $c_1, c_1$ ) : with two examples, the first is oval shaped  $c_1$  and the second is circular  $c_2$ . In the first iterations, we note that the region growing follows properly the shape of the cell. And when there is an overlap of the cells, the criterion of circularity prevents deformation of the cell.

These results are significantly better compared to our previous work in terms of segmentation quality and in terms of time and computing [7], [8].

### III. CHARACTERIZATION OF PLASMA CELL

In order to identify the plasma cell in the final stage of our work, it is necessary to quantify with attributes. Most work characterizing the WBC consider two types: Texture and Form.

#### A. Texture attributes

Texture is defined as a spatial repetition of the same pattern in different directions in space. The concept of texture is used to translate a homogeneous appearance of the surface of an object over an image. So the texture is manifested by visual information that can qualitatively describe using the following adjectives: coarse, fine, smooth, mottled, grainy, mottled, regular or irregular.

GLCMs co-occurrence matrices contain the average space of second order. Fourteen indices defined by Haralick [12] corresponding to descriptive features of textures. After a selection of relevant attributes, we present here only five of these indices:

- **Energy**: it is a parameter that measures the uniformity of the texture. It reached high values when the distribution of the gray levels is constant or periodic shape.

$$Energy = \sum_i \sum_j P_d^2(i, j) \quad (5)$$

- **Homogeneity**: when we find the same pair of pixels, this index is high, for example uniform image, texture or periodic in the direction of translation.

$$Homogeneity = \sum_i \sum_j \frac{P_d(i, j)}{1 + |i - j|} \quad (6)$$

- **Entropy**: measures the complexity of the image. When the values of the co-occurrence matrix are almost all the same, the entropy is high. It helps to characterize the degree of granulation of the image.

$$Entropy = - \sum_i \sum_j P_d(i, j) \cdot \log P_d(i, j) \quad (7)$$

- **Correlation**: describes the correlations between the rows and columns of the co-occurrence matrix.

$$Correlation = \frac{(i - \mu)(j - \mu) P_d(i, j)}{\sigma_x \sigma_x} \quad (8)$$

- **Contrast**: This parameter measures the local variations, if these variations are significant, so the contrast is high. This parameter is strongly uncorrelated with energy.

$$Contrast = \sum_{n=0}^{Ng-1} n^2 \left\{ \sum_{i=1}^{Ng} \sum_{j=1}^{Ng} P_d(i, j) \right\}, |i - j| = n \quad (9)$$

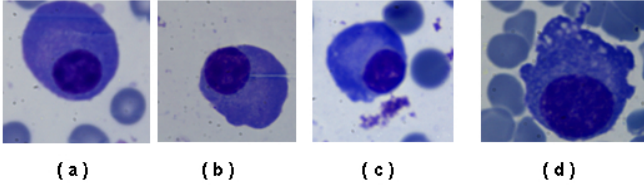


Fig. 5. Plasma cell examples:(a, b, c) normal, (d) dystrophic

### B. Shape attributes

We used several attributes of shape from the nucleus region of each leukocyte, these attributes have been selected for their ability to define the morphological aspects of the plasma cell:

- nucleocytoplasmic ratio : is the ratio between the number of pixels of the nucleus and the cytoplasm.
- Area : has been determined by counting the total number of pixels nucleus within the wbc.
- Perimeter: was measured by calculating the distance between two successive pixels belonging to the border.
- Compactness: or roundness of a nucleus.

$$Compactness = \frac{Perimeter^2}{Area} \quad (10)$$

- Solidity : is the ratio between the effective area and the area of the convex hull.

$$Solidity = \frac{Area}{ConvexArea} \quad (11)$$

- Eccentricity : this parameter is used to measure the center's shift of nucleus compared the center of the cell. This is an important feature because the plasma cells possess this property.

$$Eccentricity = \frac{\sqrt{a^2 + b^2}}{a} \quad (12)$$

where a and b are the major and minor axis of the ellipse representing the nucleus region.

- Formfactor: it is a dimensionless parameter which changes with surface irregularities.

$$Formfactor = \frac{4 \times \pi \times Area}{Perimeter^2} \quad (13)$$

In the end, each image containing a single WBC, will be characterized with 12 attributes, they will be used to build a classifier cells.

## IV. PLASMA CELL IDENTIFICATION

Once the characterization of each image done, and in order to assist in the diagnosis of myeloma, we will now proceed to the recognition of plasma cells (Figure 5).

We designed this classification with the idea of separating leukocytes into two classes: plasma cells and other cells (not plasma cells), ie the output of a classification into image(containing a single WBC ), will be accounted for as plasma

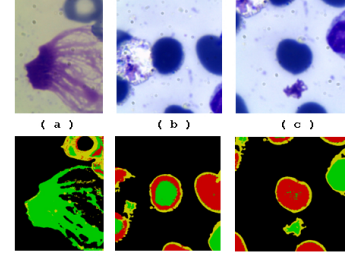


Fig. 6. Example of misfiled images: (a) cell damaged by staggering, (b) badly segmented cell, (c) cell cut on the edge of the image.

cell or as "other" (neutrophils, basophils, eosinophils, monocytes, lymphocytes, megakaryocytes, erythroblasts, proerythroblast, myelocyte, myeloblast, mtamyloocyte, prolymphocytes, promonocyte).

In other words, we aimed particularly plasma cell counts from all leukocytes, since this report is a major criterion for diagnosis of myeloma.

We start by building a classifier based on a learning database with 43 white blood cells belonging 27 images. In the purpose of achieving a good training, we took care to eliminate 6 cells, 3 represent segmentation faults, and the 3 others are non-entire cell (located on the edge of the image). A new classifier SVM is designed with 12 input parameters issue from leukocyte characterization step. This classifier is trained on the remaining 37 white blood cells of this base learning (including 15 plasma and 22 non-plasma cells). The classification on the same database of learning led us to recognize the 15 plasma and 22 non-plasma, this is equivalent to saying 100 % of recognition and accuracy rate.

The test database, is made up of 81 spinal cord image after segmentation and characterization, a total of 157 cells is obtained. With the help of an expert, in this case a doctor hematologist, we carried out a labeling of 157 cells that will serve as a test base, we got to annotate 97 plasma and 60 non-plasma cells. After classification, among the 97 plasma cells, 73 were recognized correctly, while 21 dystrophic plasma cells (atypical) and 3 poorly segmented plasma cells were identified as non-plasma; the recognition rate is 75.25 %. Note that this decrease from the results obtained on the learning base is justified primarily by the dystrophic cells, knowing that even the human expert hardly recognize this type of cells. Figure 6 represents images that have been misinterpreted; they would not received to the classifier, or rather should have been eliminating during segmentation. By cons, no false positives were recorded, i.e. 60 cells classified as non-plasma cells really are, this can be considered an advantage because the counting of plasma cells is not impaired.

## V. CONCLUSION

In this paper, we have proposed an automated segmentation, characterization and identification of Leukocytes namely plasma cells. The proposed system firstly individuates in the bone marrow image the WBC from the others blood cells, through the method otsu applied to the green channel of the image, obtaining a segmentation of the nucleus. This latter will be used as seed point for the segmentation of the

cytoplasm by the region growing. Our second contribution is in the integration of shape constraints into the region growing algorithm to have round or oval cells, and we end up extracting shape and texture indexes in order to classify plasma cells.

Our system is completely automated applicable to varying cell appearance and image quality.

The experimental results are compared with the manual results obtained by hematologists. The proposed method is more reliable and computationally less expensive, but there is a lacuna in the recognition of distrophiques cells, we will try to correct in future work.

#### ACKNOWLEDGMENT

We thank Professor Mrs. Nadia Benmansour (Hemobiology service, Tlemcen Hospital) for help and advice during images acquisition, preparation of ground truth images and especially labeling of plasma cells.

#### REFERENCES

- [1] D. M. U. Sabino, L. da Fontoura Costa, E. Gil Rizzatti, and M. Antonio Zago, "A texture approach to leukocyte recognition," *Real-Time Imaging*, vol. 10, no. 4, pp. 205–216, 2004.
- [2] N. Theera-Umpon and S. Dhompangsa, "Morphological granulometric features of nucleus in automatic bone marrow white blood cell classification," *Information Technology in Biomedicine, IEEE Transactions on*, vol. 11, no. 3, pp. 353–359, 2007.
- [3] T. Bergen, D. Steckhan, T. Wittenberg, and T. Zerfass, "Segmentation of leukocytes and erythrocytes in blood smear images," in *Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE*. IEEE, 2008, pp. 3075–3078.
- [4] D.-C. Huang and K.-D. Hung, "Leukocyte nucleus segmentation and recognition in color blood-smear images," in *Instrumentation and Measurement Technology Conference (I2MTC), 2012 IEEE International*. IEEE, 2012, pp. 171–176.
- [5] L. B. Dorini, R. Minetto, and N. J. Leite, "Semiautomatic white blood cell segmentation based on multiscale analysis," *Biomedical and Health Informatics, IEEE Journal of*, vol. 17, no. 1, pp. 250–256, 2013.
- [6] O. Kasim and A. Kuzucuoglu, "Identification of diseased leukocytes cells from blood smear," in *Signal Processing and Communications Applications Conference (SIU), 2013 21st*. IEEE, 2013, pp. 1–4.
- [7] M. Benazzouz, I. Baghli, and M. A. Chikh, "Microscopic image segmentation based on pixel classification and dimensionality reduction," *International Journal of Imaging Systems and Technology*, vol. 23, no. 1, pp. 22–28, 2013.
- [8] I. Baghli, A. Nakib, E. Sellam, M. Benazzouz, A. Chikh, and E. Petit, "Hybrid framework based on evidence theory for blood cell image segmentation," in *SPIE Medical Imaging*. International Society for Optics and Photonics, 2014, pp. 903 815–903 815.
- [9] A. R. KATZ, "Image analysis and supervised learning in the automated differentiation of white blood cells from microscopic images, rmit," Ph.D. dissertation, Diplomarbeit, February, 2000.
- [10] R. Adams and L. Bischof, "Seeded region growing," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 16, no. 6, pp. 641–647, 1994.
- [11] V. Tuset, I. Lozano, J. Gonzalez, J. Pertusa, and M. García-Díaz, "Shape indices to identify regional differences in otolith morphology of comber, *serranus cabrilla* (L., 1758)," *Journal of Applied Ichthyology*, vol. 19, no. 2, pp. 88–93, 2003.
- [12] R. M. Haralick, K. Shanmugam, and I. H. Dinstein, "Textural features for image classification," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 6, pp. 610–621, 1973.