Segmentation and Identification of Peripheral Blood Cells Using Image Analysis

Carlos Eymard Zamacona López[1], Raúl Bruno Sánchez[1], Arnulfo Catalán Villegas[1], Gustavo A. Alonso[1], Ma. Elena Moreno-Godínez[2], Antonio Alarcón-Paredes[1]

Universidad Autónoma de Guerrero, Unidad Académica de Ingeniería, Chilpancingo, México[1]
Universidad Autónoma de Guerrero, Unidad Académica de Ciencias Químico-Biológicas, Chilpancingo, México[2]

{cezamacona[1], rbruno[1], catalanvillegas[1], gsilverio[1], aalarcon[1]}
@uagro.mx
emoreno20[2]@hotmail.com

Abstract. One of the preferred techniques to measure genetic damage is the Micro-Nucleus (MN) Test in peripheral blood cells. The counting and identification of micronucleus in binucleated cells, is traditionally performed manually and directly on a microscope slide by an expert, which results in a highly time consuming task. Although it is a standard method to monitor the genetic damage, there are some significant differences in laboratories which perform the MN test. One of the main differences lie in the use of different compounds to stain the cells, and the use of sophisticated hardware and cameras to obtain high-detailed images. In this paper, a method for automatic segmentation of peripheral blood cells is presented and can serve as an alternative to the issues before mentioned. Results show that the proposed algorithm is able to identify MN in images obtained with a mobile phone camera, which results in a simple and flexible method.

Keywords: Cell segmentation, Micronucleus Test, Image analysis, Mathematical Morphology

1 Introduction

For several years, there is a growing industrial activity in several different areas. Recently, some of the growing industrial activity involves the use of genotoxic agents that endanger and threaten the genetic integrity in human beings.

In addition to this, there are different factors that influence the genetic damage, such as: feeding habits, medical therapies that involve radiation, as well as the increase of solar radiation to which we are exposed due to uncontrollable climatic change.

Whatever the factor of genetic damage is, it is important to quantify and determine an acceptable level of damage in human population[1].

A technique to measure genetic damage is the MicroNucleus (MN) Test in peripheral blood cells. These studies are useful when there is a bad or deficient cell division and there is a break or chromosome loss [2] giving as a result the generation of MN. For that reason, it is crucial to provide a reference range for a "normal" quantity of MN present in a blood cell of a person.

The counting and identification of MN in binucleated cells, is traditionally performed manually by an expert.

Different automatic methods for counting (or scoring) MN exist, whose results appear to be quite good and can be consulted in [1, 3]. However, many of these works involve the acquisition of specialized hardware equipment; in addition, the commercial software code is proprietary and it is tuned to work only with that hardware [4-6].

Although this test has been consolidated as a standard method to monitor the level of damage in chromosomes, there are some differences when it is carried out. In [7] a study called "HUMN PROJECT" was made, in this study 25 different laboratories in 16 countries were involved, where significant differences were found in methods used among the participating laboratories to perform the MN Test.

Some of the most significant differences lie in the use of different compounds to stain cells to be analyzed which hinders an algorithm for automatic processing to function properly in different samples whose staining is very different; as well as the use of specialized cameras that allow image acquisition whose quality or resolution is very different from the equipment that can be found in laboratories with low budget.

In this paper, a method of automatic segmentation in peripheral blood cells for scoring MN is presented; it can serve as an alternative to the issues which are mentioned in the previous paragraph. The proposed method uses images taken with a mobile phone camera pointing in the microscope, and with a compound of staining which differs from those clear and detailed images presented in the literature [5, 8].

The rest of this paper is organized as follows: In Section 2 the necessary techniques to get the best information from imaging microscopy are described, in Section 3 the proposed Algorithm is presented. Section 4 shows the results obtained by the use of the proposed algorithm. The conclusions and ideas for future work of the developed system are shown in Section 5. Finally, acknowledgement and references are shown.

2 Methods

In this section the techniques used to perform the Algorithm to obtain image information are described.

2.1 Otsu's Binarization Method

Binarization is an image processing technique which consists in an image information quantization into two values: 0 (black) and 255 (white). The main goal is to obtain an appropriate binarization threshold; a good option is to use the Otsu Method.

The Otsu Method is a nonparametric and unsupervised method of automatic threshold optimal selection [9], the Otsu Method can discriminate the pixels of an image in several classes $C_0 = \{1,...,k_1\}$, $C_1 = \{k_1,...,k_{2\},...,}$ $C_n = \{k_n,...,L\text{-}1\}$ where L is the maximum value of a pixel and each class includes all the pixels to a threshold detected by Otsu Method, and $U = \{k_1,k_2,...,k_n\}$ the thresholds set detected.

Probability distribution of the normalized gray levels could be expressed as $p_i = n_i/N$ where p_i is the probability of occurrence, n_i is the gray level intensity, and N is the number of pixels.

The zeroth- and the first-order cumulative moments are defined as $\omega_0 = \sum_{i=1}^{k} pi = \omega(k)$ and $\omega_1 = \sum_{i=k+1}^{L} pi = 1 - \omega(k)$, respectively, with $\mu_0 = \mu(k) / \omega(k)$ and $\mu_1 = \mu T - \mu(k) / 1 - \omega(k)$ where $\omega(k) = \sum_{i=1}^{k} p_i$ and $\mu(k) = \sum_{i=1}^{k} ip_i$.

The total mean level of the original picture is obtained with the equation defined as $\mu_{T =} \mu(L) = \sum_{i=1}^{L} ip_i$ and $\sigma_B^2 = \omega_0 (\mu_0 - \mu_T)^2 + \omega 1 (\mu_1 - \mu_T)^2$.

The optimal threshold in a search for the maximum value in σ_B^2 is

$$\sigma_B^2(k^*) = \max_{1 \le k < L} \sigma_B^2(k)$$

For problems involving multilevel threshold, it is assumed that measurement criteria σ_B^2 are based on multiple variables $k_1, k_2, ..., and\ k_n$, and an optimum set of threshold $k_1^*, k_2^*, ..., and\ k_n^*$ is selected maximizing σ_B^2

$$\sigma_B^2(k_1^*, k_2^*, \dots, k_n^*) = \max_{1 \le k1 < k2 < \cdots < kn < L} \sigma_B^2(k_1, k_2, \dots, k_n)$$

2.2 Morphological Image Processing

Mathematical morphology has simple techniques for extracting components in an image, which are useful in the representation and description of region shapes. The two main operations, dilation and erosion, are performed over a set representation of an image A using a structuring element denoted by B, which defines the shape and size of the neighborhood of the pixel that will be analyzed subsequently to modify its value [10, 11]. The morphological operations used in this paper are described next.

- **Dilation**. Dilation of A by B is the Minkowski sum [12] of A and B, it is defined as:

$$A \oplus B = \{x = a + b \in X \mid a \in A \wedge b \in B\}$$

- **Erosion**. Erosion of A by B is the Minkowski subtraction [13], and it is defined as:

$$A \ominus B = \{x \in X \mid x + b \in A, \wedge \forall b \in B\}$$

- **Opening**. Opening of A by B is the erosion of A by B, followed by dilation of the result by B, it is defined as:

$$A \circ B = (A \ominus B) \oplus B$$

- **Region Filling**. Region Filling is based on a set of dilations, complements and intersections, it is defined as:

$$Xk = (Xk - 1 \oplus B) \cap Ac\ k = 1, 2, 3, \ldots$$

Where $X_0 = \mathrm{p}$, and B is a symmetrical structural element. The algorithm ends in step k if $x_k = X_k - 1$.

2.3 Distance transform

For each pixel in a binary image, this transformation assigns a number that is the distance between the pixel and the nearest nonzero pixel, the Euclidean distance is used, it is defined as:

$$d(\mathrm{x}, \mathrm{y}) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}$$

2.4 Watershed Transform

In mathematical morphology, Watershed is a technique where an image is considered as a topographic relief [14]. There are several different methods of segmentation that can be used to perform a Watershed Transform, however, the method used in this paper is carried out as follows:

1. A pre-processing filter for noise removal is done, and then the image is binarized.
2. A distance transform is performed and its complement is obtained.
3. The areas with the minimum values are identified; these areas are placed as markers in the binarized image.
4. A Morphological Reconstruction is performed.
5. Finally Watershed Transform is carried out.

3 Algorithm

A Data flow diagram of the proposed algorithm is shown in Figure 8. This implements the methodology mentioned in the previous section, and it is able to perform cell segmenta-

tion in order to count the amount of MN present in a microscope slide, by means of a mobile phone camera. The algorithm is described as follows:
1. Obtain an RGB picture of the peripheral blood cell slide, and get only the green plane, as shown in Figures 1a and 1b.

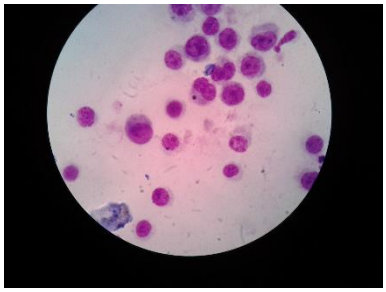

Fig. 1a. Original RGB image

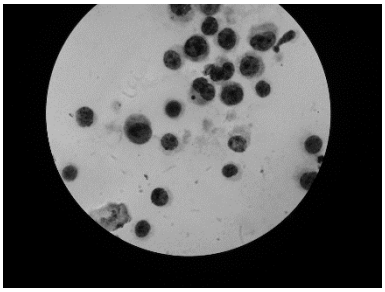

Fig. 1b. Green plane

2. Obtain four thresholds values in the green plane image by means of Otsu multi-threshold technique. Then binarize the image using the third threshold, and get the negative image, as shown in Figure 2.
3. Here, a cleaning step is performed over the binarized image. This step consists in applying an algorithm to fill holes [15] in the detected objects, as well as the application of an opening function in order to delete several small objects as showed in Figure 3.
4. The distance transform is now performed over the complement of the Figure 3. This algorithm computes the (euclidian) distance between every zero-valued pixel and the nearest non-zero pixel. The complement of the resulting image is a gray-scaled gradient where the innermost pixel in the objects, has the darkest value, as shown in Figure 4.

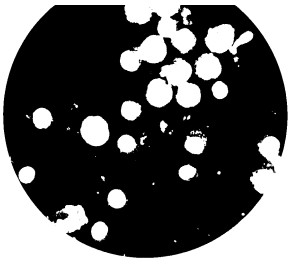

Fig. 2. Binarized

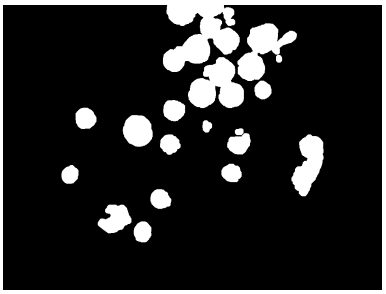

Fig. 3. Image after cleaning process

5. Obtain the morphological reconstruction of the negative distance transformed image. In this case the morphological reconstruction selects the darkest regions of the distance trans-

formed image, and masks the gray scale image so these regions are the minima regions of reconstructed image. The darkest regions could be used as markers representing the basins, and the reconstructed image is the topographic surface to be flooded with the watershed transform. Figures 5a and 5b show the result of this process.

Fig. 4. Result of applying the distance transform

6. Compute the watershed transform to the image resulting of the step 5 and could be seen in Figure 6.
7. Obtain the object labeling using an 8-connected neighbourhood (see Figure 7).

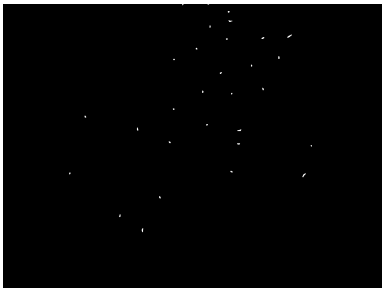

Fig. 5a. Markers (basins)

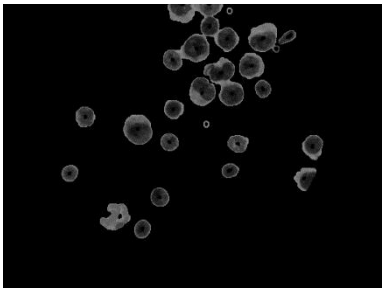

Fig. 5b. Topographic surface (morphological reconstruction)

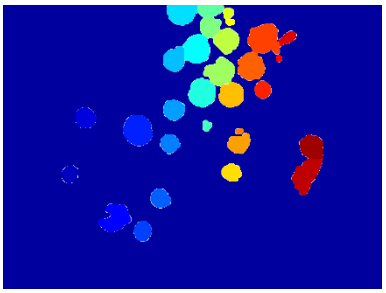

Fig. 6. Results of watershed transformation

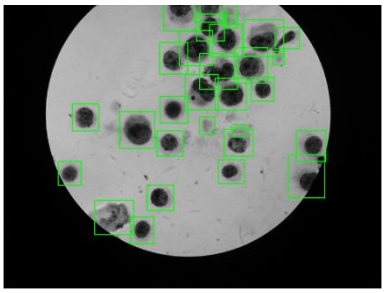

Fig. 7. Object labeling

4 Results

The proposed algorithm was implemented using Octave on a computer with an Intel core i3 processor and 4 GB of RAM, running on Windows 8.1, and was applied to a set of 16 peripheral blood cell images obtained with a cell phone camera. The cellular phone used to capture the images was a Samsung gt-i8190l, running over Android 4.1 OS with a 5.0 MP camera resolution, with image size of 2560 pixels width and 1920 pixels length.

The segmentation provides very good results in cell identification, since the images are obtained using camera phones. In Figure 9, a visual example of the results obtained is shown. In addition, Figure 10 shows the data results of the cells identified by the proposed method versus manual identification. Nonetheless there are some differences between the manual and automatic scoring processes, often caused due to border located cells and some waste of the stain compound, this problem will be overcome in future work by means of a further step to discriminate between true cells and the spurious ones.

In addition, images immerse in the literature have clearer boundaries than ours, which yields to an easier image segmentation. This is caused due to the use of a more expensive compound to stain the cells, and the hardware they use to obtain the image since it is very robust and permits increasing the zoom of the pictures, so the images to be analyzed have much more details.

Here, the results show that despite of the drawbacks of having compounds that could not make the boundaries of cells too much clear, it is possible to segment the blood cells in a microscope slide image taken with a phone camera. Results of segmentation of some other cells are shown in Figures 11 to 14.

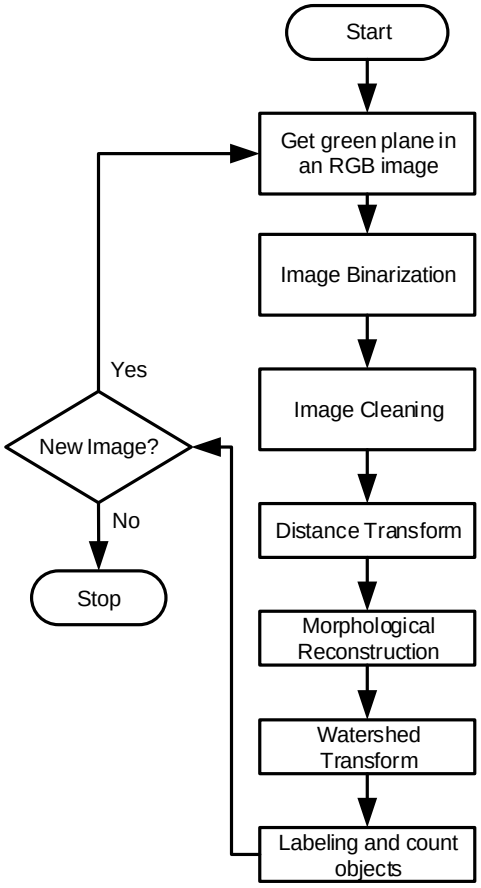

Fig. 8. Dataflow diagram of proposed algorithm

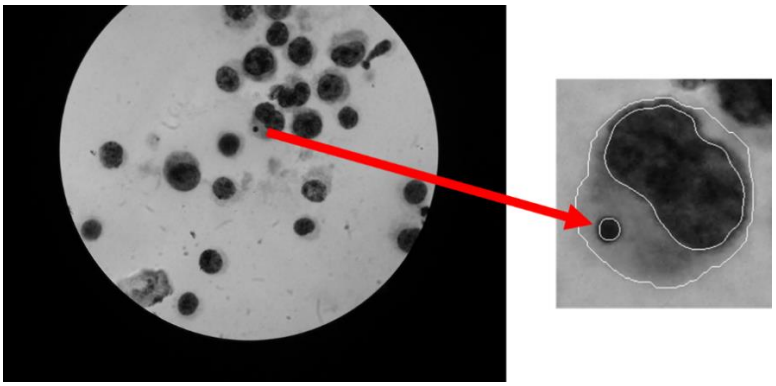

Fig. 9. A micronucleus within a binucleated cell, is shown.

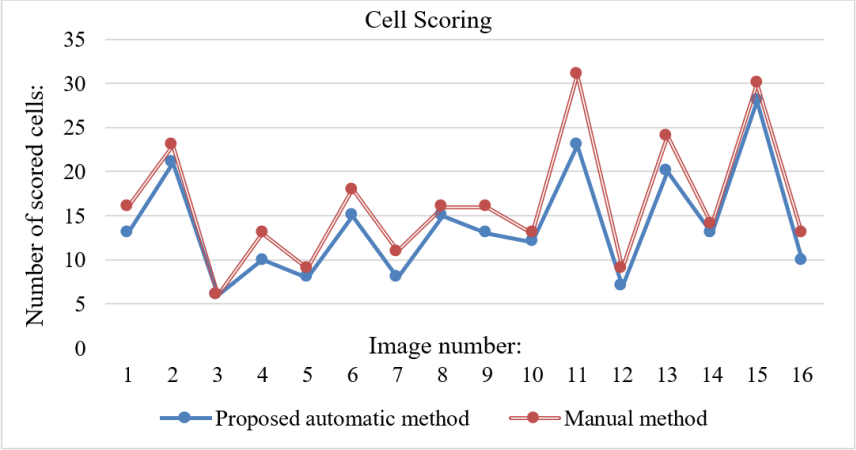

Fig. 10. Comparison of the manual vs automatic identified cells

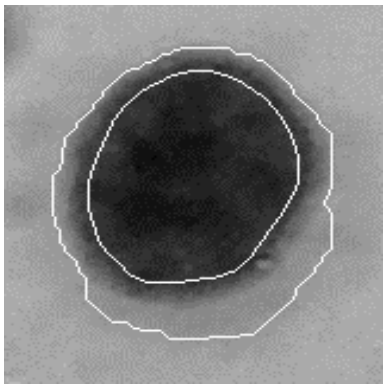

Fig. 11. Segmented cell

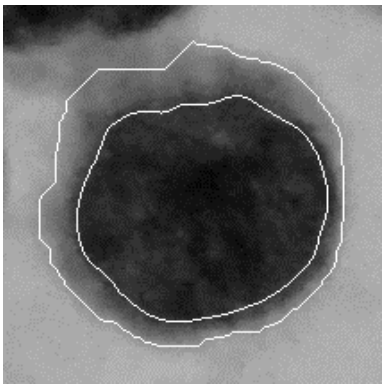

Fig. 12. Segmented cell

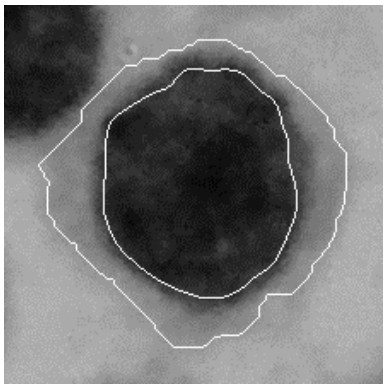

Fig. 13. Segmented cell

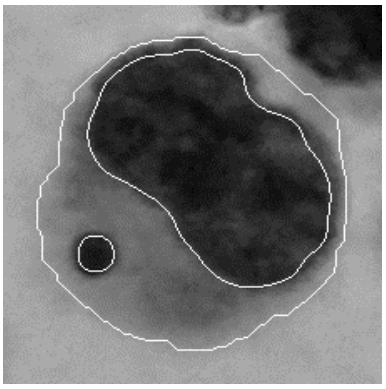

Fig. 14. Segmented cell with a MN

5 Conclusions

The automation of cell segmentation allows laboratories with no sophisticated and inexpensive equipment, and also without the need for an expert, to perform the MN test in human populations which are exposed to genotoxic compounds, to help in the genetic damage diagnosis.

The proposed algorithm, results in a reliable method for peripheral blood cell segmentation, since the results obtained using images with poor distinguishable boundaries between cells are competitive. It is worth to mention that images in the state of the art have a clear cytoplasmic boundary due to specific compound used to stain the slides, which rep-

resents a difference in the images used among laboratories worldwide. For this reason, we can conclude that if this method is applied over images in the literature, the results would be even better.

In contrast with the use of sophisticated hardware for image acquisition in the related works, here, a mobile phone camera was used to obtain the blood cell images, which represents that proposed algorithm is a very flexible method and could be used in laboratories where the high technology devices for image acquisition are not affordable.

The results presented in this work are assumed to be improved by means of merging some other segmentation methods. With the aim of enhance the hardware and time consumption, a migration from Octave to another programming language such as C or C++ are also expected.

Acknowledgments

The authors of this work appreciate the support given by CONACYT for their economic support to develop this work.

References

1. Fenech, M.: The cytokinesis-block micronucleus technique: a detailed description of the method and its application to genotoxicity studies in human populations. Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis 285, 35-44 (1993)
2. Schmid, W.: The micronucleus test. Mutation Research/Environmental Mutagenesis and Related Subjects 31, 9-15 (1975)
3. Fenech, M.: The in vitro micronucleus technique. Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis 455, 81-95 (2000)
4. Böcker, W., Müller, W.U., Streffer, C.: Image processing algorithms for the automated micronucleus assay in binucleated human lymphocytes. Cytometry 19, 283-294 (1995)
5. Frieauff, W., Martus, H., Suter, W., Elhajouji, A.: Automatic analysis of the micronucleus test in primary human lymphocytes using image analysis. Mutagenesis 28, 15-23 (2013)
6. Rodrigues, M., Beaton-Green, L., Kutzner, B., Wilkins, R.: Automated analysis of the cytokinesis-block micronucleus assay for radiation biodosimetry using imaging flow cytometry. Radiation and environmental biophysics 53, 273-282 (2014)
7. Bonassi, S., Fenech, M., Lando, C., Lin, Y.-p., Ceppi, M., Chang, W.P., Holland, N., Kirsch-Volders, M., Zeiger, E., Ban, S.: HUman MicroNucleus project: international database comparison for results with the cytokinesis-block micronucleus assay in human

lymphocytes: I. Effect of laboratory protocol, scoring criteria, and host factors on the frequency of micronuclei. Environmental and molecular mutagenesis 37, 31-45 (2001)
8. Fenech, M., Chang, W.P., Kirsch-Volders, M., Holland, N., Bonassi, S., Zeiger, E.: HUMN project: detailed description of the scoring criteria for the cytokinesis-block micronucleus assay using isolated human lymphocyte cultures. Mutation Research/Genetic Toxicology and Environmental Mutagenesis 534, 65-75 (2003)
9. Otsu, N.: A threshold selection method from gray-level histograms. Automatica 11, 23-27 (1975)
10. Díaz-de-León-Santiago, J.L., Yáñez-Márquez, C.: Introducción a la morfología matemática de conjuntos. Instituto Politécnico Nacional (2003)
11. Matheron, G., Serra, J.: Image analysis and mathematical morphology. London: Academic Press (1982)
12. Minkowski, H.: volumen and Oberfläche. Math. Ann., vol. Vol. 57, pp. 447-495. (1903)
13. Hadwiger., H.: Vorslesunger über Inhalt, Oberfläche und Isoperimetrie. Berlin: Springer (1957)
14. Beucher, S., Meyer, F.: The morphological approach to segmentation: the watershed transformation. OPTICAL ENGINEERING-NEW YORK-MARCEL DEKKER INCORPORATED- 34, 433-433 (1992)
15. Soille, P. Morphological image analysis: principles and applications. Springer Science & Business Media. (2013)