Bull. Nov. Comp. Center, Comp. Science, 36 (2014), 97–101

An algorithm for registration of erythrocytes on low-contrast images

M.S. Tarkov, D.V. Vytnova, A.F. Zadorozhny

Abstract. An algorithm for counting the erythrocytes on low-contrast images of cytological preparations is proposed. The algorithm deals with low-contrast images: brightness histogram normalization, contrast stretching, and background alignment by the "top of hat" method. Erythrocytes are detected by the similarity criterion of reference samples and an area in the picture. The Pearson correlation coefficient is used as a similarity criterion. The algorithm has been successfully tested and proved to be efficient.

1. Introduction

The clinical analysis of blood today is one of the first tests, prescribed to the patient to assess the state of his health. This analysis allows us to estimate the hemoglobin content in the blood, red blood cell count or white blood cell count. This can be done either manually, which is extremely-time consuming and costly, and demands using modern specialized software systems.

In addition to a special software, such complexes include special equipment, namely, a microscope and a high-resolution digital camera. The analysis becomes more complicated if we have only a series of low-contrast images [1–5]. This paper is aimed at developing an efficient algorithm for counting erythrocytes in low-contrast images. We propose an algorithm for the search and enumeration of erythrocytes in the image on the reference samples set created by the user. The similarity criterion for a reference sample and a region of the image is the Pearson correlation coefficient [4]. The image preprocessing is included into the algorithm in order to improve its qualitative characteristics, the contrast, in particular.

The algorithm includes:

- preliminary image processing (image conversion from the RGB color space to the HSV color space (Hue, Saturation, Value) and extraction of Value, median filtering, brightness histogram normalization, contrast stretching, background alignment by the "top of hat" method [1, 2]);
- extraction of reference samples on an image and search for pixel groups on it, similar to the reference samples, and in the case of a successful search, marking them as appropriate erythrocytes.

2. Preliminary image processing

The median filter replaces the pixel value on the median value for brightness of all pixels in the neighborhood including the source. Median filtering is used to smooth the image and to remove impulse noise. Based on the experiment, a filter mask of the size was selected for smoothing the image to maintain clear erythrocyte boundaries.

A brightness histogram was normalized by the formula

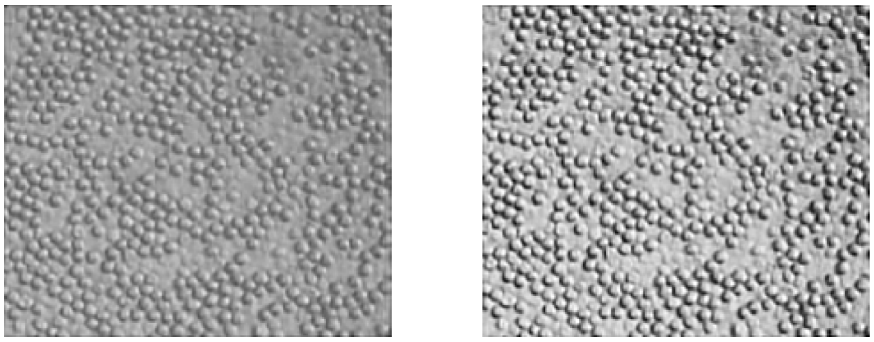
$$r_{\text{normal}} = 255\frac{r - r_{\min}}{r_{\max} - r_{\min}}, \tag{1}$$

where $r_{\min}$ and $r_{\max}$ are minimum and maximum values of the image brightness r, respectively.

The experiments show that it is not sufficient to modify the histogram by formula (1) for obtaining high-contrast images. It is necessary to strengthen a difference between light and dark pixels in the image, making a border of red blood cells more pronounced as compared to the background, and there is no significant contrast without stretching. The function of contrast stretching is the following:

$$s = T(r) = \frac{1}{1 + (m/r_{\text{normal}})^E}, \tag{2}$$

where s is the brightness of the output image, E is a parameter that controls the slope of the function, m is the threshold of stretching [3]. The parameters $m = 160$, $E = 5$ are adjusted empirically. This function compresses the input value that is smaller than m to a narrower subband of the gray-level image on the output, and, respectively, the values larger than m — to a narrow strip of bright levels. The result of the transform (2) is a higher contrast image (Figure 1).



a b

Figure 1. Transformation example: a) an original image, b) the image after enhancement of the contrast

Approximation of the image background is implemented by an open image, which is constructed by erosion and dilatation operations [1–3] successively applied to the image.

Gray-scale dilatation of the image f by the structural element b is denoted by $(f + b)$ and defined by the formula

$$(f + b)(x, y) = \max\big\{f(x - x', y - y') + b(x', y') : \ (x', y') \in D_b\big\}, \qquad (3)$$

where D_b is a domain of definition of b, and it is supposed that $f(x, y) = -\infty$ out the definition of f. In practice, the gray-scale dilatation is often the operation of taking a local maximum in which the maximum is computed over a set of the pixels neighborhood, whose shape is defined by the area D_b:

$$(f + b)(x, y) = \max\{f(x - x', y - y') : \ (x', y') \in D_b\}. \qquad (4)$$

Gray-scale erosion of the image f by the structural element b is denoted by $(f - b)$ and often defined by the formula

$$(f - b)(x, y) = \min\{f(x - x', y - y') : \ (x', y') \in D_b\}, \qquad (5)$$

i.e. the gray-scale erosion is often the operation of taking a local minimum, in which the minimum is computed over a set of the pixels neighborhood, whose shape is defined by the area D_b.

Extension of the image f to the element b is denoted by $f \circ b$ and defined by the expression

$$f \circ b = ((f - b) + b). \qquad (6)$$

First, the erosion of f occurs with b, followed by the dilation of f with b.

In general, the opening procedure (3)–(6) is used to remove narrow bursts of brightness, while maintaining average halftone background and broad areas of brightness remain relatively unchanged. As a structure-forming element b, a circle with a radius of 25 pixels is used.

The “top of hat” transform is reduced to the subtraction of the open image from the original one. The result of the “top of hat” transform is an image with aligned background.

3. The search for erythrocytes in the image

As a criterion of similarity of a reference sample and an area in the picture, the Pearson correlation coefficient is used [4]. For the samples $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$, the Pearson coefficient is

$$R(x, y) = \frac{\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \overline{x})^2 \sum_{i=1}^{n}(y_i - \overline{y})^2}}, \qquad (7)$$

where $\overline{x}$ and $\overline{y}$ are mean values in the sets of samples $\{x\}$ and $\{y\}$, respectively, $R(x, y) \in [-1, 1]$. If $|R(x, y)| = 1$, then x and y are linearly dependent, and if $|R(x, y)| = 0$, then x and y are linearly independent. Several reference samples are used to detect erythrocytes by evaluating the correlation. A maximum of correlation coefficients (7) as compared to a detection threshold, and if it turned out to be greater than threshold or equal to it, the corresponding cell on the image was marked indicating to an erythrocyte.

4. Experiments

In the software implementation, we chose three samples having good correlations with cells on the image. In addition to labeling, the number of erythrocytes detected was counted. To evaluate the effectiveness of the algorithm, a series of 18 low-contrast images of blood smears was analyzed. These images were recorded in the experiment in which it was necessary to track a change in the number of red blood cells over a time period.

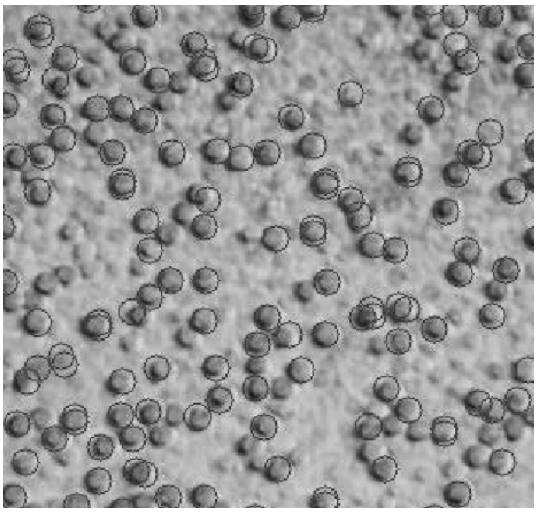


Figure 2. An example of the image with detected erythrocytes (marked by circles)

An example of the results obtained with the use of the algorithm is presented in Figure 2. Based on these data, a chart was constructed showing a change in the number of red blood cells over time. The accuracy of automatic counting erythrocytes is on the average about 86 %. The processing of one image takes less than 7 seconds on the processor Pentium (R) Dual-Core CPU E 52000, 2.5 GHz. An experimental sample of such a system is developed in Visual Studio C++ using the image processing library OpenCV [5].

5. Conclusion

An algorithm for counting the erythrocytes on low-contrast images of cytological preparations is proposed. The algorithm includes methods dealing with low-contrast images: brightness histogram normalization, contrast stretching, and background alignment by the "top of hat" method. The search for erythrocytes realized by the criterion of similarity of a reference

sample and an area in the picture. The Pearson correlation coefficient is used as a similarity criterion.

The algorithm has been successfully tested and proved to be effective. From the standpoint of the time required for image processing, the algorithm has quite an acceptable rate. The accuracy is required from the user in selecting the reference samples. It is also desirable to select a threshold of cell detection for every new group of images.

Acknowledgements. The authors thank Vadim Nurminsky (Siberian Institute of Plant Physiology and Biochemistry (SIPPB SB RAS) for placing at our disposal images of cytological preparations.

References

[1] Gonzalez R.C., Woods R.E. Digital Image Processing. — Prentice Hall, 2002.

[2] Gonzalez R.C., Woods R.E., Eddins S.L. Digital Image Processing using MATLAB. — Prentice Hall, 2004.

[3] Tarkov M.S., Tikhonov N.V., Polovinkin V.G. Evaluation of images of cell preparations for biomedical research // Izvestia TPU. — 2012. — Vol. 320, No. 5. — P. 130–133 (In Russian).

[4] Shapiro L., Stockman G. Computer Vision. — Prentice Hall, 2001.

[5] Bradski G., Kaehler A. Learning OpenCV. — O'Reilly Media, Inc., USA, 2008.