Biomedical Signal Processing and Control 45 (2018) 10–21

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

journal homepage: www.elsevier.com/locate/bspc

A novel algorithm for segmentation of leukocytes in peripheral blood

Haichao Cao, Hong Liu*, Enmin Song

Huazhong University of Science and Technology, School of Computer Science & Technology, China

ARTICLE INFO

Article history:
Received 2 February 2018
Received in revised form 6 May 2018
Accepted 16 May 2018

Keywords:
Concave-convex iterative repair method
Fuzzy divergence
Interval-valued fuzzy sets
Leukocyte segmentation
Stepwise averaging method

ABSTRACT

In the detection of anemia, leukemia and other blood diseases, the number and type of leukocytes are essential evaluation parameters. However, the conventional leukocyte counting method is not only quite time-consuming but also error-prone. Consequently, many automation methods are introduced for the diagnosis of medical images. It remains difficult to accurately extract related features and count the number of cells under the variable conditions such as background, staining method, staining degree, light conditions and so on. Therefore, in order to adapt to various complex situations, we consider RGB color space, HSI color space, and the linear combination of G, H and S components, and propose a fast and accurate algorithm for the segmentation of peripheral blood leukocytes in this paper.

First, the nucleus of leukocyte was separated by using the stepwise averaging method. Then based on the interval-valued fuzzy sets, the cytoplasm of leukocyte was segmented by minimizing the fuzzy divergence. Next, post-processing was carried out by using the concave-convex iterative repair algorithm and the decision mechanism of candidate mask sets. Experimental results show that the proposed method outperforms the existing non-fuzzy sets methods Among the methods based on fuzzy sets, the interval-valued fuzzy sets perform slightly better than interval-valued intuitionistic fuzzy sets and intuitionistic fuzzy sets.

1. Introduction

Medical image processing occupies a leading position in the hematology, a branch of medical science that focuses on the study of blood and hematopoietic tissue diseases. In pathological studies, the number of erythrocytes (red blood cells), leukocytes (white blood cells), platelets and other blood cells is very important for the detection of diseases such as anemia, leukemia, cancer and other infectious diseases. Among those blood cell parameters, the number of leukocyte plays an essential role in the body's immune system [1,2]. Blood leukocytes are mainly divided into five categories, namely eosinophils, basophils, neutrophils, lymphocytes and monocytes. The five types of leukocytes can be identified by the size of cell, the type of nucleus lobes, the ratio of nucleus to cytoplasm, cytoplasmic granules and the staining properties of granules, etc. [3].

Particularly the numbers and types of leukocytes are essential for the detection of blood type disease. Traditional medical experts count the cells directly, though it is not only quite time-consuming but also frequent to miss and repeat. Nowadays, automated techniques have been introduced for the diagnosis of medical images [4]. However, under different backgrounds, methods of staining, and degrees of dyeing, light conditions and so on, it remains difficult to accurately extract related features and count the number of cells [5,6]. Therefore, in order to adapt to various complex situations, a method based on stepwise averaging method, interval-valued fuzzy sets and concave-convex iterative repair algorithm for segmentation of leukocytes is proposed in images of Wright's stained [7,8] blood smears preparations.

In recent years, many researchers have proposed different segmentation solutions for blood cells. We divide the leukocyte segmentation method into two categories, i.e., one is based on non-fuzzy set methods and the other is based on fuzzy set methods.

The non-fuzzy set approaches for leukocytes segmenting are based on threshold [9–11], morphological operations [12], region growing [13], neural network [14], and clustering [15–17]. These methods, however, may lead to unsatisfactory segmentation results due to the existence of the red blood cells, and its performance is not good for overlapping cells or cells with boundaries that are not smooth enough (rough edge). Some methods have been developed based on deformable models, such as watershed algorithm [18,19], level set [20] and parametric active contour [3,21]. However, these methods require good contour initialization to obtain good segmentation results, and the processing time is long, and cannot meet the requirement of engineering application. In addition, several methods that can be categorized into color-

* Corresponding author.
E-mail addresses: dr.caohaichao@163.com (H. Cao), hongliu@hust.edu.cn (H. Liu), esong@hust.edu.cn (E. Song).

https://doi.org/10.1016/j.bspc.2018.05.010
1746-8094/

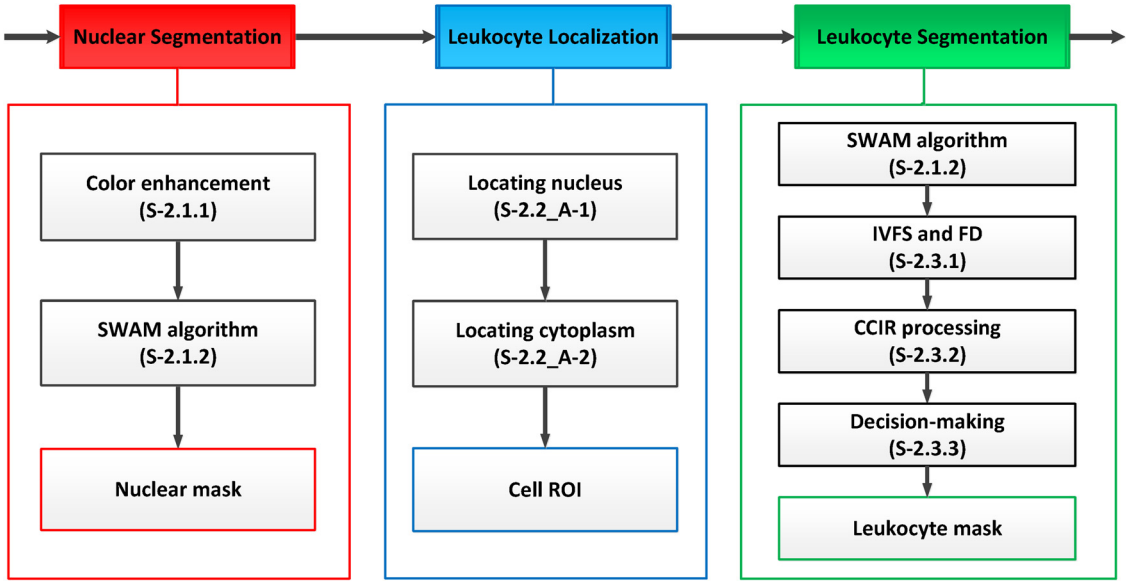

Fig. 1. The flowchart of the proposed algorithm. Three modules are shown within the red, green and blue boxes, respectively. Note that, "S-c" denotes section c; "A-n" denotes algorithm n; "SWAM (stepwise averaging method)denotes proposed threshold-based segmentation algorithm. For details, see Section 2.1.2; 'CCIR' denotes a concave-convex iterative repair algorithm; "IVFS" denotes interval-valued fuzzy sets; "FD" denotes fuzzy divergence. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

based methods have been proposed in the literature to exploit the advantages of the color variation between different cell components in segmentation [22–27]. However, these methods in [23–25] are unable to segment the cytoplasm, and the nuclei segmentation results still need to be improved. There are also some recent approaches proposed, for example, in [28], the authors uses multiple windows obtained by scoring multiscale cues to locate the leukocytes, and then the GrabCut algorithm based on dilation is iteratively run to segment the leukocytes.

In the fuzzy set approaches, the fuzzy divergence is used to obtain the optimal threshold. In [29], the authors proposed a Yager's method based on fuzzy sets, which respectively processed three channel images of R, G and B to get three segmented images. Finally, the three segmentation images are merged to obtain the final segmentation results. However, it cannot be applied to the larger connected cells. In [30], the authors proposed an automatic segmentation by intuitionistic fuzzy divergence based thresholding, that can still ensure better robustness in the case of noise interference. However, it is not good for the low staining and weak edge information, and the scale of test sample set is small, which needs to be further tested. In [31,32], the authors introduced a general version of the classical fuzzy sets [33] known as intuitionistic fuzzy sets. Extending the concept further, the authors used Atanassov's intuitionistic fuzzy sets and interval II fuzzy sets theory to carry out the segmentation of blood leukocyte image [34]. Compared with other algorithms, it has a better performance on the segmentation of an image containing more than one cell. In [35], the authors developed a new leukocyte segmentation methodology based on intuitionistic fuzzy divergence to achieve robust segmentation performance automatically without training set. However, it cannot adapt to changes in background and brightness.

The leukocyte segmentation technique proposed in this paper has the following three innovations:

(1) We propose a novel thresholding method and color transformation methods.
(2) The G, H and S three single-channel images are considered, which improves the adaptability of the algorithm to various backgrounds, and has better segmentation performance on the samples with lighter staining.
(3) In order to deal with the situation of connected cells, an improved concave-convex iterative repair algorithm is proposed.

2. Methods

The number and type of peripheral leukocytes are essential for the diagnosis of blood diseases. Because doctors cannot focus their attention for a long time, manual counting and classification are prone to omissions and repetitions. Therefore, there is an urgent need for an automated and accurate leukocyte analysis technique that generally involves three steps, namely cell segmentation, feature extraction, and cell classification. We focus on the first step: cell segmentation, since the accuracy of subsequent steps mainly depend on the cell segmentation in automated leukocyte analysis.

In this paper, the leukocyte segmentation algorithm is divided into three modules: the segmentation of nucleus, sub-image cropping, and the segmentation of cytoplasm. The flow chart is shown in Fig. 1.

2.1. Nuclear segmentation

The segmentation process of leukocyte nucleus is firstly, constructs the transformed image, HSG; secondly, the SWAM segmentation method was used to calculate the thresholds for segmentation of background, erythrocytes, cytoplasm and nucleus respectively, namely T_b, T_r, T_c and T_n. Finally, nucleus can be segmented by using the average of T_c and T_n as the threshold value. Note: the cytoplasm and nucleus are referring to the cytoplasm and nucleus of leukocytes.

The construction of the transformed image HSG and the SWAM segmentation method are described in Sections 2.1.1 and 2.1.2, respectively.

2.1.1. Transformed image

In general, after Wright's staining of the blood smear, the color of nucleus is dark purple, which is darker than that of cytoplasm and mature red blood cells. The blood smear image is usually stored in the RGB (Red-Green-Blue) format. In order to enhance the contrast between the nucleus of leukocytes and other cell tissues, it is