World Scientific
www.worldscientific.com

A counting method for complex overlapping erythrocytes-based microscopic imaging

Xudong Wei, Yiping Cao*, Guangkai Fu and Yapin Wang

Department of Optical Electronics
Sichuan University, Chengdu
Sichuan 610064, P. R. China
*ypcao@scu.edu.cn

Red blood cell (RBC) counting is a standard medical test that can help diagnose various conditions and diseases. Manual counting of blood cells is highly tedious and time consuming. However, new methods for counting blood cells are customary employing both electronic and computer-assisted techniques. Image segmentation is a classical task in most image processing applications which can be used to count blood cells in a microscopic image. In this research work, an approach for erythrocytes counting is proposed. We employed a classification before counting and a new segmentation idea was implemented on the complex overlapping clusters in a microscopic smear image. Experimental results show that the proposed method is of higher counting accuracy and it performs much better than most counting algorithms existed in the situation of three or more RBCs overlapping complexly into a group. The average total erythrocytes counting accuracy of the proposed method reaches 92.9%.

Keywords: Cell counting; image processing; image segmentation; overlap erythrocyte; cell classification; K-means.

1. Introduction

The erythrocytes, which also called red blood cells (RBCs), are a crucial kind of blood cells in human blood. Normal human blood consists of red blood cells (RBCs or erythrocytes, $\pm 4.500 \times 10^3/\mu L$ blood), white blood cells (WBCs or leukocytes, $\pm 8 \times 10^3/\mu L$ blood), platelets (thrombocytes, $\pm 300 \times 10^3/\mu L$ blood) and blood plasma (containing water, electrolytes, plasma proteins, carbohydrates, lipids and various extracellular vesicles).[1] RBC deformability plays an essential role in the delivery of oxygen to tissues. Biconcave disk-shaped RBCs have dimensions of approximately $8\,\mu m$ in diameter and $\sim 2\,\mu m$ in thickness, through capillaries with diameters not more than 3–5 μm, thus supplying oxygen to tissues of the body.[2] After blood smear being examined under microscope, the number of RBCs per unit volume can be counted. Counting process, when performed manually, have been proved to be tedious, time consuming and subjective. Results may vary by

*Corresponding author.

different inspectors or even at different turns though by the same inspector. Precision of the process may also be low because of the hematologist's subjective evaluation. Erythrocytes are much more common than the other blood particles: WBCs and thrombocytes. Pathological alterations in RBC have been associated with various diseases such as malaria, sickle cell anemia, diabetes, hereditary disorders and so on. Erythrocyte counting is one of the most commonly performed medical test as it plays a vital role in the diagnosis of various diseases like Alport syndrome and anemia.[3] Therefore, some automatic and accurate mechanism is required to cope with the aforementioned problems. Automatic blood cells inspection may be performed with the help of automatic flow cytometry, but algorithms based on image processing and computer vision techniques reduce analysis cost. RBC segmentation is the crucial part of taking count of RBCs. Erythrocyte automatic segmentation technology has made a huge progress with the development of computer science and image processing technology, and various automatic, computer-aided blood cell counting techniques have been proposed in recent decades.[5–12]

In this work, an automatic method to detect and count RBCs in microscope images of blood smears was proposed. Leukocytes and PLTs will be detected based on the H and S component, and be wiped off in order to obtain images that only contain erythrocytes. Then we use the area size and circularity features to classify all the closed regions into two classes: individual ones and overlapping clusters. Individual ones will be counted by watershed rapidly and the clusters will be segmented before counting. In the segmentation, K-means algorithm is applied. This method relied on RBCs' typical biconcave shape, even irregularly shaped ones can also be detected. Unfortunately, it is helpless while taking count of quasi-spherical and crenate shapes.[4] Results reveal that the proposed scheme method outperforms some other image segmentation algorithms by possessing less calculation, reducing processing time and higher validity.

The main contributions of this paper are as follows:

(1) We propose a classification of different structures of erythrocytes clusters that can be straightforward and accomplished at a high-speed.
(2) To improve the efficiency and precision of the cluster segmentation, K-means algorithm is applied.
(3) We present a complete approach for taking count of RBCs in smear blood images.

2. Related Works

Counting of RBCs in a microscopic blood smear image is a potential research issue due to its medical significance. Researchers have been putting serious efforts for this purpose, and much encouraging and accurate results have been produced as a result of these endeavors. Following text states a brief survey of literature related to the computer vision-based blood cells segmentation and counting. Gao *et al.*[5] propose a new marker based on watershed algorithm, which finds watershed lines that represent low-intensity areas around high-intensity objects. Xu *et al.*[6] presented a modified watershed algorithm with markers and a novel marking function, whose markers are extracted by adaptive H-minima transformation and the marking function based on Quasi-Euclidean distance transform is introduced for accurate and rapid separation. Dervieux and Thomasset[7] established the idea of level set in 1980, which uses an active contour model and curve evolution to calculate the new contour. On this basis Li *et al.*[8] define a new variation called distance regularized level set (DRLS) that not only does away with the burden for re-initialization but also permits greater flexibility in the selection of initial level set function. Nosrati and Hamarneh[9] augmented the level set framework with two important geometric constraints, containment and exclusion, along with a distance prior for segmenting spatially recurring multi-region objects. They showed that only adding the containment and exclusion terms into the level set framework can improve the segmentation. They also proposed a method that is able to handle multiple instances of multi-part objects defined by these geometrical constraints using a single labeling function while maintaining global optimality.[10] Saveliev and Pahwa proposed a topological approach to cell counting and three years later he extended the idea and exploited the topological relationship between different objects in the image to identify and count the number of cells in the image.[11] Rathore *et al.*[12] employed a few existing segmentation techniques, and also proposed a new scheme to count total blood cells in a smear microscopic image, called capture largest included circles (CLIC), which is a parameterized segmentation algorithm that captures largest possible circles in an object boundary.

3. Objects Classification and K-Means Methods

3.1. *Classification of erythrocytes*

Repeated experiments corroborate that most of the proposed methods and automatic blood cell counting techniques can detect separate erythrocytes perfectly, the accuracy rate of which is very close to 100% ordinarily. But in practice, due to the great amount of erythrocytes and the insufficient manufacture of blood smear, there would ineluctably be some overlapping clusters formed by at least two erythrocytes in the blood smear images. The fatal error demonstrated by traditional RBC counting methods (like watershed algorithm) is terrible as the clustered erythrocytes would be detected as an individual. Thus, segmentation of cells was introduced into erythrocytes detection. Researchers tend to exploit varied segmentation which aimed at overlapping clusters that made a groundbreaking progress, hence the precision is greatly increased. However, to a certain extent, all the existing methods disregard objects' feature in the image: they apply their own complex and lengthy segmentation algorithm on every closed region even those regions that could be judged easily as individual erythrocytes (the majority in a blood smear image). Additionally, one irregularly shaped RBC would be segmented into two or three RBCs because of its warped edge. Maybe some morphological operations is applied to smooth the details of the whole smear image in pre-processing to avoid this mistake, but the erythrocyte clusters' edge will also be ruined which is crucial in segmentation.

It has become apparent that in order to detect erythrocytes exactly, we have to pay attention to the different shape characteristics between individual and overlapped RBSs. It stands to reason that individuals and clustered to be disposed, respectively will exhibit a good performance in counting. Also, the runtime will be reduced greatly. Details on classification are discussed in later sections. This targeted ideology deserved much more attention in the future research in this field.

3.2. *K-means methods applied on image processing*

The K-means clustering algorithm is a widely used partitioning method in many fields. It is a method of vector quantization, originally from signal processing, that is popular for cluster analysis in data mining. K-means clustering aims to partition n observations into k clusters in which each observation belongs to the cluster with the nearest mean, serving as a prototype of the cluster.[13]

In K-means clustering k points are chosen as the initial cluster centers. The choice of the number k of clusters is usually based on some heuristic. Each data point is assigned to the group that has the closest center. The cluster centers of mass are then recomputed. The assignment and re-computation steps are iterated until the intra-cluster variance converges to a minimum.[14] This algorithm belongs to the class of squared error algorithms which aim at minimizing some objective function in order to cluster the data. As with any clustering algorithm of partition, it is highly sensitive to the initial parameters: the number k of clusters and their centroids, respectively. Given an initial set of k means $m_{1,...,m_k}$, the Standard algorithm proceeds by alternating between two steps[15]:

(1) Assignment step: Assign each observation to the cluster whose mean yields the least within-cluster sum of squares. Since the sum of squares is the squared Euclidean distance, this is intuitively the "nearest" mean.[16]

$$S_i^{(t)} = \{x_p : \|x_p - m_i^{(t)}\|^2$$

$$\leq \|x_p - m_j^{(t)}\|^2 \forall j, 1 \leq j \leq k\}, \quad (1)$$

where each x_p is assigned to exactly one $s^{(t)}$, even if it could be assigned to two or more of them.

(2) Update step: Calculate the new means to be the centroids of the observations in the new clusters.

$$m_i^{(t+1)} = \frac{1}{|s_i^{(t)}|} \sum_{x_j \in s_i^{(t)}} x_j. \quad (2)$$

When applying K-means cluster algorithm on image processing, the gray value of each pixel constitute an array. After assigning the image based on K-means with number k, which is chosen the most appropriate to obtain a best partition of processing, values of the pixels in the gray image are replaced by the k centroids. Then the region we are interested in is obvious and can be segmented easily. Details on clustering are shown in later sections.

4. Proposed Complete Erythrocytes Counting Solutions

4.1. *Pre-processing*

After screening the blood smear by microscope, image with 1024×768 sizes was obtained (see Fig. 4(a)). An original image usually contains RBCs, WBCs, PLTs and maybe some stains on the blood smear. Hence, WBCs and PLTs should be wiped off before taking the count of erythrocytes. RGB and HSI are commonly used to describe color space of images. RGB color space (the red, green, and blue components of an image) is usually used in devices for color display, while HSI is suitable for human visual characteristic, so HSI color space is considered for image processing.[17] The HSI color space is developed to provide an intuitive representation of color and to approximate the way in which humans perceive and manipulate color. H represents the dominant color in its pure form, as in green, red or yellow. Adding white to the pure color changes the color; the less white, the more S the color is. This corresponds to S. The I corresponds to the brightness of the color.[18] The transformation from RGB to HSI can be obtained as follows:

$$H = \begin{cases} \theta, & B \leq G \\ 360 - \theta, & \text{else} \end{cases}, \tag{3}$$

$$S = 1 - 3^* \min(R, G, B)/(R + G + B), \tag{4}$$

$$I = (R + G + B)/3, \tag{5}$$

where

$$\theta = \arccos\left\{ \frac{[(R - G) + (R - B)]/2}{[(R - G)^2 + (R - G)(G - B)]^{1/2}} \right\}, \tag{6}$$

In the S component image, gray values of background pixels are almost lower than 0.1 as shown in Fig. 1, so objects can be easily detected by evaluating 0 to pixels whose gray value smaller than the threshold while 1 to pixels of others. With some morphological operations, noise and thrombocytes will be wiped out and it only contains regions of erythrocytes and leukocytes. By comparing R, G, B component in RGB color space and H, S, I component in HSI color space, regions of both the nucleus and the cytoplasm of leukocytes in H component image belong to similar gray level (see Figs. 2 and 4(d)). Hence, while applying K-means algorithm on
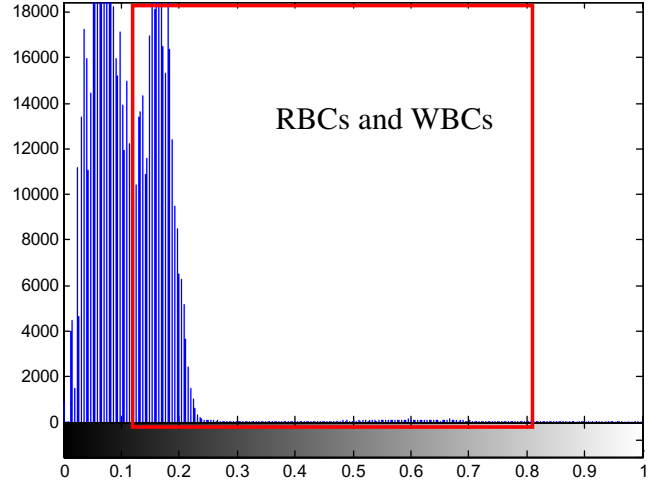

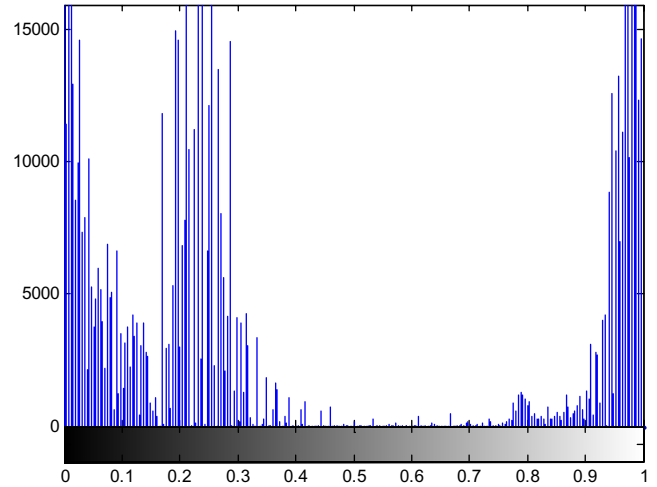Fig. 1.   *S* component histogram.

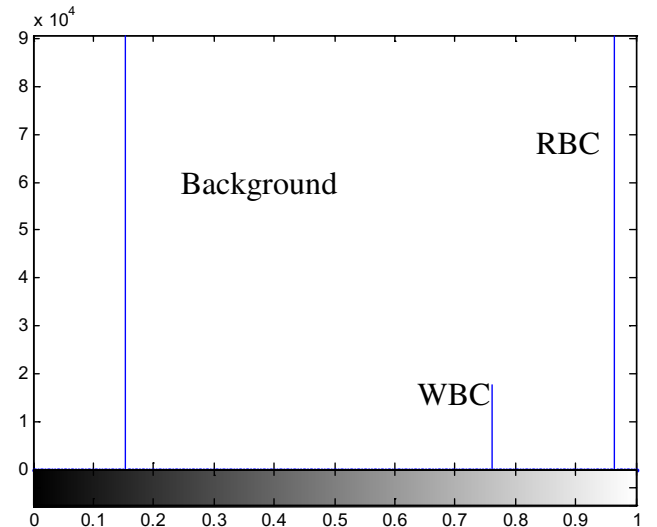Fig. 2.   *H* component histogram.

Fig. 3.   K-means processing.

Fig. 4(d) by setting the index k as 3 (RBCs, WBCs and background), leukocyte will be separated from erythrocytes and background (see Fig. 3). Changing the index would not encourage a better outcome. So regions of leukocytes can be detected easily and then replaced by backdrop, thus after denoising there are only erythrocytes in the image. Figure 4(f) shows the result after executing pre-processing.

Sequential steps of pre-processing are:

Step 1. Read the input microscope blood smear image.

Step 2. Extract the S and H component image.

Step 3. Pixels whose gray value is bigger than a threshold assigned by 1 than others assigned by 0 in S component image.

Step 4. Dislodge noise by removing areas of image smaller than an area threshold.

Step 5. Perform morphological operation of open and hole-filling (see Fig. 4(c)).

Step 6. Detect the regions of leukocyte based on K-means in H component image (see Fig. 4(e)).
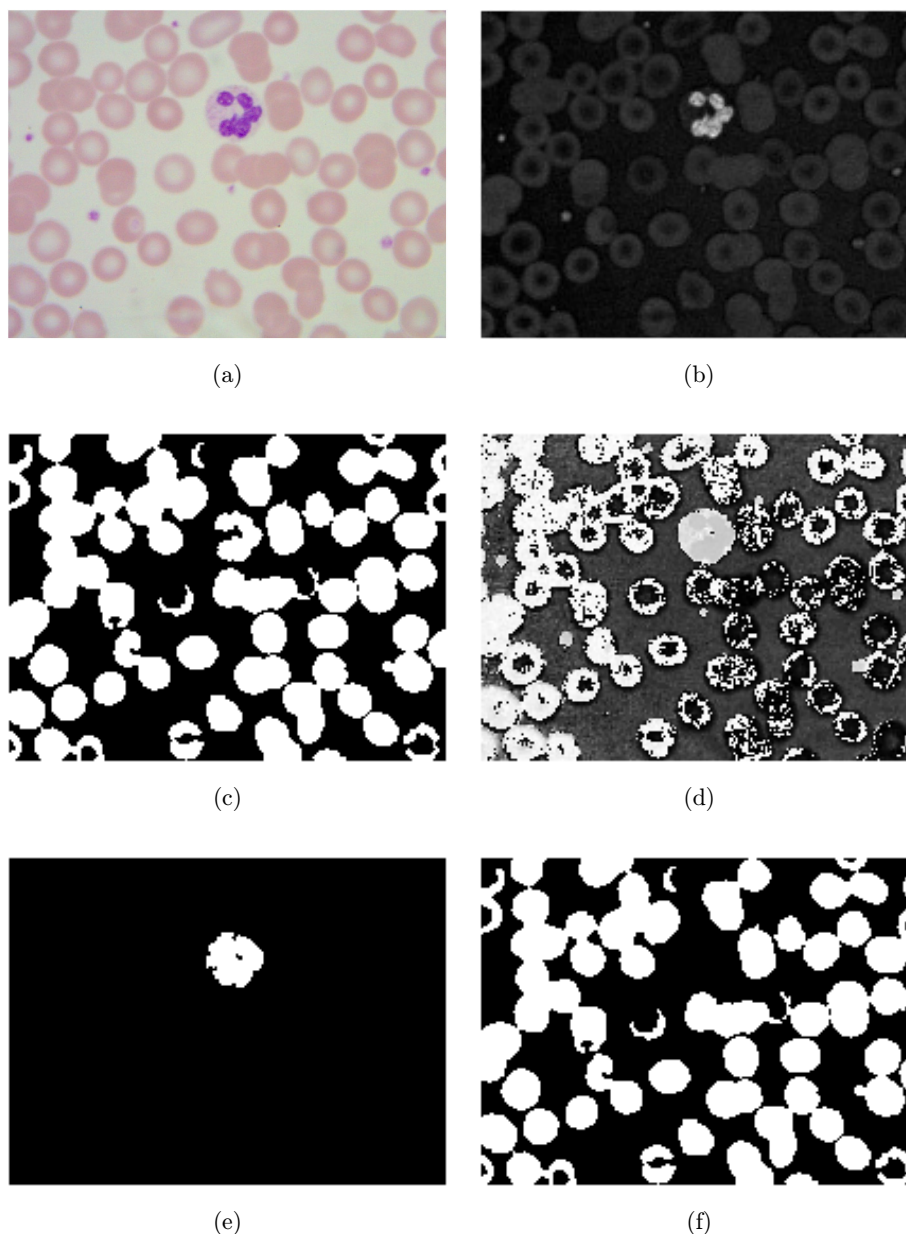

(a)

(b)

(c)

(d)

(e)

(f)

Fig. 4.   Image pre-processing (a) original blood smear image, (b) S component image, (c) detection of RBCs and WBC, (d) H component image, (e) segmentation of WBC and (f) erythrocytes only.

Step 7. By doing "IMAGE SUBTRACT" operation between images obtained in Steps 3 and 5, there are only erythrocytes in the image (see Fig. 4(f)).

4.2. *Classification*

According to the discussion in Sec. 3.1, before taking count of erythrocytes it is necessary to divide all the closed regions into two classes: individual ones and overlapping clusters. Notice that the area sizes of the two classes is obviously different. Areas of clusters are at least 1.5 times of individuals. There is no need to discuss this apparent discrepancy in detail. Another conspicuous difference between individuals and clusters is the circularity index, which was calculated using the following formula:

$$\text{CI} = 4\pi \frac{A}{P^2}, \qquad (7)$$

where CI, P and A represent the circularity index, perimeter and area, respectively. This results in a range of CI values between 0 and 1, with a CI of 1 indicating a perfect circle.[20] One erythrocyte's shape is very close to a circle and the average CI of individual ones is 0.8733 (relied on 1000 samples of individual erythrocytes in this study). In stark contrast to individuals' high CI, closed regions of overlapping clusters have a low CI due to their irregular shape and screwy edge. In the classification, the first step should be the calculation of every closed region's area and CI. While one object met both the qualifications: its area is bigger than the average size and its CI is lower than the threshold, then it will be judged as a cluster. Then this region will be covered by background. After applying this judgment on each object in the image which only contains erythrocytes (see Fig. 4(f)), images of the two classes will be presented. Results of classification are shown in Fig. 5.

Sequential steps of classification are:

Step 1. Calculate the area size and circularity of each closed region.
Step 2. Calculate the average area size and average circularity as the criteria of individuals.
Step 3. Execute Judgment No. 1: Mark all the regions whose area is more than 1.5 times of the average. Regions that do not meet the guideline will be removed.

(a)
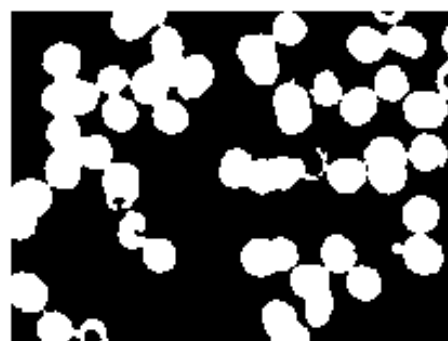
(b)

Fig. 5. Classification of erythrocyte (a) individual ones and (b) overlapping erythrocytes clusters.

Step 4. Execute Judgment No. 2: Mark all the regions left (in Step 3) whose CI is smaller than the average. Regions that do not meet the guideline will be removed (see Fig. 5(b)).
Step 5. By doing "IMAGE SUBTRACT" operation between Fig. 4(f) and image obtained in Step 4, individual erythrocytes will be displayed (see Fig. 5(b)).

4.3. *RBC counting*

The above preliminary processing has provided remarkable image of different classes of erythrocytes to be measured in the next step: Segmentation of overlapping clusters with K-means and taking count of individuals and clusters, respectively.

4.3.1. *Aim at individual ones*

From an implementation standpoint, the accuracy and precision are pretty good while applying most counting algorithm on simple structure objects like

individual erythrocytes. In this counting link, we chose watershed algorithm which has an easy operate processing on image and a high speed.

4.3.2. *Aim at overlapping clusters*

In the actual microscopic blood smear images, some erythrocytes overlapping into a group is a normal phenomenon. These RBCs crowd together and appear as a big close region in the 2D plane original images. As a matter of fact, they pile up like toy bricks in the 3D space and their own structure feature would not change. We define RBCs overlying on others as upper strata and RBCs be under pressure by the upper ones as lower strata. Since the shooting angle of the microscopic images is perpendicular to the blood smears, part of lower strata ones are ward off by upper ones so that only the RBC located in the upper strata has a complete imaging in the original images. So it can be inferred that upper strata's contour profiles are more obvious and complete, and they can be captured much more easily than lower ones. Human being's RBCs are biconcave in shape as is known to all and color of RBCs' middle pit is lighter than the border due to this feature. It is easy to find that regions of upper strata's border-rings are more conspicuous and complete, the color of these areas are darker than other regions in original images. There is one more point that should be explained that lower strata lay close to background so their color is blended with background and they will turn to pink while upper strata's degree of being blended is lower than lower strata in original images. So erythrocytes belong to upper strata have a complete border-ring which is impossible to lower ones, and these border-rings' color are darker due to the overlap. This border-ring

is the key in the proposed method. Computing the R, G, B, H, S and I component of overlapping regions respectively, R component can illustrate all the differences discussed above. But the gray level of the closed regions is a gradient in which the gray value of the middle concavity is lowest while the border has a bigger gray value. To improve the vague dividing line between the border-ring and the concavity, K-means algorithm was applied to amplify this difference of gray value between different parts in one same closed cluster region. The summing-up through the above analysis can be summed as: The gray value of the border-rings of erythrocytes in the upper strata is the highest while lower ones' incomplete borders have smaller gray value and the concavities take the third place due to the biconcave structure feature. The background's gray value is the smallest (0). Hence, we apply K-means algorithm on cell segmentation and intercalate the index k as 4 to partition pixels in image into four clusters, so that the erythrocytes lay on the upper strata with a complete contour profile in the overlapping clusters were segmented successfully, and other ones would be captured by calculating the complementary set of Fig. 5(b). Taking count of these two classes will be carried out one by one simultaneously. It is worthy to illustrate that there will still be a small amount of overlapping erythrocytes which cannot be separated successfully. In this situation, we execute a judgment of the sizes of these regions: if the area size is bigger than 1.5 times of average which has been calculated in classification, and meanwhile smaller than 2.5 times of average, we deem that there are two erythrocytes in this area. The two classes are shown in Figs. 6(b) and 6(c).

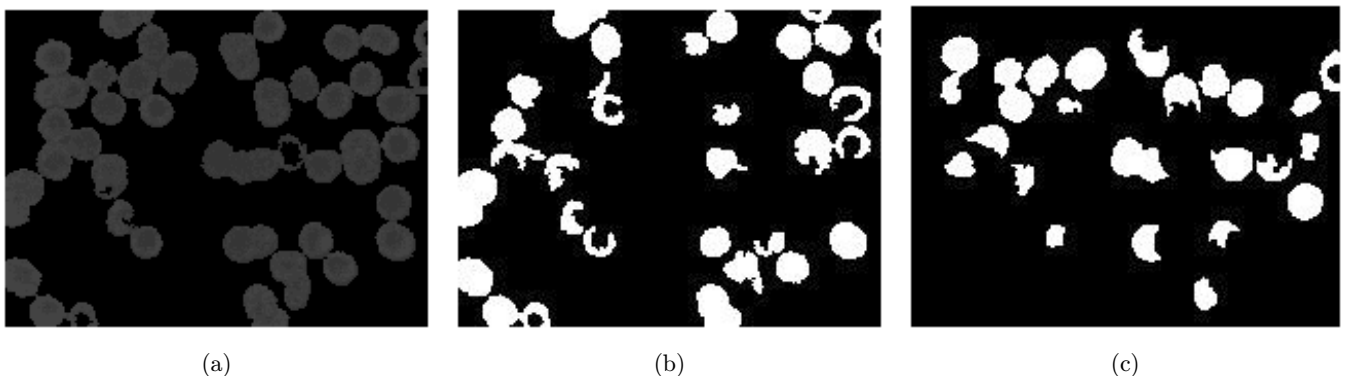(a)   (b)   (c)

Fig. 6.   Complex overlapping erythrocyte clusters counting. (a) Cluster with K-means, (b) upper strata ones and (c) lower strata ones.

This algorithm to compute number of overlapping RBCs can be described step by step as follows:

Step 1. Count individual erythrocytes (closed regions) with watershed based on Fig. 5(a).

Step 2. Do "IMAGE SUBTRACT" operation between R component of original image and Fig. 5(b). So we acquire an image that only contains R component of cluster areas.

Step 3. Cluster with K-means.

Step 4. Detect the regions of upper strata erythrocytes and remove others.

Step 5. Perform morphological operation of open and hole-filling (see Fig. 6(b)).

Step 6. Remove upper strata in the image obtained in Step 3 (see Fig. 6(c)).

Step 7. Count erythrocytes in Figs. 6(b) and 6(c).

Step 8. Sum the calculation achieved in Steps 1 and 7.

5. Experiments and Result Analysis

Peripheral blood smears are prepared with Wright–Giemsa stain. Degrees of dye absorption are different for every blood smear that depend on some influence factors, such as environment, features of erythrocytes and so on. Lots of microscopic images with 1024×768 sizes are counted. Eight representative samples are chosen to validate the experimental results.

Figure 7(a) are eight original microscopic images while the experimental processes are shown in Figs. 7(b)–7(f): Figure 7(b) shows the detection of WBCs in each original image. Figure 7(c) contains RBCs only. Figure 7(d) are individual erythrocytes which are detected after classification. Overlapping erythrocytes clusters are shown in Fig. 7(e). We taking count of different classes of RBCs step by step we use different colors to mark different categories in order to make a legible exhibit of our algorithm (see Fig. 7(f)).

A dataset consisted of 100 test images was divided into 10 sets and the experimental results were shown below to prove the performance of this proposed method. Precision (PR), Recall (RC) and F-measurements (FM) were used to determine segmentation and counting accuracy with the following

equations:

$$PR = \frac{TP}{TP + FP}, \qquad (8)$$

$$RC = \frac{TP}{TP + FN}, \qquad (9)$$

$$FM = \frac{2}{\frac{1}{PR} + \frac{1}{RC}}, \qquad (10)$$

where TP expresses the true positive value which is an agreement between the expert and the proposed method. The false negative (FN), which is when the proposed method was unable to detect the cell but the expert detected a cell, and the false positive (FP), which is when the proposed method detected a cell but the expert did not.[20]

Table 1 summarizes the segmentation and counting accuracy results of the proposed method for RBCs. The following is the segmentation and counting accuracies of RBCs that were calculated using our proposed method: PR = 98.4%, RC = 92% and F-measure = 95.2%. Some discussions were made as follows:

Precision and accuracy: The total accuracy of the counting result is satisfactory which reaches 95%. Precision of images whose RBC clusters only be constituted by 2–3 RBCs present a quite high level which reaches almost 99%. Some of the experimental results counted by the proposed method present a high FN rate and leads to 92% average RC. The high FN rate means a certain number of RBCs still cannot be segmented successfully. The main reason maybe the degree of erythrocyte clusters' overlapping situation is too complex to be separated. Notice that the more erythrocytes an original image contains, the lower counting accuracy is present. It is mainly because that with the increase of RBCs' amount, the distribution of RBCs will be more intensive and the degree of overlapping will be more complex. Considering that even experts cannot distinguish some RBC clusters specifically, the error is acceptable.

Parameter selection: In this method, some morphological operations are used in order to obtain better images during image processing. They will bring no substantial changes to every object. But the index k in K-means algorithm must be set up as their current values are according to the explanation in Secs. 4.1 and 4.3.2, changing their values cannot encourage a better outcome.
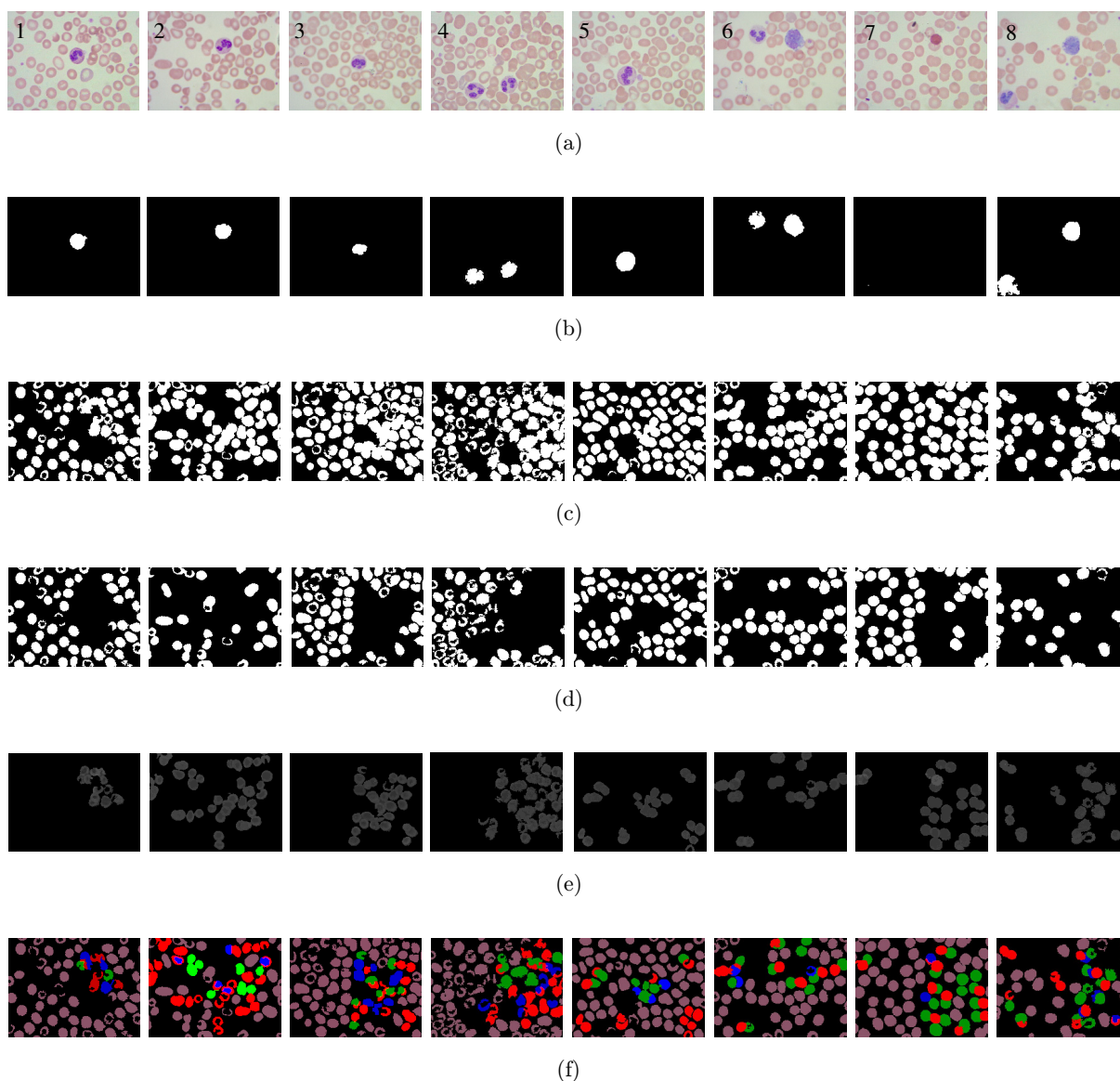
(a)

(b)

(c)

(d)

(e)

(f)

Fig. 7.   Representative experimental results. (a) original microscopic smear blood images, (b) detection of WBC, (c) erythrocytes only, (d) classification of individual ones, (e) clustering with K-means of overlapping ones and (f) different classes of erythrocytes.

Table 1.   Experimental results of the dataset.

Set	Artificial counting	Counting by proposed method	TP	FN	FP	PR (%)	RC (%)	FM (%)
1	561	549	549	12	0	100	98	98.92
2	847	823	807	40	16	98.1	95	96.65
3	1043	998	986	57	12	98.8	95	96.62
4	977	956	943	34	13	98.6	97	97.57
5	1174	1068	1043	131	25	97.7	89	93.04
6	1198	1114	1095	103	19	98.3	91	94.72
7	894	883	877	17	6	99.3	98	98.71
8	1437	1297	1254	183	43	96.7	87	91.73
9	1376	1225	1216	160	9	99.3	88	93.50
10	1271	1194	1172	99	22	98.2	92	95.09
Total	10,778	10,107	9942	836	165	98.4	92	95.21

Table 2.   Results by different counting algorithm.

Image serial	Methods automated	Number of separate RBCs	Counting Results	Separate RBC accuracy (%)	Number of overlapping RBCs	Counting results	Overlap RBC accuracy (%)	Number of all the RBCs	Counting results	Total accuracy (%)
1	Watershed	58	58	100	13	3	23.1	71	61	85.9
	Level set		58	100		4	30.8		62	87.3
	CLIC		57	98.3		5	38.5		62	87.3
	Proposed		58	100		11	84.6		69	97.2
2	Watershed	30	30	100	40	11	27.5	70	41	58.6
	Level set		30	100		13	32.5		43	61.4
	CLIC		27	90		16	40.0		43	61.4
	Proposed		30	100		35	87.5		65	92.9
3	Watershed	67	64	95.5	30	6	20.0	97	70	72.2
	Level set		64	95.5		9	30.0		73	75.3
	CLIC		65	97.0		11	36.7.		76	78.4
	Proposed		64	95.5		27	90.0		91	93.8

Restrictions: So as long as an erythrocyte is biconcave in shape, its border-ring can be obtained, even the irregularly shaped RBCs can be captured by the proposed method. Unfortunately, this proposed method cannot do well in taking count of hypotonic or hypertonic cases of erythrocytes, as their smooth surface cannot be used in segmentation. The essentiality of erythrocytes' biconcave shape to the segmentation has been explained in Sec. 4.3.2.

Runtime: Unlike some other segmentation algorithms, there is no tedious data operation in the proposed method. The longest time-consuming link is applying K-means on segmentation. Using non-optimized computer program code on standard 2.3 GHz CPU, the average running time of this method is 49.86 s/image. Automated Segmentation of Overlapping Cervical Cells proposed by Lu *et al.*[21] takes 56 s per cell on average. The running time of Masoud's algorithm[22] ranged between 60 and 300 s/image, which depended primarily on the number of cells per image. Generally speaking, the proposed method of taking count of erythrocytes saves a pretty substantial amount of time compared with other methods.

We also make a comparison with other algorithms dealing with three images randomly selected. Though there is only a slight improvement in case of two or three erythrocyte clusters into a group, this algorithm outclasses other segmentation algorithms for the cases while more than three RBCs are overlapping together. Table 2 shows the comparison results.

From Table 2, it is evident that one of the major advantages of this new measure of erythrocytes counting is that it can detect overlapping clusters efficiently, which provides a powerful foundation of exactitude counting. Compared to the existing algorithms, the accuracy of total counting results are generally advanced by 10% to 15% and the accuracy of clusters counting results have a great improvement by almost 50% on average.

Although much effort has been made to this approach, it is far from completion. For example, it is deeply affected by erythrocytes shape feature. Segmentation is still the most difficult choke point which ought to be kept on deep research.

6.   Conclusion

In this paper, a complete erythrocyte counting method based on blood smear microscope imaging is proposed. After extracting erythrocytes based on H and S components of original images, all the closed regions will be classified into two categories: individual ones and overlapping clusters. K-means algorithm is applied in the segmentation of clusters, and then individuals and clusters will be counted, respectively. Results have proven the superiority of our proposed scheme especially in case of overlapping erythrocyte clusters. Though counting results are quite promising, our approach still needs

to be improved. It can further be fine-tuned to produce more robust results. Counting the number of blood cells in a blood smear image is a standard medical test that plays vital role in diagnosis of various diseases. The method can be used in blood test for a medical use. With a high accuracy and a fast runtime, it will provide a better support to health workers.

Acknowledgments

References

1. N. Bosschaart, G. J. Edelman, M. C. G. Aalders, "A literature review and novel theoretical approach on the optical properties of whole blood," *Lasers Med. Sci.* **29**, 453–479 (2014).

2. G. Tomaiuolo, "Biomechanical properties of red blood cells in health and disease towards microfluidics," *Biomicrofluidics* **8**, 051501 (2014).

3. http://www.nlm.nih.gov/medlineplus/ency/article/003644.htm (accessed on 28 March 2015).

4. P. Memmolo, L. Miccio, "3D morphometry of red blood cells by digital holography," *Cytometry A* **85A**, 1030–1036 (2014).

5. H. Gao, P. Xue, W. Lin, "A new marker based watershed algorithm," *Proc. Int. Symp. Circuits and Systems*, Vol. 2 II-81-4 (2004).

6. L. Xu, H. Lu, M. Zhang, "Automatic segmentation of clustered quantum dots based on improved watershed transformation," *Digital Signal Process.* **34**, 108–115 (2014).

7. A. Dervieux, F. Thomasset, "A finite element method for the simulation of Rayleigh-Taylor instability," *Lecture Notes Math.* **771**, 145–158 (1980).

8. C. Li, C. Xu, C. Gui, M. D. Fox, "Distance regularized level set evolution and its application to image segmentation," *IEEE Trans. Image Process.* **19**(12), 3243–3254 (2010).

9. M. Nosrati, G. Hamarneh, "Local optimization based segmentation of spatially-recurring, multi-region objects with part configuration constraints," *IEEE Trans. Med. Imaging (IEEE TMI)*, **33**(9), 1845–1859 (2014).

10. M. Nosrati, S. Andrews, G. Hamarneh, "Bounded labeling function for global segmentation of multi-part objects with geometric constraints," *IEEE Int. Conf. Computer Vision (IEEE ICCV)*, pp. 2032–2039 (2013).

11. P. Saveliev, "A graph, non-tree representation of the topology of a gray scale image," *Proc. SPIE*, J. T. Astola, K. O. Egiazarian, Eds., — The International Society for Optical Engineering, Bellingham, WA 2011 Image Processing: Algorithms and Systems IX, Vol. 7870, pp. 1–19 (2011).

12. S. Rathore, A. Iftikhar, A. Ali, M. Hussain, A. Jalil, "Capture largest included circles: An approach for counting red blood cells," *IMTIC 2012*, CCIS 281, pp. 373–384 (2012).

13. http://en.wikipedia.org/wiki/K-means_clustering#cite_ref-8 (accessed on 28 March 2015).

14. L. Galluccio, O. Michel, P. Comon, A. O. Hero III, "Graph based k-means clustering," *Signal Process.* **92**, 1970–1984 (2012).

15. D. J. C. MacKay, "An example inference task: Clustering," *Information Theory, Inference and Learning Algorithms*, Chapter 20, Cambridge University Press. pp. 284–292 (2012).

16. M. Mahajan, P. Nimbhorkar, K. Varadarajan, "The planar k-means problem is NP-hard," *Lecture Notes Comput. Sci.* **5431**, 274–285 (2009).

17. Y. Yang, Y. Cao, W. Shi, "A method of leukocyte segmentation based on S component and B component images," *J. Innov. Opt. Health Sci.* **7**(1), 1450007:1–8 (2014).

18. N. Li, R. Wang, "Unsupervised polarimetric synthetic aperture radar classification of large-scale landslides caused by Wenchuan earthquake in hue-saturation-intensity color space," *J. Appl. Remote Sens.* **8**, 083595-1–8 (2014).

19. S. Nakano, T. Nakano, "Change in circularity index of cell lumen in a cross-section of wood induced by aqueous NaOH," *J. Wood Sci.* **60**, 99–104 (2014).

20. Y. M. Alomari, S. N. H. S. Abdullah, "Automatic detection and quantification of WBCs and RBCs using iterative structured circle detection algorithm," *Comput. Math. Methods Med.*, Vol. 2014, pp. 1–17 (2014)

21. Z. Lu, G. Carneiro, A. P. Bradley, "Automated nucleus and cytoplasm segmentation of overlapping cervical cells," *MICCAI 2013, Part I*, Lecture Notes in Computer Science, Vol. 8149, pp. 452–460 (2013).

22. M. S. Nosrati, G. Hamarneh, "Segmentation of cells with partial occlusion and part configuration constraint using evolutionary computation," *MICCAI 2013, Part I*, Lecture Notes in Computer Science, Vol. 8149, pp. 461–468 (2013).