# The Milk Somatic Cell Image Segmentation Method Based on Dimension Reduction and Fusion

Jie Bai, Heru Xue[(✉)], and Yanqing Zhou

College of Computer and Information Engineering,
Inner Mongolia Agricultural University, Huhhot 010010, China
`baijie9554@l63.com, xuehr@imau.edu.cn,`
`974864l34@qq.com`

**Abstract.** Milk somatic cell image segmentation has a very important effect on milk quality analysis. In this paper, a new segmentation method is proposed for color milk somatic cell image. Aiming at the milk somatic cell image, a new method for image segmentation is proposed. First of all, selecting the appropriate color components by reducing the dimension of RGB color image; Then using the k-means algorithm to segment the low dimension image; Finally, the segmented images are fused by the region splitting and merging process. The experimental results show that the proposed method is better than the original three-dimensional (3D) color space segmentation method. The method improves in the performance and running time, the correct segmentation rate reached 98.4 %. Therefore, this method has certain feasibility.

**Keywords:** Milk somatic cell image · Dimension reduction · k-means algorithm · Image fusion

## 1 Introduction

With the development of computer application technology, graphics and image processing has penetrated into all walks of life. the research on somatic cell bovine milk is a very important research field. The number of milk somatic cells is an important index to judge the quality of milk [1]. The more somatic cell counts, the higher the incidence of mastitis is. But the image segmentation is the first step of the milk somatic cell counting, which has very important significance to improve the quality of milk cow mastitis detection and diagnosis.

At present, domestic and foreign scholars for the segmentation of cell image are also put forward a series of effective methods, such as Threshold segmentation method, Edge detection method and Region growing method. Threshold segmentation method has the advantages of simple algorithm and is easy to realize, which can use boundary definition closed and connected overlapping regions. The disadvantage is difficult to obtain the accurate object boundary for the small difference between object and background [2]. The advantage of edge detection method is precise contour location, and the disadvantage is the inability to guarantee the closed contour [3]. Region growing method can get a closed contour, but it is hard to determine the growth of

termination conditions [4]. In biological cells, the cell shape varies, so there is no universal method for all cell images.

By comparing the existing algorithms, selecting reasonable low dimensional space and using K-means algorithm and image fusion technology to segment the milk somatic cell images can get better segmentation results, and it provides a basis for the further analysis of milk somatic cell image.

The remaining of the paper is organized as follows. Section 2 describes how to select the appropriate color components and implement the segmentation in the low dimensional space. Section 3 presents a fusion of these segmentation results and gets a final segmentation. Section 4 evaluates the segmentation results. conclusion is given in Sect. 5.

## 2 The Milk Somatic Cell Image Segmentation Base on the Low Dimensional Space

The color image contains more information than the gray image, so in many applications of computer vision and pattern recognition, the color image can obtain better results [5]. This paper uses the color milk somatic cell image. The processed color image will produce the problem of large computation and slow processing speed. So in order to overcome the above problems, it is necessary to reduce the dimensionality of 3D color space. In this paper, the 3D color space is projected to the 2D color space [6], and the k-means segmentation is implemented in 2D color space.

### 2.1 Select the Low Dimensional Space for the Milk Somatic Cell

In this paper, the milk somatic cell image is RGB color image. Therefore, it is necessary to analyze its color components, In the process of analysis, through calculating the standard deviation and the histogram of each color component, to select the spectral component with the most abundant informations. In the final analysis, the R component of the standard deviation is 0.1865, G is 0.1469, B is 0.0216. The histogram of different color components is shown in Fig. 1.
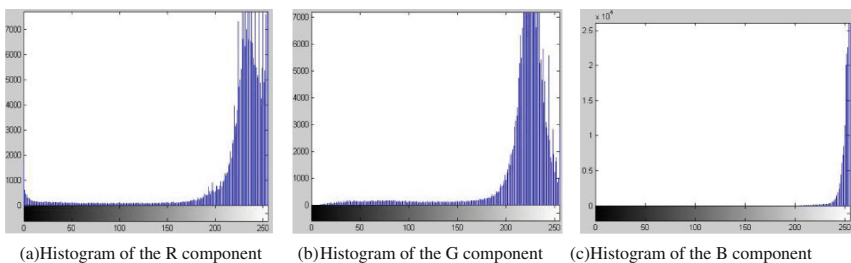


(a)Histogram of the R component    (b)Histogram of the G component    (c)Histogram of the B component

**Fig. 1.** Histograms of different color components (Color figure online)

By observing the color component histogram and the compute results of the standard deviation, the standard deviation of R is the largest, B is the smaller. It can be

concluded that the distinguishing effect is from strong to weak is R, G, B. It also can be say that the R and G components have the significant difference in the nucleus, cytoplasm and the background. In this paper, the segmentation is performed in RG channel and RB channel by using k-means method.

### 2.2    Using K-Means to Segment the Milk Somatic Cell in Low Dimensional Space

Using the k-means algorithm, first of all, selecting the initial cluster centers, and then classifying all data points. In the end, the average value is calculated to adjust the cluster center for the each cluster, and continuous loop iteration, which makes the evaluation of clustering criterion function to achieve optimal performance, in order to reach the compactness within class and the independence in class of each cluster [7]. The image segmentation steps are described as follows:

1. Select the $k$ sample points in the sample data sets $N$, and the k sample values are assigned to the initial clustering center ($\mu_1^{(1)}$, $\mu_2^{(1)}$, …, $\mu_k^{(1)}$);
2. When the $j$ iteration, to all points in the sample points $P_t$(t = 1, …, n), followed by calculating the Euclidean distance d ($t$, $i$) of each cluster center $\mu_i^{(j)}$;

$$d(t,i) = \sqrt{(P_t - \mu_i^{(j)})^2} \tag{1}$$

3. Find out the minimum distance of $P_t$ and $\mu_i^{(j)}$, put $P_t$ into the minimum distance cluster of $\mu_i^{(j)}$;
4. Update the clustering center of each cluster

$$\mu_i^{(j+1)} = \frac{1}{n_i} \sum_{t=1}^{n_i} p_{it} \tag{2}$$

5. Calculate the square error $E_i$ of all the points in the data sets $N$, and compared with previous error $E_{i-1}$

$$E_i = \sum_{i=1}^{k} \sum_{t=1}^{n_i} |P_{it} - \mu_i^{j+1}|^2 \tag{3}$$

If $|E_{i+1} - E_i| < \delta$, the algorithm ends, or go to (2) iterative again. The segmentation results are shown in Fig. 2.

Figure 2 shows that it has different segmentation results by using k-means algorithm in the RG space, RB space and the original three-dimensional space. And the original image segmentation results (Fig. 2(f)) compared to in RG space segmentation results (Fig. 2(b)) the background is misclassification into cytoplasm, in RB space segmentation results (Fig. 2(d)) the nucleus is misclassification into cytoplasm. It can be seen that the effect of segmentation in the two two-dimensional subspaces is not the best, which is due to loss a color component. Therefore, In order to overcome the above problems, it is need to use the image fusion technology to fusion the segmentation results of subspace, and obtain the more accurate segmentation results.
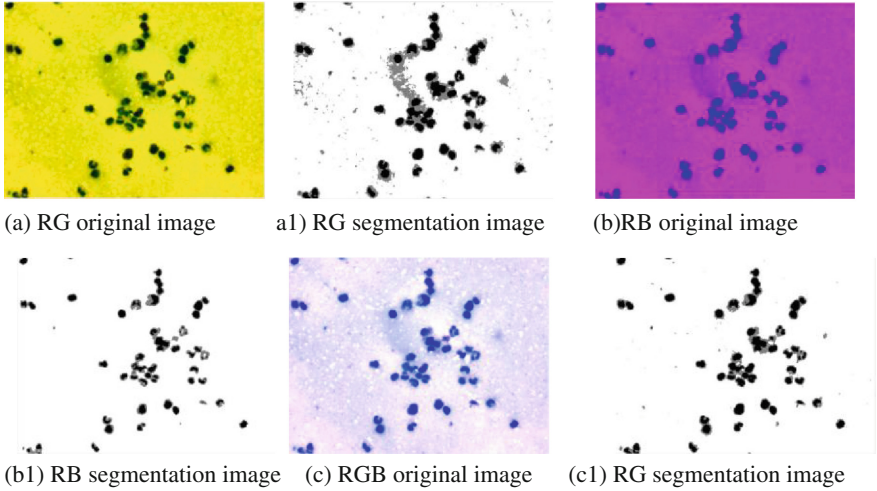
(a) RG original image    a1) RG segmentation image    (b)RB original image



(b1) RB segmentation image    (c) RGB original image    (c1) RG segmentation image

**Fig. 2.** Segmentation results of milk somatic cell image by k-means algorithm

## 3    Fusion Results of Low Dimensional Space Segmentation

Image fusion is the spatial matching of image data from different sources, and using complementary organic combination and producing new image data. This new data has a description of the research object and better representation of information, compared with the single information source, it can reduce or inhibit the ambiguity, incompleteness, uncertainty and error to explain the perceived object or environment and maximize the use of the information provided by the various sources of information [8, 9].

The fusion of milk somatic cell image is divided into the following steps. The first step is to apply the k-means segmentation process in two 2D subspaces, and then use the region splitting and merging process to segment again. Finally it gets a better segmentation image. The specific segmentation process as shown in Fig. 3.

The region splitting and merging process is briefly described as follows:

Inputing the 3D color space image I, Extracting it 2D subspace and implementing k-means algorithm in two 2D subspaces, the segmentation image are denoted by $I_{RG}$ and $I_{RB}$. During the region splitting, The image I is divided into class L in the RG subspace, we denote by $I_{RG}^{(l)}$, $l = 1, 2, \ldots, L_{RG}$ and $I_{RG}^{(1)} \cup I_{RG}^{(2)} \cup \ldots \cup I_{RG}^{(1)} = I_{RG}$, $I_{RG}^{(1)} \bigcap I_{RG}^{(t)} = \varphi(l \neq t)$, The RB subspace can also be obtained in segmentation result and expression of symbols for $\forall l \in \{1, 2, \ldots, L_{RG}\}$ and $t \in \{t_1, t_2, \ldots, t_k\} \subset \{1, 2, \ldots, L_{RB}\}$, The region $I_{RG}^{(1)}$ is subdivided into $I_{RG}^{(1)} \cap I_{RB}^{(t_1)}$, $I_{RG}^{(1)} \cap I_{RB}^{(t_2)}$, $\ldots$, $I_{RG}^{(1)} \cap I_{RB}^{(t_k)}$ and $I_{RG}^{(l)} - I_{RG}^{(t)} \cap \left( I_{RB}^{t_1} \cup I_{RB}^{t_2} \cup \ldots \cup I_{RB}^{t_k} \right)$. Above is the splitting process, this process has realized the secondary segmentation of $I_{RG}$ relative to $I_{RB}$, the RG subspace image is further divided into many subclasses, the results of segmentation is denoted by $I_{RG+RB}$. Computing the respective regional of the average pixel components R, G, B which is reclassified, to
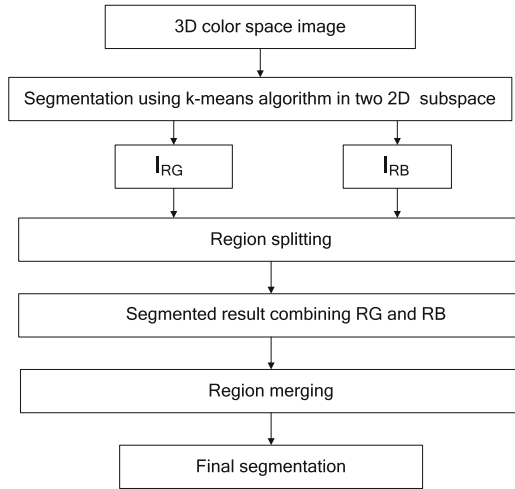
**Fig. 3.** The flow chart of milk somatic cell image segmentation

compute the average pixel components is assigned to all the pixels belonging to this class. At the end of the splitting process, the number of sub class may be larger than L. So it is necessary to carry out the above segmentation image region merging process. In this paper, by comparing the two color points of Euclidean distance $D = \sqrt{(r1 - r2)^2 + (g1 - g2)^2 + (b1 - b2)^2}$, the smallest color distance of two classes are first merged and the mean color values for new regions are updated. This process is iterative, until the image to achieve a specified class number date. In this experiment, the milk somatic cell image should be divided into three classes, namely the nucleus, cytoplasm and background.

The experiment is performed in Matlab2011b environment to achieve the above algorithm. This paper uses the milk somatic cell image that comes from the Inner Mongolia Agricultural University laboratory animal pathology. The image using
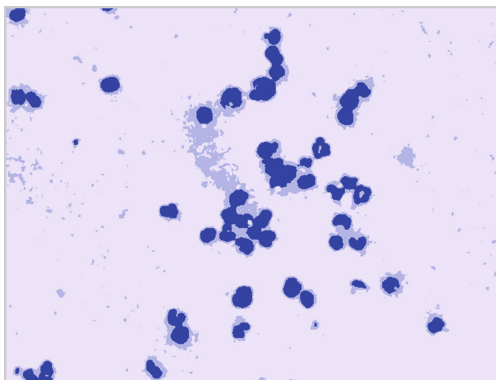


**Fig. 4.** The results of subspace images segmentation and fusion

Newman's staining and it is magnified by 40× oil lens, which is stored as a 24 bit color image. In this experiment, it can be divided into three categories, namely the nucleus, cytoplasm and background. The final result using k-means and fusion approach in two 2D subspaces as depicted by Fig. 4.

## 4 Experimentation and Results

In this paper, through computing the image segmentation correct rate, the quantitative results are evaluated. Firstly, a standard and stable method is used to segment the image as a reference image segmentation. The FCM [10] method is used to segment the milk somatic in the experiment, then the comparison calculation of segmented image and the reference image are made. If the nucleus is divided into the pixel numbers of nucleus is $a$, the cytoplasm is divided into the pixel numbers of cytoplasm is $b$, and the background is divided into the pixel numbers of background is $c$. The accuracy will be calculated: $(a + b+c)/n$, and n is the total number of the image.

Table 1 is the time comparison with the milk somatic cell image by k-means method in different color spaces. Tables 2 and 3 is the accuracy of the segmentation which is used the paper method and directly segment in 3D space using k-means method. As Table 2 shows, the segmentation accuracy is 98.4 %, and the Table 3 shows that the segmentation accuracy is 95.4 %. Through many times of experimenting, we can know that both segmentation methods in 3D color space and in 2D subspace can get the correct results. However the correct rate of 2D method is higher than using 3D method directly and the running time is shorter. So the method proposed in this paper has some maneuverability and practicability.

**Table 1.** Comparison between different methods

|                | 3D method | 2D method |
|----------------|-----------|-----------|
| Running time(s) | 7.4 s     | 3.9 s     |

**Table 2.** Segmentation accuracy using 2D method

| 2D method |            |         |           |            |        |                              |
|-----------|------------|---------|-----------|------------|--------|------------------------------|
| Reference method | Classes    | Nucleus | Cytoplasm | Background | Total  | Accuracy                     |
|           | Nucleus    | 15060   | 0         | 0          | 15060  | (15060 + 18632 + 278358)/    |
|           | Cytoplasm  | 0       | 18632     | 5150       | 23782  | 317200 = 98.4 %              |
|           | Background | 0       | 0         | 278358     | 278358 |                              |

**Table 3.** Segmentation accuracy using 3D method

| 3D method |            |         |           |            |        |                              |
|-----------|------------|---------|-----------|------------|--------|------------------------------|
| Reference method | Classes    | Nucleus | Cytoplasm | Background | Total  | Accuracy                     |
|           | Nucleus    | 12960   | 2100      | 0          | 15060  | (15060 + 11306 + 278358)/    |
|           | Cytoplasm  | 0       | 11306     | 12376      | 23782  | 317200 = 95.4 %              |
|           | Background | 0       | 0         | 278358     | 278358 |                              |

## 5    Conclusions

Milk Somatic cell image segmentation is an important part of milk somatic cell image processing, due to color image with large amounts of data, segmentation directly in 3D color space will produce the problem of slow processing speed. However in this paper, we have presented an efficient method for the milk somatic cell image segmentation. Firstly, through analyzing the effect of each color component, the 3D color space is projected onto the low dimensional space. Secondly, the k-means algorithm is implemented in the low dimensional space to segment. Finally using region split and merge methods of the image fusion to obtain the final segmentation result. By comparison the method, the method proposed in the paper not only overcomes the disadvantages of the large amount of calculation and lower speed in 3D color space segment directly, but also improves the accuracy of segmentation. Therefore, the proposed method is feasible and has certain theoretical and practical significance.

## References

1. Hai-xia, Liu: Screening test of somatic cell count in bovine milk. China Dairy Ind. **06**, 31–34 (2004)
2. Han, S., Wang, L.: A survey of thresholding methods for image segmentation. Syst. Eng. Electron. **06**, 91–94 + 102 (2002)
3. Mohamed, H.T., Bouhlel, S., Derbel, N., Kamoun, L.: A Surrey and Evaluation of Edge Detection Operators Application to Medical Images. IEEE (2002)
4. Chang, Y.L., Li, X.B.: Adaptive image region-growing. IEEE-IP **3**(6), 868–872 (1994)
5. Xue, H., Ma, S., Pei, X.: Color image segmentation based on mathematical morphology and fusion. J. Image Graph. **11**(12), 1764–1767 (2006)
6. Kurugollu, F., Sankur, B., Harmanci, A.E.: Color image segmentation using histogram multithresholding and fusion. J. Image Vis. Comput. **19**(13), 915–928 (2001)
7. Zhou, A., Yu, Y.: The research about clustering algorithm of K-means. Comput. Technol. Dev. **21**(2), 2 (2011)
8. Xue, H.R., Pei, X.C.: Color image segmentation using fuzzy sets and fusion. In: The Sixth International Conference on Electronic Measurement & Instruments. Taiyuan ISTP (2003)
9. Zhao, X.: Modern digital image processing technology to improve and applications comments (MATLAB version), vol. 04, pp. 113–114. Bei Hang University Press (2012)
10. Fang, M.: The Application of FUZZY C-Means Algorithm in the Segmentation of Milk Somatic Cell Color Image. Inner Mongolia Agriculture University (2008)