FCM Image Segmentation Algorithm Based on Color Space and Spatial Information

Zhengjian Ding, Jin Sun, and Yang Zhang

***Abstract*—This paper proposed a fuzzy C-means clustering (FCM) algorithm which based on color space and spatial information. First, the color histogram is applied to fuzzy clustering algorithm, to determine the initial number of clusters and initial cluster centers of fuzzy clustering. Then bringing spatial information into FCM, to reconstruct the new objective function contains neighborhood information. Finally achieve the image segmentation, evaluate and compare the algorithm. The experimental results show: this algorithm has a high quality and effect on image segmentation, and has a stronger anti-nose ability.**

***Index Terms*—Fuzzy clustering, color space, spatial information, image segmentation**

I. Introduction

In the research and application of image, people often only interest in certain parts of the image. These parts are called objective, the other parts are called background, they generally correspond to some areas have unique properties. In order to identify and segment the objective, need to extract and isolate these areas, only on this basis, there is possible to a further use of the objective. Image segmentation is a technology and process to divide the image into different characteristic regions and extract the interested objective [1].

Image segmentation has been widely used in practical application, for example, industrial automation, on line product inspection, the production process control, document image processing,Remote sensing and biomedical image analysis,Security surveillance and military,Sports,Agricultural projects, etc. In general, in a variety of image application, as long as need extraction on target image and measurement.these are inseparable from the image segmentation [2]. The development of image segmentation is related closely to many other disciplines and fields.In recent years, with the emergence of many new theories and new methods in various disciplines, people also proposed many new image segmentation technique. For example, emerging of Markov random field, expert system, Gibbs random filed, Bayesian theory, wavelet modulus maxima, fractal, Brown chain, simulated annealing, clustering algorithm and genetic algorithm,etc.

Manuscript received July 15, 2012; revised September 12, 2012.

The authors are with the LanZhou University of Technology, LanZhou, China (e-mail: 1206398472@qq.com).

II. Standard Fuzzy C-Means Clustering Segmentation Algorithm

Fuzzy c-means clustering algorithm realize set partition through iterative optimization of the objective function, it can show the degree of each pixel belongs to a different category. Let n be the number of pixels to be clustering, *C* is the number of categories ($2 \le c \le n$), m is the fuzzy weighted index ($m \in [1, \infty]$), and it controls the degree of membership among the various types. The value of objective function is weighted distance square which is each pixel in image to the C cluster centers, and can be expressed as:

$$J_m(U,V) = \sum_{i=1}^{c} \sum_{k=1}^{n} u_{ik}^{m} (d_{ik})^2 \tag{1}$$

where' u_{ik} is the membership k-th pixel belongs to the i-th class. d_{ik} is the distance k-th to i-th class. U is fuzzy classification matrix. V is a collection of cluster centers [3].

The clustering criterion is to seek the best team (*U, V*) to make $J_m(U,V)$ minimum. The minimization of A can be realized by the following iterative algorithm.

1) To determine the number of clustering categories, c and the weighted index m.Take d_{ik} is Euclidean distance,Set the iterative stop threshold ε is a positive number, initial iterative times l = 0 and fuzzy classification matrix.
2) To bring $U^{(L)}$ into the formula (2), to calculate the clustering center matrix $V^{(L)}$.

$$v_i = \frac{1}{\sum_{k=1}^{n} (u_{ik})^m} \sum_{k=1}^{n} (u_{ik})^m x_k \tag{2}$$

3) According to the formula (3), using $V^{(l)}$ to update $U^{(l)}$, obtained a new fuzzy classification matrix.

$$u_{ik} = \frac{1}{\sum_{j=1}^{c} \left(\frac{d_{ik}}{d_{jk}} \right)^{\frac{2}{m-1}}} \tag{3}$$

4) If $\left\| U^{(L)} - U^{(L+1)} \right\| < \varepsilon$, end; else, set l=l+1, return to Step 2 and continue.

Visible, the FCM algorithm is an unsupervised fuzzy clustering method,In the algorithm implementation process without human intervention.Especially,when the image

exists ambiguity and uncertainty, this algorithm can reflect the advantage of fuzzy mathematics, it has been applied in the field of image segmentation and achieved a good results. But this segmentation algorithm itself has defects. The defect is that using the gray information of image to establish clustering criterion function without taking the spatial information of pixels into account. So, the segmentation model is not complete [4]. In this paper, the standard fuzzy C-means clustering algorithm is improved. Using fuzzy c-means clustering (FCM) segmentation algorithm which combined the spatial information.

III. The Influence of Color Histogram Determination on the Initial Parameters and the Spatial Information

In the clustering algorithm, the number of clusters is often given as initial conditions. However, the number of data classes often is unknown. So need an algorithm to judge the number of clusters. If the selection of cluster number is not appropriate, will make the results of the data set (characteristic vector sets to be classified) and the real structure of the data set do not match and result in the failure of the clustering. The original algorithm deficiency is that the cluster number is set in advance [5]. And it needs a lot of human factor to participate in and clustering is not completely without supervision. So, using the number of peaks in the color histogram of the image to determine the cluster number, and realize the unsupervised cluster.

A. Color Space of RGB

According to the structure of the human eye,all colors can be as a different combination of three basic color, red (R), green (G) and blue (B).The RGB model is as follows:

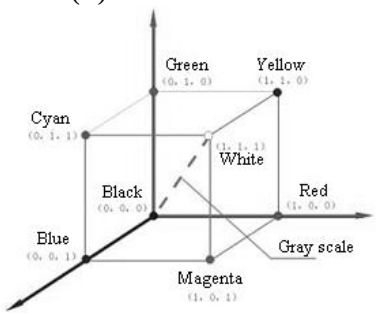

Fig. 1. RGB color space.

As the Fig. 1 shows, the line connected the black and white called gray line, it is obtained by the same amount of three colors mixture. Three vertexes in the coordinates are red, green, blue.The other three vertexes are yellow, cyan and magenta.

Each component has a 0-255 level. That is eight bit. So, this combined together, total number of the color is 256×256×256 (about 16 million), with 24 bit depth. RGB model is suitable for color display, but due to the high correlation of the three components, it is not suitable for color image segmentation and analysis.

B. Color Space of HIS

HIS color space conforms with perception of the human eye.Among, H is called hue, representing the basic color, reflecting the wavelength of color; S is called saturation, reflecting the degree of color depth, indicating the number of that white light and tonal mixed. I is called brightness. H and S contain the color information. While I component has nothing to do with the color information. Three components can be handled separately, and they are independent. HIS model is showed in Fig. 2.

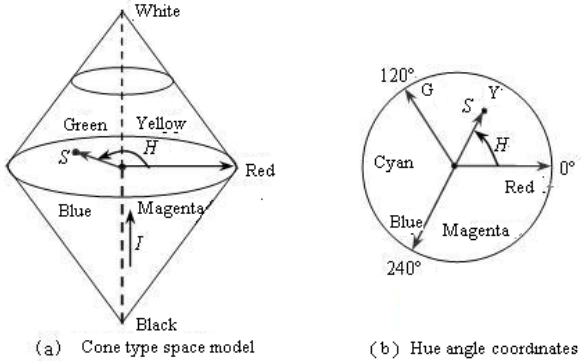

Fig. 2. HIS model.

Hue H expresses angle, the range of value is 0 ~ 360. Saturation S is the radius of the circle to the color point. Strength I is expressed by height in the direction of the axis, which describes the gray level. HIS color space is obtained by nonlinear transformation of the RGB color space.The transformation formula is:

$$H=\begin{cases}\theta, G\geq B\\ 2\pi-\theta, G\leq B\end{cases},\qquad across\theta=\cos^{-1}\left(\frac{(R-G)+(R-B)}{2\sqrt{(R-G)^2+(R-B)(G-B)}}\right) \tag{4}$$

$$S=1-\frac{3\min(R,G,B)}{R+G+B} \tag{5}$$

$$I=\frac{R+G+B}{3} \tag{6}$$

C. FCM Algorithm is Impacted by Increasing Spatial Information

Clustering algorithm should consider the spatial relationship between the pixel and its pixel in adjacent areas. However,the traditional FCM only use the gray characteristics of the pixels, did not consider the space feature into account. In the iterative process of algorithm, if adding neighborhood information of pixels, the clustering effect should be better.

Membership of pixel refers to the degree of each pixel attributable to different class. The iterative process of traditional FCM algorithm is completely making use of pixel gray information, ignoring any neighborhood information. However,the degree of a pixel belongs to the i-th class is influenced by whether its neighborhood pixel belongs to the i-th class.

In order to make clustering model of the algorithm contains space neighborhood information, can define space neighborhood information function:

$$h_{ij}=\frac{1}{l}\sum_{R_s(x_j)}u_{ik} \tag{7}$$

where, R is the pixels of the neighborhood, X_j as the center pixel. The size of neighborhood is l, generally take 4,8,24.

In order to make the algorithm has space constraints in the iterative process, defined a weighted membership of pixels belongs to a clustering space information:

$$u'_{ij} = \frac{u_{ij}^{p} h_{ij}^{q}}{\sum_{k=1}^{c} u_{kj}^{p} h_{kj}^{q}} \quad (8)$$

where, h_{ij} is spatial information of the j-th pixel belongs to *i*-th clustering space, u_{ij} is membership of *j*-th pixel belongs to *i-th* clustering. According to the above the spatial information weighted membership which is the pixel belongs to a clustering. We can obtain a formula to calculate the cluster center V_i:

$$v_i = \sum_{j=1}^{N} \left(u'_{ij}\right)^{m} x'_{ij} / \sum_{j=1}^{N} \left(u'_{ij}\right)^{m} \quad (9)$$

D. The Determination of Initial Parameter

This paper mainly uses the 3 dimensional color histogram of image in RGB and HIS color space. Respectively,to obtain the peak number and the pixel point on the peak as the initial clustering number and clustering center of FCM algorithm.There are many algorithm to obtain the value of peak.such as particle swarm optimization algorithm, climbing method,etc. This subject intends to use a simple particle swarm optimization algorithm.

Using mean-shift algorithm for image filtering, comprehensive utilization of pixel color and spatial information, and to realize the image smoothing and eliminate noise interference, and retain the main edge of image, so as to facilitate the image segmentation.

First, to extract the three basic color components of the image and respectively drawn out the histogram of three color components. To determine the initial clustering number use the color histogram. And then transformed the image into HIS color space, H is called hue and it represents the basic color. Then draw out the histogram of H component. And use this color histogram to determine the initial clustering number and clustering center. This method avoids the blindness of image segmentation brought by the initial parameter[7].

E. The Detail Steps of the Algorithm are as Follows:

1) Transformation of color space from the RGB space to the HIS space. And calculating the mean of 8-neighborhood. Then saving the processed data.
2) According to formula above proposed, to calculate the number of clustering category C, initial clustering center. To set fuzzy degree m and iteration cut-off error ε, Set initial iterative time t=100 and completing the initialization of the membership matrix.
3) According to formula (9) to calculate clustering center v_i.
4) According to formula (8) to update the membership matrix u'_{ij}.
5) If $\|v_i(t+1) - v_i(t)\| \le \varepsilon$, the system reaches a certain steady state, then end the iteration. Otherwise, set t=t+1, return to the step (3) and continue.
6) According to the final membership to judge the ownership of every pixel[8].

F. The Effective Evaluation Function of the Algorithm

This paper uses two different evaluation methods to evaluate image segmentation results. These two methods complement each other and have a good evaluation function and robustness. The one is to define dividing coefficient and dividing entropy through the given clustering center number and the given membership matrix. The other one is discrimination on the basis of the degree of association between the fuzzy dividing classfication[9].

The definition of dividing coefficient and dividing entropy are as follows:

$$V_{pc} = \frac{1}{n} \sum_{i=1}^{c} \sum_{j=1}^{n} u_{ij}^{2} \quad (10)$$

$$V_{pe} = -\frac{1}{n} \sum_{i=1}^{c} \sum_{j=1}^{n} \left(u_{ij} \log u_{ij}\right) \quad (11)$$

When clustering reaches the best results, value of the partition coefficient achieved the maximum, while value of the partition entropy achieved minimum. Because of these two functions lack direct relation on the space geometric characteristics of the pixel. According to the correlation between the fuzzy dividing classification to define the effective evaluation function:

$$V_{xb} = \sum_{i=1}^{c} \sum_{j=1}^{n} u_{ij}^{2} \left\| x_j - v_i \right\|^{2} // n\left[\min i \neq k\left(\left\| v_k - v_i \right\|^{2}\right)\right] \quad (12)$$

For a good clustering segmentation results, the distribution of pixels within the class is compact, the degree of fuzzy correlation between class and class should as small as possible.When pixel clustering achieved the best, the value of the V_{xb} achieved minimum[10].

IV. Experimental Results and Analysis

In order to verify the effectiveness of the method proposed in this paper, respectively, using threshold segment algorithm based on the direct histogram and standard FCM clustering algorithm to conduct test. We select several images to explain.

In this paper, we use matlab to program, the segmentation results shown as pseudo-color. The size of image is 152×152, we chose the color image "lena.bmp". The experimental results are as the Fig. 4.

(a) The original color visible image, the size is 512*512; (b) the result of threshold segmentation based on color histogram; (c) the result of the traditional FCM segmentation; (d) the segmentation result using the algorithm this paper proposed. From the segmentation results, it's easy to say that these three all can obtain the objective. Among these,(b)(c)exist certain noise in the segmentation.(d)the result is relatively clean.

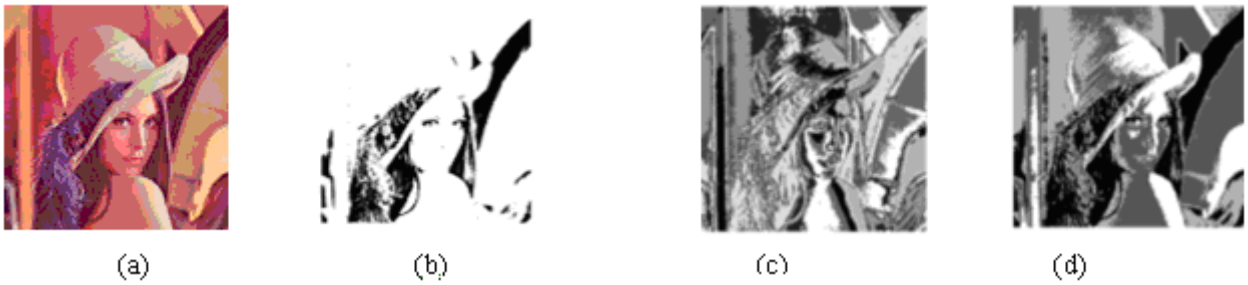

Fig. 3. Segmentation results based on "lena.bmp"

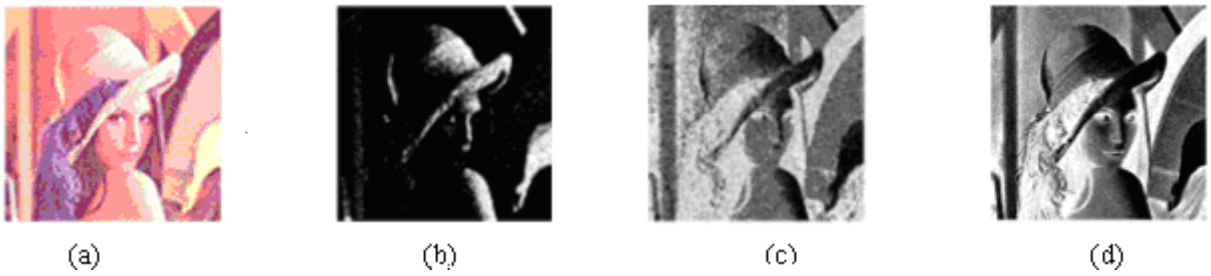

Fig. 4. Segmentation results based on the noisy image"lena.bmp"

(a) the image is added the Gaussian noise,(noise mean is 0,variance is 0.2);(b) the result of threshold segmentation using the color histogram;(c) the result of the traditional FCM segmentation;(d) the segmentation result using the algorithm this paper proposed. From the segmentation results, it's easy to say that (b), (c)exist a lot of noise. while, the result of (d) is relatively better. The objective outline is clear and the noise points are relatively less.

TABLE I: THE EXPERIMENTAL CONTRAST BETWEEN NON-NOISE AND NOISE

algorithm	V_{pc}	V_{xb}	V_{pe}
FCM	0.8578	0.0834	0.2785
improved algorithm	0.8598	0.0829	0.2781
add noise	0.8242	0.0927	0.2876
improved add noise	0.8256	0.0921	0.2871

V. CONCLUSION

This paper combined the basic fuzzy mean algorithm and the advantages of color histogram, using the color histogram to determine the clustering number of image segmentation and using particle swarm optimization algorithm to determine the initial clustering center. At last, we compared the traditional fuzzy clustering algorithm and improved algorithm, and adding noise to the original image to be segmented, repeating the above processing. Finally, obtained a conclusion, the improved algorithm is better than the traditional algorithm on the image segmentation. The segmentation results are more accurate.

REFERENCES

[1] J.-C. Fan, M. Han, and J. Wang, "Single point itenative weighted fuzzy C-means clustering algorithm for remote sensing image segmentation," *Patterm Recogmation*, vol. 42, no. 11, pp. 2527-2540, 2009.

[2] L. Gang, Y. Fan, and L.-L.Wang, "An algorithm for remote sensing image edge detection based on fuzzy sets," *Proc of the 2nd International Symposium on Intelligent Information Technology Application*. 2008, pp. 1087-1090.

[3] X. Li, X. Lu, J. Tian, *et al.*, "Application of fuzzy c-means clustering in data analysis of metabolomics," *Analytical Chemistry*, vol. 80, no. 11, pp. 4468-4475, 2009.

[4] Y.-X. Liu, T. Ikenag, and S. Goto, "A fully automatic approach of color image edge detection," *International Conference on Systems, Man and Cybernetics, IEEE*. 2006, pp. 1609-1612.

[5] J. Zou, H.-S. Li, B. Lin, *et al.*, "Color edge detection based on morphology," *Proc of the 1st international Conference on Communications and Electronics, IEEE*, 2006, pp. 291-293.

[6] S. Bandyopadhyay and U. Maulik, "An evolutionary technique based on K-means algorithm for optional clustering in RN," *Information Sciences*, vol. 146, pp. 221-237, 2002.

[7] M. N. Ahmed, S. M. Yanlany, and N. Mohamed, "A modified fuzzy c-means algorithm for bias field estimation and segmentation of MRI data," *IEEE Transactionon Medieal Imaging*, vol. 21, no. 3, pp. 193-199, 2002.

[8] D. L. Pham, "Fuzzy clustering with spatial constraints," *Proeeeding of the International Conference Image Proeessing*, 2002, pp. 65-68.

[9] W. Yang, Z. Zeng, and S. Zhang, "Application of combining watershed and fast clustering method in image segmentation," *Computer Modeling and Simulation*, 2010, pp. 170-174.

[10] J. Wu, H. Wu, and S. Zhou, "The study of particle swarm optimizer algorithm based on simulated annealing," *Chinese Journal of Scientific Instrument*, vol. 29, no. 8, pp. 50-53, 2008.