A Novel Fuzzy C-Means Clustering Algorithm for Image Thresholding

Y. Yong, Z. Chongxun, L. Pan

Key Laboratory of Biomedical Information Engineering of Education Ministry,
Institute of Biomedical Engineering, Xi'an Jiaotong University, Xi'an 710049,
China, greatyyy765@sohu.com

Abstract: *Image thresholding has played an important role in image segmentation. In this paper, we present a novel spatially weighted fuzzy c-means (SWFCM) clustering algorithm for image thresholding. The algorithm is formulated by incorporating the spatial neighborhood information into the standard FCM clustering algorithm. Two improved implementations of the k-nearest neighbor (k-NN) algorithm are introduced for calculating the weight in the SWFCM algorithm so as to improve the performance of image thresholding. To speed up the FCM algorithm, the iteration is carried out with the gray level histogram of image instead of the conventional whole data of image. The performance of the algorithm is compared with those of an existed fuzzy thresholding algorithm and widely applied between variance and entropy methods. Experimental results with synthetic and real images indicate the proposed approach is effective and efficient. In addition, due to the neighborhood model, the proposed method is more tolerant to noise.*

Keywords: *Image thresholding; fuzzy c-means; k-nearest neighbor; fuzzy thresholding*

1. Introduction

Thresholding is an important technique for image segmentation based on the assumption that the objects can be distinguished and extracted from the background by their gray levels. The output of the thresholding operation is a binary image whose gray level 0 (black) will indicate the foreground and gray level 255 (white) will indicate the background, and vice versa. Many thresholding methods have been developed. A detailed survey can be found in references [1] and [2]. In general, threshold selection can be categorized into two classes, local methods and global methods. The global thresholding methods segment an entire image with a single threshold using the gray level histogram of image, while the local methods partition the given image into a number of sub-images and select a threshold for each of sub-images. The global thresholding techniques are easy to implement and computationally less involved, therefore they are superior to local methods in terms of many real image processing applications. The global thresholding methods select the threshold based on different criterions, such as Otsu's method [3], minimum error thresholding [4], and entropic method which was first proposed by Pun [5] and then modified and extended by Kapur et al. [6], etc.

Otsu [3] selects the optimal thresholds by maximizing the between-class variance of gray values. Kittler and Illingworth [4] assume that the gray values of object and background are normally distributed. In their methods, threshold was chosen by a minimum error rate scheme for resultant classes. Kapur et al. [6] proposed a thresholding method by maximizing the entropy of the histogram of gray levels of object and background. Generally, all these conventional one-dimensional (1D) histogram thresholding techniques work well when the two consecutive gray levels of the image are distinct. However, all above 1D thresholding techniques did not combine the spatial information and the gray-level information of the pixels into the process for image segmentation. This drawback will lead to serious misclassification in the case of image thresholding, since the data in the image are inherently correlated. In addition, when the image is corrupted by noise and other artifacts the performance of these thresholding techniques will be poor or even fail. To compensate this drawback, Abutaleb et al. [7] extended 1D method to 2D thresholding method by considering the joint entropy of two random variables, namely, the image gray value and the average gray value, but it is very time consuming. Brink [8] refined Abutaleb's method and later Chen et al. [9] improved Brink's method and proposed a fast two-stage approach to search for the optimal threshold. Gong et al. [10] proposed a recursive algorithm for 2D entropic thresholding to further reduce the computation complexity. However, all these methods are still more complex than 1D entropic method proposed by Kapur et al [6].

Another important issue for image thresholding is that in real life situations a number of images are ambiguous and usually have indistinguishable histogram. In these cases, for thresholding it is not easy to find a criterion of similarity or closeness for thresholding. Since the fuzzy set theory was introduced, it has become a powerful tool to tackle this difficulty in image thresholding. Fuzzy set theory has been successfully applied to image thresholding to partition the image space into meaningful regions [11], [12], [13]. Jawahar et al. [11] proposed different fuzzy thresholding schemes based on fuzzy c-means (FCM) clustering. Cheng et al. [12] introduced the concept of fuzziness into the maximum entropy technique. More recently, Zhao et al [13] presented a more straightforward solution in the search for fuzzy thresholding parameters by exploiting the relationship between the fuzzy c-partition and the probability partition. However, all the mentioned algorithms still do not include the contextual information in image thresholding.

In this paper, we proposed a new global image thresholding named spatially weighted fuzzy c-means (SWFCM) clustering. It is formulated by incorporating the spatial neighboring information into the FCM algorithm. The weight in the algorithm, inspired by k-nearest neighbor [14] (k-NN) pattern classifier by considering the neighborhood influence on the central pixel, is then modified in this paper to improve the performance of image thresholding. The method is a 1D thresholding approach. Since we utilize the gray level histogram of image instead of the whole data of image to calculate the parameter for the FCM algorithm, the method is as fast as the conventional 1D techniques. Due to considering the neighborhood information, the method is also noise resistant.

The rest of this paper is organized as follows. Section 2 describes the fast FCM algorithm. The SWFCM clustering algorithm is presented in Section 3. Experimental results and comparisons are given in Section 4. Finally, some conclusions are drawn in Section 5.

2. Fast Fuzzy C-means algorithm

The Fuzzy C-means (FCM) algorithm is an iterative clustering method that produces an optimal c partition, which minimizes the weighted within group sum of squared error objective function $J_q(U,V)$ [15]:

$$J_q(U,V) = \sum_{k=1}^{n} \sum_{i=1}^{c} (u_{ik})^q d^2(x_k, v_i) \tag{1}$$

where $X = \{x_1, x_2, \cdots, x_n\} \subseteq R^p$, n is the number of data items, c is the number of clusters with $2 \le c < n$, u_{ik} is the degree of membership of x_k in the i^{th} cluster, q is a weighting exponent on each fuzzy membership, v_i is the prototype of the centre of cluster i, $d^2(x_k, v_i)$ is a distance measure between object x_k and cluster centre v_i. A solution of the object function Jq can be obtained via an iterative process, which is carried as follows:

1) Set value for c, q and ε.
2) Initialize the fuzzy partition matrix U.
3) Set the loop counter $b = 0$.
4) Calculate the c cluster centers $\{v_i^{(b)}\}$ with $U^{(b)}$:

$$v_i^{(b)} = \frac{\sum_{k=1}^{n} \left(u_{ik}^{(b)}\right)^q x_k}{\sum_{k=1}^{n} \left(u_{ik}^{(b)}\right)^q} \tag{2}$$

5) Calculate the membership $U^{(b+1)}$. For $k = 1$ to n, calculate the following:

$I_k = \{i \mid 1 \le i \le c, d_{ik} = \|x_k - v_i\| = 0\}$,

$\tilde{I}_k = \{1,2,\cdots,c\} - I_k$;

for the k^{th} column of the matrix, compute new membership values:

a) if $I_k = \phi$, then

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

b) else $u_{ik}^{(b+1)} = 0$ for all $i \in \tilde{I}_k$ and $\sum_{i \in I_k} u_{ik}^{(b+1)} = 1$; next k.

6) If $\left\|U^{(b)} - U^{(b+1)}\right\| < \varepsilon$, stop; otherwise, set $b = b + 1$ and go to step 4.

Since FCM algorithm is an iterative operation, it is very time consuming which makes the algorithm impractical used in image segmentation. To cope with this problem, the gray level histogram of image is applied to the algorithm. Define the non-negative integrate set $G = \{L\min, L\min+1, \cdots, L\max\}$ as gray level, where $L\min$ is the minimum gray level, $L\max$ is the maximum gray level, so the grayscale is $L\max - L\min$. For image size $S \times T$, at point (s,t), $f(s,t)$ is the gray value with $0 \le s \le S-1$, $0 \le t \le T-1$. Let $His(g)$ denote the number of pixels having gray level g, $g \in G$. The statistical histogram function is as follows:

$$His(g) = \sum_{s=0}^{S-1} \sum_{t=0}^{T-1} \delta(f(s,t) - g) \tag{4}$$

where $g = \{L\min, L\min+1, \cdots, L\max\}$, $\delta(0) = 1$ and $\delta(g \ne 0) = 0$. With the gray level histogram the membership function is still by (3), while the cluster centers are updated by:

$$v_i^{(b)} = \frac{\sum_{g=L_{\min}}^{L_{\max}} \left(u_{ig}^{(b)}\right)^q His(g)g}{\sum_{g=L_{\min}}^{L_{\max}} \left(u_{ig}^{(b)}\right)^q His(g)} \tag{5}$$

It is important to note that k in (3) now denotes the gray level as g. Since the FCM algorithm now only operates on the histogram of the image, it is faster than the conventional version which processes the whole data.

3. Spatially Weighted Fuzzy C-means clustering algorithm

The general principle of the techniques presented in this paper is to incorporate the neighborhood information into the FCM algorithm. Since in the standard FCM algorithm for a pixel $x_k \in I$ where I is the image, the clustering of x_k with class i only depends on the membership value u_{ik}, if we consider a noisy image, since clustering process is related only to gray levels independently on pixels, FCM is noise sensitive. Considering the influence of the neighboring pixels on the central pixel, the fuzzy membership function given in (3) can be extended to

$$u_{ik}^{*} = u_{ik}\, p_{ik} \tag{6}$$

where $k = 1, 2, \cdots, n$, n is the index of each pixel, and p_{ik} is the probability of data point k belonging to cluster i, referred to as weight in this paper which can be determined by the following neighborhood model. Therefore the degrees of membership u_{ik}^{*} and the cluster centers v_i are now updated via:

$$u_{ik}^{*\,(b)} = \frac{p_{ik}}{\sum_{j=1}^{c} \left(\frac{d_{ik}}{d_{jk}}\right)^{2/(q-1)}} \tag{7}$$

$$v_i^{(b+1)} = \frac{\sum_{k=1}^{n} \left(u_{ik}^{*\,(b)}\right)^q x_k}{\sum_{k=1}^{n} \left(u_{ik}^{*\,(b)}\right)^q} \tag{8}$$

The core idea now is to define the auxiliary weight variable p_{ik}, which is a priori information to guide the outcome of the clustering process. This paper proposes a method for determining the weight based on the neighborhood information inspired from k-nearest neighbor (k-NN) algorithm [14].

$$p_{ik} = \frac{\sum_{x_n \in N_k^i} 1/d^2(x_n, k)}{\sum_{x_n \in N_k} 1/d^2(x_n, k)} \tag{9}$$

where N_k is the dada set of the nearest neighbors of central pixel k, and N_k^i is the subset of N_k composed of the data belonging to class i which is got after defuzzyfying the result of the FCM algorithm in our method. In order to give an appropriate method to describe the probability of a data point belonging to any cluster, two improved implementations of the k-NN algorithm are introduced. First, the equation (9) is extended by considering the potential function of each feature vector [16].

$$K(x, x_k) = \frac{1}{1 + \alpha \|x - x_k\|^2} \tag{10}$$

where α is a positive constant, and $\|x - x_k\|^2$ is the norm of the vector $(x - x_k)$. Then the potential is modified by assigning the proximity of feature vector to each prototype instead of the potential for feature vector to feature vector. Hence the new equation for the weight value is defined as:

$$p_{ik} = \frac{\sum_{x_n \in N_k^i} 1 \big/ 1 + \alpha \cdot d^2(x_n, v_i)}{\sum_{x_n \in N_k} 1 \big/ 1 + \alpha \cdot d^2(x_n, v_i)} \tag{11}$$

where v_i is the prototype of cluster i. After the a-priori weight is determined, a new iteration step starts with this auxiliary variable p_{ik}. To prevent that the SWFCM gets trapped in a local minima, the SWFCM algorithm is initialized with the above fast FCM algorithm. Once the FCM is stopped, the SWFCM algorithm continues with the values for the prototypes and membership values obtained from the fast FCM algorithm. When the algorithm has converged, a defuzzification process then takes place in order to convert the fuzzy partition matrix U to a crisp partition. A number of methods have been developed to defuzzify the partition matrix U, in which the maximum membership procedure is the most important. The procedure assigns object k to the class C with the highest membership:

$$C_k = \arg_i \{\max(u_{ik})\} \quad , \quad i = 1, 2, \cdots, c \ . \tag{12}$$

With this procedure, the fuzzy images are then converted to crisp image. For image thresholding, $c = 2$ in equation (12). We call this method soft threshoding scheme contrary to conventional hard threshold scheme which has been proven to be associated with loss of structure details on thresholding [11]. Although Jawahar et al. [11] has proposed a fuzzy thresholding method with FCM algorithm by finding the hard threshold in the intersection of both membership distributions (see equation (10) in [11]), it is easily verified that this technique is almost equivalent to thresholding the image using the maximum membership procedure.

4. Experimental results

In this section, the results of the application of the SWFCM algorithm are presented. The performance of the proposed method is compared with fuzzy thresholding method introduced by Jawahar et al. (see Section 3) and two well-known thresholding methods, including Otsu and Kapur et al. algorithms. For all cases, unless otherwise stated, the weighting exponent $q = 2.0$ and $\varepsilon = 0.0001$. We tried several values for α and found that a value of $\alpha = 1$ gives a convenient result. A 3×3 window of image pixels is considered in this paper, thus the spatial influence on the centre pixel is through its 8-neighborhood pixels. All the algorithms are coded in Microsoft Visual C++ Version 6.0 and are run on a 1.7GH$_Z$ PentiumⅣ personal computer with a memory of 256 MB. Since we use the fast FCM algorithm, all the algorithms are implemented within 1 seconds.

In the first example, we generate a synthetic image with gray levels 0 and 255 for background and foreground respectively. The image was then corrupted by additive Gaussian noise such that the $SNR = 5$. Fig. 1 (a) is the original image and Fig .1(b) is the degraded noisy images. Fig.1 (f) shows the result of proposed method. The results of Jawahar's method, Otsu's method and Kapur's method are displayed in Figs. 1 (c), (d) and (e), respectively. The results show the proposed method is an effective method and outperforms the other methods in the noisy situation. The number of misclassified pixels for different thresholding methods is counted during the experiments and is listed in Table 1. It can be seen from Table 1. that the total number of misclassification pixels for the proposed method is the least in the four different methods, and the total misclassification number for Jawahar's method, Otsu's method and Kapur's method is nearly the same which is about 18 times than that of the proposed method.

The test image in the second example, which is given in Fig2 (a), is a true Chinese character received by a camera and blurred by the Gaussian noise. The results of thresholding by applying the different algorithms to the image appear in Figs. 2 (b)-(e). From these images, we can see that the Jawahar's method, the Otsu's method and the Kapur's method can not threshold the image correctly in the case of noise, while our proposed method can do this with the least errors in such a case.

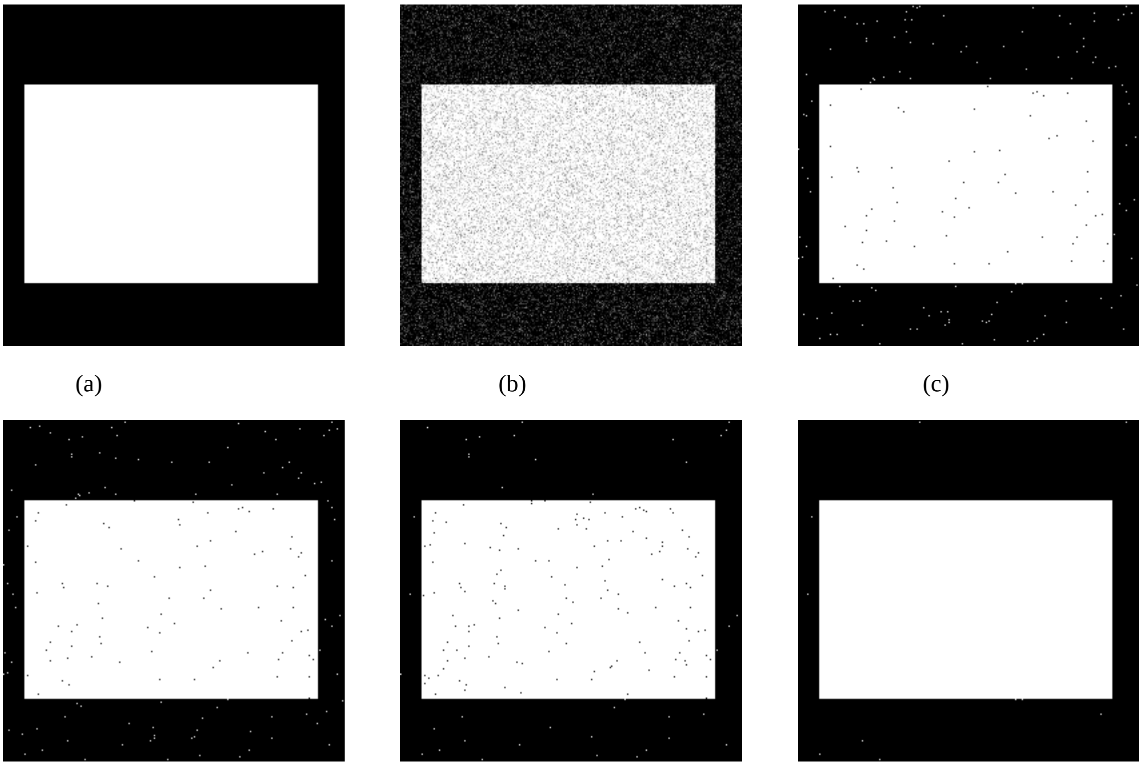
The last example is a standard test image named *camerman*, which is illustrated in Fig.3 (a). The result of the proposed method is presented in Fig.3 (e). The results for comparison are given in Figs.3 (b)-(d). As can be seen, the Jawahar's method and Otsu's method give almost the same result, while the Kapur's method can not accurately extract the object from the background. However, it can be seen the proposed method performs the best for segmenting the object from the background with the least spurious components and noise in the four methods.

5. Conclusions

In this paper, a novel spatially weighted fuzzy c-means (SWFCM) clustering algorithm for image thresholding is presented. The method not only takes into account of the advantage of the fuzzy framework, but also considers the spatial relation among pixels. The weight plays a key role in this algorithm, which is derived from k-NN algorithm and is modified to improve the property in the SWFCM algorithm. The performance of the proposed method is compared with the Otsu's method, Kapur's method, and a fuzzy thresholding method proposed by Jawahar et al. Experiments with synthetic and real image show that SWFCM can effectively extract object from background. Since the algorithm is initialized with fast FCM algorithm, the presented approach is as fast as the conventional 1D techniques. Also, owing to the incorporation of spatial information, the SWFCM algorithm is less prone to noise. In fact, if the result of thresholding is an image with two gray values, the process can also be called bilevel segmentation. Future work will extend the algorithm to multi-level thresholding or segmentation.

Table 1. Number of misclassified pixels for foreground and background with different methods

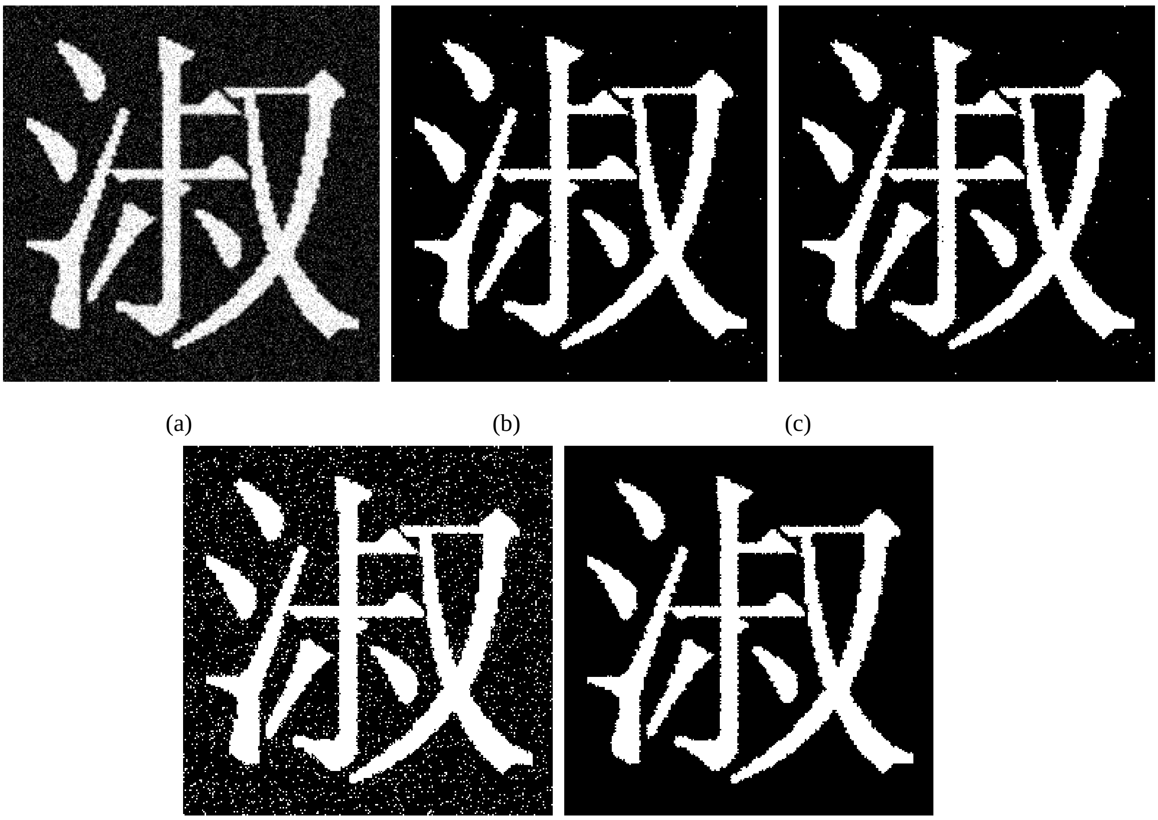
Methods	Jawahar	Otsu	Kapur	Our method
Foreground	61	83	141	0
Background	118	96	39	10
Total	179	179	180	10



(a) (b) (c)

(d) (e) (f)

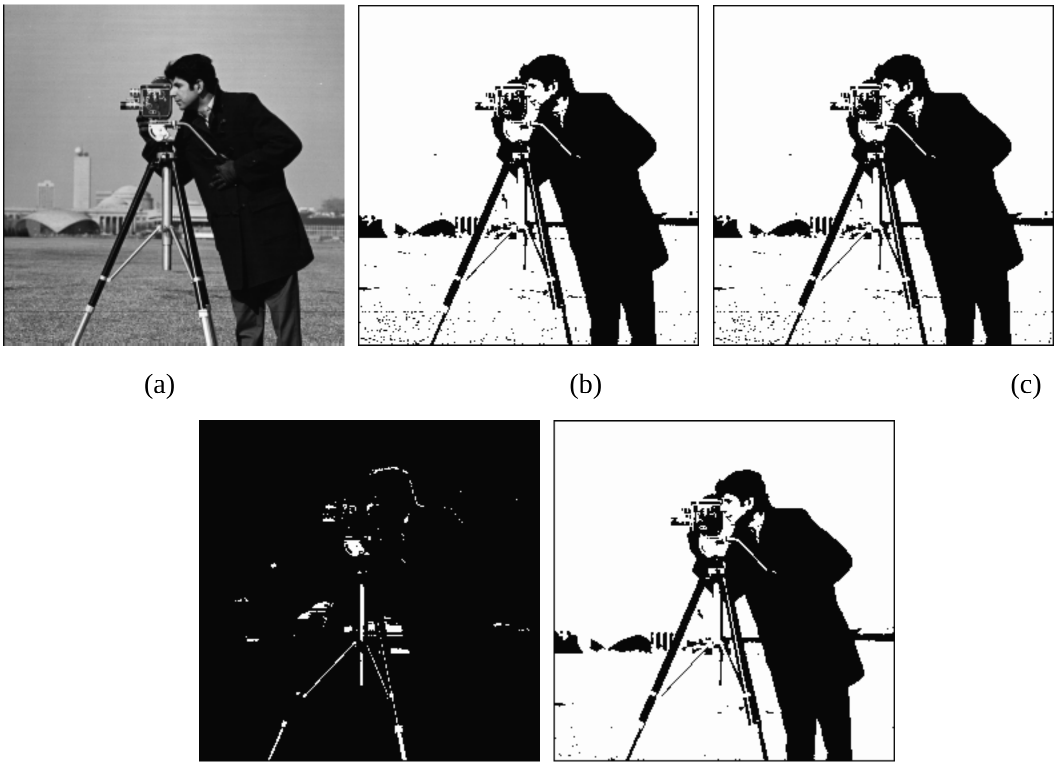
Fig.1. Results of thresholding: (a) the original image; (b) noisy image with $SNR = 5$; (c) by the Jawahar's method; (d) by the Otsu's method; (e) by the Kapur's method; (f) by the proposed method.



(a) (b) (c)

(d) (e)

Fig.2. Results of thresholding: (a) the original image; (b) by the Jawahar's method; (c) by the Otsu's method; (d) by the Kapur's method; (e) by the proposed method.



(a) (b) (c)

(d) (e)

Fig.3. Results of thresholding: (a) the original image; (b) by the Jawahar's method; (c) by the Otsu's method; (d) by the Kapur's method; (e) by the proposed method.

References

[1] S. K. Fu and J. K. Mu, "A survey on image segmentation," Pattern Recognit., Vol 13, pp. 3–16, 1981.

[2] P. K. Sahoo, S. Soltani, and A. K. C. Wong, "A survey of thresholding techniques," Comput. Vis. Graph. Image Process., Vol 41, pp. 233–260, 1988.

[3] N. Otsu, "A threshold selection method from gray level histogram," IEEE Trans. Syst., Man, Cybern., Vol SMC-9, pp. 62–66, 1979.

[4] J. Kittle and J. Illingworth, "Minimum error thresholding," Pattern Recognit., Vol 19, pp. 41–47, 1986.

[5] T. Pun, "A new method for gray level picture thresholding using the entropy of the histogram," Signal Process., Vol 2, pp. 223–237, 1980.

[6] J. N. Kapur, P. K. Sahoo, and A. K. C. Wong, "new method for gray level picture thresholding using the entropy of the histogram," Comput. Vis. Graph. Image Process., Vol 29, pp. 273–285, 1985.

[7] A. S. Abutaleb, "Automatic thresholding of gray-level pictures using two-dimensional entropy," Comput. Vis. Graph. Image Process., Vol 47, pp. 22–32, 1989.

[8] A. D. Brink, "Thresholding of digital images using two-dimensional entropies," Pattern Recognit., Vol 25, pp. 803-808, 1992.

[9] W. T. Chen, C. H. Wen, and C. W. Yang, “A fast two-dimensional entropic thresholding algorithm,” Pattern Recognit., Vol 27, pp. 885-893, 1994.

[10] J. Gong, L. Li, and W. Chen, “Fast recursive algorithms for two-dimensional thresholding,” Pattern Recognit., Vol 31, pp. 295-300, 1998.

[11] C. V. Jawahar, P. K. Biswas, and A. K. Ray, “Investigations on fuzzy thresholding based on fuzzy clustering,” Pattern Recognit., Vol 30, pp.1605-1613, 1997.

[12] H. D. Cheng, J. Chen, and J. Li, “Thresholding selection based on fuzzy c-partition entropy approach,” Pattern Recognit., Vol 31, pp. 857–870, 1998.

[13] M. S. Zhao, A. M. N. Fu, and H. Yan, “A technique of three level thresholding based on probability partition and fuzzy 3-partition,” IEEE Trans. Fuzzy Systems, Vol 9, pp. 469–479, 2001.

[14] T. M. Cover and P. E. Hart, “Nearest neighbor pattern classification,” IEEE Trans. Inform. Theory, Vol IT-13, pp. 21–27, 1967.

[15] J. C. Bezdek, Pattern Recognition with Fuzzy Objective Function Algorithms, Plenum Press, New York, 1981.

[16] T. Tou Julius, Pattern Recognition Principles, Addison–Wesley Company, 1974.