Gaussian Kernelized Fuzzy c-means with Spatial Information Algorithm for Image Segmentation

Cuiyin Lui1,2,3
1College of Computer Science, Sichuan University, Chengdu 610064, China
2State Key Laboratory of Fundamental Science on Synthetic Vision, Chengdu 610064, China
3College of Computer Science, Panzhihua University, Sichuan Panzhihua 610007, China
Email: liucuiyin@163.com

Xiuqiong Zhang1,2, Xiaofeng Li1,2, Yani Liu4, Jun Yang1,2,3
1College of Computer Science, Sichuan University, Chengdu 610064, China
2State Key Laboratory of Fundamental Science on Synthetic Vision, Chengdu 610064, China
3College of Computer Science, Sichuan Normal University, Chengdu, 610066, China
4College of Computer Science, Panzhihua University, Sichuan Panzhihua 610007, China
Email: zxq_03@tom.com, lixiaofeng2009@scu.edu.cn

Abstract—FCM is used for image segmentation in some applications. It is based on a specific distance norm and does not use spatial information of the image, so it has some drawbacks. Various kinds of improvements have been developed to extend the adaptability, such as BCFCM, SFCM and KFCM. These methods extend FCM from two aspects, one is replacing the Euclidean norm, and the other is considering the spatial information constraints for clustering. Kernel distance can improve the robustness for multi-distribution data sets. Spatial information can help eliminate the sensitivity to noises and outliers. In this paper, Gaussian kernel-based fuzzy c-means algorithm with spatial information (KSFCM) is proposed. KSFCM is more robust and adaptive. The experiment results show that KSFCM has the better performance.

Index Terms—kernel, spatial information, segmentation, FCM

I. INTRODUCTION

Image segmentation plays an important role in a variety of applications in high level image processing. Segmentation of structures from images is an important step in image analysis that can help in visualization, automatic feature detection, image-guided surgery, and also for registration of different images. There are various kinds of methods for image segmentation, such as gray threshold method, region growing method, watershed method, gradient method, geodesic active contour, energy method, and clustering method [1].

Clustering is a popular method used for image segmentation for its simplicity and easiness to implement. The commonly used clustering algorithms are hard c-means and the fuzzy c-means algorithm. Hard c-means clustering is based on classical set theory, and it assigns an object to one cluster or not. Fuzzy c-means (FCM) clustering reported by C.J. Dunn in [1], and proposed by Bezdek J C in [2], is an unsupervised method that has been successfully applied to image segmentation. Fuzzy clustering methods allow objects to belong to several clusters simultaneously with different degrees of membership [4]. In many real applications, fuzzy clustering is more natural than hard c-clustering, in segmentation of an unclear image. It is difficult to decide a pixel belonging to which one cluster, but rather membership degrees between 0 and 1, indicating its partial memberships.

The original intensity-based FCM algorithm functions well in image segmentation, however, it fails to segment images corrupted by noises, outliers, other imaging artifacts and intensity in-homogeneity. The failures are induced by two problems. The first problem is detailed in [5]. Fuzzy algorithm requires a specific distance function given, which is known that different distance function yield different data structures, thus FCM clustering is not robust for multiple distribution data set. The second problem is that FCM disregards spatial information of image, thus it is sensitive to noises and outliers.

There are two approaches to mitigate the impact of the two problems to improve the adaptively and robustness of FCM algorithm. On the one hand, many researchers have concentrated on extending distance functions in FCM [6-7]. A. F. Gomez-Skarmeta and M. Delgado [8] developed a Gustafson-Kessel algorithm employing an adaptive
distance norm, to detect clusters of different geometrical shapes in one data set. Gath and A. B. Geva [9] developed an algorithm that combines the fuzzy K-means algorithm and the fuzzy maximum likelihood estimation (FMLE), which performs well in situations of large variability of cluster shapes, densities, and number of data points in each cluster. However, the cluster covariance matrix is used in conjunction with an “exponential” distance, and the clustering is not constrained in volume, so it leads to converge to a local optimum. On the other hand [10], many other algorithms have also been proposed to deal with the second problem by incorporating spatial information into original FCM objective function [11, 12]. Ahamed [11] et al. modified objective function of FCM by incorporating spatial constraints (called as BFCM) [2]. Although the iterative process partitions the biased information but the computation of spatial constraints takes much time and lowers the algorithm speed. Chen and Zhang [6] proposed two modifications for reducing the computation complexity. The constraints information of a pixel neighborhood can be computed in advance. For different noise models use the different constraints, the calculation can be mean-filtered image or median-filtered image. Only for eliminating the noises and regardless of the bias field, Keh-Shih Chuang et. al. [18] proposed a modified FCM algorithm (called SFMC) with neighborhood constraints that is the sum of membership in neighborhood. The method is not only simple and fast but also is more robust to noises and outliers.

The modified algorithms have achieved some good results, but the two problems still exist in some degree. One problem is that is sensitive to outliers, and the other problem is that is difficult to identify multiple data distribution for the norm based on Euclidean distance (or some specified distance norm). In this paper a kernel method is induced into the algorithm SFCM algorithm to improve self-adaptively to multiple distribution data sets. The fundamental idea of the kernel method is to transform the original low-dimensional inner-product input space into a higher(possibly infinite)dimensional feature space through some nonlinear mapping where complex nonlinear problems in the original low-dimensional space can be more likely solved linearly in the transformed space according to the well-known Cover’s theorem. In this paper we modified the SFMC algorithm proposed by Keh-Shih Chuang in [18]. The kernel FCMS algorithm will be more robust to noises and outliers in image segmentation than the algorithms without the kernel substitution and be more adaptive to various kinds of data distribution.

The rest of this paper is organized as follows. In section II, the conventional Fuzzy C-means (FCM) algorithm and the modified algorithm (SFCM) with spatial constraint are presented and analyzed. In section III, kernel method is described and induced to the SFCM algorithm to construct the new algorithm (KSFCM). A detailed analysis and process is presented. In section IV, Some experimental comparisons are presented. We present several experimental comparisons to assess the performance of KSFCM. The comparisons are made between FCM, SFCM, KFCM and KSFCM. Finally, some conclusions are given in Section V.

II. CONVENTIONAL SEGMENTATION WITH FCM ALGORITHM

A. the conventional FCM for segmentation

Clustering method is a statistical technique used for image segmentation. The first kind of clustering method is hard clustering, which is based on classical set theory, it assigns an object to one cluster or not, based on the nearest distance from the data to the cluster center. One object can only belong to a cluster. Fuzzy Clustering algorithm one of most popular clustering algorithm used in variety of domains. The hard partition is defined as follow. The Hard partitioning space for X is the set [4].

\[ M_{hc} = \{U \in R^{CN} | \mu_{ik} \in \{0,1\}, \forall i,j \leq c, \sum_{k=1}^{c} \mu_{ik} = 1, \forall i; 0 \leq \sum_{k=1}^{c} \mu_{ik} < N \} \]

(1)

Let X=[x_1, x_2, ..., x_N] be a finite data set in an s-dimensional Euclidean space R^s with its norm ||.||. Let 2 ≤ C ≤ N be an integer, and the X can be clustered into c prototype. The clustering number must be known in advance. Where, the \( \mu_{ik} \) shows the i-th data belonging to k cluster or not, which allows to attain integer value 0 or 1;

Hard clustering algorithm minimizes the following objective function iteratively to complete the classification process.

\[ J(X) = \sum_{i=1}^{c} \sum_{k=1}^{N} (\mu_{ik}) \| x_k - v_i \|^2. \]

(2)

Given the data set X, choose the number of cluster 2 ≤ C ≤ N . The Hard clustering Algorithm is briefly described as follows:

Step1: Initialize K cluster centers (c_1,...,c_K).

Step2: Compute the distances

\[ D_{ik} = (x_k - v_i)^T (x_k - v_i), 1 \leq i \leq c, 1 \leq k \leq N \]

(3)

And then assign each x_i to its nearest cluster center v_k.

Step3: Calculate new cluster centers

\[ C_i = \frac{1}{n_i} \sum_{x_j \in C_i} x_j, i = 1,2,..,k \]

(4)

Step4: Calculate the sum of interclass variation,

\[ J = \sum_{i=1}^{c} \sum_{k=1}^{N} \mu_{ik} \| x_k - v_i \|^2 \]

(5)

The iterative process stops when the maximum number of iterations is reached or when the objective function improvement between two consecutive iterations
is less than the minimum amount of improvement specified.

In many real situations, fuzzy clustering is more natural than hard clustering, as objects on the boundaries between several classes are not forced to fully belong to one of the classes, but rather are assigned membership degrees between 0 and 1 indicating their partial memberships. FCM defined by Dunn in 1974 and extended by Bezdek in 1981 is an unsupervised clustering algorithm, which has been successfully applied to image segmentation by minimizing following objective function (6).

\[ J = \sum_{i=1}^{C} \sum_{k=1}^{N} (\mu_{ik})^m ||x_k - v_i||^2 \]  

\[ V = \{v_1, v_2, ..., v_C\}, v_i \in \mathbb{R}^n \]  

\[ dist_{(x_i, v_j)} = ||x_i - v_j||^2 \]  

Where, \( \{V_i\} \) is the i-th cluster center. The parameter m is weighting exponent on each fuzzy membership and control the degree of “fuzziness” of the resulting classification. In real application, m=2 is often adopted. \( \mu_{ik} \) represents the different degree membership of pixel \( x_k \) in the group of \( v_i \) cluster. \( ||.|| \) is a norm metric for the distance between two feature vectors, Euclidean distance is often used. U is fuzzy matrix and with the constraint is that sum of membership belong to one prototype is 1.

\[ \sum_{i=1}^{C} \mu_{ik} = 1, 1 \leq k \leq N \]  

The problem can be seen as a nonlinear optimization problem. The object function is defined as (10) with the membership constraint.

\[ J = \sum_{i=1}^{C} \sum_{k=1}^{N} (\mu_{ik})^m ||x_k - v_i||^2 + \sum_{k=1}^{N} \lambda_k (\sum_{i=1}^{C} \mu_{ik} - 1) \]  

The objective cluster center and partition matrix points can be found by adjoining the constraint (8) to J by means of Lagrange multiplier method.

Getting the partial derivative of \( \mu_{ik} \) and \( v_i \) to zero, the membership functions and clusters are updated by the following equations.

\[ v_i = \frac{\sum_{k=1}^{N} (\mu_{ik})^m x_k}{\sum_{k=1}^{N} (\mu_{ik})^m} \]  

\[ \mu_{ik} = \frac{(1/||x_i - v_j||)^{1/(m-1)}}{\sum_{k=1}^{N} (1/||x_i - v_k||)^{1/(m-1)}} \]  

By iteratively updating the fuzzy membership with (11) and the centers with (12), the algorithm converges to a local minimum of \( J \). The original algorithm is described as follows:

Step1: Initialize the clustering center and the partition matrix randomly with the membership constraints.

Step2: Compute the new cluster prototypes by (11).

Step3: Compute the distance and update the partition matrix by equation (12).

Repeat steps 2 to 3.

When the objective function improvement between two consecutive iterations is less than the minimum amount of improvement specified, the iteration stops.

In fuzzy clustering used for image segmentation, the gray value and other local characteristic all can be used as the character vector. The conventional FCM algorithm is sensitive to noises and the artifacts, so it is not popularly used in real applications.

B. The improved FCM algorithm (SFCM)

The conventional FCM segmentation algorithm classifies only based on the character vector without considering the spatial information of the image, and it is sensitive to noises. In [12], an improved FCM is presented to exploit the spatial information to correct drawbacks. A spatial function is incorporated and considered in the FCM. The Spatial information is defined as:

\[ h_{ik} = \sum_{m \in NB(x_i)} u_{im} \]  

Where, \( NB(x_i) \) represents a square or round window centered on pixel \( x_i \) in the spatial domain. The spatial function \( h_{ik} \) represents the probability that pixel \( x_i \) belongs to i-th cluster. The spatial function of a pixel for a cluster is larger if the majority of its neighborhood pixels belong to the same cluster. The value \( h_{ik} \) is the sum of pixel membership in neighborhood except current pixel.

The fuzzy membership matrix is modified with spatial information as

\[ u'_{ik} = \frac{u_{ik}^p h_{ik}}{\sum_{j=1}^{C} u_{ij}^p h_{ij}} \]  

The membership is not only decided by distance from the pattern to the clustering center, but also is affected by the neighborhood pixel labeling. In the homogeneity, the membership of a pixel is increases when spatial function value increases, and decreases when spatial function value decreases.

The SFCM algorithm is divided into two steps actually. The first step is the same as the conventional FCM. The fuzzy membership matrix and cluster
prototype are firstly calculated, and then incorporated with the spatial information to get the new fuzzy membership value. The modified FCM has been tested to be insensitive to the noises in real application. Many researchers have been done on the modification, and various kinds of extension-FCM algorithm have been proposed for different applications, e.g. BFCM is especially for correcting the biased in MRI image.

Only considering the spatial information is still not enough, since the distribution of data set is diversified. The distance norm based on Euclidean is only effective for detecting circle and ellipse data. The algorithm will fail when the data follows other distributions. Researchers have also been done in solving this problem, in which the Euclidean distance replaced by kernel mapping. The algorithm maps the patterns from the low-dimension space to high-dimension space to enlarge the discrimination among the patterns.

III. KENERL METHOD AND THE IMPROVED KSFMC

A. the kernel method
In recent years, the kernel method has been developed and widely used in image processing. Kernel method can transform a linear problem to a high-dimensional space to resolve. The kernel method has been used in some FCM and modified FCM algorithms to improve these algorithms effectiveness and get better clustering results. Many researches and experiments have been developed, which are reviewed in next paragraph.

SFCM incorporates spatial information into the iteration to eliminate noises and the outliers. It uses Euclidean distance used as the basis for classifying, which is only suitable for spherical and ellipsoidial cluster. It needs to improve the robustness to noises and outliers, and also to increase the adaptability for various distributions of data sets. Dao-Qiang Zhang [16] induced the kernel into the BFCM by modifying the objective function to use a kernel-induced distance metric and added a spatial penalty in the membership functions to correct the intensity inhomogeneous region. Miin-Shen Yang [17] induced the kernel into the fast BFCM algorithm for two different preliminary processing of different noises, which are BFCM_S1 and BFCM_S2 (called GFCM). Miin-Shen Yang [17] found that the parameters \( \sigma \) and \( \alpha \) are the key parameters that affects the experiment results, and make a detailed research in the selection of a suitable parameters. The experiment results showed that GKFCM is more robust to noises than conventional algorithm.

The kernel method realizes the clustering in the feature space. First, a nonlinear map is applied to map the data space to the feature space, \( \Phi : \chi \rightarrow F \) (\( x \in R^q \rightarrow \)), \( \Phi(x) \in R^q \), \( q > p_c \) the Then, the problem is solved in the high dimensional feature space easily, which is proved by Cover [18].

\[ x \in R^q, k = 1,2,...,q \] are patterns from the input space, and are mapped into the feature space H, the results are \( \Phi(x_k), \Phi(x_1),\ldots,\Phi(x_c) \). The inner product in input space can be replaced by the kernel function \( k(x, y) \), and need not to know the specific mapping function \( \Phi \). \( k(x, y) = < \Phi(x), \Phi(x) > \), when satisfying the Mercer condition. The commonly used kernel functions are Gaussian kernel, polynomial kernel, and sigmoid kernel.

1) Gaussian kernel
\[
K(x_j, a_i) = \exp(-\|x_j - a_i\|^2 / \sigma^2)
\]

2) Polynomial kernel
\[
K(x_j, a_i) = (1+<x_j, a_i>)^d
\]

3) Sigmoid kernel
\[
K(x_j, a_i) = \tanh(\alpha <x, y> + \beta)
\]

In this paper the Gaussian kernel function is adopted and incorporated in SFCM algorithm. Other kernel functions may also be useful; the discussion is not developed in this paper.

B. Kernelized Fuzzy C-means clustering with spatial constraints (KSFCM)
There are two steps in the SFCM algorithm. The kernel is induced in the object function at the first step.

\[
J = \sum_{i=1}^{c} \sum_{j=1}^{n} (\mu_{ij})^m \| x_j - v_i \|^2
\]

After embedded kernel, the equation is defined as follow:

\[
J(U) = \sum_{i=1}^{c} \sum_{k=1}^{n} u_{ik}^m \| \Phi(x_k) - \Phi(v_i) \|^2
\]

| \( \Phi(x_k) - \Phi(v_i) \|^2 = \| \Phi(x_k)\Phi(x_k) - 2\Phi(x_k)\Phi(v_i) + \Phi(v_i)\Phi(v_i) \|
\]

So, in the input space, \( d(x,y) \) can be defined as follow:

\[
d(x,y) = ||\Phi(x_i)\Phi(x_i) - 2\Phi(x_i)\Phi(v_i) + \Phi(v_i)\Phi(v_i) || = k(x_i, x_i) - 2k(x_i, v_i) + k(v_i, v_i)
\]

Object function is then transformed to the form as follow:

\[
J(U) = \sum_{i=1}^{c} \sum_{k=1}^{n} u_{ik}^m \| \Phi(x_k) - \Phi(v_i) \|^2
\]

\[
= \sum_{i=1}^{c} \sum_{k=1}^{n} u_{ik}^m d(x,y)
\]

\[
= \sum_{i=1}^{c} \sum_{k=1}^{n} u_{ik}^m (k(x_i, x_i) - 2k(x_i, v_i) + k(v_i, v_i))
\]
Gaussian RBF kernel is adopted in this paper, so \( k(x,x) = 1 \). Objection function can be simplified to
\[
J = 2 \sum_{i=1}^{c} \sum_{k=1}^{n} u_{ik}^n (1 - 2k(x_k, v_j))
\]
(22)
The constraint is as follows,
\[
\sum_{j=1}^{c} u_{ik} = 1, \quad 0 < \sum_{k=1}^{n} u_{ik} < n
\]
(23)
Using Lagrange multiplier method, take the derivative of \( J \) with respect to \( u_{ik} \) and \( v_i \) and set the result equal to zero, the update equations of the membership function \( \mu_{ik} \) and the cluster center \( v_i \) are given in (24) and (25).
\[
\mu_{ik} = \frac{(1-k(x_k, v_j))^{-1(n-1)}}{\sum_{j=1}^{c} (1-k(x_k, v_j))^{-1(n-1)}} \quad (24)
\]
\[
v_i = \frac{\sum_{k=1}^{n} u_{ik}^n k(x_k, v_j) x_k}{\sum_{k=1}^{n} u_{ik}^n k(x_k, v_j)} \quad (25)
\]
When facing multi-distribution problem in real applications, the FCM algorithm will be modified to suit different problems, the model will be more complex and difficult to get the proper solution. Aiming at developing a simple and effective method, we proposed our algorithm (KFCMS), which induces the kernel into the SFCM algorithm.

Figure 1. (a) source image with 10% Gaussian noise. (b) Segmentation result of a standard FCM algorithm. Fig.1(c) Segmentation result of the SFCM with the spatial information with \((p=1,q=1)\). (d) Segmentation result of KFCM. (e) Segmentation result of the proposed KSFCM. The experiment shows the segment results based on the KSFCM with spatial are better than the FCM than the conventional FCM.
The algorithm KSFCM includes two steps. A kernel inducing was finished at the first step, which is identical to the conventional KFCM introduced in Section II. The second step in SFCM is to modify the membership value based on neighborhood information as shown in equation (14).

The proposed kernelized fuzzy c-means algorithm (KSFCM) with spatial information can be summarized in the follow steps:

Step1: Set $c$ (cluster number), $m$ (weighting parameter), $\varepsilon$ and (the stop threshold a very small constant).

Step 2: Initialize the membership $U$ matrix.

Step 3: Repeat

(1) Update all prototypes $v$ with equation (25).

(2) Update all membership $U$ with equation (24).

(3) Calculate the difference of $v_i$ in consecutive steps.

Step 4: Modify the membership matrix with equation (14).

The iteration is stopped when the maximum difference between cluster centers of consecutive iterations is less than the given threshold.

KSFCM is constructed based on SFCM proposed in [18] instead of BFCM algorithm, because BFCM algorithm is complex and time-consuming. Cheng and Song reduced the computation complexity and proposed the BFCM_S algorithm, which is still also more complex than the KFCMS. In segmentation process, noises should be eliminated and variation among the pixel will be classified. The use of spatial information is beneficial for eliminating the noises and outliers. The use of kernel is beneficial for enlarging the variance between pixels.

Figure 2  (a) source IR image with Gaussian noises. (b) Segmentation result of a standard FCM algorithm. (c) Segmentation result of SFCM with spatial information ($p=1, q=1$). (d) Segmentation result of KFCM. (e) Segmentation result of the proposed KSFCM method.
IV. EXPERIMENT AND CONCLUSION

In this section, we performed some experiments to compare the performances of these algorithms with some synthetic and natural images. FCM, FCM with spatial constraints [12] (SFCM), KFCM, and the proposed KSFCM are compared. We tested the four methods on two different datasets. One is a simple synthetic image, another one is an infrared image. Gaussian RBF kernel is used for KFCM and KSFCM, use the range of neighborhood recommended in [18], and the control parameters p and q are same respectively in all experiments, the parameter of $\sigma$ used in Gaussian kernel $K(x,y) = 1- \exp(- ||x-y||^2/\sigma^2)$ is optional. We use the value 50 in experiments.

In the first experiment shown in Fig.1, we considered a synthetic image to validate the effect of handling outliers of FCM, SFCM, KFCM and KSFCM. The synthetic image shown in Fig.1a corrupted by 5% Gaussian noise. The neighborhood is a $5 \times 5$ window. The control parameter values of p and q are set to 1. As show in the Fig.1b and Fig.1d, without the spatial information, FCM and KFCM can identify two classes, but some of the noises cannot be eliminated. Fig.1e shows all the result of the SFCM effect, Fig.1e shows the result of the KSFCM effect. It is obvious that KSFCM can segment image much better and eliminate Gaussian noises compared to the FCM, SFCM, KFCM algorithms. In the second experiment shown in Fig.2, segmentation results of an infrared image are compared. The results show comparison of the methods on infrared images. Fig.2a is the original image. Figs. 2b to Fig. 2e are the results of FCM, SFCM, KFCM and KSFCM, respectively. Note that KSFCM get better performance than other algorithms. The text down the lower left corner is clearly segmented in Fig.2e. In other three results images, the text is not clearly segmented and edge blurring occurs.

From the experiment1 results, it can be concluded that FCM with spatial information has better performance than without using spatial information. It can also be concluded that SFCM with kernel has good segmentation result than without using kernel.

V. Summary

In this paper, we proposed a novel segmentation method which combines the kernel method and used the spatial information of neighborhood. Compared to other modified fuzzy C-means method, the proposed method is simple, fast and easy to realize implement. Keh-Shih proposed SFCM by improving FCM membership matrix with spatial information of neighborhood. Chen and Zhang improved the BFMC by replacing the Euclidean distance with a kernel-induced distance. In this paper, we proposed an algorithm (KSFCM) that adopts a kernel method and uses neighborhood spatial information. This method is not only robust and not sensitive to noises and outliers, but also transforms the patterns to the feature space for easy clustering since it is enlarges the differences of the patterns in the feature space. For low contrast image, such as infrared image, this algorithm has better performance comparing to FCM, KFCM, and SFCM. The experiment results demonstrated the effectiveness of the proposed KSFCM algorithm.

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for their detailed review and constructive comments. This work was partially supported by National Natural Science Foundation of China (No. 60736046), the National Program 973 of China (Grant No. 2009CB320803).

REFERENCES

[4] Balazs Balasko, Janos Abonyi and Balazs Feil, “Fuzzy clustering and Data Analysis Toolbox” (wwwfmt.vein.hu sofcompat)


Cuiyin Liu, female, received the MS degrees from the Southwest JiaoTong University, Chengdu, China, in 2007. Since 2009, she has been working towards the Ph.D. degree at Sichuan University. Her research interests include pattern recognition, image segmentation, image registration, and virtual reality. She is a member of the ACM.

Xiu-Qiong Zhang received the MS and PhD degrees from Sichuan University, Chengdu, China, in 2006 and 2011, respectively. She is currently an associate professor in the School of computer science at the Leshan Normal University of China. Her research interests include image fusion and computer vision. She is a member of the ACM.

Xiaofeng Li, Corresponding author, received MS degree from Tsinghua University, Beijing, China, in 1999. He is currently an associate professor in school of computer science in Sichuan University, Chengdu, China. His research interests include image registration, image fusion and computer vision.

Yani Liu, born in 1981, received the ME degrees from the Electronic Science and Technology University, Chengdu, China, in 2007. Her research interests include pattern recognition and image segmentation.