2014 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)
July 6-11, 2014, Beijing, China

Kernel Non-Local Shadowed C-Means for Image Segmentation

Long Chen, Jing Zou and C. L. Philip Chen
Faculty of Science and Technology, University of Macau
Macau, China
longchen@umac.mo, mb15535@umac.mo, philipchen@umac.mo

Abstract—In order to apply successfully the fuzzy clustering algorithms like shadowed C-means (SCM) to image segmentation problems, the spatial information related with each pixel in the image should be carefully calculated and appended to the clustering algorithms. In this paper, the non-local spatial information calculation is introduced to SCM. Because the data in the kernel space demonstrate more linearly-separable shape and the distances calculated in it shows the property of robust to noise and outliers, the proposed clustering algorithm is conducted in the kernel space (aka feature space) mapped from the original space by some implicit mapping functions defined in the kernel functions. Simulations results on some noise images and the comparison with traditional methods demonstrate the efficiency and superiority of the proposed new approach.

Keywords: Fuzzy clustering; Image segmentation; Shadowed c-means; Non-local spatial information; Kernel method.

I. INTRODUCTION

Image segmentation works as a basic task in image and video understanding [1], computer vision [2], medical imaging [3], and pattern recognition [4]. The major target of image segmentation is to split the image into homogeneous regions in which pixels in the same region demonstrate similar characteristics while the pixels in the different regions show noticeable dissimilarities. Fuzzy clustering algorithms are widely used in image segmentation tasks because of their simple implementation, fast convergence and the benefit of offering soft boundary between different clusters[5-6,12-14].

The shadowed set [9] is a derivation of fuzzy set that offers a new approach to deal with the outliers in data clustering. The shadowed set is defined over a fuzzy set by split the universe of discourse of the fuzzy set into three parts: the core region, the shadowed region and exclusive region. Then, the membership values in core region are boost to one and the membership values in exclusive region are suppressed to zero. By applying the concept of shadowed set in the iteration of fuzzy clustering, the Shadowed C-Means (SCM) algorithm and its variants have been studied [10-11]. Because the outliers in the boundary of cluster or in the exclusive region of fuzzy set usually take the suppressed membership values, their contribution to the cluster is limited. Therefore, the SCM has a good capability in reducing the outliers' influence. In [11], SCM is recommended as a better image segregation problems approach than FCM variants.

For fuzzy clustering based image segmentation, researchers have noticed that simply applying the intensities of pixels as the input data of clustering algorithm discarded the valuable spatial information pertained to each pixel because pixels are positioned at specific place of an image and the pixels neighbored to a pixel also give the pixel specific characteristics differed to other pixels [7]. Inspired by the success of nonlocal image denoise algorithms [8], the non-local spatial information calculation methods and the corresponding integration with SCM, the Non-Local SCM (NLSCM), will be explained in this paper.

Kernel methods have drawn many researchers' interests since the support vector machines (SVM) are proposed and widely used in a range of applications [15]. The kernel method maps the data from the original data space to a high dimensional Hilbert space via some mapping functions. The data in the kernel or feature space will show higher linear separatibility. The advantages of kernel methods boost the quick kernelization of traditional fuzzy clustering algorithms and their applications in image segmentation problems [12-14]. Encouraged by good performance of different Kernel Fuzzy C-means(KFCM) based image segmentation [12-14], this paper also introduce the kernel version of NLSCM, Kernel NLSCM (KNLSCM) for image segmentation problems. The experiments on some synthetic and real images demonstrate the efficiency and advantages of the proposed methods.

The remainder of this paper is organized as follows. Section 2 reviews algorithms like FCM, KFCM and SCM. Then the NLSCM and its kernel version KNLSCM are detailed in Section 3. The comparison results between proposed algorithms and the traditional approaches are summarized in Section 4. Finally, in Section 5, we conclude this paper with future work.

II. RELATED WORK

A. Fuzzy C-Meansand Kernel Fuzzy C-Means

Through the minimization of an objective function that summarizes the weighed distances between the data and the prototypes, the Fuzzy C-means (FCM) partitions N patterns X_k (k=1,2, …, N) into C clusters [16]. The objective function is:

$$J=\sum_{k=1}^{N}\sum_{i=1}^{C}(u_{ik})^m\left\|X_k-V_i\right\|^2 \tag{1}$$

where $m>1$ is the fuzzy degree, $u_{ik}\in[0,1]$ is the membership of k th pattern in cluster i . The prototype or cluster center of cluster i is V_i . The $\|\bullet\|$ is the Euclidean distance.

The FCM algorithm continually updates u_{ik} and V_i as

$$V_i = \frac{\sum_{k=1}^{N}(u_{ik})^m X_k}{\sum_{k=1}^{N}(u_{ik})^m} \quad (2)$$

and

$$u_{ik} = \frac{1}{\sum_{j=1}^{C}(\frac{d_{ik}}{d_{jk}})^{2/m\text{-}1}} \quad (3)$$

Where d_{ik}=$||X_k - V_i||^2$. The iteration usually stop when the stop criterion like maximum iteration number or $max|u_{ik}(t)\text{-}u_{ik}(t\text{-}1)|<\varepsilon$, is satisfied.

The kernel version FCM was proposed in [17-19]. By mapping the data from the original data space to a much higher dimensional space using a transform function, the kernel fuzzy c-means (KFCM) algorithms try to solve the partition problem in the mapped space. Normally, the mapping function is not given directly. On the other hand, we have the kernel function is defined with the mapping function as

$$\kappa(X,Y) = \langle \varphi(X), \varphi(Y) \rangle \quad (4)$$

where φ is the non-linear mapping, and $\langle , \rangle$ is the inner product in the Hilbert space. Similar to FCM, the partition of data points X_k (k=1,2, …, N) in C clusters is conducted by minimize the weighed distances between the data and the prototypes, but this time in the mapped space. As a result, we have the new objective function of KFCM like

$$\begin{aligned} J_m &= \sum_{i=1}^{C}\sum_{K=1}^{N} u_{ik}^m \left\|\varphi(X_k) - \varphi(V_i)\right\|^2 \\ &= \sum_{i=1}^{C}\sum_{K=1}^{N} u_{ik}^m \left(\kappa(X_k, X_k) + \kappa(V_i, V_i) \text{ - } 2\kappa(X_k, V_i)\right) \end{aligned} \quad (5)$$

where $u_{ik} \in [0,1]$ is the membership of the k th pattern to cluster center V_i, and $\|\bullet\|$ is the distance in the kernel space. In this paper, we apply the Guassian kernel $\kappa(X,Y) = \exp(-\|X-Y\|^2/\sigma^2)$ in kernel methods. Because $\kappa(X,X)$=1 for Gaussian kernels, equation (5) can be reformulated as

$$J_m = \sum_{i=1}^{C}\sum_{K=1}^{N} u_{ik}^m \left(2 \text{ - } 2\kappa(X_k, V_i)\right) \quad (6)$$

The new objective function can be minimized by iteratively updating the memberships and cluster centers as below [20]:

$$u_{ik} = \frac{1}{\sum_{j=1}^{C}(\frac{d_{ik}}{d_{jk}})^{2/m\text{-}1}} = \frac{\left(1-\kappa(X_k, V_i)\right)^{-1/(m-1)}}{\sum_{j=1}^{C}\left(1-\kappa(X_k, V_j)\right)^{-1/(m-1)}} \quad (7)$$

$$V_i = \frac{\sum_{k=1}^{N} u_{ik}^m \kappa(X_k, V_i) X_k}{\sum_{j=1}^{N} u_{ik}^m \kappa(X_k, V_i)} \quad (8)$$

Here d_{ik}=$||X_k - V_i|| = 2 \text{ - } 2\kappa(X_k, V_i)^{\,2}$

B. Shadowed C-Means

As pointed out in Section I, the concept of shadowed set can be combined with fuzzy c-means and bring us the shadowed c-means (SCM). In other words, the SCM can be regarded as a variant of fuzzy c-means (FCM) [10].

The shadowed set is a modified fuzzy set [9], in which, the original membership values close to 1 are changed to exactly 1 and the original membership values close to 0 are changed to 0 evenly. The rationale behind such modifications is that for data points in the core region, their member values are close to 1; therefore, the difference between 0.99 and 1 in membership is useless trivia. Similarly, for the points in the exclusive region, the membership difference between 0.01 and 0 also can be discarded. After all, the unchanged region is defined as the uncertainty/shadow zone to maintain the entire level of ambiguity. Formally, the shadowed set is defined as

$$J : X \to \{0, 1, [0, 1]\} \quad (4)$$

Here X is the universe of discourse. As illustrated in Fig. 1, for a shadowed set, the core region is the group the points with grade 1. The exclusive region contains all the points with grade 0. In the middle, the shadowed region includes points with $J(x)$ = [0, 1].

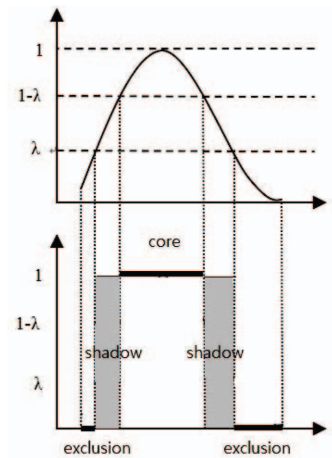


Figure 1. The shadowed set derived from a fuzzy set by some threshold.

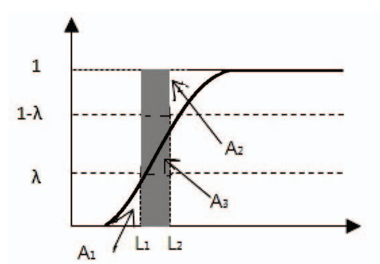


Figure 2. The threshold λ is calculated by: A_1+A_2= card {A_3}.

Now the determining of specific regions of core and exclusive is a problem. In other words, if we increase original fuzzy membership values larger than 1-λ to 1 and decrease the membership values smaller than λ to 0. How should we set the value of λ? This problem is solved in [10] by the basic idea of keeping the vagueness in the original fuzzy set in the produced shadowed set. That is, the λ is selected to minimize the different between the vagueness removed by reduction and enhancement of membership values in core and exclusive regions and the increased uncertainty in membership values in the form of [0, 1] in the shadowed region. Mathematically, the λ is determined by minimizing the term expressed as following [10-11].

$$O(\lambda)=\left|\int_{-\infty}^{L1}J(x)dx+\int_{L2}^{\infty}(1-J(x))dx-\int_{L1}^{L2}dx\right| \quad (9)$$

Here $\lambda \in (0,1/2)$ such that $O(\lambda)=0$. The three integrals at the right side of (9) represent regions A_1, A_2 and A_3 in Fig.2. The parameters L_1 and L_2 denote the boundaries in the integrals, which are computed by λ. They specify the regions in the universe of discourse where the membership values are below the threshold λ and above the threshold1-λ.

In the iteration of FCM, when we get the estimated membership matrix U= $[u_{ik}]$, we have a discrete fuzzy set $[u_{i1}, u_{i1}, \ldots, u_{iN}]$ for cluster i. For each fuzzy set, we can define a shadowed set over it, that is, we split the universe of discourse into three regions by some parameter λ_i.The λ_i. is derived by minimizing the discrete version of (9)

$$O(\lambda_i)=\left|\sum_{X_k|u_{ik}\le\lambda_i}(u_{i_{\max}}-u_{ik})+\sum_{X_k|u_{ik}\ge u_{i_{\max}}-\lambda_i}(u_{i_{\max}}-u_{ik})\right.$$

$$\left.-card\{X_k\mid\lambda_i<u_{ik}<(u_{i_{\max}}-\lambda_i)\}\right| \quad (10)$$

$$\lambda_i=\lambda_{opt}=argminO(\lambda_i) \quad (11)$$

Here u_{imin} and u_{imax} denote the lowest and the highest membership values of data points X_k to the i th cluster [11,23].

With the core, shadowed and exclusion regions of a shadowed set, we modify the updating rule of cluster centers as below [11]:

$$V_i=\frac{\sum_{X_k|u_{ik}\ge u_{i_{\max}}-\lambda_i}X_k+\sum_{X_k|\lambda_i<u_{ik}<(u_{i_{\max}}-\lambda_i)}(u_{ik})^m X_k+\sum_{X_k|u_{ik}\le\lambda_i}(u_{ik})^{m^m}X_k}{\phi_i+\eta_i+\varphi_i} \quad (12)$$

where

$$\phi_i=card\{X_k\mid u_{ik}\ge u_{i_{\max}}-\lambda_i\} \quad (13)$$

$$\eta_i=\sum_{X_k|\lambda_i<u_{ik}<(u_{i_{\max}}-\lambda_i)}(u_{ik})^m \quad (14)$$

$$\varphi_i=\sum_{X_k|u_{ik}\le\lambda_i}(u_{ik})^{m^m} \quad (15)$$

Here, m>1 is the fuzzy coefficient and λ_i is the corresponding threshold for cluster center V_i. According to (8), we will see the data points in the core region have the enhanced weights of 1; therefore, they contribute more to the final value of cluster centers. On the contrary, the data points in the exclusion region have little impact due the double exponenetialed weights $(u_{ik})^{m^m}$. Such kind of processing benefits to the suppression of outliers because the outliers are usually in the boundary of clusters, and their impact on the cluster center calculation is decreased in (12).

In a summary, the shadowed c-means (SCM) works similar to the fuzzy c-means and updates the cluster centers and membership values in an iterative way. The main procedure of the algorithm is listed as below.

1. Initialize cluster centers.
2. Repeat Steps3–5 by incrementing t until convergence or the maximum number of iterations.
3. Compute u_{ik} as (3).
4. Compute threshold λ_i for the i th cluster according to (11).
5. Update cluster centers V_i using (12).

III. Local Spatial Shadowed C-Means and Non-local Spatial Shadowed C-Means

A. Non-local Spatial Shadowed C-Means

The noticeable downside of classical SCM algorithm for image segmentation problem is it does not make full use of the spatial information in the image. In [11], the intensity of each pixel is the data points to be partitioned by SCM. As a result, the noisy pixel of the image is easy to be classified incorrectly due to its noised intensity.

In order to reduce the influence of the noise in the image on image segmentation, the local spatial information obtained from the image, like the average of neighborhood can be incorporated into the Shadowed C-Means clustering algorithm. However, the non-local pixels in the image may contain valuable spatial information as well. Therefore, we introduced the non-local spatial information into the SCM clustering algorithms. The non-local means algorithm (NL-means) is first proposed by Buades et al. in [8] as an image denoising method. In NL-means, we try to find a set of pixels with similar neighborhood configurations or with similar patches taking those pixels as the centers. Then the pixel under consideration could be denoised by the weighted averaging over these patches. Under the same rationale, the non-local spatial information derived from similar patches will be included in the non-local spatial SCM (NLSCM).

In NLSCM, the distance measurement between the cluster center and the pixel is influenced by non-local information, d_{nl} is computed as a weighted average of all the pixel to center distances in the given image defined by LB

$$d_{nl}^2(X_j,V_i)=\sum_{x_k\in LB}w_{nl}(X_k,X_j)d^2(X_k,V_i)\Big/\sum_{x_k\in LB}w_{nl}(X_k,X_j) \quad (16)$$

Here the weight $w_{nl}(x_k,x_j)$ is the similarity between the pixel x_k and x_j in the given image LB. While the similarity between two pixels x_k and x_j located in the given image LB, depends on the difference between the patch N_k and patch N_j, where N_k is a square window of fixed size and centered at a pixel x_k. The difference between patch N_k and N_j is measured as a weighted Euclidean distance $\|v(N_k)-v(N_j)\|_\rho$, where ρ>0 is the standard deviation of the Gaussian kernel [24]. In the distance, we use a Gaussian weighting function to give more weights to the pixels near the center. Specifically, the similarity of two pixels is represented by the distance between two patches as:

$$w_{nl}(X_k,X_j)=e^{-\frac{\|v(N_k)-v(N_j)\|_\rho}{h^2}} \quad (17)$$

Here the parameter h acts as a degree of filtering. It controls the decay of the exponential function and therefore the level of impacts of the nonlocal patches to the considered patches.

Based on the distances between the data points (image pixels) and the cluster centers described in (16), we compute and update the cluster center V_i and the membership values u_{ij} just like the SCM.

We list the major steps of the NLSCM algorithm as below:

1. Initialize cluster centers

2. Compute u_{ij} by (3) and (16).

3. Compute threshold λ_i for the i th class, using (11).

4. Compute new cluster centers, V_i, using (12).

5. Repeat Steps 2–4 by incrementing t until convergence or the maximum number of iterations.

It should be pointed out that after step 2, we can have an optional step that includes the local spatial information into the calculation of the membership values u_{ij} by filtering the membership values in the local window of considered pixel. For example, we can weighted-average u_{ij} to:

$$u_{ij}^{'} = \sum_{k \in NB(X_j)} w_k u_{ik} \ . \tag{18}$$

Here $NB(x_j)$ denotes the neighborhood of pixel x_j.

B. *Kernel Non-Local Spatial Shadowed C-Means*

By considering the spatial influence of a pixel, we can enhance the SCM's robustness to the noise problem of image segmentation. Because the distances in Hilbert space induced by the Gaussian kernel are more robust to noises [12] [21], we further kernelize the NLSCM by mapping the data from the original data space into a higher dimensional kernel space. After this reproduction in the kernel Hilbert space, the data are more easily to be separated or clustered. Kernelization of other fuzzy clustering algorithms and their good performance [12-14] are the major motivation of the kernel version of NLSCM.

Compared to NLSCM, before computing the influence of memberships of these nonlocal similar patches on the membership of each pixel point, we should first to map the original data space into the higher dimensional Hilbert space.

Similar to the KFCM, we update the membership values according to the (7) as well. However, due to the impact of the spatial information, the distance calculation in (7) is changed to the nonlocal version as formulated in (16). After that, we update the center like the SCM. That is, we calculate the threshold λ_i for each cluster and organize the data points into core, shadowed and exclusion regions, and enhance the weights of core points and suppress the weights of exclusive points in the cluster center updating rule of KFCM, i.e. the (12). Specifically, we compute the updated center V_i as

$$V_i = \frac{\sum_{X_k | u_{ik} \geq u_{i_{\max}} - \lambda_i} \kappa(X_k, V_i) X_k + \sum_{Xk | \lambda_i < u_{ik} < (u_{i_{\max}} - \lambda_i)} \kappa(X_k, V_i)(u_{ik})^m X_k + \sum_{X_k | u_{ik} < \lambda_i} \kappa(X_k, V_i)(u_{ik})^{m^m} X_k}{\kappa(X_k, V_i)(\phi_i + \eta_i + \varphi_i)} \tag{19}$$

where

$$\phi_i = card\{X_k \kappa(X_k, V_i) \,|\, u_{ik} \geq u_{i_{\max}} - \lambda_i\} \tag{20}$$

$$\eta_i = \sum_{X_k | \lambda_i < u_{ik} < (u_{i_{\max}} - \lambda_i)} (u_{ik})^m \kappa(X_k, V_i) \tag{21}$$

$$\varphi_i = \sum_{X_k | u_{ik} \leq \lambda_i} (u_{ik})^{m^m} \kappa(X_k, V_i) \tag{22}$$

Clearly in (19), the contribution of data in the exclusive region will be decreased from $\kappa(X_k,V_i)(u_{ik})^m$ in (12) to $\kappa(X_k,V_i)(u_{ik})^{m^m}$. By doing this, the outliers' contribution to the center calculation is decreased.

In conclusion, the major iterations of the Kernel NLSCM (KNLSCM) algorithm are similar to NLSCM algorithm as below.

1. Initialize cluster centers.

2. Compute u_{ik} by (7) and (16), and conduct (18) optionally.

3. Compute threshold λ_i for the i th class, using (11).

4. Compute new cluster centers, V_i, using (19).

5. Repeat Steps 2–4 until convergence or the maximum number of iterations.

IV. Experimental Results

In this section, we present a comparative study on FCM, SCM, NLSCM and KNLSCM to validate the efficiency and superiority of the proposed NLSCM and KNLSCM. The different clustering algorithms are performed on several synthetic and medical images with different types of noises.

For the NLSCM and KNLSCM, the parameter like the variance σ^2 of Gaussian kernel and the degree of filtering, h, in non-local spatial information calculation, are varied in a fixed range and the best settings are selected by the performance index discussed later. The first testing images include a synthetic noised two-value image and a synthetic image named 'Trin'. The synthetic noised two-value image is similar to the one used in [12, 14]. It is in the size of 50×50 pixels. It contains two clusters with two values 0 and 1. The synthetic image named 'Trin' is an image having size 32×32 pixels and has four regions. In real image acquisition process, except the "Gaussian noise", another widely appeared noise is the "Rician noise", which usually contaminates the medical images [25]. So here two different noises including "Gaussian noise" and "Rician noise" (generated by a code obtained from Ged Ridgway [26]), are added to the synthetic images. The sampling noise images and the segmentations are shown as Fig.3 and Fig.4.

In presence of ground truth information, the segmentation accuracy is used to evaluate the goodness of clustering corresponding to a given number of segments. The segmentation accuracy is measured as

$$Seg.\ accuracy = \frac{\text{num. of correctly classfied pixels}}{\text{total num. of pixels}} \tag{18}$$

The segmentation accuracies (SAs) of the 4 methods for the two kinds of noised images are listed in Table I and Table II. According to the last two columns of the Tables, the NLSCM or KNLSCM provides the best results - with better SA than the FCM and SCM.

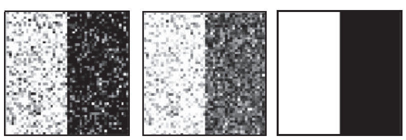
Another testing image is a medical images obtained from database 'BrainWeb' [27]. Fig 5 (a) illustrate one sample. The image is T1-weighted MR phantom with slice thickness of 1 mm. The 5% and 7% Rician noise are added to the image and the noise images will be partitioned into three regions corresponding to White Matters (WMs), Gray Matters (GMs) and Cerebrospinal Fluid (CSF) with 4 segmentation methods. For the MR image with ground truth segmentation, the SA of algorithm i on class j, is used to evaluate the goodness of clustering and calculated as

$$S_{ij} = \frac{A_{ij} \cap Aref_j}{A_{ij} \cup Aref_j} \tag{24}$$

where A_{ij} stands for the set of pixels belonging to class j that are found by algorithm i and $Aref_j$ stands for the set of pixels belonging to class j that is in the reference segmented image. The SAs of FCM, SCM, NLSCM and KNLSCM about the images having 5% and 7% Rician noise, are listed in Table III and IV, in which S1 means the SA for the cluster of CSF, S2 is for the cluster of GM and S3 is for the cluster of WM. According to these two tables, NLSCM and KNLSCM's performance are evidently better.

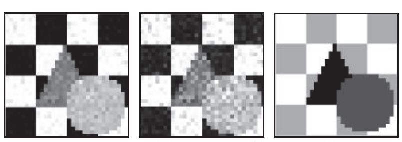
The last testing image is the MR brain image obtained from Internet Brain Segmentation Repository (IBSR) data set [28], in which the GMs in center of brain are obtained using some contour mapping algorithm, and different segmentation algorithms are mainly applied to segment the CSF and WMs. The image is 7% Rician noised and shown in Fig. 5(b). For this image, we also use the S_i (the SA on class i) to evaluate the performance of different algorithms. Table V shows the segmentation results with FCM, SCM, NLSCM and KNLSCM. From the Table V, the KNLSCM is superior, that is, the KNLSCM obtains the maximum SAs in comparison. Except KSSCM, the spatial clustering method NLSCM shows decent results as well. This demonstrates the spatial information's capability in coping with noise in images.

In a short summary, the comparison experiments presented in this section reveal that NLSCM and KNLSCM are good choices for segregation problems of noise images.



(a) (b) (c)

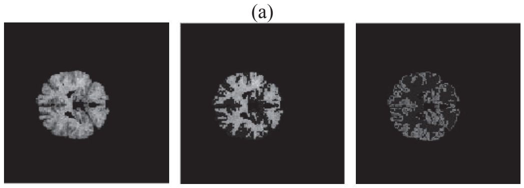
Figure 3. (a) Gaussian noised two-value image, (b) Recisan noised two-value image and (c) the ground truth of segemntation



(a) (b) (c)

Figure 4. (a) Gaussian noised image 'Trin', (b) Recisan noised image 'Trin' and (c) the ground truth of segemntation

(a)



(b)

Figure 5. (a) One brainweb image and its ground truth segmenation (b) One ISBR image and its ground truth segmenation

TABLE I. SEGMENTATION ACCURACIES OF DIFFERENT METHODS ON NOISED TWO VALUE IMAGES

	FCM	SCM	NLSSCM	KNLSFCM
3%Gaussian noise	0.9976	0.9976	1.0000	1.0000
5%Gaussian noise	0.9892	0.9892	1.0000	1.0000
10%Gaussian noise	0.9472	0.9472	1.0000	1.0000
10%Rician noise	1.0000	1.0000	1.0000	1.0000
20% Rician noise	0.9828	0.9832	1.0000	1.0000
30% Rician noise	0.8948	0.8944	0.9980	1.0000

TABLE II. SEGMENTATION ACCURACIES OF DIFFERENT METHODS ON NOISED IMAGES OF 'TRIN'

	FCM	SCM	NLSSCM	KNLSFCM
3%Gaussian noise	1.0000	1.0000	1.0000	1.0000
5%Gaussian noise	0.9958	0.9958	0.9985	1.0000
10%Gaussian noise	0.9077	0.9111	0.9939	0.9954
10% Rician noise	0.8113	0.8279	0.9956	0.9985
20% Rician noise	0.6201	0.6165	0.8579	0.9180
30% Rician noise	0.4734	0.4709	0.7515	0.7681

TABLE III. SEGMENTATION ACCURACIES OF DIFFERENT METHODS ON 'BRAINWEB5'. S1: SEGMENTATION ACCURACY FOR THE CLUSTER OF CSF, S2: SEGMENTATION ACCURACY FOR THE CLUSTER OF GM, S3: SEGMENTATION ACCURACY FOR THE CLUSTER OF WM

5%Ricernd noise	FCM	SCM	NLSSCM	KNLSFCM
S1	0.8760	0.8822	0.9143	0.9121
S2	0.8270	0.8329	0.9160	0.9213
S3	0.9022	0.9058	0.9609	0.9646
SA	0.8707	0.8756	0.9390	0.9426

TABLE IV. SEGMENTATION ACCURACIES OF DIFFERENT METHODS ON 'BRAINWEB7'. S1:SEGMENTATION ACCURACY FOR THE CLUSTER OF CSF, S2: SEGMENTATION ACCURACY FOR THE CLUSTER OF GM, S3: SEGMENTATION ACCURACY FOR THE CLUSTER OF WM

7%Ricernd noise	FCM	SCM	NLSSCM	KNLSFCM
S1	0.7791	0.7919	0.8844	0.8908
S2	0.6989	0.7099	0.8847	0.8934
S3	0.8137	0.8188	0.9455	0.9503
SA	0.7637	0.7723	0.9163	0.9227

TABLE V. SEGMENTATION ACCURACIES OF DIFFERENT METHODS ON 'ISBR'. S1:SEGMENTATION ACCURACY FOR THE CLUSTER OF CSF, S2: SEGMENTATION ACCURACY FOR THE CLUSTER OF GM, S3: SEGMENTATION ACCURACY FOR THE CLUSTER OF WM

7%Ricernd noise	FCM	SCM	NLSSCM	KNLSFCM
S1	0.7470	0.7409	0.7470	0.7823
S2	0.7527	0.7437	0.7488	0.8055
SA	0.7410	0.7381	0.7451	0.7529

V. CONCLUSION

In the paper, the traditional clustering algorithms, FCM, KFMC and SCM are revisited first. Based on the consideration of spatial information, we introduce the non-local shadow c-means (NLSCM) and its kernel version KNLSCM for image segmentation problems. It is demonstrated that the NLSCM and KNLSCM are possible to suppress irrelevant information and outlines. They take care of noise in image very well and can derive better segmentation results than other conventional clustering algorithms like FCM and SCM. NLSCM and KNLSCM are good choices to the clustering based segmentation problems of noise images. But we should be point out that the iteration steps of SCM type algorithms including NLSCM and KNLSCM modified the traditional optimization steps used to derive some local minima. The validity of these SCM type algorithms' capability to obtain local minima becomes an open problem now.

There are some future work can be conducted to improve the methods proposed in this paper. For example, to derive better segmentation results, we can test other approaches to embed the non-local spatial information in the clustering procedure. The multiple kernel approach is a candidate [14]. Another tough problem is the theoretical convergence of SCM and its variants like NLSCM and KNLSCM proposed in this paper. Currently in the experiments, the convergence is not a problem. Nevertheless, the theoretical proof of convergence is still an open problem to the authors' knowledge.

REFERENCE

[1] I. Biederman, "Recognition-by-components: a theory of human image understanding," Psychological review, vol. 94, p. 115, 1987.

[2] N. Paragios, Y. Chen, and O. D. Faugeras, Handbook of mathematical models in computer vision: Springer Publishing Company, Incorporated, 2010.

[3] D. L. Pham, C. Xu, and J. L. Prince, "Current methods in medical image segmentation 1," Annual review of biomedical engineering, vol. 2, pp. 315-337, 2000.

[4] C. M. Bishop and N. M. Nasrabadi, Pattern recognition and machine learning vol. 1: springer New York, 2006.

[5] J. C. Bezdek, R. J. Hathaway, M. J. Sabin, and W. T. Tucker. "Convergence theory for fuzzy c-means: Counterexamples and repairs," *IEEE Transactions on Systems, Man, and Cybernetics*, vol.17, no. 5, pp.873-877, 1987.

[6] Miyamoto, Sadaaki. "Different Objective Functions in Fuzzy c-Means Algorithms and Kernel-Based Clustering," *International Journal of Fuzzy Systems*, vol.13, no.2, pp.89-97, 2011.

[7] K.-S. Chuang, H.-L. Tzeng, S. Chen, J. Wu, and T.-J. Chen, "Fuzzy c-means clustering with spatial information for image segmentation," computerized medical imaging and graphics, vol. 30, pp. 9-15, 2006.

[8] A. Buades, B. Coll, and J.-M. Morel, "A non-local algorithm for image denoising," in Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on, 2005, pp. 60-65.

[9] W. Pedrycz, "Shadowed sets: Representing and processing fuzzy sets, " *IEEE Transactions on Systems, Man, and Cybernetics-Part B: Cybernetics,* vol.28, pp.103–109, 1998.

[10] S. Mitra, W. Pedrycz, B. Barman, "Shadowed C-Means: Integrating fuzzy and rough clustering," Pattern Recognition 43 1282–1291. (2010)

[11] S. Mitra, "Satellite image segmentation with Shadowed C- Means", Information Sciences, vol.181, pp.3601-3613, 2011.

[12] S. C. Chen and D. Q. Zhang, "Robust image segmentation using FCM with spatial constraint based on new kernel-induced distance measure," *IEEE Trans. Syst., Man, Cybern. B, Cybern.,* vol.34, no. 4, pp. 1907-1916, 2004.

[13] M. Gong, Y. Liang, J. Shi et al., "Fuzzy C-Means Clustering With Local Information and Kernel Metric for Image Segmentation," *IEEE Transactions on Image Processing*, vol. 22, no. 2, pp. 573-584, 2013.

[14] L. Chen, C. P. Chen, and M. Lu, "A multiple-kernel fuzzy c-means algorithm for image segmentation," *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 41, no. 5, pp.1263-1274, 2011.

[15] V. Vapnik, The nature of statistical learning theory: springer, 2000.

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

[17] Shen. H ., Yang. J., Wang. S., and Liu. X., "Attribute weighted Mercer kernel based fuzzy clustering algorithm for general non-spherical datasets," *Soft Comput.,* vol.10, no.11, pp.1061-1073, 2006.

[18] Zhang. D. Q. and Chen. S. C., "Clustering incomplete data using kernel-based fuzzy c-means algorithm," *Neural Process. Lett.,* vol.18, no.3, pp.155-162, 2003.

[19] W. L. Cai, S. C. Chen, and D. Q. Zhang, "Fast and robust fuzzy c-means clustering algorithms in kernel-induced feature space," *Pattern Recognit*., vol.40, no.3, pp.825-838, 2007.

[20] D. Graves and W. Pedrycz, "Kernel-based fuzzy clustering and fuzzy clustering: A comparative experimental study," Fuzzy Sets and Systems, vol. 161, pp. 522-543, Feb 16 2010.

[21] P. J. Huber, *Robust Statistics*, New York: Wiley, 1981.

[22] L. Liao, T. S. Lin, and B. Li, "MRI brain image segmentation and bias field correction based on fast spatially constrained kernel clustering approach," *Pattern Recognit.Lett.,* vol.29, no.10, pp.1580-1588, 2008.

[23] W. Pedrycz, "Interpretation of clusters in the framework of shadowed sets," *Pattern Recognition Letters,* vol.26, pp.2439–2449, 2005.

[24] A. Buades, B. Coll, and J.-M. Morel, "A review of image denoising algorithms, with a new one," Multiscale Modeling & Simulation, vol. 4, pp. 490-530, 2005.

[25] H. Gudbjartsson, S. Patz, "The Rician distribution of noisy MRI data Magnetic Resonance in Medicine", vol.34, pp.910–914, 1995

[26] M. Mathworks and Natick, Image Processing Toolbox, [Online]. http://www.mathworks.com/matlabbcentral/fileexchange/14237

[27] C. A. Cocosco, V. Kollokian, R. K.-S. Kwan, and A. C. Evans, "BrainWeb: Online interface to a 3D MRI simulated brain database," NeuroImage, vol. 5, no. 4, pp.S425, 1997. [Online]. http://mouldy.bic.mni.mcgill.ca/brainweb

[28] The Internet Brain Segmentation Repository (IBSR), [Online]: http://www.nitrc.org/projects/ibsr