ISSN: 0976-3104
SPECIAL ISSUE (SCMDSA)
Hudedagaddi and Tripathy

ARTICLE OPEN ACCESS

KERNEL BASED SPATIAL FUZZY C-MEANS FOR IMAGE SEGMENTATION

Deepthi P Hudedagaddi[1*] and Balakrushna Tripathy[2]

[1,2]SCOPE, VIT University,Vellore,Tamil Nadu-632014, INDIA

ABSTRACT

An extension of various available clustering algorithms has been serving as a solution to serve many current problems by the researchers. The Fuzzy C Means (FCM) algorithm that has been in use all these days is extremely noise sensitive. Hence it fails to provide the desired results. This was solved to an extent with the introduction of spatial fuzzy c means. This included a spatial function which was the summation of all the membership values of the neighbors of the pixel considered for study. This paper proposes a new and better modification of the spatial fuzzy c means(sFCM) by introducing kernel distance metric. This groups the objects into clusters which are not separable linearly.Here radial basis kernel function is applied for sFCM clustering. The proposed clustering algorithm is tested on MRI image and noise induced MRI image. The results reveal that kernel based spatial fuzzy c means (sKFCM) is better than Euclidean based spatial fuzzy c means

.

Received on: 30th-Nov-2015
Revised on: 11th-March-2016
Accepted on: 26– March-2016
Published on: 20th–May-2016

KEY WORDS

Clustering, spatial,kernel, fuzzy sets, DB and D index, image segmentation

***Corresponding author:** Email: deepthiph@gmail.com **Tel:** +91-9986387435

INTRODUCTION

Image processing is a rapidly growing field of which image segmentation forms a major part. It has diverse field of applications. Some of them are object recognition, machine vision and medical imaging. Development of segmentation algorithms which are efficient and insensitive to noise has become a challenge. Hence, it has become the need of the hour to develop better algorithms in the field of image segmentation. They must also be capable of solving real world applications and which are sensitive to noise. Noise in real world images are inevitable[1].

Conventional FCM algorithm fails to provide appropriate results on images which have noise. Spatial FCM (sFCM) [2], is a two step procedure. It includes the neighborhood information of pixel taken for study. Though it fails in completely removing the distortion of noise, the algorithm proves to handle noise more efficiently than conventional FCM[3].

The sFCM uses Euclidean distance formula for finding the spatial data point distances. It is found in literature that Euclidean distance fails to provide good results in situations where clustering algorithms are distance based. However, kernel methods provide better results as compared to Euclidean. This paper uses kernel distance formula and compares it with the results from Euclidean. This is hence an extension to the existing spatial FCM.

DISTANCE METHODS

EUCLIDEAN DISTANCE

Suppose $a=(a_1,a_2,...a_n)$ and $b=(b_1,b_2,...b_n)$ are two points in the n-dimensional Euclidean space. Then the Euclidean distance d(a, b) between a and b is given by

COMPUTER SCIENCE

THE IIOAB JOURNAL ISSN: 0976-3104

$$d(a,b)=\sqrt{\sum_{i=1}^{n}(a_i-b_i)^2} \quad (1)$$

KERNEL DISTANCE

Let 'a' denote a data point. Then transformation of 'a' to the feature plane which possess higher dimensionality be denoted by $\phi(a)$. Description of inner product space is given by $K(a,b)=\langle\phi(a),\phi(b)\rangle$. Let $a=(a_1,a_2,...a_n)$ and $b=(b_1,b_2,...b_n)$ are two points in the n-dimensional space. Kernel functions used in this paper are stated as follows

- Radial Basis Kernel

$$R(a,b)=\exp\left(-\frac{\sum_{i=1}^{n}(a_i{}^p-b_i{}^p)^q}{2\sigma^2}\right) \quad (2)$$

Implementations of all the algorithms corresponding to Radial Basis Kernel have been done using p=2 and q=2 in equation (2).

- Gaussian Kernel (RBF with p=1 and q=2)

$$G(a,b)=\exp\left(-\frac{\sum_{i=1}^{n}(a_i-b_i)^2}{2\sigma^2}\right) \quad (3)$$

- Hyper-tangent Kernel

$$H(a,b)=1-\tanh\left(-\frac{\sum_{i=1}^{n}(a_i-b_i)^2}{2\sigma^2}\right) \quad (4)$$

where $\sigma^2=\frac{1}{N}\sum_{i=1}^{N}\|a_i-a'\|^2$ and $a'=\frac{1}{N}\sum_{i=1}^{N}a_i$

For all kernels functions, N denotes total number of existing data points and ||x-y||denotes Euclidean distance between points x and y which pertain to Euclidean metric space. By D(a, b) denotes the complete form of kernel distance function where D(a, b) = K(a, a) + K(b, b) – 2K(a, b) and when similarity property (i.e. K(a, a) =1) is applied, the following is obtained

$$D(a,b)=2(1-K(a,b)) \quad (5)$$

EXISTING METHODS

Fuzzy models are incorporated in analysing spatial data.

Fuzzy Clustering

James C Bezdek developed fuzzy set based Fuzzy c-mean algorithm[4,5]. In this clustering method, each element can belong to more than one cluster. Each element is also associated with a set of membership values. Fuzzy clustering process invloves assigning every data element to one or more than one cluster by taking into account their membership values[6,7].

THE IIOAB JOURNAL ISSN: 0976-3104

1. Assign initial centers for c clusters.
2. Calculate distance d_{ik} between data objects x_k and centroids v_i using Euclidean formula

$$d(x,y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \cdots + (x_n - y_n)^2}$$

3. Generate the fuzzy partition matrix or membership matrix U:

   If $d_{ij} > 0$ then

$$\mu_{ik} = \frac{1}{\sum_{j=1}^{C}\left(\frac{d_{ik}}{d_{jk}}\right)^{\frac{2}{m-1}}} \tag{6}$$

   Else

$$\mu_{ik} = 1$$

4. The cluster centroids are calculated using the formula

$$V_i = \frac{\sum_{j=1}^{N}(\mu_{ij})^m x_j}{\sum_{j=1}^{N}(\mu_{ij})^m} \tag{7}$$

5. Calculate new partition matrix by using step 2 and 3
6. If $\|U^{(r)} - U^{(r+1)}\| < \varepsilon$ then stop else repeat from step 4.

Usually, for all experimental purpose, m is considered to be 2 and ε to be 0.02.

SPATIAL KERNEL FUZZY C MEANS (sKFCM)

Chuang. et al[2]developed spatial Fuzzy C means(sFCM) which incorporated spatial data of the image.It was a modification to the conventional FCM. The spatial function is given by the summation of the membership values in the neighboring of every pixel under consideration. The advantages were that they yielded more homogeneous clusters, diminished spurios blobs and boisterous spots and is insensitive to noise when compared to other systems. On comparable lines, spatial IFCM was additionally developed by Tripathy et.al [8] by presenting the intuitionsitic feature.

Kernel functions add an added advantage to clustering[9,10].In general, when two beside pixels are considered, correlation between them is relatively high. Since the neighboring pixels share similar intensity, the probability of them getting grouped into a same cluster is extremely high. The spatial FCM algorithm exploits this criteria. A spatial function is defined as:

$$h_{ij} = \sum_{k \in NB(x_j)} u_{ik} \tag{8}$$

where NB(x_j) gives neighborhood pixels of x_j. A mask of 5x5 which is equally weighted is used having pixel x_j as it's center. Spatial function h_{ij} portrays the likeliness degree of x_j is in i^{th} cluster. The spatial function values is usually high if most of the pixels in the neighborhood of a particular pixel belong to the same cluster. It is included in the membership function as:

$$u'_{ij} = \frac{u_{ij}^{p} h_{ij}^{q}}{\sum_{k=1}^{c} u_{kj}^{p} h_{kj}^{q}} \tag{9}$$

Here p and q denote the relative weightage of the initial membership and the spatial function respectively. The spatial kernel FCM with parameters p and q is denoted by $sKFCM_{p,q}$. In noisy image conditions, the spatial function reduces the number of misclassified pixels by taking the neighboring pixels into account.

The sKFCM clustering algorithm has two-steps. For every iteration, the conventional FCM algorithm is followed in the first step. Here the distance formula being used is radial-based kernel distance. Later, the centroid and the membership functions are updated. In the second step, the spatial function h_{ij} is calculated and then the new membership function (6) is computed.

RESULTS

A 225x225 dimension brain MRI image has been considered for implementation of sKFCM. We have considered the number of clusters, c=3. **Figure- 1** denotes the original image and **figure- 2** is the image induced with speckle noise with mean 0 and variance 0.04.

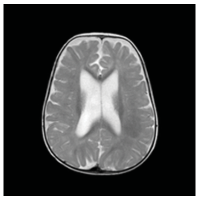

Fig:1. Original MRI image **Fig:2. MRI image with speckle noise**

……..

V_{pc} and V_{pe} indicate fuzzy partition coefficient and partition entropy respectively. Maximum V_{pc} and minimum V_{pe} indicate better clustering[2]. DB and D indices are used to measure the cluster quality[11,12].Higher V_{pc} and lower V_{pe} indicate a good clustering. FCM, sFCM and sKFCM have been applied to both the images. V_{pc} and V_{pe} is given by

$$V_{pc} = \frac{\sum_j^N \sum_i^c u_{ij}^2}{N} \qquad (10)$$

and

$$V_{pe} = -\frac{\sum_j^N \sum_i^c [u_{ij} \log u_{ij}]}{N} \qquad (11)$$

DAVIS-BOULDIN (DB) INDEX

The DB index is defined as the ratio of sum of within-cluster distance to between-cluster distance. It is formulated as given.

$$DB = \frac{1}{c}\sum_{i=1}^{c} \max_{k\neq i}\left\{\frac{S(v_i)+S(v_k)}{d(v_i,v_k)}\right\} \quad \text{for } 1 < i,k < c \ (12)$$

This index tries to minimize the within cluster distance and maximize the between cluster separation. Therefore a good clustering procedure should give value of DB index as minimum as possible[11].

DUNN(D) INDEX

Similar to the DB index the DD index is used for the identification of clusters that are compact and separated. It is computed by using

$$Dunn = \min_i\left\{\min_{k\neq i}\left\{\frac{d(v_i,v_k)}{\max_l S(v_l)}\right\}\right\} \quad \text{for } 1 < k,i,l < c \ (13)$$

This tries in maximizing the between-cluster distance and minimizing the within-cluster distance. Hence a larger value for the D index proves clustering to be more efficient[13].

The results of the validity measures on original image are shown in **Table- 1**.

THE IIOAB JOURNAL ISSN: 0976-3104

Table:1.Cluster Evaluation Results On Normal Image

METHOD	RESULTS ON ORIGINAL IMAGE			
	V_{pc}	V_{pe}	DB	D
FCM	0.7107	$1.5350x10^{-4}$	0.2581	5.0521
$sFCM_{1,1}$	0.7151	$3.4602x10^{-09}$	0.2553	5.3562
$sFCM_{2,1}$	0.7191	$1.7579x10^{-13}$	0.2603	5.1039
$sFCM_{1,2}$	0.7159	$6.4532x10^{-14}$	0.2592	5.399
$sKFCM_{1,1}$	0.727	$5.2001x10^{-32}$	0.2492	5.6926
$sKFCM_{2,1}$	0.7013	$4.3964x10^{-43}$	0.2537	5.3698
$SKFCM_{1,2}$	0.7276	$4.7799x10^{-48}$	0.2501	5.7399

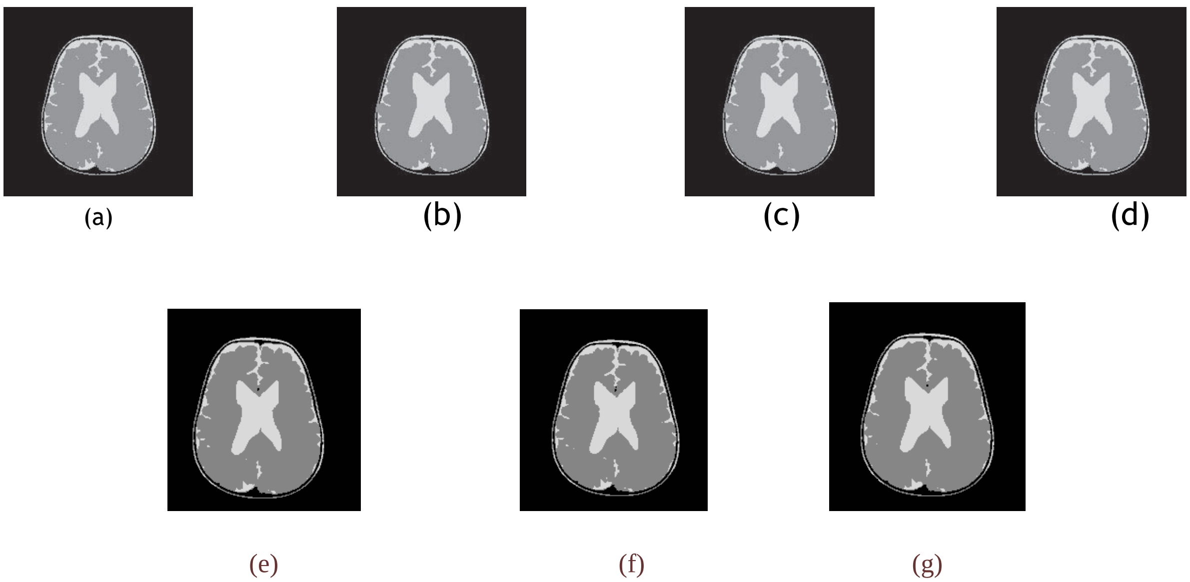

Fig: 3. Image segmentation of original image. (a) FCM. (b) $sFCM_{1,1}$(c)$FCM_{1,2}$.(d)$sFCM_{2,1}$.(e)$sKFCM_{1,1}$.(f)$sKFCM_{1,2}$.(g)$sKFCM_{2,1}$

………..

From the above table and images, it can be seen that sKFCM has better partition coefficient and also possesses less partition entropy. sKFCM also has lower DB and higher D value when compared to conventional FCM and sFCM, thereby, sKFCM provides better clustering.

In the scenario of image with noise, the results are proved to be much better. Conventional FCM does not cluster the image to the expected level in presence of noise. Hence, leads to misclassification. The table below shows the performance of the sKFCM with other techniques implemented on the noisy image.

Table: 2. Cluster Evaluation Results On Noisy Image

METHOD	RESULTS ON NOISY IMAGE			
	V_{pc}	V_{pe}	DB	D
FCM	0.6975	$2.8195x10^{-4}$	0.4517	3.4183
$sFCM_{1,1}$	0.7101	$5.9541x10^{-9}$	0.4239	3.6734
$sFCM_{2,1}$	0.6922	$7.7585x10^{-12}$	0.4326	3.4607

THE IIOAB JOURNAL ISSN: 0976-3104

$sFCM_{1,2}$	0.6874	4.2711×10^{-12}	0.4412	3.6144
$sKFCM_{1,1}$	0.7432	1.3488×10^{-25}	0.3708	4.0343
$sKFCM_{2,1}$	0.7309	3.7309×10^{-31}	0.3839	4.1718
$sKFCM_{1,2}$	0.7086	7.3218×10^{-37}	0.3832	3.9454

From the **Table- 2** and **Figure- 4**, it can be seen that sKFCM produces better results. For noisy image, sKFCM overpowers FCM and all other forms of sFCM.sKFCM reduces the number of spurious spots and blobs to a large extent. It produces a segmented image with a good homogeneity. Smoother segmentation is achieved by taking a high value of *q*.But disadvantage is that, it may blur some of the finer details. The below figures show the segmentation results of sKFCM on the image induced with specklenoise.

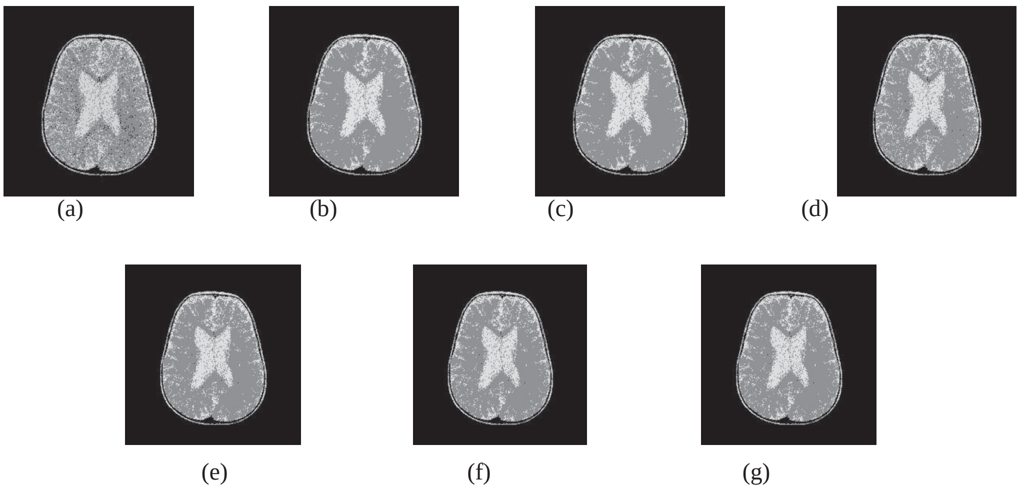

Fig: 4.Image segmentation of noisy image (a)FCM(b)$sFCM_{1,1}$(c)$sFCM_{1,2}$(d)$sFCM_{2,1}$(e)$sKFCM_{1,1}$(f)$sKFCM_{1,2}$(g)$sKFCM_{2,1}$

CONCLUSION

The proposed method adds a kernel approach to the conventional sFCM algorithm. It can be seen that the results of image segmentation and clustering by this approach has brought in better results. The Euclidean distance fails when clustering is to be done where distance is the major parameter. It is when kernel distance provides better results. Proposed method provides a novel way of clustering. The other kernel techniques like Gaussian and Hyper Tangent were also applied during the study. But only radial basis kernel's results were significantly different. However, the reason for unsatisfactory results of Gaussian and hyper tangent kernels is an area open for study. Likeways, this area calls research for developing hybrid clustering algorithms based on uncertainty.

FINANCIAL DISCLOSURE
No financial support was received to carry out this research.

ACKNOWLEDGEMENT
None.

CONFLICT OF INTERESTS
Authors declare no conflict of interest.

ISSN: 0976-3104

REFERENCES

[1] EH Ruspini.[1969] A new approach to clustering, *Information and control*, 15(1): 22-32.

[2] Eman KS Chuang et al.[2006] Fuzzy c-means clustering with spatial information for image segmentation, *Computerized Medical Imaging and Graphics*, 30.

[3] P Swarnalatha, BK Tripathy, PL Nithin and D Ghosh.[2014] Cluster Analysis Using Hybrid Soft Computing Techniques, CNC- 2014International Conference of Network and Power Engineering ,Proceedings of Fifth CNC-2014, 516-524.

[4] BKTripathy, A.Ghosh, GK Panda.[2012] Kernel Based K-Means Clustering Using Rough Set, Proceedings of 2012 International Conference on Computer Communication and Informatics (ICCCI -2012), Jan. 10 – 12, (2012), Coimbatore, INDIA, pp.1 -5.

[5] B K Tripathy, A Tripathy, K Govindarajulu, R Bhargav.[2014] On Kernel Based Rough Intuitionistic Fuzzy C-means Algorithm and a Comparative Analysis. In Advanced Computing, *Networking and Informatics*- 1: 349-359). Springer International Publishing.

[6] R Bhargav, BK Tripathy, A Tripathy, R Dhull, E Verma, and P Swarnalatha.[2013] Rough Intuitionistic Fuzzy C-Means Algorithm and a Comparative Analysis, ACM conference, Compute 2013, (2013), ACM 978-1-4503-2545-5/13/08.

[7] S Mitra, H Banka, and W Pedrycz. [2006] Rough-Fuzzy Collaborative Clustering, IEEE Transactions on System, Man, and Cybernetics, Part B: *Cybernetics*, 36(4):795-805.

[8] Tripathy BK, Avik Basu, and Sahil Govel.[2014]] Image segmentation using spatial intuitionistic fuzzy C means clustering." Computational Intelligence and Computing Research (ICCIC), 2014 IEEE International Conference on. IEEE,

[9] BKTripathy and P Swarnalatha.[2014] A Comparative Study of RIFCM with Other Related Algorithms from Their Suitability in Analysis of Satellite Images using Other Supporting Techniques, Kybernetes,.43(1):53-81.

[10] B.K.Tripathy and R. Bhargav: Kernel Based Rough-Fuzzy C-Means, PReMI, ISI Calcutta, December, LNCS 8251, (2013), pp.148-157.

[11] D L Davis, and DW Bouldin. [1979] A cluster separation measure, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.PAMI-1,.2:224 – 227.

[12] D Zhang, and S Chen.[2002]Fuzzy Clustering Using Kernel Method, Proceedings of the international conference on control and automation, Xiamen, China, 123 – 127

[13] JC Dunn.[1973] A fuzzy relative of the ISODATA , Process and its use in detecting compact well-separated clusters, 32-57.

ABOUT AUTHORS

Deepthi P Hudedagaddi is pursuing her Masters at Vellore Institute of Technology. She is working on fuzzy clustering techniques on spatial data.

Dr. B K Tripathy, a triple gold medalist, is a senior professor in VIT University. He has supervised 19 PhD s, 13 M. Phil s and 02 M.S degrees. He is a senior member of IEEE, ACM, ACEEE and CSI. and is associated with over 60 international journals,published 320 articles, two research volumes and two books.He is working on Rough sets, Fuzzy sets, Social networks, Data mining, Soft Computing, E-Learning, Granular computing, Multi criteria decision making, Neighbourhood systems, SIoT and Soft Sets.