ANALYSIS OF FUZZY C MEANS AND SUBTRACTIVE CLUSTERING USING FUZZY INFERENCE SYSTEM FOR HYPERSPECTRAL IMAGE PROCESSING

M.R. Thiyagupriyadharsan[1] and S. Suja[2]
[1]Department of Electronics and Communication Engineering, Sri Krishna College of Technology, Coimbatore, India. Email: thiyagu.priyadharsan@gmail.com
[2]Department of Electrical and Electronics Engineering, Coimbatore Institute of Technology, Coimbatore, India. Email: ssujacit@yahoo.com

Abstract: *This paper presents the clustering techniques fuzzy C-means clustering(FCM) and subtractive clustering for Hyperspectral image dataset. The FCM should be done only on known number of clusters and subtractive clustering can be done on unknown data also. First FCM is done on known hyperspectral image dataset and for the dataset the subtractive clustering is done. Both the algorithms yielded good results. Then both the algorithms performance are evaluated by using Fuzzy Inference system (FIS) model. The model shows improvement in the test data over the training data.*

Keywords: *FCM clustering, Subtractive Clustering, Hyperspectral image dataset and FIS.*

1.Introduction

Hyperspectral image processing is widely used in remote sensing, military, urban planning, medical, vegetation and mineral mapping. The hyperspectral image dataset has several hundreds of bands in the electromagnetic spectrum [1]. The dataset will have several components or spectra most of them is unknown. Therefore classifying the hyperspectral image is a difficult task [2].

To simplify the classification process, first clustering is performed. Clustering is the process of grouping similar data. The clustering algorithm is of two categories, first the number of clusters is known and second the number of clusters is unknown [3]. The conventional algorithms kmeans and FCM are of the first category and subtractive clustering and mountain clustering are of the second category.

For kmeans and FCM algorithms the number of clusters has to be given for clustering the data. Both the algorithms are well proven for their efficient clustering provided the number of clusters is given. If the number of clusters is unknown then these algorithms will not perform well. For those applications the subtractive clustering and mountain clustering algorithm is used [4].

In the proposed work, the FCM algorithm and the subtractive clustering algorithm is applied to hyperspectral image dataset. The proposed work first FCM algorithm is applied to known dataset and for the same dataset the subtractive clustering is applied.

After applying both the algorithm, the performance is validated by using FIS. The FIS is implemented separately for FCM and subtractive clustering algorithm.

This paper is organized as follows. Section 2 explains FCM algorithm. Section 3 explains the subtractive clustering algorithm. Section 4 explains about the FIS. Section 5 explains about the dataset used in the algorithm. Section 6 describes the results and discussion. The conclusion is given in section 7.

2. Fuzzy C-means Clustering algorithm

Fuzzy C-means clustering (FCM), is based on fuzzy partitioning, in that a data point may belong to several clusters specified by grades of membership in the range of 0 and 1 [10].In order to partition a dataset the cost function has to be minimized.

The membership matrix **U** should have the values in the range of 0 and 1. The cumulative sum of belongingness degree of the data point in all clusters should equal unity:

$$\sum_{i=1}^{c} \mathbf{U_{ij}} = \mathbf{1}, \forall\, \mathbf{j} = \mathbf{1}, \dots, \mathbf{n}. \qquad (\mathbf{1})$$

The cost function of FCM is given by the equation (2):

$$\mathbf{J(U,c_1\ldots\ldots c)} = \sum_{i=1}^{c} J_i = \sum_{i=1}^{c} \sum_{j=1}^{n} U_{ij}^{m} d_{ij}^{2} \qquad (\mathbf{2})$$

where u_{ij} is in the range of 0 and 1; c_i is the center of cluster in fuzzy group i ; $d_{ij} = \|c_i - x_j\|$,$\mathbf{c} - \mathbf{x}$ is the Euclidean distance between i and j; and $m\in[1,\infty]$ is a weighting exponent.

The necessary conditions for Equation (2) to reach its minimum are

$$\mathbf{c_i} = \frac{\sum_{j=1}^{n} u_{ij}^{m}}{\sum_{j=1}^{n} u_{ij}^{m}} \qquad (3)$$

and

$$\mathbf{u_{ij}} = \frac{1}{\sum_{k=1}^{c} \left(\frac{d_{ij}}{d_{kj}}\right)^{2l(m-1)}} \qquad (4)$$

The algorithm works continuously till further update is made. FCM works in the following steps:

Step 1:Begin with random values in the range of 0 and 1 for the membership matrix U which satisfies Equation (1).
Step 2: The fuzzy cluster center c is find using Equation (3).
Step 3: The cost function is computed by using Equation (5). Based on the threshold value it is terminated.
Step 4: New membership matrix U is computed using Equation (4).

Then jump to step 2.

FCM has to be iterate for many times for various values of membership grades since it is based on the starting values of membership matrix.

3. Subtractive Clustering algorithm

Subtractive clustering finds the center of clusters by using data points. Therefore the computation will be based on the size of problem [8]. However, it is not mandatory the center of clusters should lie on any of the data points. Even though it's a good optimization as it reduces the number of computations in many cases [5].

Hence every data point is a member for the center of clusters c, the *density measure* of data point $\mathbf{x}_i$ is given by

$$\mathbf{D_i = \sum_{j=1}^{n} exp\left(-\frac{\|x_i - x_j\|^2}{\left(\frac{r_a}{2}\right)^2}\right)}, \qquad \textbf{(5)}$$

Where, r_a represents positive constant with neighborhood radius. Therefore, for data point with more neighboring data points will acquire high density value.

The largest density value D_{c1} will be selected as the first center of cluster $\mathbf{x}_{c1}$. Then, each data point density measure $\mathbf{x}_i$ is changed by:

$$\mathbf{D_i = D_i - D_{c1} exp\left(-\frac{\|X_i - X_{c1}\|^2}{\left(\frac{r_b}{2}\right)^2}\right)} \qquad \textbf{(6)}$$

Where, *rb* represents positive constant with neighborhood that has reduced density measures. Hence, there will be reduced density measures around the first center of cluster $\mathbf{x}_{c1}$.

After the revision of density function, the highest density value is chosen as the next cluster center. The process continues till a sufficient cluster numbers are obtained.

4. Fuzzy Inference System (FIS)

Fuzzy inference system has the ability of intuition and analyzes human perceptions. Therefore it has more applications in economics, science and engineering applications. FIS has an expert knowledge to capture the changes in the environment and can easily put in the fuzzy systems [7].

Fuzzy inference system algorithm has three main steps:

A.Fuzzification

Fuzzy system will convert the input into fuzzy input based on the membership function. The membership function types are Gaussian, trapezoidal, and triangular membership function.

B.Inference

The fuzzy input and output variables are used to construct fuzzy rules. The output is got by the fuzzy rules inference.

C.Defuzzification

This is used to convert fuzzy output to crisp output.

.

Methods of FIS

The types of FIS are

- Mamdani Fuzzy Inference System
- Takagi-Sugeno Fuzzy Model (TS Method)

5 .Dataset

The datasets which are used in the proposed work are as follows [9] :

Indian Pines

This image was captured by AVIRIS sensor in the Indian Pines site of North western Indiana and consists of 145x145 pixels and 224 spectral reflectance bands. The Indian Pines image consists of two-thirds agriculture, and one-third forest or other natural vegetation. It consists of a rail track, two dual lane highways, and some houses, other buildings, and small roads. The sample image of Indian Pine is shown in figure 1.

Figure 1.Sample image of Indian Pine dataset

Salinas-A scene

A small sub image of Salinas image. It comprises 86x83 pixels and has six classes. The sample image of SalinasA scene is shown in figure 2.

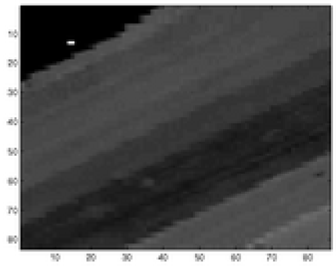

Figure 2.Sample image of SalinasA dataset

Pavia University

This scene is captured over Pavia, northern Italy by ROSIS sensor. The Pavia University is 610x610 pixels. The sample image Pavia University is shown in figure 3.

Figure 3.Sample band of SalinasA dataset

6.Results and Discussions

The simulations are done using the three datasets mentioned above. First the FCM algorithm is implemented with 100 iterations for each datasets. The output of the FCM clustering is depicted in the figures 4,5,6 respectively. Figures 7,8,9 shows the output of the objective function of FCM which will indicate the stopping point of the clustering.

Then FIS-Subtractive clustering and FIS-FCM clustering simulations are done by separating the datasets into test and training data. In FIS-subtractive clustering [7], it is implemented only in Mamdangini model and FIS-FCM it is implemented in Mamdangini and Sugeno model The results are depicted in figures 10,11,12,13,14, and 15.

The Table1 shows the RMSE results of the models [6]. The end results indicate that the model has improved a lot with respect to training data. The model has produced good results with test data.

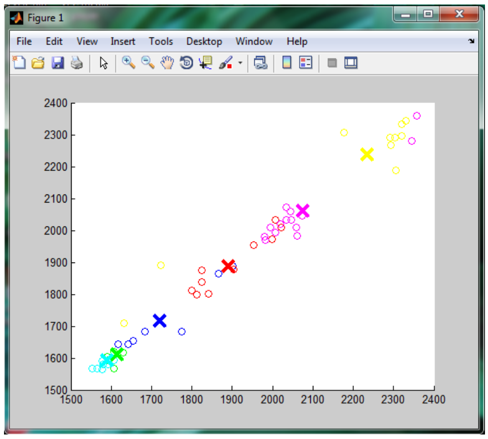

Figure 4.FCM clustering output of SalinasA dataset

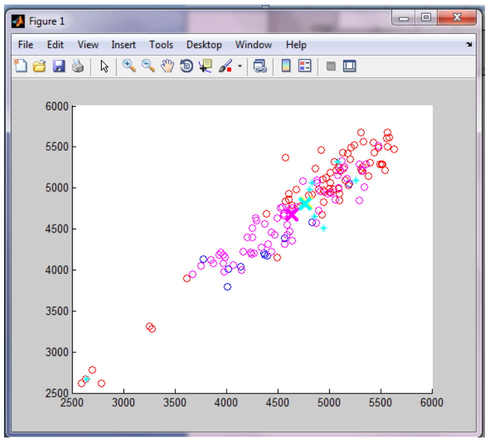

Figure 5.FCM clustering output of Indian Pines dataset

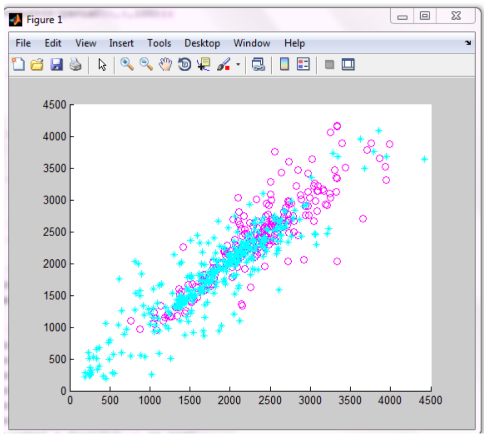

Figure 6.FCM clustering output of Pavia University dataset

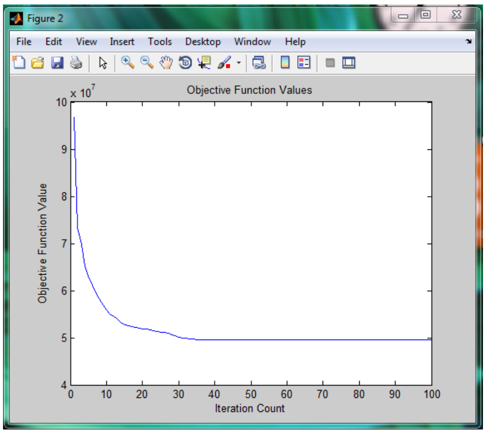

Figure 7.FCM objective function output of SalinasA dataset

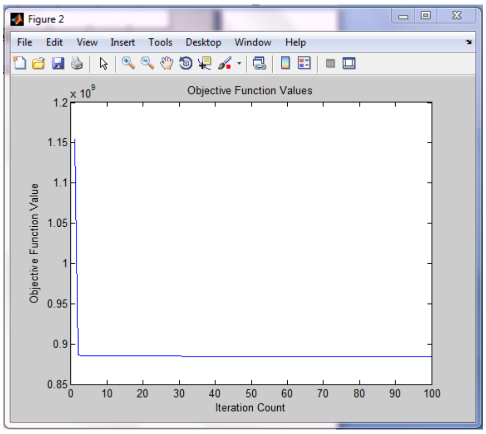

Figure 8.FCM objective function output of Indian Pines dataset

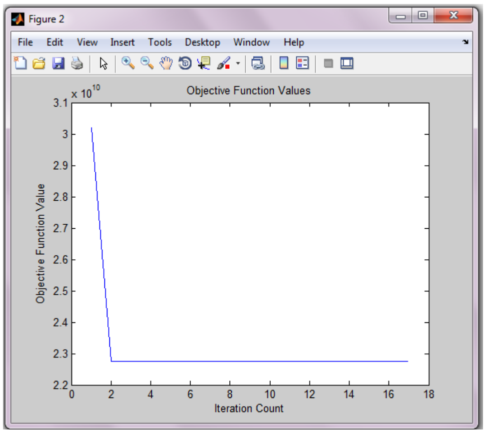

Figure 9.FCM objective function output of Pavia University dataset

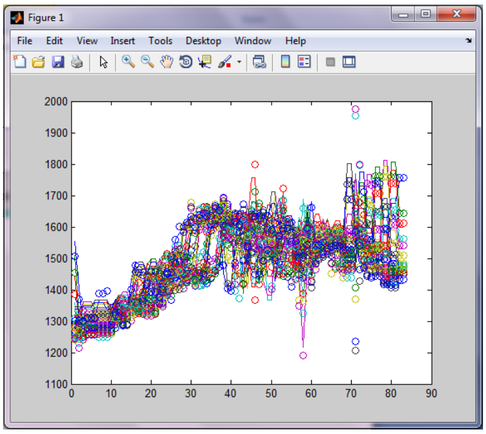

Figure 10.FIS-subtractive clustering plot output of Salinas dataset

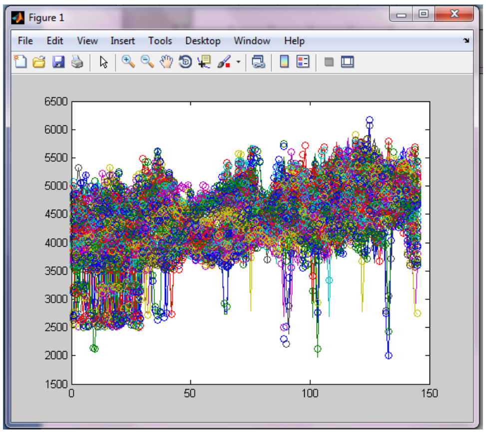

Figure 11.FIS-subtractive clustering plot output of Indian Pines dataset

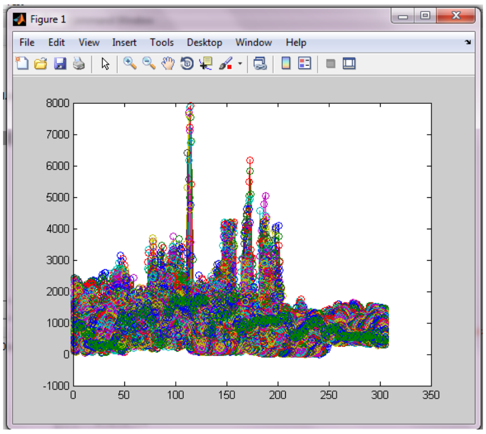

Figure 12.FIS-subtractive clustering plot output of Pavia University

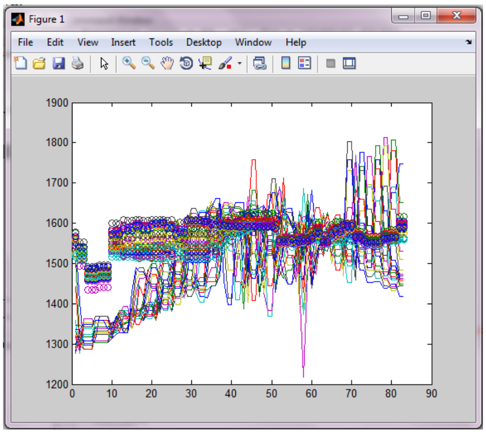

Figure 13.FIS-FCMclustering plot output of Salinas

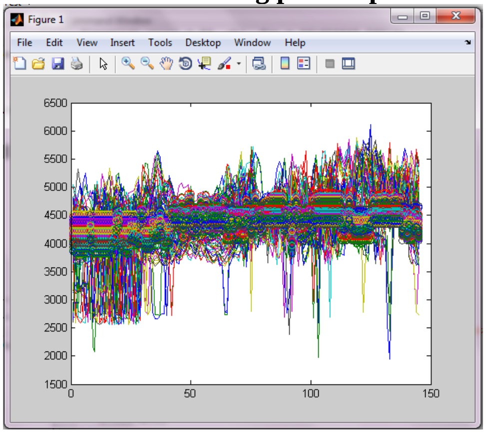

Figure 14.FIS-FCMclustering plot output of Indian Pines

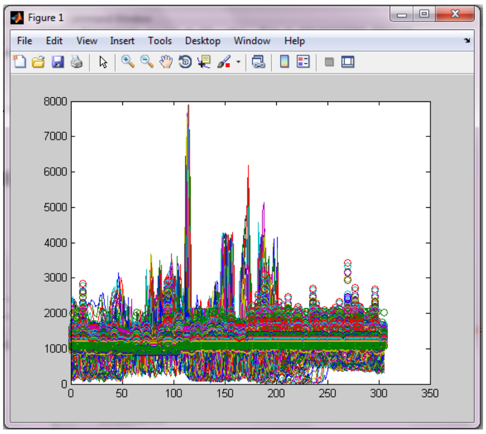

Figure 15.FIS-FCMclustering plot output of Pavia University

Table1 Result of FIS subtractive and FCM models

Dataset	FIS-Subtractive Clustering		FIS-FCM - Mamdangini		FIS-FCM -Sugeno	
	Train RMSE	Test RMSE	Train RMSE	Test RMSE	Train RMSE	Test RMSE
Salinas	0.21	0.23	0.40	0.44	0.71	0.67
Indian pine	0.20	0.31	0.28	0.25	0.23	0.19
Pavia University	0.80	0.17	0.55	0.55	0.64	0.63

4. Conclusion

The proposed method FCM and subtractive clustering algorithms performs good and finds the clusters available. First the FCM algorithm is implemented to the known dataset and for the dame dataset the subtractive clustering is done .Both the methods got good output. Then both algorithms performance are evaluated by FIS model. On testing with the FIS model it shows improvement in the test data over the training data.

References

1. G. Shaw and D. Manolakis, “Signal processing for hyperspectral image exploitation,” IEEE Signal Process. Mag., vol. 19, no. 1, pp. 12–16, Jan. 2002.
2. D. Landgrebe, “Hyperspectral image data analysis,” IEEE Signal Process. Mag., vol. 19, no. 1, pp. 17–28, Jan. 2002.
3. Khaled Hammouda, “A Comparative study of Data Clustering technique” Department of System Design Engineering, University of Waterloo, Canada.
4. Veronica S. Moertini, “Introduction to five data clustering” , Integral, Vol. 7, No. 2, October 2002.
5. Nameirakpam Dhanachandra, Khumanthem Manglem and Yambem Jina Chanu, “Image segmentation K-means clustering algorithm and subtractive clustering algorithm” Procedia Computer Science , vol. 54, pp. 764 – 771, 2015.
6. Hussein A. Lafta,” Solving overvitting problem in Fuzzy subtractive clustering” Global Conference on Power Control and Optimization, Gold Coast, Australia, 2-4, February 2010.
7. Swati R. Chaudhari, Manoj E. Patil,”Comparative analysis of fuzzy inference systems for air conditioner”, International Journal of Advanced computer research, vol.4, pp.922-927,December 2014.
8. Gökhan Bilgin, Sarp Ertürk, Tülay Yıldırım, “Segmentation of Hyperspectral Images via Subtractive Clustering and Cluster Validation using One-Class Support Vector Machines”, IEEE Transactions on Geoscience and Remote sensing, vol. 49, pp.2936-2944, August 2011.
9. [Online]:Available:http://www.ehu.eus/ccwintco/index.php?title=Hyperspectral_Remote_Sensing_ScenesC.
10. J. C. Bezdek and R. Ehrlich, “FCM: The fuzzy c-means cluster- ing algorithm,” Computer Geoscience, vol. 10, no. 22, pp. 191–203, 1981.