

# FUZZY C-MEANS ALGORITHMS IN REMOTE SENSING

**Andrej Turčan, Eva Ocelíková, Ladislav Madarász**

Dept. of Cybernetics and Artificial Intelligence  
Faculty of Electrical Engineering and Informatics  
Technical University of Košice  
Slovakia

*Abstract: Fuzzy clustering is a widely applied method for obtaining fuzzy models from data. It has been applied successfully in various fields including geographical surveying, finance or marketing. A brief overview on Fuzzy C-Means based algorithms and detailed views on Fuzzy C-Means (FCM) and its improvement by Gustafson-Kessel (GK) are shown below. Experiments on artificial made-up data and data from remote sensing gathered from probe LANDSAT TM7 are made using FCM and GK.*

*Keywords: fuzzy clustering, fuzzy c-means, remote sensing*

## 1 Introduction

### 1.1 Clustering

Clustering is a division of data into groups of similar objects. Each group, called cluster, consists of objects that are similar between themselves and dissimilar to objects of other groups. Representing data by fewer clusters necessarily loses certain fine details, but achieves simplification. It represents many data objects by few clusters, and hence, it models data by its clusters. Data modelling puts clustering in a historical perspective rooted in mathematics, statistics, and numerical analysis. From a machine learning perspective clusters correspond to hidden patterns, the search for clusters is unsupervised learning, and the resulting system represents a data concept. Therefore, clustering is unsupervised learning of a hidden data concept.

There is a close relationship between clustering techniques and many other disciplines. Clustering has always been used in statistics and science. Typical

applications include speech and character recognition. Machine learning clustering algorithms were applied to image segmentation and computer vision. Clustering can be viewed as a density estimation problem. This is the subject of traditional multivariate statistical estimation. Clustering is also widely used for data compression in image processing, which is also known as vector quantization.

Clustering algorithms, in general, are divided into two categories:

- Hierarchical Methods (agglomerative algorithms, divisive algorithms)
- Partitioning Methods (probabilistic clustering, *k*-medoids methods, *k*-means methods ...)

Hierarchical clustering builds a cluster hierarchy. Every cluster node contains child clusters; sibling clusters partition the points covered by their common parent. Such an approach allows exploring data on different levels of granularity. Hierarchical clustering methods are categorized into agglomerative (bottom-up) and divisive (top-down). An agglomerative clustering starts with one-point (singleton) clusters and recursively merges two or more most appropriate clusters. A divisive clustering starts with one cluster of all data points and recursively splits the most appropriate cluster. The process continues until a stopping criterion (frequently, the requested number *k* of clusters) is achieved.

Data partitioning algorithms divide data into several subsets. Because checking all possible subset possibilities may be computationally very consumptive, certain heuristics are used in the form of iterative optimization. Unlike hierarchical methods, in which clusters are not revisited after being constructed, relocation algorithms gradually improve clusters.

## 1.2 Remote Earth's survey

Satellite remote sensing is an evolving technology with the potential for contributing to studies of the human dimensions of global environmental change by making globally comprehensive evaluations of many human actions possible. Satellite image data enable direct observation of the land surface at repetitive intervals and therefore allow mapping of the extent, and monitoring of the changes in land cover. Evaluation of the static attributes of land cover and the dynamic attributes on satellite image data may allow the types of change to be regionalized and the proximate sources of change to be identified or inferred. This information, combined with results of case studies or surveys, can provide helpful input to informed evaluations of interactions among the various driving forces.

From a general perspective, remote sensing is the science of acquiring and analyzing information about objects or phenomena from a distance. As humans, we are intimately familiar with remote sensing in that we rely on visual perception to provide us with much of the information about our surroundings. As sensors, however, our eyes are greatly limited by sensitivity to only the visible range of

electromagnetic energy, viewing perspectives dictated by the location of our bodies, and the inability to form a lasting record of what we view. Because of these limitations, humans have continuously sought to develop the technological means to increase our ability to see and record the physical properties of our environment.

## 2 Fuzzy Clustering Algorithms

In classical cluster analysis each datum must be assigned to exactly one cluster. Fuzzy cluster analysis relaxes this requirement by allowing gradual memberships, thus offering the opportunity to deal with data that belong to more than one cluster at the same time. Most fuzzy clustering algorithms are objective function based. They determine an optimal classification by minimizing an objective function. In objective function based clustering usually each cluster is represented by a cluster prototype. This prototype consists of a cluster centre and maybe some additional information about the size and the shape of the cluster. The size and shape parameters determine the extension of the cluster in different directions of the underlying domain.

The degrees of membership to which a given data point belongs to the different clusters are computed from the distances of the data point to the cluster centres with regard to the size and the shape of the cluster as stated by the additional prototype information. The closer a data point lies to the centre of a cluster, the higher is its degree of membership to this cluster. Hence the problem to divide a dataset into  $c$  clusters can be stated as the task to minimize the distances of the data points to the cluster centres, since, of course, we want to maximize the degrees of membership. Most analytical fuzzy clustering algorithms are based on optimization of the basic  $c$ -means objective function, or some modification of it.

### 2.1 Fuzzy C-Means

The Fuzzy  $C$ -means (FCM) algorithm proposed by Bezdek [1] aims to find fuzzy partitioning of a given training set, by minimizing of the basic  $c$ -means objective functional:

$$J(\mathbf{Z}; \mathbf{U}, \mathbf{V}) = \sum_{i=1}^c \sum_{k=1}^N (\mu_{ik})^m \|\mathbf{z}_k - \mathbf{v}_i\|_A^2$$

where:

- $\mathbf{U} = [\mu_{ik}] \in M_{fc}$  is a fuzzy partition matrix of  $\mathbf{Z}$
- $\mathbf{V} = [\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_c]$ ,  $\mathbf{v}_i \in \mathbf{R}^n$  is a vector of cluster prototypes, to be determined

- $\|\mathbf{z}_k - \mathbf{v}_i\|^2$  is dissimilarity measure between the sample  $\mathbf{z}_k$  and the center  $\mathbf{v}_i$  of the specific cluster of the specific cluster  $i$  (Euclidean distance)
- $m \in (1, \infty)$  is a parameter, that determines the fuzziness of the resulting clusters

The minimization of  $J(\mathbf{Z}; \mathbf{U}, \mathbf{V})$ , under the constraint  $\sum_{i=1}^c \mu_{ik} = 1$ , leads to the iteration of the following steps:

$$\mathbf{v}_i^{(l)} = \frac{\sum_{k=1}^N (\mu_{ik})^m \mathbf{z}_k}{\sum_{k=1}^N (\mu_{ik})^m}, \quad 1 \leq i \leq c$$

and

$$u_{ik} = \begin{cases} \frac{1}{\sum_{j=1}^c (D_{ik} / D_{jk})^{2/(m-1)}} \\ \mu_{ik} = 0 \end{cases} \quad \text{if } D_{ik} > 0 \text{ and } \mu_{ik} \in \langle 0, 1 \rangle, \quad \sum_{i=1}^c \mu_{ik} = 1$$

The iteration stops when the difference between the fuzzy partition matrices in two following iterations is lower than  $\varepsilon$ .

## 2.2 Gustafson-Kessel Algorithm

Gustafson and Kessel extended the standard fuzzy  $c$ -means algorithm by employing an adaptive distance norm, in order to detect clusters of different geometrical shapes in one data set. Each cluster has its own norm-inducing matrix  $\mathbf{A}_i$ . Here we have to employ the fuzzy covariance matrix  $\mathbf{F}_i$  of the  $i$ -th cluster:

$$\mathbf{F}_i = \frac{\sum_{k=1}^N (\mu_{ik})^m (\mathbf{z}_k - \mathbf{v}_i)(\mathbf{z}_k - \mathbf{v}_i)^T}{\sum_{k=1}^N (\mu_{ik})^m}$$

Algorithm is again based on iteration of the next steps

- computing of the cluster covariance matrices:

$$\mathbf{F}_i = \frac{\sum_{k=1}^N (\mu_{ik})^m (\mathbf{z}_k - \mathbf{v}_i)(\mathbf{z}_k - \mathbf{v}_i)^T}{\sum_{k=1}^N (\mu_{ik})^m} \quad 1 \leq i \leq c$$

- computing of the distances

$$D_{ikA_i}^2 = (\mathbf{z}_k - \mathbf{v}_i)^T \left[ \det(\mathbf{F}_i)^{1/n} \mathbf{F}_i^{-1} \right] (\mathbf{z}_k - \mathbf{v}_i), \quad 1 \leq i \leq c, 1 \leq k \leq N$$

updating of the partition matrix

$$- \mu_{ik} = \begin{cases} \frac{1}{\sum_{j=1}^c (D_{ik} / D_{jk})^{2/(m-1)}} \\ \mu_{ik} = 0 \end{cases} \quad \text{if } D_{ik} > 0 \text{ and } \mu_{ik} \in \langle 0, 1 \rangle, \sum_{i=1}^c \mu_{ik} = 1$$

### 2.3 Differences between FCM an GK

First experiments were made on the self-made data. The reason was to examine differences in approach of both algorithms and to see the differences in the shape of the clusters.

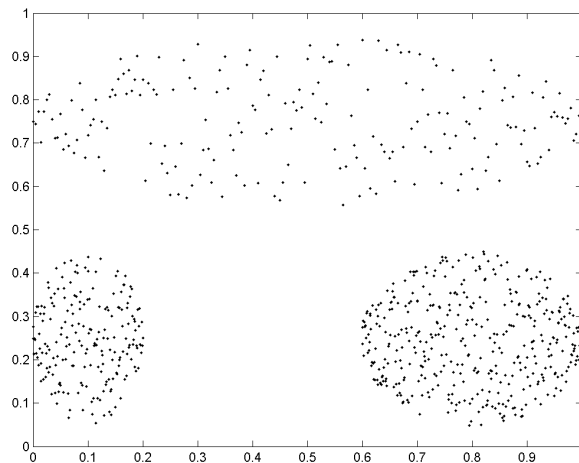


Figure 1: self-made data set

Next figures show how the data set was clustered using the FCM and GK. It is possible to see that clusters after the FCM clustering have spherical shape, while clusters after the GK clustering adopted the shape of particular subset of points.

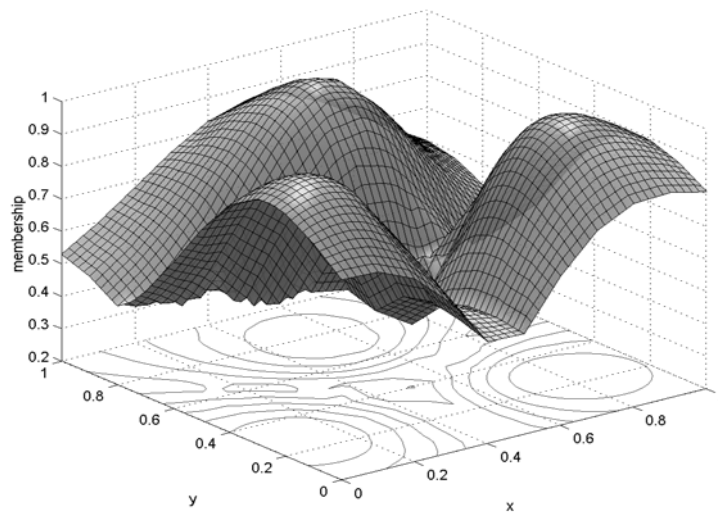


Figure 2: Outcome of the FCM

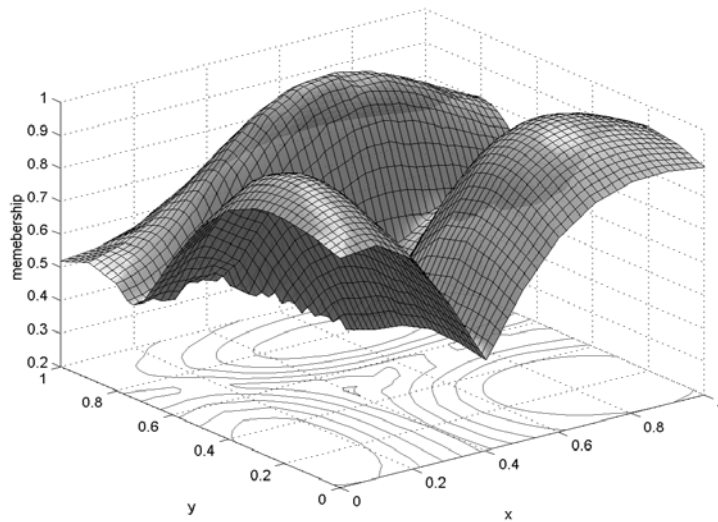
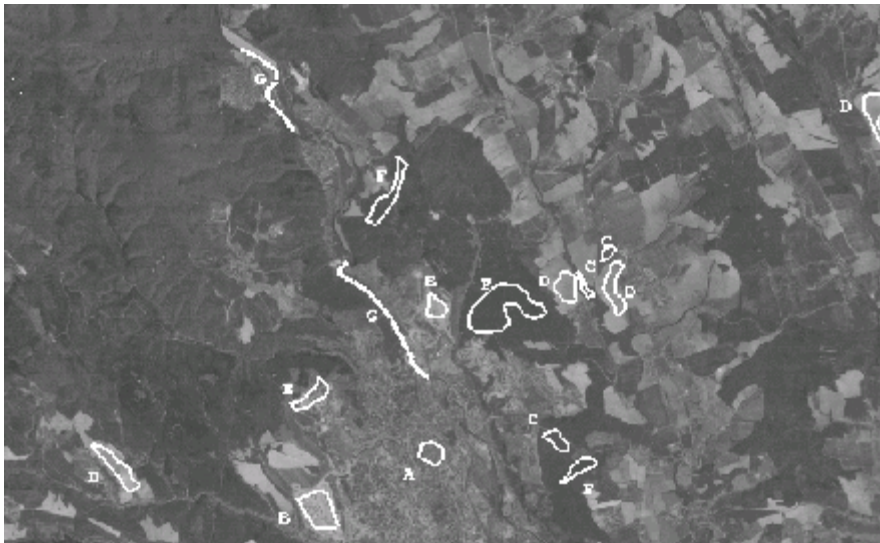


Figure 3: Outcome of the GK algorithm

### 3 Experiments with real-world data

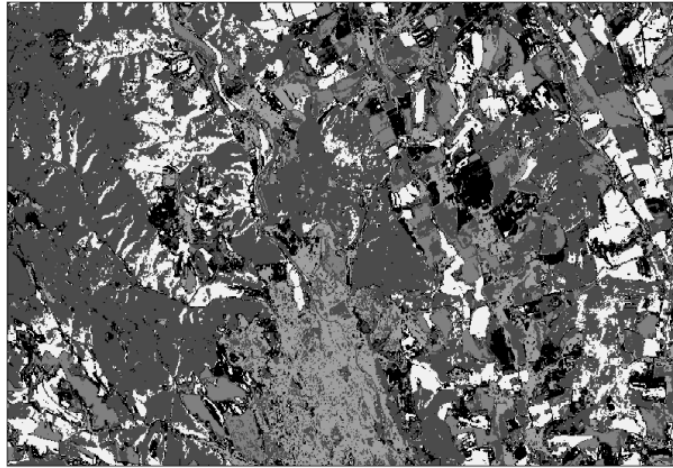
The data set consists of multi-spectral Landsat images (7 dimensional data). The selected geographical area is located in the north part of the city Kosice, Slovakia. The goal was to divide image into 7 particular types of land: A) urban area, B) rural area, C) barren land, D) agricultural land, E) mines, F) forest and G) water.



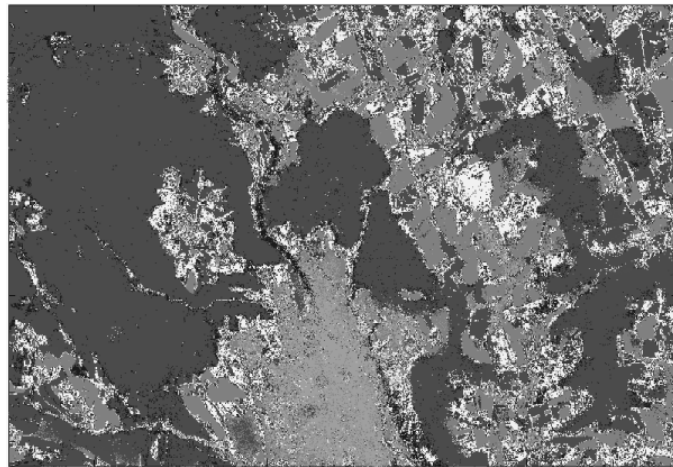
*Figure 4: Original image – Kosice*

The data set consists of 368125 samples, where one sample represents area of size 30 x 30 meters and it represents area of approximately 332 km<sup>2</sup>.

Experiments were made using both fuzzy *c*-means and Gustafson-Kessel algorithms. Fuzziness parameter was chosen  $m=2$ . As a computational tool was chosen Matlab installed on PC with 650 MHz processor and 320 MB RAM.










*Figure 5: Segmentation obtained using FCM*



*Figure 6: Segmentation obtained using GK*

LEGEND:

	urban area		rural area		barren land
	agricultural area		mines		forest
	water				



## 4 Results and conclusion

The results obtained by classification with fuzzy  $c$ -means and Gustafson-Kessel algorithm are shown in Fig. 5 and Fig 6. As shown, the results generated by the Gustafson-Kessel algorithm outperform those generated with fuzzy  $c$ -means.

The fuzzy clustering methods allow classification of the data, where no a priori information is or content is not known. In particular, the fuzzy methods allow to identify data in more flexible manner, assigning to each datum degree of membership to all classes.

Experiments show, that the areas labeled as “mines”, are problematic to classify. Reason for this is probably in size of the image area which is covered by this class. Or in other words area covered by mines is relatively small according to the area covered by one pixel. Other possibility for this, what is also a big disadvantage of  $c$ -means based algorithms, that they tend to stuck in a local extremes. On the other hand these algorithms offer a good tradeoff between accuracy and speed.

## References

- [1] Bezdek, J.C.: *Pattern Recognition with Fuzzy Objective Function.*, Plenum Press, New York, 1981.
- [2] Jain, A.K, Murty, M.N, Flynn, P.J.: *Data Clustering: A Review*
- [3] Bonner, R.E.: *On Some Clustering Techniques.*IBM, 1964.
- [4] Ball, G.H., Hall, D.J.: *ISODATA, A Novel Method of Data Analysis and Pattern Clasification.* Standford Res. Institute, Menlo Park, 1965.
- [5] Fromm, F.R., Northouse, R.A.: *CLASS, A Nonparametric Clustering Algorithm. Pattern Recognition.*
- [6] Raja, A.,Meister, A., Martverk., P.: *Fuzzy Classification Algorithms with Some Applications*
- [7] Baraldi, A., Blonda. P.: *A Survey of Fuzzy Clustering Algorithms for Pattern Recognition.* ICSI, TR-98-038, 1998
- [8] Siddheswar, R., Turi, R.H.: *Determination of Number Clusters in K-Means Clustering and Application in Colour Image Segmentation.*
- [11] Setnes, M., Kaymak, U.: *Fuzzy Modeling of Client Preference in Data-Rich Marketing Enviroments.*

- [12] Setnes, M., Kaymak, U.: *Extended Fuzzy Clustering Algorithms*.
- [13] McBratney, A.B., De Gruijter, J.J.: *A Continuum approach to soil classification by modified fuzzy k-means with extragrades*.
- [14] Yeung, K.Y., Ruzzo, W.L.: *An Empirical Study on Principal Component Analysis for Clustering Gene Expression Data*.
- [15] Faber, V.: *Clustering and the Continuous K-Means Algorithm*.
- [16] Halkidi, M., Batistakis, Y., Vazirgiannis, M.: *On Clustering Validation Techniques*
- [16] Babuška, R.: *Fuzzy and Neural Control*.