IMAGE SEGMENTATION BY FUZZY C MEANS CLUSTERING: SURVEY

R Mohan1  Dr. M Venkatesan2
1Assistant Professor, Department of Electrical and Electronics Engineering, Vidyaa Vikas College of Engineering and Technology, Tiruchengode.
2Principal, K S R Institute for Engineering and Technology, Tiruchengode.

Abstract—This paper presents a latest survey of different technologies using fuzzy clustering algorithms. Clustering approach is widely used in biomedical field like image segmentation. A different methods are used for medical image segmentation like Improved Fuzzy C Means(IFCM), Possibilistic C Means(PCM), Fuzzy Possibilistic C Means(FPCM), Modified Fuzzy Possibilistic C Means(MFPCM) and Possibilistic Fuzzy C Means(PFCM). In this paper, we give a survey methods based on fuzzy clustering and segmentation techniques.

Index Terms—Image segmentation, clustering, fuzzy c-means

I. INTRODUCTION

Image segmentation is one of the most widespread means to classify correctly the pixels of an image in a decision oriented applications. Image segmentation is a technique that partitions an image into uniform and non-overlapping regions based on some likeness measure (G.Dong, 2005).

Cluster analysis is one of the major techniques in pattern recognition. It is an approach to unsupervised learning. The importance of clustering in various areas such as taxonomy, medicine, geology, business, engineering systems and image processing, etc., Clustering and classification are both fundamental tasks in Data Mining. Classification is used mostly as a supervised learning method, clustering for unsupervised learning (some clustering models are for both). The goal of clustering is descriptive, that of classification is predictive (Veyssieres and Plant, 1998).

Clustering of objects is as ancient as the human need for describing the salient characteristics of men and objects and identifying them with a type. Therefore, it embraces various scientific disciplines: from mathematics and statistics to biology and genetics, each of which uses different terms to describe the topologies formed using this analysis. From biological “taxonomies”, to medical “syndromes” and genetic “genotypes” to manufacturing “group technology” the problem is identical: forming categories of entities and assigning individuals to the proper groups within it.

Medical diagnosis and prognosis problems are prime examples of decision making in the face of uncertainty. Dealing with uncertainties is a common problem in pattern recognition and the use of fuzzy set theory has given rise to a lot of new methods of pattern recognition for medical diagnosis. Fuzzy set theory plays a key in formalizing uncertainties for medical diagnosis and prognosis (Zadeh 1965, Bezdek 1981, Adlassing 1986, Sterimann 1997, Kuncheva et al. 1999, Steinmann 2001). For the diagnosis and prognosis through the medical imaging, the supervised classification and unsupervised clustering are common pattern recognition techniques.

In the medical application domain, there are usually imprecise conditions and therefore fuzzy methods seem to be more suitable than crisp one. The major groups of fuzzy methods are represented by fuzzy clustering, fuzzy rule-based, fuzzy pattern matching methods and methods based on fuzzy relations. In fuzzy clustering, fuzzy c-means (FCM) clustering algorithms are the best known and most powerful methods used in cluster analysis in medical application. In the field of medical application, mining the medical images is the task of searching and retrieving the images and the use of pattern recognition methods for abstraction, indexing and retrieval of images is presented by (Anatani et al. 2002). Uncertainties also affect image analysis and the most challenging problem in image analysis and pattern recognition research is segmentation (Souza et al 2008, Hasanzadeh et al. 2008, Yang 2009). In medical image analysis, magnetic resonance (MR) image segmentation is the most popular imaging technique. The advantages of MRI are its high spatial resolution and soft-tissue contrast (Siyal et al. 2005, Bezdek et al. 1993) presents a review of MR image segmentation techniques using pattern recognition.

We discuss different types of fuzzy c means algorithms such as Improved Fuzzy C Means(IFCM), Possibilistic C Means(PCM), Fuzzy Possibilistic C Means(FPCM), Modified Fuzzy Possibilistic C Means(MFPCM) and Possibilistic Fuzzy C Means(PFCM)

II. IMPROVED FUZZY C MEANS ALGORITHM (IFCM)

The improved FCM algorithm is based on the concept of data compression where the dimensionality of the input is highly reduced. The data compression includes two steps: quantization and aggregation (D Jude Hemanth, 2009). The quantization of the feature space is performed by masking the lower 'm' bits of the feature value. The quantized output will result in the common intensity values for more than one feature vector. In the process of aggregation, feature vectors which share common intensity values are grouped together. A representative feature vector is chosen from each group and they are given as input for the conventional FCM algorithm. Once the clustering is complete, the representative feature vector membership values are
distributed identically to all members of the quantization level. Since the modified FCM algorithm uses a reduced dataset, the convergence rate is highly improved when compared with the conventional FCM.

The improved FCM algorithm uses the same steps of conventional FCM except for the change in the cluster updation and membership value updation criterions. The modified criterions are showed below

\[
\text{modified criterions are showed below}
\]

where

\[
x = \text{Reduced Dataset}
\]

III. POSSIBILISTIC C-MEANS (PCM)

To overcome difficulties of the FCM, (Krishnapuram and Keller, 1993) proposed a new clustering model named Possibilistic c-Means (PCM).

\[
r_i = \text{determines distance at which the membership value of a point in a cluster becomes 0.5.}
\]

\[
n_i = \frac{\text{data to a cluster must be equal to}}{\text{to avoid this problem, (Pal, 2005) etc propose a new algorithm called Possibilistic Fuzzy c-means (MFPCM) that integrates the features of both FCM and PCM.}}
\]

The advantage of PCM is clustering noisy data samples and disadvantage is very sensitive to good initialization.

IV. POSSIBILISTIC C-MEANS (PCM)

To overcome difficulties of the PCM, Pal defines a clustering technique that integrates the features of both Fuzzy a Possibilistic c-means called Fuzzy Possibilistic c-Means (FPCM). Membership and Typicality’s are very significant for the accurate characterization of data substructure in clustering difficulty (Vuda Sreenivasarao, 2010). An objective function in the FPCM depending on both membership and typicality’s are represented as: Memberships and typicalities is represented as:

\[
I_{\text{POCM}}(U, T, V) = \sum_{i=1}^{n} \sum_{j=1}^{c} (u_{ij}^{e} + t_{ij}^{e}) d^{u} \left( x_{j}, v_{i} \right)
\]

FPCM generates Memberships and possibilities at the same time, together with the usual point prototypes or cluster center for each cluster. The advantage of the FPCM is, it ignores the noise sensitivity deficiency of FCM and overcomes the coincident clusters problem of PCM. The disadvantage is the row sum constraints must be equal to one.

V. MODIFIED FUZZY POSSIBILISTIC C-MEANS ALGORITHM (MFPCM)

The objective function in (A Rajendran, 2011) is called the modified fuzzy possibilistic c-means (MFPCM) function composed of two expressions: the first is the fuzzy function and it uses a fuzziness weighting exponent, the second is possibilistic function and it uses a typical weighting exponent; but the two coefficients in the objective function are only used as exhibitor of membership and typicality (MN Ahmed, 2002). The objective function of the modified fuzzy possibilistic c-means (MFPCM) can be formulated as follows

\[
I_{\text{MFPCM}} = \sum_{i=1}^{n} \sum_{j=1}^{c} \left( u_{ij}^{e} + t_{ij}^{e} \right) d^{u} \left( x_{j}, v_{i} \right)
\]

Where \( U = \{ \{ u_{ij}^{e} \} \} \) represents a fuzzy partition matrix and is defined by:

\[
u_{ij}^{e} = \left( \frac{\sum_{k=1}^{c} \left( d^{2}(x_{j}, c_{i}) \right)^{2m/(m-1)} }{ \sum_{k=1}^{c} \left( d^{2}(x_{j}, c_{k}) \right)^{2m/(m-1)} } \right)^{1/(m-1)}
\]

Where \( U = \{ \{ u_{ij}^{e} \} \} \) represents a fuzzy partition matrix and is defined by:

\[
u_{ij}^{e} = \left( \frac{\sum_{k=1}^{c} \left( d^{2}(x_{j}, c_{i}) \right)^{2m/(m-1)} }{ \sum_{k=1}^{c} \left( d^{2}(x_{j}, c_{k}) \right)^{2m/(m-1)} } \right)^{1/(m-1)}
\]

represents c centers of the clusters, and is defined by:

\[
W_{ij}^{e} = \exp \left( - \frac{d^{2}(x, c_{i})^{2m/(m-1)}}{\sum_{k=1}^{c} d^{2}(x, c_{k})^{2m/(m-1)}} \right)
\]

The above equations indicate that membership is influenced by all C cluster centers, while possibility \( t_{ij}^{e} \) is influenced just by the i-th cluster center \( c_{i} \). The possibilistic term distributes the fit with respect to every n data points, but not by means of every C clusters. Thus, membership can be described as relative typicality, it determines the degree to which a data fit in to cluster in accordance with other clusters and is helpful in correctly labeling a data point. Possibility can be observed as absolute typicality, it determines the degree to which a data point belongs to a cluster correctly, and it can decrease the consequence of noise. Joining both membership and possibility can yield to good clustering result.

VI. POSSIBILISTIC FUZZY C MEANS ALGORITHM (PFCM)

In FPCM, the constraint corresponding to the sum of all typicality values of all data to a cluster must be equal to one cause problems particularly for a big data set. In order to avoid this problem (Pal, 2005) et al propose a new algorithm called Possibilistic Fuzzy c means algorithm (PFCM). The objective function is defined by

\[
I_{\text{PFCM}}(U, T, V, Z) = \sum_{i=1}^{n} \sum_{j=1}^{c} (u_{ij}^{e} + t_{ij}^{e}) d^{u} \left( x_{j}, v_{i} \right)
\]

Where

\[
W_{ij}^{e} = \exp \left( - \frac{d^{2}(x, c_{i})^{2m/(m-1)}}{\sum_{k=1}^{c} d^{2}(x, c_{k})^{2m/(m-1)}} \right)
\]
The advantages of the PFCM is, it ignores the noise sensitivity deficiency of FCM, overcomes the coincident clusters problem of PCM and eliminates the row sum constraints of FPCM.

VII. CONCLUSION

FCM algorithm is a distinctive and extensively used in engineering, scientific disciplines, medical image, pattern detection and bio-informatics. The different clustering techniques used for image segmentation based on fuzzy pattern recognition. Fuzzy c means provide better results than other clusters. FCM provides flexibility and change easily.

REFERENCES