TWO-PHASE C-MEANS CLUSTERING WITH NOISE REDUCTION USING FUZZY RULES

K. Kadambavanam¹, T. Senthilnathan²

¹Associate Professor and Head, Department of Mathematics, Sri. Vasavi College, Erode, Tamil Nadu, (India).
²Associate Professor, Department of Mathematics, Erode Arts and Science College, Erode, Tamil Nadu, (India).

ABSTRACT

The objective of this paper is to develop an effective fuzzy clustering method for segmentation of images. The conventional fuzzy c-means clustering with spatial feature (FCM_S) and its variations have their own limitations if the images are corrupted by heavy noise. If noise exists in the spatial neighbourhood information, then it affects the clustering. This article proposes a new mask to provide better spatial information to overcome the above problem in clustering, using conventional fuzzy c-means.

In the conventional defuzzification method used in fuzzy c-means clustering with spatial feature, there are two disadvantages. First, as the pixels’ partition membership values are used to calculate the pixels’ gray scale values, the noise less pixels’ gray scale values are changed by the inference calculations. So the peak signal to noise ratio between the output data and the original (noise less) data is very low. Second, as the output image loses its sharpness, the non-Euclidian structures are not revealed well. For better performance, a new defuzzification method is proposed in this article. It presents algorithms for the proposed new mask, clustering, and the defuzzification. The main properties of the proposed method are illustrated by using synthetic, real, and magnetic resonance (MR) images. A quantitative evaluation of this method is also presented.

Keywords: Defuzzification, Filter Window, Fuzzy C-Means Clustering, Grey Value, Spatial Feature

I INTRODUCTION

Clustering is the process of dividing data elements into classes or clusters so that items in the same class are as similar as possible. In clustering research, many methods are invoked. The applications are in various engineering and scientific branches like medical image segmentation (Yong, et al. [1]), remote sensing, marketing, analysing public opinion about an issue, elections etc. (Pham, et al. [2], Bezdek, et al. [3], Wells, et al. [4]). But most of the clustering methods are crisp in which a databelongs to only one cluster. But in real life many cases are related to the fact that a pattern (or data) often cannot be thought of as belonging to a single cluster only. So, a description in which the membership of a pattern is shared among clusters is necessary. To rectify the problem of being crisp (Pham, et al. [5]), the fuzzy set theory, proposed by Zadeh [6] gives an effective method of soft clustering. The fuzzy c-means (FCM) clustering algorithm was first introduced by Dunn [7] and later extended by Bezdek [8]. Chen et al. [9] improved the fuzzy c-means clustering by introducing kernel-induced metric, to reveal non-Euclidian structures in clustering. Ahmed et al. [10] incorporated spatial...
neighbourhood information into fuzzy c-means clustering to make it insensitive to noise in input image. But these methods are having some disadvantages, that noise in spatial feature creeps into miss classification of data. Noiseless data are changed unnecessarily by the inference calculations in defuzzification. It results in blurred output image. In order to rectify these drawbacks, this paper introduces (i) a new mask to filter the noise in spatial information, (ii) a new clustering method based on the standard FCM, and (iii) a new defuzzification method.

II SOME BASIC TERMINOLOGIES

2.1 Fuzzy c-means clustering (FCM)

Bezdek [8] introduced the concept of fuzzy partition in order to extend the notion of membership of data to clusters. The FCM algorithm identifies clusters as fuzzy sets, in which each datum is assigned to partition membership in each cluster. It attempts to partition a set of N data \( \{x_1, x_2, \ldots, x_N\} \) into C (\( 2 \leq C \leq N \)) fuzzy clusters based on some criterion. The algorithm returns a set of C cluster centers \( \{v_1, v_2, \ldots, v_c\} \) and a partition matrix \( U = [u_{ij}] \) when \( u_{ij} \in [0,1] \), \( i = 1, 2, \ldots, C \), \( j = 1, 2, \ldots, N \). The element \( u_{ij} \) specifies the degree to which element \( x_j \) belongs to the cluster \( c_i \).

A mathematical structure for the problem is

\[
\begin{align*}
\text{minimize} & \quad J_u = \sum_{i=1}^{C} \left[ \sum_{j=1}^{N} u_{ij}^{m} \|x_j - v_i\|^2 \right]^{\frac{1}{m-1}} \\
\text{subject to} & \quad \sum_{i=1}^{C} u_{ij} = 1 \quad \text{for } j = 1, 2, \ldots, N, \\
& \quad 0 < u_{ij} < 1 \quad \text{for } i = 1, 2, \ldots, C, \ j = 1, 2, \ldots, N, \\
\text{where} \quad & m - \text{fuzziness of the resulting partition}, (1 \leq m \leq \infty), \\
& \|x_j - v_i\| - \text{the difference between } x_j \text{ and } v_i, (x_j, v_i \in \mathbb{R}^p), \\
& u_{ij} - \text{partition membership of } x_j \text{ in the cluster } i, \\
& v_i - \text{prototype or centroids of the cluster } i, \\
& N - \text{number of data}, \\
& C - \text{number of clusters}.
\end{align*}
\]

In image clustering, the gray scale value of the image pixels is used as feature. Thus the nature of the objective function is of minimization type. The high membership values are assigned to the pixels when gray scale values are close to the centroid of their clusters. The low membership values are assigned to the pixels when the gray scale value is far from the centroid. In FCM algorithm, the degree of membership depends on the distance between the pixel and each individual cluster center. The necessary conditions on \( u_{ij} \) and \( v_i \) to minimize the objective function (1) are derived by Bezdek [8] as follows:

\[
u_{ij} = \frac{\left(\frac{\|x_j - v_i\|^2}{\sum_{k=1}^{C}[\|x_j - v_k\|^2]^{m}}\right)^{\frac{1}{m-1}}}{\sum_{i=1}^{C}\left(\frac{\|x_j - v_i\|^2}{\sum_{k=1}^{C}[\|x_j - v_k\|^2]^{m}}\right)^{\frac{1}{m-1}}} \tag{4}
\]
and 
\[ v_i = \frac{\sum_{j=1}^{N_j} w_{ij} x_j}{\sum_{j=1}^{N_j} w_{ij}} \] 
for \( i = 1, 2, \ldots C \).

Using iteration technique, the partition memberships and cluster centres are updated to optimize the objective function. Starting with an initial guess for each cluster centre, the iteration processes will be terminated when
\[ |J_n^{k+1} - J_n^k| < \epsilon \]

where \( 0 < \epsilon < 1 \) represents the precession of accuracy and \( k \) represents the iteration number.

### 2.2 Introduction to Fuzzy c-means clustering with Spatial Features (FCM_S)

One of the important characteristics of an image is that neighbouring pixels are highly correlated, i.e. the pixels in the immediate neighbourhood possess nearly the same feature data. The probability that they belong to the same cluster is great. So, the spatial relationship of neighbouring pixels is an important characteristic that can be of great aid in image clustering. Utilizing this characteristic in FCM, fuzzy c-means clustering with Spatial Features (FCM_S) was developed (Ahmed, et al. [10]). In this method, the spatial information (which is formed by using the distribution statistics of the neighbouring pixels) and the prior probability are used to form a new membership function for clustering. So in the mathematical structure of the objective function of FCM given in equation (1) is modified as given below:

\[ J_n = \sum_{i=1}^{C} \sum_{j=1}^{N_j} u_{ij}^m \| x_j - v_i \|^2 + \mu \frac{\sum_{i=1}^{C} \sum_{j=1}^{N_j} u_{ij}^m \| x_j - v_i \|^2}{N_j} \| x_j - v_i \|^2 \]

where \( \alpha \) stands for the controlling parameter of the neighbourhood feature, \( N_j \) - the cardinality of the neighbourhood feature, \( N_j \) - the set of neighbours of \( x_j \), \( x_j \) - the neighbouring data point around \( x_j \) falling in the window, with centre \( x_j \).

As in the standard FCM algorithm, the objective is to minimize \( J_n \) subject to the constraints on \( u_{ij} \) as in the equations (2) and (3). Taking the first derivatives of \( J_n \) with respect to \( u_{ij} \) and \( v_i \) and zeroing them, respectively, two necessary but not sufficient conditions for \( J_n \) to be at its local extrema are obtained as follows:

\[ u_{ij} = \left( \frac{\| x_j - v_i \|^2 + \frac{\alpha}{N_j} \sum_{i=1}^{C} \sum_{j=1}^{N_j} \| x_j - v_i \|^2}{\sum_{l=1}^{N_j} \left( \| x_j - v_l \|^2 + \frac{N_l \alpha}{N_j} \sum_{i=1}^{C} \sum_{j=1}^{N_j} \| x_j - v_l \|^2 \right)^{-1}} \right)^{-1} \]

and

\[ v_i = \frac{\sum_{j=1}^{N_j} w_{ij} x_j}{\sum_{j=1}^{N_j} w_{ij}} \]

### 2.3 Variations of FCM_S

As FCM_S takes much time for computation, to reduce it, Chen, et al.[9] replaced the set of neighbourhood pixels by its mean or median in the equations (7). The objective function is modified as

\[ J_n = \sum_{i=1}^{C} \sum_{j=1}^{N_j} u_{ij}^m \| x_j - v_i \|^2 + \mu \sum_{i=1}^{C} \sum_{j=1}^{N_j} u_{ij}^m \| x_j - v_i \|^2 \]

where \( \overline{x}_j \) represents the mean in FCM_S1 and median in FCM_S2 respectively.
By an optimization way similar to the standard FCM, $J_s$ is minimized under the constraint of $U$ and $V$ same as in equations (4) and (5) are derived as follows:

$$U_{ij} = \frac{\sum_{j=1}^{N} \| x_j - v_i \|^2 + a \| x_j - v_i \|^2 \cdot \alpha}{\sum_{j=1}^{N} \| x_j - v_i \|^2 + \alpha \| x_j - v_i \|^2 \cdot \alpha}$$

and

$$V_{i} = \frac{\sum_{j=1}^{N} u_{ij} \| x_j + \alpha x_j \|}{(1 + \alpha) \sum_{j=1}^{N} u_{ij} \| x_j \|}$$

### 2.4 Introduction to kernelized fuzzy c-means clustering (KFCM)

Euclidean distance ($\ell_2$-norm) is used in FCM, FCM_S and in its variations. So they are not efficient to reveal non-Euclidean structure of the input data. To overcome this disadvantage, a kernelized version of fuzzy-clustering method has been introduced by Chen, et al.[9]. The basic idea of kernelizing is, first transforming the low-dimensional inner product input space into a higher dimensional feature space through some nonlinear mapping. Computing a linear partitioning in this feature space results in a nonlinear partitioning in the input space (Chen, et al.[9]). By this technique nonlinear structures in input data are preserved after clustering.

Chen, et al.[9] derived a nonlinear transformation $\Phi: \mathbb{R}^d \rightarrow \mathbb{R}^k$ where $d \leq k$ and the metric (norm) is expressed by an inner product space as

$$\| \Phi(v_i) - \Phi(x_j) \|^2 = \langle \Phi(v_i), \Phi(x_j) \rangle + \langle \Phi(x_j), \Phi(x_j) \rangle - 2 \langle \Phi(v_i), \Phi(x_j) \rangle.$$

A kernel is defined as

$$K(v, x) = \langle \tau(v_i), \tau(x_j) \rangle = \exp \left(-\frac{1}{\sigma^2} \| v_i - x_j \|^2 \right)$$

where $a \geq 0, 1 \leq b \leq 2$.

By introducing the kernel, the complexity of dimensions in the calculation of the inner products in $\mathbb{R}^k$ is overcome.

So, $$\| \Phi(v_i) - \Phi(x_j) \|^2 = K(v_i, v_i) + K(x_j, x_j) - 2K(v_i, x_j)$$ where $i = 1, 2, \ldots C$ and $j = 1, 2, \ldots N$.

As $K(x_i, x_i) = 1$ for any $x_i$, the above equation can be reformed as

$$\| \Phi(v_i) - \Phi(x_j) \|^2 = 2[I - K(v_i, x_j)]$$

(14)

By using the mapping $\Phi$, the objective function is re-written as follows:

$$J^k_s = \sum_{i=1}^{C} \sum_{j=1}^{N} u_{ij} \| \Phi(v_i) - \Phi(x_j) \|^2$$

By introducing this kernel (14), the objective function is modified as,

$$J^k_s = \sum_{i=1}^{C} \sum_{j=1}^{N} u_{ij} [I - K(v_i, x_j)]$$

(15)

The necessary constraints on $u_{ij}$ and $v_i$ for the equation (15) to obtain its local minimum are derived by Chen, et al.[9] as follows.

$$u_{ij} = \frac{[I - K(v_i, x_j)] \cdot \alpha \cdot \alpha}{\sum_{i=1}^{C} [I - K(v_i, x_j)] \cdot \alpha \cdot \alpha}$$

(16)
\[
\nu_j = \frac{\sum_{i=1}^{N} \tilde{u}_{ij} f([v_j, x_j])}{\sum_{j=1}^{N} \tilde{u}_{ij} f([v_j, x_j])}
\]

(17)

2.5 KFCM with spatial feature

Analogous to the work in FCM_S, the spatial features are utilized in the KFCM and derived KFCM_S1, KFCM_S2 by Chen, et al. [9].

The objective function of KFCM with spatial feature (KFCM_S) is defined by Chen, et al. [9] as,

\[
J_{m}^{\Phi} = \sum_{i=1}^{c} \sum_{j=1}^{N} \tilde{u}_{ij} f([I - K(v, x)]) + \left(\frac{\alpha}{\sum_{j=1}^{N} u_{ij} f([v_j, x_j])}\right) \sum_{i=1}^{c} \sum_{j=1}^{N} u_{ij} f([v_j, x_j]) \sum_{k \in V_{i,j}} [I - K(v_k, x_k)]
\]

(18)

In an analogous way the two necessary conditions on \( u_{ij} \) and \( v_j \) for the equation (18) to obtain its local minimum are derived by Chen, et al. [9]. They are

\[
u_j = \frac{\left[ I - K(v_j, x_j) \right] + \left(\frac{\alpha}{\sum_{i=1}^{N} u_{ij}}\right) \sum_{i=1}^{c} \sum_{j=1}^{N} u_{ij} f([v_j, x_j]) \sum_{k \in V_{i,j}} [I - K(v_k, x_k)]}{\sum_{j=1}^{N} \tilde{u}_{ij} f([v_j, x_j])}
\]

(19)

and

\[
u_j = \frac{\sum_{i=1}^{c} \sum_{j=1}^{N} u_{ij} f([v_j, x_j]) + \left(\frac{\alpha}{\sum_{i=1}^{N} u_{ij}}\right) \sum_{i=1}^{c} \sum_{j=1}^{N} u_{ij} f([v_j, x_j]) \sum_{k \in V_{i,j}} [K(v_k, x_k)]}{\sum_{j=1}^{N} \tilde{u}_{ij} f([v_j, x_j])}
\]

(20)

2.6 The variations KFCM_S1 and KFCM_S2

To reduce the computational time of KFCM_S, Chen, et al. [9] replaced the set of neighboring pixels in equation (18) by its mean or median. The modified objective function is as follows:

\[
J_{m}^{\Phi} = \sum_{i=1}^{c} \sum_{j=1}^{N} \tilde{u}_{ij} f([I - K(v, x)]) + \alpha \sum_{i=1}^{c} \sum_{j=1}^{N} u_{ij} f([I - K(v, x)])
\]

(21)

Fuzzy partitioning is carried out through an iterative optimization of the \( J_{m}^{\Phi} \) with the update of membership \( u_{ij} \) and the cluster centres \( v_j \) by

\[
u_j = \frac{\left[ I - K(v_j, x_j) \right] + \alpha \left[ I - K(v_j, x_j) \right]}{\sum_{i=1}^{c} \sum_{j=1}^{N} u_{ij} f([v_j, x_j])}
\]

(22)

and

\[
u_j = \frac{\sum_{i=1}^{c} \sum_{j=1}^{N} u_{ij} f([v_j, x_j]) + \alpha \sum_{i=1}^{c} \sum_{j=1}^{N} u_{ij} f([v_j, x_j]) \sum_{k \in V_{i,j}} [K(v_k, x_k)]}{\sum_{i=1}^{c} \sum_{j=1}^{N} u_{ij} f([v_j, x_j])}
\]

(23)

where \( \bar{x}_j \) is the mean and median of the neighboring pixels within the neighboring window around \( x_j \) in KFCM_S1 and KFCM_S2 respectively.

2.7 Some of the disadvantages in using spatial feature

There are two disadvantages in using spatial feature in clustering. They are

(a) If noise exists in the spatial information, then it will affect the calculations.
Genuine pixels’ gray scale values are changed unnecessarily by adding the spatial feature. To eliminate the above disadvantages, this article proposes (i) a new mask for spatial feature, (ii) a new “two-phase kernelized fuzzy c-means (TKFCM) algorithm”(iii) a new set of fuzzy rules in defuzzification.

The rest of this paper is organized as follows. In section 3, a “dynamic threshold” is suggested to reduce the noise in the filter widow. In section 4 the new filter window is used as spatial feature in clustering. In section 5 a “two-phase kernelized fuzzy c-means clustering” and new fuzzy rules for defuzzification are introduced. In section 6 the relationship between the methods are discussed. The experimental results are presented in section 7. Section 8 explains the advantages of the above proposed algorithms.

### III DYNAMIC SIZED SPATIAL FEATURE WINDOW

For spatial feature in FCM_S, FCM_S_1 and FCM_S_2, a window of neighbouring pixels (generally of size 3x3 or 5x5) around each pixel is used (Ahmed, et al. [10]). When any of these neighbouring pixels have noise, it will affect the clustering calculations.

When α- trimmed neighbouring data (Bansal, et al. [11], Jampour, et al. [12] and Taguchi, et al. [13]) are used in the second term of the equations (7) - (12), the sorted neighbouring data are trimmed on both ends equally. Alkhazaleh et al. [14] specified that when trimming the neighbouring pixels, if the distribution of neighbouring data is not smooth and having skewness on one side, then the genuine data (without noise) will be trimmed on one end and noisy data will be included on the other end.

Adaptive threshold (dynamic threshold) has been used in recent research works on denoising. (Aborisade[15], Huang et al. [16], Sadeghipour[17], Singh[18], Sun et al. [19], Yuan et al. [20]). But in these articles the entire process is done in wavelet format. Moreover these articles are considering the noise in the central pixel of the spatial window but not in neighbourhood of a pixel. To eliminate the time taken to convert the image to wavelet format, and to eliminate the noise in the spatial information of each pixel this article proposes a new dynamic threshold which trims the spatial information of each pixel reducing noise. The procedure is explained below:

First, for each pixel value \( x_j \), trimmed mean filter value \( \tilde{N}_T(x_j) \) is calculated using the equation

\[
\tilde{N}_T(x_j) = \left[ \frac{1}{N(x_j)} \sum_{k=1}^{\left\lfloor \frac{N(x_j)}{2} \right\rfloor} x_k \right]
\]

where,

\[
T_j < \frac{\left\lfloor \frac{N(x_j)}{2} \right\rfloor}{2}
\]

is an integer representing the “trimming size”

\( x_k \in \{N(x_j)\} \) is the sorted neighboring pixel of \( x_j \).

Second, the mean deviation of \( \tilde{N}_T(x_j) \) is calculated. It is denoted as \( \sigma_j \). It is calculated by using the equation

\[
\sigma_j = \frac{1}{\left\lfloor \frac{N(x_j)}{2} \right\rfloor} \sum_{x_k \in \tilde{N}_T(x_j)} \left\| x_k - \tilde{N}_T(x_j) \right\|
\]

Next by using \( \sigma_j \) as the dynamic threshold, noise less neighboring pixels are selected by using the equation

\[
N_j(x_i) = \left\{ x_i \middle| \left\| x_i - \tilde{N}_T(x_j) \right\| \leq \beta \sigma_j \text{ where } x_k \in N(x_j) \right\}
\]

\[
(25)
\]
This set \( \{N_{s}(x_j)\} \) is proposed to use as spatial information in clustering. The main concept of this process is, if a pixel, either in centre or in the neighbourhood of the filter window, has noise, it will lie in the region of rejection. The quantity \( \beta \) multiplied with \( \sigma_j \) together with \( N_{s}(x_j) \) determines the region of acceptance. The spatial information of the neighbouring points \( \{N_{s}(x_j)\} \) which are lying within the region of acceptance are taken into account of clustering calculations. The set of points in \( \{N_{s}(x_j)\} \) is called “selected neighbouring pixels”. When the median of \( N_{s}(x_j) \) is used as spatial information in clustering, it serves better than median filter value used in Chen, et al.[9].

**IV HYBRID METHOD OF KFCM WITH SELECTED NEIGHBOURHOOD (KFCM_SS2)**

4.1 Mathematical model

This method of clustering is based on the Gaussian kernel function. The normed kernel function and the standard objective function of KFCM_SS2 given by Chen, et al.[9] are utilized to build the new KFCM_SS2. By introducing the new filter window, the objective function given in the equation (21) is modified as

\[
J_{KFCM_SS2} = \sum_{i=1}^{C} \sum_{j=1}^{N} u_{ij}^p \left[ I - K(v_i, x_j) \right] + \alpha \sum_{i=1}^{C} \sum_{j=1}^{N} u_{ij}^p \left[ I - K(v_i, \text{median}_{N_{s}(x_j)}) \right]
\]  

(26)

Subject to,

\[
0 \leq u_{ij} \leq \text{for} \quad i = 1, 2, \ldots, C \quad \text{and} \quad j = 1, 2, \ldots, N
\]

\[
\sum_{j=1}^{N} u_{ij} = \text{for} \quad i = 1, 2, \ldots, N
\]

where \( \text{median}_{N_{s}(x_j)} \) refers the median of \( N_{s}(x_j) \)

An iterative algorithm for minimizing equation (26) with respect to \( u_{ij} \) and \( v_i \) is derived, as given in (27-28):

\[
u_{ij} = \frac{1}{\sum_{j=1}^{N} [K(v_i, x_j) x_j + \alpha \sum_{j=1}^{N} [K(v_i, \text{median}_{N_{s}(x_j)}) x_j]}\]

(27)

\[
u_i = \frac{\sum_{j=1}^{N} u_{ij}^p \left[ K(v_i, x_j) x_j + \alpha \sum_{j=1}^{N} [K(v_i, \text{median}_{N_{s}(x_j)}) x_j]}{\sum_{j=1}^{N} u_{ij}^p \left[ K(v_i, x_j) x_j + \alpha \sum_{j=1}^{N} [K(v_i, \text{median}_{N_{s}(x_j)}) x_j]}}
\]

for \( i = 1, 2, \ldots, C \) and \( j = 1, 2, \ldots, N \).  

(28)

4.2 Defuzzification

The output data values are approximated using fuzzyinference formula given below

\[
Z = \left\{ z_j \mid z_j = \frac{\sum_{i=1}^{C} v_{ij}}{\sum_{j=1}^{N} u_{ij}^p} \right\}
\]

(29)

where \( j = 1, 2, \ldots, N \) and \( V = \{v_1, v_2, \ldots, v_C\} \).
V TWO-PHASE KERNELIZED FUZZY C-MEANS CLUSTERING WITH SPATIAL FEATURE (TKFCM_Ss2)

The proposed method consists of two phases. In phase I, the input image is fuzzified by KFCM. A partition membership matrix \( U = \{u_{ij}\} \) of order \( C \times N \) and a set of cluster centers \( V = \{v_j\} \) are calculated. In phase II, a new partition membership matrix \( U' = \{u'_{ij}\} \) of order \( C \times N \) and a set of cluster centers \( V' = \{v'_j\} \) are introduced. \( U' \) and \( V' \) are initialized by the final iteration values of \( U \) and \( V \) obtained in phase I. The image is further fuzzified by KFCM_Ss2 and the values of \( U' \), \( V' \) are calculated by iteration method proposed in section 4.

5.1 Defuzzification

During defuzzification process in KFCM_Ss2, the gray scale values of pixels are calculated by fuzzy inference method. This set of gray scale values are an approximation to the original values. There are many defuzzification techniques used in practice (Chaudhuri, et al.[21], Roventa[22], Nejad, et al.[23], Udupa, et al.[24], Sladoje, et al.[25], Lowen, et al.[26] and Leekwijck, et al.[27]), and different rules are applied to find a suitable crisp representation for the fuzzy set.

5.2 Limitations of existing defuzzification methods

In the defuzzification methods referred above, the noise less pixels’ gray scale values are unnecessarily changed by inference calculations. The gray scale values in an original crisp image usually match with some physical units in the real world (such as Hounsfield units in computed tomography) or they are relative to some known quantity (e.g., giving the value 0 to water in MRI). Thus, to obtain a better approximation for the original crisp set, defuzzification plays a vital role in fuzzy clustering.

To overcome this problem, in this article, a new set of fuzzy rules are proposed for defuzzification. The fact behind this process is that KFCM is more sensitive to noise than KFCM_Ss2. It reflects in the corresponding cluster assignment, partition membership values of each pixel (clustered by KFCM and KFCM_Ss2). The noise pixels can be identified by comparing the outcomes (i.e., cluster assignment, partition membership and gray scale value calculated by inference method) of KFCM and KFCM_Ss2. The difference in the above said outcomes of KFCM and KFCM_Ss2 of a noisy pixel will be higher than that of its neighboring noiseless pixels. Based on this, the following fuzzy rules are introduced.

If any pixel \( x_j \) satisfies at least any one of the following three conditions (i), (ii), and (iii) given in section 5.3 is identified as a noisy pixel, and its fuzzy values are mapped to the spatial feature \( N_{\text{f}}(x_j) \).

5.3 Fuzzy rules to identify the noisy pixels

(i) If the pixel assigned to one cluster by KFCM is assigned to some other cluster by KFCM_Ss2 then the pixel is a noisy pixel.

(ii) If \( \|u_{ij} - u'_{ij}\| > \|N(u_{ij}) - N(u'_{ij})\| \) then \( x_j \) is a noisy pixel, where

\[ N(u_{ij}) \] is the partition membership value of neighbouring pixels of \( x_j \) in cluster \( i \) in KFCM.

www.ijarse.com
$N(u'_{ij})$ is the partition membership value of neighbouring pixels of $x_j$ in cluster $i$ in KFCM_Ss.

$\|N(u_{ij}) - N(u'_{ij})\|$ is the mean of the differences between the partition memberships of neighbouring pixels in KFCM and in KFCM_Ss.

(iii) if $\|x_j - z_j\| > \|N(x_j) - N(z_j)\|$ then $x_j$ is a noisy pixel,

where $\|N(x_j) - N(z_j)\|$ is the mean of the differences between the neighbouring pixel values in input image and their corresponding inference values in KFCM_Ss.

5.4 Proposed Fuzzy rules for defuzzification

The new set of fuzzy rules induces two mappings, one from the fuzzy set to the crisp input set and another from the fuzzy set to the crisp spatial feature set. If the pixel is a noisy pixel, the new set of fuzzy rules constructs an optimal mapping from the fuzzy set to the crisp spatial feature set. Otherwise it makes a map from the fuzzy set to the crisp input set. By this proposed method the quantitative features of the pixel value are preserved as in the original noiseless image. In other words, the Peak Signal to Noise Ratio (PSNR) is preserved, and non-Euclidian structures are revealed well when clustering an image. The experiments on synthetic, real and on MR images illustrate that TFCM_Ss is more efficient than the existing fuzzy c-means clustering methods.

5.5 Algorithm for Two-phase Kernelized Fuzzy C-means (TKFCM_Ss)

Phase I

Step 1: Get the data from the image.
Step 2: Fix the number of clusters.
Step 3: Initialize the cluster centres $V = \{v_j\}_j$ using random numbers.
Step 4: Initialize the partition membership matrix $U = [u_{ij}]$ where $i = 1, 2 \ldots C$, $j = 1, 2 \ldots N$ by random numbers satisfying the equations (2) and (3).
Step 5: Fix the precession value $\epsilon$, ($0 < \epsilon < 1$).
Step 6: Update the partition membership matrix $U = [u_{ij}]$ using the equation (16).
Step 7: Update the prototype centres $V = [v_j]$ using the equation (17).
Step 8: Calculate the value of the objective function $J_m$ using the equation (15).
Step 9: Repeat steps from (5) to (7) until $|J_m^{k+1} - J_m^{k}| < \epsilon$ is satisfied, where $J_m^{k}$ and $J_m^{k+1}$ are the values of the objective function obtained in the $k^{th}$ and $(k+1)^{th}$ iterations respectively.

Phase II

Step 1: Initialize the new partition membership matrix $U' = [u'_{ij}]$ and the cluster centre $V' = [v'_{j}]$ of phase II with the final values of $U$ and $V$ respectively obtained from the step 9 in phase I.
Step 2: Initialize the controlling parameter $\alpha$ ($0 < \alpha < \infty$) in the neighbourhood feature.
Step 3: Fix a (3x3) neighbourhood window on each pixel $x_j$ ($j = 1, 2 \ldots N$) with $x_j$ as the centre of the window.
Step 4: Update the partition membership matrix $\mathbf{U}$ by giving modification (i.e. replacing $u_{ij}$ by $u'_{ij}$) in equation (27).

Step 5: Update the prototype $\mathbf{v}'$ by giving modification (i.e. replacing $v_{j}$ by $v'_{j}$) in equation (28).

Step 6: Calculate the value of the objective function $J_{m}^{(k)}$ using the equation (26).

Step 7: Repeat steps from (12) to (14) until $|J_{m}^{(k+1)} - J_{m}^{(k)}| < \epsilon$ is reached, where $J_{m}^{(k)}$ and $J_{m}^{(k+1)}$ are the values of the objective function obtained in $k^{th}$ and $(k+1)^{th}$ iterations respectively.

5.6 Algorithm for defuzzification using the proposed fuzzy rules

Step 1: Create two binary matrices $B$ and $B'$ of same order $(C \times N)$, with

$B_{ij} = \begin{cases} 1 & \text{if } x_{j} \text{ has maximum membership in } \text{ith cluster of } \mathbf{U}, \\ 0 & \text{otherwise} \end{cases}$

and

$B'_{ij} = \begin{cases} 1 & \text{if } x'_{j} \text{ has maximum membership in } \text{ith cluster of } \mathbf{U}', \\ 0 & \text{otherwise} \end{cases}$

Step 2: Calculate the set $Z$ of inference values by using the equation (29)

$Z = \left\{ z_{j} \mid z_{j} = \frac{J_{m}^{(k)}(v_{j})^{p}}{\sum_{i=1}^{C} (B_{ij})^{p}} \right\}$

where $j = 1, 2, N$ and $V' = \{ v_{1}', v_{2}', \ldots, v_{N}' \}$ obtained from the modified equation used in step (5) of phase II in section (5.5).

Step 3: Construct the new crisp set $\{y_{j}\}$ as

$y_{j} = \begin{cases} x_{j} & \text{if } B_{ij} \text{ and } B'_{ij} \text{ are equal,} \\ \text{otherwise} & \text{for } i = 1, 2, \ldots, C \text{ and } j = 1, 2, \ldots, N. \end{cases}$

where $N_{a}(x_{j})$ is the median of the selected neighbourhood pixels of $x_{j}$.

Step 4: Convert $\mathbf{U}$ and $\mathbf{U}'$ into three dimensional matrices of order $(C, H, W)$.

Convert $\{y_{j}\}$ and $\{z_{j}\}$ into $(H, W)$ matrix, where $H, W$ is the height and width of the input image.

Step 5: If $|u_{i,k,q} - u'_{i,k,q}| \geq |u_{i,k,q} - u_{i,k,q}^{*}|$ then $y_{k,q} = \overline{N_{a}(x_{k,q})}$

where $|u_{i,k,q} - u'_{i,k,q}|$ is the mean of the set of absolute differences$\{|N(u_{i,k,q}) - N(u'_{i,k,q})|\}$

$i = 1, 2, \ldots, C, k = 1, 2, \ldots, H, q = 1, 2, \ldots, W$

Step 6: Fix $N(z_{k,q})$ and $N(y_{k,q})$ for $k = 1, 2 \ldots, H, q = 1, 2 \ldots, W$

Step 7: Calculate $|N(z_{k,q}) - N(y_{k,q})|$

If $|z_{k,q} - y_{k,q}| \geq |N(z_{k,q}) - N(y_{k,q})|$ then $y_{k,q} = \overline{z_{k,q}}$

where $|N(z_{k,q}) - N(y_{k,q})|$ is the mean of the set of absolute differences$\{|N(z_{k,q}) - N(y_{k,q})|\}$

Step 8: Assign $y_{k,q}$ to the cluster in which $x_{k,q}$ has maximum membership value in $\mathbf{U}'$ for $k = 1, 2 \ldots, H, q = 1, 2 \ldots, W$. 

www.ijarse.com
VI RELATIONSHIP AMONG THE VARIOUS METHODS.

Two phase fuzzy c-means clustering algorithm can be considered as a general framework. Besides TKFCM_Ss, some other typical clustering algorithms for image clustering can be derived from the framework as follows:

1) By setting $N_s(x_j)$ as the mean of $N_s(x_j)$ in equations (26), (27) and (28), the KFCM_Ss reduces to KFCM_Ss.

Similarly, TKFCM_Ss reduces into TKFCM_Ss.

2) By replacing $N_s(x_j)$ with $N_s(x_j)$ in equations (19), (20), and (21) and replacing the equations (26), (27), (28) with the modified form of equations (19), (20), and (21), KFCM_Ss reduces to KFCM_Ss. In an analogous way TKFCM_Ss reduces to TKFCM_Ss.

3) When $\beta \to \infty$ in equation (24), $N_s(x_j) \to N(x_j)$. Then KFCM_Ss reduces into KFCM_Ss and TKFCM_Ss becomes TKFCM_S which is a special case of two-phase clustering.

4) By setting the trimming size $T_x=0$, in equation (24), KFCM_Ss reduces to KFCM_Ss and TKFCM_Ss reduces to TKFCM_Ss.

VII EXPERIMENT RESULTS AND ANALYSIS

This section compares the efficiency of the proposed algorithms TKFCM_Ss and KFCM_Ss, with KFCM_Ss (with $\alpha$-trimmed mean), and KFCM_Ss on synthetic, real and simulated MR images. The results are graded by measuring the Segmentation Accuracy (S.A) and Peak Signal to Noise Ratio (PSNR), and the number of misclassified pixels. As FCM, FCM_Ss, FCM_Ss, KFCM_Ss are inferior to KFCM_Ss (with $\alpha$-trimmed mean), and KFCM_Ss they are not used for comparison. To provide the best methods only, here the new variations KFCM_Ss, KFCM_Ss, TKFCM_Ss, and TKFCM_Ss are not used for comparison as they are performing inferior to TKFCM_Ss and TKFCM_Ss.

7.1 Segmentation Accuracy (S.A.)

The Segmentation Accuracy is calculated as $S_A = \frac{A_i \cap A_{refj}}{A_i \cup A_{refj}}$

where $A_i$ refers the set of pixels of the $i^{th}$ cluster found by the $i^{th}$ algorithm, and $A_{refj}$ represents the set of pixels of the $j^{th}$ cluster in the reference (original) segmented image.

7.2 Peak Signal to Noise Ratio (PSNR)

When spatial feature is used in clustering, even though the S.A is increased, the output clusters are blurred. So S.A is not the only scale to measure the quality of clustering. PSNR is an approximation to human perception of reconstruction quality. PSNR is generally used to measure the quality of reconstruction from the noisy data. Higher PSNR indicates that the reconstruction is of higher quality. It is calculated in decibels (dB) and is defined via the mean squared error (MSE) as follows:

If $X$ is the noise-free ($HxS$) monochrome image, and $Y$ is the output from the noise induced version then, the MSE is given by the equation
\[ MSE(x,y) = \frac{1}{T \times S} \sum_{i=1}^{S} \sum_{j=1}^{T} |x_{ij} - y_{ij}|^2 \]

and the PSNR is calculated as

\[ PSNR = 20 \log_{10} \left( \frac{\frac{B}{2}-1}{MSE} \right) \]

where \( B \) represents the bits per sample pixel. In 8-bit synthetic and real images \( B = 8 \), in dicom format MRI, it is 16. In the present experiments the PSNR value is calculated for the sum of the clusters.

### 7.3 Parameter settings

In fact, all these algorithms have some crucial parameters needed to be adjusted for clustering and the parameters are noise dependent. Therefore their selection will trivially influence the clustering results. This section focuses on discussion on the parameter setting.

In all these experiments the parameters are set as follows: \( m = 2, \epsilon = 0.00001 \). The algorithms are tested on images corrupted by “Gaussian”, “Salt & Pepper” and “Mixed” noises respectively.

### 7.4. Results on synthetic image

A synthetic image of size [128x128] pixels includes two classes of grey values {0, 90} is used for this experiments.

**7.4.1 The first experiment** is conducted to test the effect of \( \beta \) in PSNR of output clusters. In this experiment \( \beta \) is assigned values 0.75 through 10 in steps of 0.25. Fig. (1a) and (1b) show the PSNR of results of KFCM_Ss, and TKFCM_Ss2 varying with the parameter \( \beta \) on the synthetic image corrupted by Gaussian” and “Salt & Pepper” noise respectively. Similar result is obtained in the case of mixed noise. As there is no significant change in the clustering performance after \( \beta = 3.5 \) the value of \( \beta \) is set as 3.5 in equation (24) for the second and third experiments.

**7.4.2 The second experiment** is conducted on synthetic image to test the effect of various values for the parameter \( \alpha \) in segmentation performance. In this experiment \( \alpha \) is assigned values 0 through 10 in steps of 0.25. The results of the proposed methods are compared with the existing methods KFCM_Ss, and KFCM_Ss2. The result is presented in Fig. (1c) and in Fig.(1d). The results show that the segmentation performance is varying with the value of \( \alpha \). As there is no significant change in the clustering performance after \( \alpha = 3.8 \) the value of \( \alpha \) is set as 3.8 for third experiments. In the first and second experiments, the noise level is fixed as 10%.

**7.4.3. The third experiment** is conducted on synthetic image to compare the PSNR together with the classification error in the results obtained by KFCM_Ss, KFCM_Ss2, and TKFCM_Ss2. The value of \( \epsilon \) is fixed as 3.8 (the optimal value used in Chen, et al.[9]) and C as 2. The noise level is set at various levels ranging from 3% through 15%. “Gaussian”, “Salt & Pepper” and “Mixed noises” are used. Simulation of noise was performed by 100 independent runs on each level and each type of noise. The simulated different structures of noise in synthetic image are tested by KFCM_Ss, KFCM_Ss2, and TKFCM_Ss2. The average segmentation accuracy and average PSNR% are
presented in Table (1). The graphical representations of the results are given in Fig.(1.e) and Fig.(1.f). From the results it is observed that as noise level increases the clustering performance is decreasing. But in all the cases KFCM_Ss2 and TKFCM_Ss2 is performing better than KFCM_Ss (with α-trimmed mean) and KFCM_Ss. The output images are presented in Fig. (2).
Fig (2). Comparison of segmentation results of synthetic image.
(a) Original image, (b) image with ‘Salt & Pepper’ noise, (c) KFCM_S2 result, (d) KFMC_S1 (with Trimmed mean) result, (e) KFCM_Ss2 result, (f) TKFCM_Ss2 result.
7.5. Experiment on real image

To examine the robustness of the algorithms, the real image “eight” of size 242 x 308 with gray values 0 to 255 corrupted simultaneously by Gaussian white noise N(0,180) with unit dispersion, and salt & pepper noise. There are three types of experiments conducted on the experimental object.

7.5.1 The first experiment is conducted on real image to test the effect of various values for the parameter $\beta$ in segmentation performance. The graphical representation of the results are presented in Fig.(3.a) show that the PSNR of the output of KFCM_Ss, and TKFCM_Ss is varying with the parameter $\beta$. As there is no significant change in the clustering performance after $\beta = 3.5$ the value of $\beta$ is set as 3.5 in equation (24) for the second and third experiments.

7.5.2 The second experiment is conducted on real image to test the effect of various values for the trimming size ($T_s$) in segmentation performance. As a 3x3 window is used for spatial feature, the trim size $T_s$ in equation (24) can be assigned as $T_s \in \{0,1,2,3,4\}$. The graphical representation of the segmentation performance is presented in Fig.(3b). From the results, it is observed that when $T_s = 3$, the algorithms KFCM_Ss and TKFCM_Ss are reaching their maximum performance.

7.5.3. The third experiment is conducted on real image to compare the segmentation performance of the proposed methods with existing methods. The PSNR and the classification errors in the results obtained by KFCM_S1 (with $\alpha$-trimmed mean), KFCM_S2, KFCM_Ss, and TKFCM_Ss are compared. The value of $\alpha$ is fixed as 3.8 (the optimal value used in Chen, et al.[9]) $\beta$ as 3.5. $T_s$ as 3 and C as 2. The PSNR and segmentation accuracy are presented in Table (1), and the graphical representation of the results are given in Fig.(3.c) and Fig.(3.d). From the results it is realized that KFCM_Ss and TKFCM_Ss are performing better than KFCM_S1 (with $\alpha$-trimmed mean) and KFCM_S2. The output images are given in Fig.(4). As the existing defuzzification method is used in KFCM_S1 (with $\alpha$-trimmed mean) KFCM_S2 and KFCM_Ss, the output images lose their clarity. But as the proposed “two-phase defuzzification” is used in TKFCM_Ss, the defuzzified image has better clarity and revealing non-Euclidean structures well in the output clusters (please zoom in the image to see). Similar performances are obtained in the case of “Gaussian” and “Salt & Pepper” noised images.
Fig (3). Graphical representation of effect of parameters used in clustering the real image.

(a) Effect of the parameter $\beta$ in PSNR in the real image ‘eight’ corrupted by mixed noise.
(b) Effect of various values of Trimming Size ($Ts$) in Number of misclassified pixels
(c) PSNR against the parameter $\alpha$, (d) Classification errors against the parameter $\alpha$. 

Fig(4) Comparison of clustering results of real image.
(a) Original image, (b) image with mixed noise with $r = 0.5$, (c) KFCM_S2 result, (d) KFMC_S1 (with Trimmed mean) result, (e) KFCM_Ss2 result, (f) TKFCM_Ss2 result.
7.6. Experiments on simulated MRI

In this experiment, a high-resolution T1-weighted simulated phantom image (used in Cai et al. [28]) with 181×181 pixels, 1 mm slice thickness, 9% Gaussian noise and no gray inhomogeneous is used as experimental object. The slice is in the axial plane with sequence 91. In nature, the MRIs are not affected by Gaussian noise. But it is added for experimental purpose.

In an analogous way the first two types of experiments are conducted to test the effect of various values of β and α in segmentation performance. The graphical results are presented in Fig.(5.a), (5.c) and (5.d). Similar result obtained in the case of salt& pepper noise is presented in Fig.(5.b), (5.e) and (5.f).

7.6.1 Third experiment on Simulated MRI

In this experiment the PSNR and the classification errors in the results obtained by KFCM_S1 (with α-trimmed mean), KFCM_S2, KFCM_Ss2, and TKFCM_Ss2 are compared. For this experiment the parameter α is set as 9 (the optimal value used in Chen, et al.[9]), β as 3.5 and C as 3. The output images are presented in Fig.(6).

As the existing defuzzification method is used with KFCM_S1 (with α-trimmed mean) KFCM_S2, KFCM_Ss2, the image loses its clarity and non-Euclidean structures are not revealed well and they have low PSNR values. But when the proposed “two-phase defuzzification” is used in TKFCM_Ss2, the defuzzified image is very similar to the original noiseless image and the PSNR is increased.

For Gaussian and Salt& Pepper noised images, the quantitative comparisons of “selected neighbourhood pixels” with the existing neighbourhood masks are presented in Table (1). From the figures and Table values “selected neighbourhood pixels” produce better segmentation accuracy and “two-phase defuzzification methods” are giving higher PSNR values than the corresponding existing defuzzification methods. The effect is reflecting in the output figures.
Fig(5). Graphical representation of effect of parameters used in clustering the simulated MRI.
(a) $\beta$ vs PSNR in simulated MRI corrupted by Gaussian noise.
(b) $\beta$ vs PSNR in simulated MRI corrupted by Salt & Pepper noise.
(c) $\alpha$ vs PSNR in simulated MRI corrupted by Gaussian noise.
(d) $\alpha$ vs MissClassification in simulated MRI corrupted by Gaussian noise.
(e) $\alpha$ vs PSNR in simulated MRI corrupted by Salt & Pepper noise.
(f) $\alpha$ vs MissClassification in simulated MRI corrupted by Salt & Pepper noise.
Fig (6). Comparison of defuzzification results of Simulated Brain image.
(a) Original image, (b) image with 9% Gaussian noise, (c) KFCM_S2 result, (d) KFMC_S1 (with Trimmed mean) result, (e) KFCM_Sx2 result, (f) TKFCM_Sx2 result.
VIII CONCLUSION

From the figures and table values “selected neighbourhood masks” produce better segmentation accuracy and “two-phase defuzzification method” is giving higher PSNR values than the corresponding existing defuzzification methods in synthetic, real and simulated MR images and the effect is reflecting in the output figures.

Table 1
Comparison of segmentation performance of clustering methods

<table>
<thead>
<tr>
<th>Noise Type and level (%)</th>
<th>Existing Methods</th>
<th>Proposed Methods</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>KFCM_S1 (Trimmed Mean)</td>
<td>KFCM_S2</td>
</tr>
<tr>
<td></td>
<td>PSNR</td>
<td>No. of Misclassified</td>
</tr>
<tr>
<td>Synthetic Image</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3% Gaussian</td>
<td>27.66</td>
<td>0.9999</td>
</tr>
<tr>
<td>5% Gaussian</td>
<td>27.62</td>
<td>0.9999</td>
</tr>
<tr>
<td>8% Gaussian</td>
<td>27.52</td>
<td>0.9994</td>
</tr>
<tr>
<td>10% Gaussian</td>
<td>27.51</td>
<td>0.9995</td>
</tr>
<tr>
<td>3% Salt &amp; Pepper</td>
<td>27.61</td>
<td>0.9996</td>
</tr>
<tr>
<td>5% Salt &amp; Pepper</td>
<td>27.44</td>
<td>0.9989</td>
</tr>
<tr>
<td>8% Salt &amp; Pepper</td>
<td>27.12</td>
<td>0.9974</td>
</tr>
<tr>
<td>10% Salt &amp; Pepper</td>
<td>26.15</td>
<td>0.9873</td>
</tr>
<tr>
<td>15% Salt &amp; Pepper</td>
<td>26.15</td>
<td>0.9873</td>
</tr>
<tr>
<td>Mixed noise r=0.3</td>
<td>25.01</td>
<td>0.9998</td>
</tr>
<tr>
<td>Mixed noise r=0.5</td>
<td>22.23</td>
<td>0.9996</td>
</tr>
<tr>
<td>Mixed noise r=0.7</td>
<td>20.32</td>
<td>0.9994</td>
</tr>
<tr>
<td>Real Image</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3% Mixed noise r=0.5</td>
<td>23.58</td>
<td>0.9916</td>
</tr>
<tr>
<td>Simulated MRI</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9% Gaussian</td>
<td>25.06</td>
<td>0.9719</td>
</tr>
</tbody>
</table>
REFERENCES


**Dr. K. Kadambavanam** was born in Palni, Tamilnadu, India in 1956. He did his postgraduate studies at Annamalai University, Tamil Nadu, India in 1979. He completed his M.Phil. degree in Annamalai University in 1981. He obtained his Ph.D., degree from Bharathiar University, Coimbatore, Tamil Nadu, India. His area of research is Fuzzy Queueing Models and Fuzzy Inventory Models. His research interests include Solid mechanics, Random Polynomials and Fuzzy Clustering. He has 33 years of teaching experience in both U.G and P.G level and 18 years of research experience. He is a Chairman of Board of Studies (P.G) in Mathematics and Ex-Officio Member in Board of Studies (U.G) in Mathematics in Bharathiar University, Coimbatore. He is a member in: Panel of Resource Persons, Annamalai University, Annamalai Nagar, Doctorial Committee for Ph.D., program, Gandhigram Rural University, Gandhigram.
T. Senthil nathan is an Associate Professor in Post Graduate and Research Department of Mathematics, Erode Arts and Science College, Erode, Tamil Nadu, India. He received M.Sc. from Periyar E.V.R College, Trichy, Tamil Nadu, and M.Phil. from A.V.V.M. Sri.Pushpam College, Tanjavour, Tamil Nadu, India in 1985 and 1987 respectively. He received M.S. (software) from Birla Institute of Technology and Science (BITS), Pilani, India in 1994. He has 25 years of experience in teaching. He is a member in Board of Studies (U.G. and P.G.) in Mathematics in Erode Arts and Science College. Her research interests include image processing, signal processing, fuzzy logic and its application.