CHAPTER 4

A MODIFIED ADAPTIVE FUZZY C-MEANS CLUSTERING ALGORITHM FOR BRAIN MR IMAGE SEGMENTATION

4.1 PROLOGUE

Magnetic resonance (MR) imaging has several advantages over other medical imaging modalities, including high contrast among different soft tissues, relatively high spatial resolution across the entire field of view and multi-spectral characteristics. Therefore, it has been widely used in quantitative brain imaging studies. The noise in MR images is Rician distributed and can affect significantly the performances of classification methods.

The best solutions consist of either filtering the image prior to classification or embedding spatial regularization inside the classifier itself. While stable, FCM-based algorithms are known to be vulnerable to noise. To address this, Possibilistic Fuzzy C-Means (PF CM) algorithm is developed and has been shown to be more robust to outliers as compared with FCM. However, the robustness of PFCM comes at the expense of the algorithm’s stability. In order to overcome this problem, a new method is proposed to segment MR images with intensity inhomogeneities and noise. It computes the weights for the neighborhood of each pixel in the image. This adaptive method can not only overcome the effect of the noise effectively, but also prevent the edge from blurring.
4.2 INTRODUCTION

Fuzzy c-means (FCM) clustering has been widely used in image segmentation. However, in spite of its computational efficiency and wide spread popularity, the FCM algorithm does not take the spatial information of pixels into consideration, and hence may result in low robustness to noise and less accurate segmentation. In this work, a modified adaptive Fuzzy C-Means Clustering (MAFCM) algorithm is presented for fuzzy segmentation of Magnetic Resonance Images (MRI). The new adaptive fuzzy clustering algorithm is capable of utilizing local contextual information to impose local spatial continuity, thus improving the classification accuracy and reducing the number of iterations.

To estimate the intensity in homogeneity (Meyer et al. 1995), the global intensity is introduced into the coherent local intensity clustering algorithm and takes the local and global intensity information into account. The proposed method has been successfully applied to recorded MR images with desirable results. Our results show that the proposed AFPCM algorithm can effectively segment the test images and MR images. Comparisons with other FCM approaches based on number of iterations and time complexity demonstrate the superior performance of the proposed algorithm.

Magnetic resonance (MR) imaging has several advantages over other medical imaging modalities, including high contrast among different soft tissues, relatively high spatial resolution across the entire field of view and multi-spectral characteristics. Therefore, it has been widely used in quantitative brain imaging studies. Quantitative volumetric measurement and three-dimensional (3D) visualization of brain tissues are helpful for pathological evolution analysis, where image segmentation plays an important role. The size alterations in brain tissues often accompany various diseases, such as schizophrenia. Thus, estimation of tissue sizes has become an
extremely important aspect of treatment which should be accomplished as precisely as possible. This creates the need to properly segment the brain MR images into Gray Matter (GM), White Matter (WM) and CerebroSpinal Fluid (CSF) and also to identify tumors or lesions, if present. The main difficulties in brain segmentation are the intensity inhomogeneities and noise (Clark et al. 1998).

In fact, intensity inhomogeneity occurs in many real-world images from different modalities. In particular, it is often seen in medical images, such as X-ray radiography/ tomography and MR images. For example, the intensity variation across the image, which arises from Radio-Frequency (RF) coils or acquisition sequences. Thus the resultant intensities of the same tissue vary with the locations in the image. The noise in MR images is distributed and can affect significantly the performances of classification methods. The best solutions consist of either filtering the image prior to classification or embedding spatial regularization inside the classifier itself.

In medical images, the brain has a complicated structure, and segmentation of brain tissues in MR images plays a crucial role. The brain MR images partition into three main tissue types: White matter (WM), Grey matter (GM), and Cerebro Spinal fluid (CSF), which is a topic of great importance.

Clustering is the most popular method for medical image segmentation, with artificial neural network, Expectation-Maximization (EM), and Fuzzy C-Means (FCM) algorithms being the typical methods. A common disadvantage of the EM algorithm is that the intensity distribution of brain images is modeled as a normal distribution, which is not good for noisy images. The FCM algorithms are over sensitive to noise and many extensions to FCM and have been proposed to overcome its drawback. However, the
segmentation results are deteriorated by the noise and overlapping grey-scale intensities for different tissues in MR images.

MRI technology has many advantages such as nonradioactive contamination, high resolution, without electricity radiation damage to the human body, so it is widely applied in clinical diagnosis and treatment now. MRI image processing promotes the development of medical research and has important applications value. According to the above process, the accurate medical image segmentation by computational techniques plays an essential role.

In MRI medical image segmentation, the image is parsed into a number of meaningful regions based on the consistency principle. These regions usually do not cross each other and satisfy the consistency principle. If merging any two adjacent regions, it will break this principle. Hence, the medical image segmentation can be seen as the classification of image pixels in this viewpoint.

The influence is that a slowly varying shading artifact over the MRI image can produce errors with conventional intensity-based classification. Consequently, several methods for MRI intensity inhomogeneity correction are applied before the image segmentation. Series methods on intensity inhomogeneity correction/removal have been proposed in the last two decades. Nevertheless, intensity inhomogeneity correction is still incompletely solved problem. Because of this and the evolving MRI technology and associated applications, the problem of intensity inhomogeneity correction will certainly continue to be paid more research attention in the future.

In particular, for the fuzziness of the medical image, the fuzzy theory is introduced into the medical image processing, which generates lots
of new segmentation methods and achieves good segmentation results. The FCM algorithm has many advantages such as being without supervision, simple realization, and fast processing speed, which can carry out the accurate segmentation for the image with high contrast and signal-to-noise ratio. But there are also lots of obvious disadvantages. In the process of fuzzy clustering, the gray value distance between single pixel and cluster center can be considered only, while the influence of the adjacent pixels is neglected. That is to say, the spatial information cannot be well used in image segmentation. So the large deviation will be produced when the FCM algorithm is used to segment the brain MRI image with noises and the low signal-to-noise ratio.

4.3 CLUSTER ANALYSIS

Cluster analysis or clustering is the task of assigning a set of objects into groups (called clusters) so that the objects in the same cluster are more similar (in some sense or another) to each other than to those in other clusters.

Clustering is a main task of explorative data mining (Tan et al. 2005), and a common technique for statistical data analysis used in many fields, including machine learning, pattern recognition, image analysis, information retrieval, and bioinformatics.

Cluster analysis itself is not one specific algorithm, but it is a general task to be solved. It can be achieved by various algorithms that differ significantly in their notion of what constitutes a cluster and how to efficiently find them. Popular notions of clusters include groups with low distances among the cluster members, dense areas of the data space, intervals or particular statistical distributions. Clustering can therefore be formulated as a multi-objective optimization problem. The appropriate clustering algorithm
and parameter settings (including values such as the distance function to use, a density threshold or the number of expected clusters) depend on the individual data set and intended use of the results.

Cluster analysis as such is not an automatic task, but an iterative process of knowledge discovery or interactive multi-objective optimization that involves trial and failure. It will often be necessary to modify preprocessing and parameters until the result achieves the desired properties. This often leads to misunderstanding between researchers coming from the fields of data mining and machine learning, since they use the same terms and often the same algorithms, but have different goals.

The notion of a cluster varies between algorithms and is one of the many decisions taken when choosing the appropriate algorithm for a particular problem. At first the terminology of a cluster seems obvious: a group of data objects. However, the clusters found by different algorithms vary significantly in their properties, and understanding these "cluster models" is key to understanding the differences between the various algorithms. Typical cluster models include:

- Connectivity models: for example hierarchical clustering builds models based on distance connectivity.
- Centroid models: for example the k-means algorithm represents each cluster by a single mean vector.
- Distribution models: clusters are modeled using statistical distributions, such as multivariate normal distributions used by the Expectation-maximization algorithm.
- Density models: for example DBSCAN and OPTICS define clusters as connected dense regions in the data space.
• Subspace models: in Biclustering (also known as Co-clustering or two-mode-clustering), clusters are modeled with both cluster members and relevant attributes.
• Group models: some algorithms (unfortunately) do not provide a refined model for their results and just provide the grouping information.

4.4 CLUSTERING ALGORITHMS

Clustering algorithms can be categorized based on their cluster model, as listed above. The following overview will only list the most prominent examples of clustering algorithms, as there are probably a few dozen (if not over 100) published clustering algorithms. Not all provide models for their clusters and can thus not easily be categorized.

4.4.1 Connectivity based clustering (hierarchical clustering)

Connectivity based clustering, also known as hierarchical clustering, is based on the core idea of objects being more related to nearby objects than to objects farther away. As such, these algorithms connect "objects" to form "clusters" based on their distance. A cluster can be described largely by the maximum distance needed to connect parts of the cluster. At different distances, different clusters will form, which can be represented using a dendrogram, which explains where the common name "hierarchical clustering" arises: these algorithms do not provide a single partitioning of the data set, but instead provide an extensive hierarchy of clusters that merge with each other at certain distances. In a dendrogram, the y-axis marks the distance at which the clusters merge, while the objects are placed along the x-axis such that the clusters don't mix (Karypis et al. 1999).
Connectivity based clustering is a whole family of methods that differ by the way distances are computed. Apart from the usual choice of distance functions, the user also needs to decide on the linkage criterion (since a cluster consists of multiple objects, there are multiple candidates to compute the distance to) to use. Popular choices are known as single-linkage clustering (the minimum of object distances), complete linkage clustering (the maximum of object distances) or UPGMA ("Unweighted Pair Group Method with Arithmetic Mean", also known as average linkage clustering). Furthermore, hierarchical clustering can be agglomerative (starting with single elements and aggregating them into clusters) or divisive (starting with the complete data set and dividing it into partitions).

While these methods are fairly easy to understand, the results are not always easy to use, as they will not produce a unique partitioning of the data set, but a hierarchically the user still needs to choose appropriate clusters form. The methods are not very robust towards outliers, which will either show up as additional clusters or even cause other clusters to merge known as "chaining phenomenon", in particular with single-linkage clustering as shown in the Figure 4.1.

In the general case, the complexity is $O(n^3)$, which makes them too slow for large data sets. For some special cases, optimal efficient methods (of complexity $O(n^2)$) are known: SLINK for single-linkage and CLINK for complete-linkage clustering. In the data mining community these methods are recognized as a theoretical foundation of cluster analysis, but often considered obsolete. They did however provide inspiration for many later methods such as density based clustering as shown in the Figure 4.2. (Ester et al. 1996)
Figure 4.1 Single-linkage on Gaussian data

At 35 clusters, the biggest cluster starts fragmenting into smaller parts, while before it was still connected to the second largest due to the single-link effect.

Figure 4.2 Single-linkage on density-based clusters
4.4.2 Centroid-based clustering

In centroid-based clustering, clusters are represented by a central vector, which may not necessarily be a member of the data set. When the number of clusters is fixed to k, k-means clustering gives a formal definition as an optimization problem: find the $k$ cluster centers and assign the objects to the nearest cluster center, such that the squared distances from the cluster are minimized.

The optimization problem itself is known to be NP-hard, and thus the common approach is to search only for approximate solutions. A particularly well known approximative method is Lloyd's algorithm, often actually referred to as "k-means algorithm". It does however only find a local optimum, and is commonly run multiple times with different random initializations. Variations of k-means often include optimizations such as choosing the best of multiple runs, but also restricting the centroids to members of the data set (k-medoids).

Most k-means-type algorithms require the number of clusters - $k$ - to be specified in advance, which is considered to be one of the biggest drawbacks of these algorithms. Furthermore, the algorithms prefer clusters of approximately similar size, as they will always assign an object to the nearest centroid. This often leads to incorrectly cut borders in between of clusters (which is not surprising, as the algorithm optimized cluster centers, not cluster borders).

K-means has a number of interesting theoretical properties. On one hand, it partitions the data space into a structure known as Voronoi diagram as shown in the Figure 4.3. On the other hand, it is conceptually close to nearest neighbor classification and as such popular in machine learning. Third, it can be seen as a variation of model based classification, where as Lloyd's
algorithm is a variation of the Expectation-maximization algorithm for this model and is discussed as shown in the Figure 4.4.

Figure 4.3 K-means separates data into Voronoi-cells

Figure 4.4 K-means cannot represent density-based clusters
4.4.3 Distribution-based Clustering

The clustering model is most closely related to statistics is based on distribution models. Clusters can then easily be defined as objects belonging most likely to the same distribution. A nice property of this approach is that this closely resembles the way artificial data sets are generated: by sampling random objects from a distribution.

While the theoretical foundation of these methods is excellent, they suffer from one key problem known as overfitting, unless constraints are put on the model complexity. Usually a more complex model will always be able to explain the data better, which makes choosing the appropriate model complexity inherently difficult.

The most prominent method is known as expectation-maximization algorithm (or short: EM-clustering). Here, the data set is usually modeled with a fixed (to avoid overfitting) number of Gaussian distributions that are initialized randomly and whose parameters are iteratively optimized to fit better to the data set. This will converge to a local optimum, so multiple runs may produce different results. In order to obtain a hard clustering, objects are often then assigned to the Gaussian distribution they most likely belong to, for soft clusterings this is not necessary.

Distribution-based clustering is a semantically strong method, as it not only provides you with clusters, but also produces complex models for the clusters that can also capture correlation and dependence of attributes as shown in the Figure 4.5. However, using these algorithms puts an extra burden on the user: to choose appropriate data models to optimize, and for many real data sets, there may be no mathematical model available, the algorithm is able to optimize (e.g. assuming Gaussian distributions is a rather strong assumption on the data).
Figure 4.5 Gaussian-distributed data for modelling clusters

4.4.4 Density-based Clustering

In density-based clustering, clusters are defined as areas of higher density than the remainder of the data set. Objects in these sparse areas that are required to separate clusters are usually considered to be noise and border points.

The most popular density based clustering method is Density-Based Spatial Clustering of Applications with Noise (DBSCAN). In contrast to many newer methods, it features a well-defined cluster model called "density-reachability". Similar to linkage based clustering, it is based on connecting points within certain distance thresholds. However, it only connects points that satisfy a density criterion, in the original variant defined as a minimum number of other objects within this radius. A cluster consists of all density-connected objects (which can form a cluster of an arbitrary shape, in contrast to many other methods) plus all objects that are within these objects' range as shown in the Figure 4.6.
Another interesting property of DBSCAN is that its complexity is fairly low as it requires a linear number of range queries on the database - and that it will discover essentially the same results (it is deterministic for core and noise points, but not for border points) in each run, therefore there is no need to run it multiple times. OPTICS is a generalization of DBSCAN that removes the need to choose an appropriate value for the range parameter $\varepsilon$, and produces a hierarchical result related to that of linkage clustering. DeLi-Clu, Density-Link-Clustering combines ideas from single-linkage clustering and OPTICS, eliminating the $\varepsilon$ parameter entirely and offering performance improvements over OPTICS by using an R-tree index (Ankerst et al. 1999).

The key drawback of DBSCAN and OPTICS is that they expect some kind of density drop to detect cluster borders. Moreover, they cannot detect intrinsic cluster structures which are prevalent in the majority of real life data.

A variation of DBSCAN, Embedded cluster using Density-Based Spatial Clustering of Applications with Noise (EnDBSCAN) efficiently detects such kinds of structures. On data sets with, for example, overlapping Gaussian distributions - a common use case in artificial data - the cluster borders produced by these algorithms will often look arbitrary, because the cluster density decreases continuously. On a data set consisting of mixtures of Gaussians, these algorithms are always outperformed by methods such as EM clustering that are able to precisely model this kind of data.
4.5 RELATED RESEARCH

Magnetic resonance (MR) imaging has several advantages over other medical imaging modalities, including high contrast among different soft tissues, relatively high spatial resolution across the entire field of view and multi-spectral characteristics. Therefore, it has been widely used in quantitative brain imaging studies. Quantitative volumetric measurement and three-dimensional (3D) visualization of brain tissues are helpful for pathological evolution analysis, where image segmentation plays an important role. The size alterations in brain tissues often accompany various diseases, such as schizophrenia. Thus, estimation of tissue sizes has become an extremely important aspect of treatment which should be accomplished as
precisely as possible. This creates the need to properly segment the brain MR images into Gray matter (GM), White matter (WM) and Cerebro Spinal Fluid (CSF) and also to identify tumors or lesions, if present (Sikka et al. 2009).

The main difficulties in brain segmentation are the intensity inhomogeneities and noise. In fact, intensity inhomogeneity occurs in many real-world images from different modalities (Awate et al. 2007) (Wong and Chung 2009). In particular, it is often seen in medical images, such as X-ray radiography/tomography and MR images. For example, the intensity variation across the image, which arises from Radio-Frequency (RF) coils or acquisition sequences. Thus the resultant intensities of the same tissue vary with the locations in the image. The noise in MR images is distributed and can affect significantly the performances of classification methods. The best solutions consist of either filtering the image prior to classification or embedding spatial regularization inside the classifier itself.

Segmentation subdivides an image into different regions or objects based on the information found about objects in imaging data. In the segmentation of medical images, the objective is to identify different regions, organs and anatomical structures from data acquired via MRI or other medical imaging technique. Initially segmentation has been done based manually by human experts. But manual segmentation is a difficult and time consuming task, which makes automated breast cancer segmentation (Chen et al. 2006) method desirable. The automated segmentation (Ketsetzis and Brady 2004) of MR images into anatomical tissues, fluids, and structures is an interesting field in medical image analysis. Automatic tumor segmentation based on artificial intelligence (Clark et al. 1998) techniques was proposed to improve the edge detection accuracy.

In the last decades, fuzzy segmentation algorithms, especially the fuzzy c-means algorithm (FCM), have been broadly used in the image
segmentation (Xing et al. 2007) and such a success mostly attributes to the introduction of fuzziness for the belongingness of each image pixel. Fuzzy c-means (Bezdek et al. 1999) allows for the ability to make the clustering methods able to retain more information from the original image than the crisp or hard segmentation methods (Bezdek et al. 1993).

Clustering is used to panel a set of given observed input data vectors or image pixels into clusters so that components of the same cluster are similar to one another than to members of other clusters where the number of clusters is usually predefined or set by some weight criterion or a priori knowledge. Fuzzy c-means segmentation (Chuang et al. 2006) methods are having significant profit in segmentation of medical images, because they could retain a lot more information from the original image than hard c-means segmentation methods.

The main advantage in fuzzy c-means algorithm is it allows pixels to belong to multiple clusters with reasonable degrees of membership grades. However, there are some disadvantages in using fuzzy c-means, the membership of an object has not strong enough or significantly high for a particular cluster, it means that the equation of calculating membership is not an effective, and sometimes the equation for updating prototypes has incapable to work with data which greatly affected by noise. Thus the equation for updating prototypes leads the result of clustering might be uncorrected. The main reason for underlying drawbacks of above is, fuzzy c-means employs based on existed Euclidean distance measures.

Computer aided brain tumor segmentation system is an important application in medical image analysis. Developing a medical image analysis system not only can lighten the workload and decrease the errors of the doctors, but also can provide a quantitative measure about variation of the brain tumor throughout its whole therapeutic treatment. However, it is still a
difficult problem to automatically segment brain tumor regions from MRI multi-sequences because of many existing types of tumors with morphological variability, a variety of shapes and appearance properties among individuals, the deformation near the structures in the brain which results in an abnormal geometry also for healthy tissues, and lack of prior knowledge about them. Therefore, it is practically meaningful to focus on semi-automatic or fully-automatic segmentation methods on multiple MRI scans for medical research, disease monitoring, therapeutic control and so on. Different MRI sequences from different excitations can respectively provide different and partly independent information about different tissues, and reflect pathologic information about the tumors in the brain. As a tumor consists of different biologic tissues, one type of MRI cannot give complete information about abnormal tissues. Combining different complementary information can enhance the segmentation of the tumors. Therefore, radiology experts always combine the multi-spectral MRI information of one patient to make a decision on the location, extension, prognosis and Clustering is a process of classifying objects or patterns in such a way that the samples in the same group are more similar than the samples in different groups.

Based on the fuzzy theory, Zadeh (1965) proposed the fuzzy clustering method, which produces the idea of partial membership of belonging. As a soft clustering method, fuzzy clustering has been extensively studied and successfully applied to image segmentation. One of the most important and widely used fuzzy clustering methods is the Fuzzy C-Means (FCM) algorithm, which was first proposed by Dunn (1974) and promoted as the general FCM clustering algorithm by Bezdek (1981). The main purpose of the FCM algorithm is to make the vector space of a sample point to be divided into a number of sub-spaces in accordance with a distance measure. However, the FCM algorithm does not take the local spatial property of images into consideration, and hence suffers from high sensitivity to noise. To
improve its robustness, many modifications to the FCM algorithm that incorporate spatial information into clustering have been proposed.

4.6 CONVENTIONAL FCM CLUSTERING ALGORITHM

Multiresolution segmentation is a bottom up region merging technique starting with one-pixel objects (Sheikholeslami et al. 1998). In numerous subsequent steps, smaller image objects are merged into bigger ones. Throughout this pair wise clustering process, the underlying optimization procedure minimizes the weighted heterogeneity of resulting image objects, where \( n \) is the size of a segment and \( h \) an arbitrary definition of heterogeneity. In each step, that pair of adjacent image objects is merged which stands for the smallest growth of the defined heterogeneity. If the smallest growth exceeds the threshold defined by the scale parameter, the process stops. Doing so, multi-resolution segmentation is a local optimization procedure. The entropy based methodology for segmentation of satellite images is performed as shown in the Figure 4.7.

Images are divided into square windows with a fixed size \( L \), the entropy is calculated for each window, and then a classification methodology is applied for the identification of the category of the respective windows. The classification approach can be supervised or non-supervised. Supervised classification needs a training set composed by windows whose classes are previously known (prototypes), such as rural and urban areas.

Fuzzy c-means algorithm allows data to belong to two or more clusters with different membership coefficient. Fuzzy C-Means clustering is an iterative process. First, the initial fuzzy partition matrix is generated and the initial fuzzy cluster centers are calculated. In each step of the iteration, the cluster centers and the membership grade point are updated and the objective function is minimized to find the best location for the clusters. The process
stops when the maximum number of iterations is reached, or when the
objective function improvement between two consecutive iterations is less
than the minimum amount of improvement specified.

Moreover the update in the iteration is done using the membership
degree as well as the centre of the cluster which is a two parameter change as
the steps are being repeated until a set point called the threshold is reached or
the process stops when the maximum number of iterations is reached, or when
the objective function improvement between two consecutive iterations is less
than the minimum amount of improvement specified. In addition a fuzziness
coefficient ‘m’ is chosen which may be any real number greater than 1.

This algorithm works by assigning membership to each data point
corresponding to each cluster center on the basis of distance between the
cluster center and the data point and more data is near to the cluster center.
Clearly summation of membership of each data point should be equal to one.

\[
\mu_{kj} = \frac{1}{\sum \left(\frac{d_i}{d_{ij}}\right)^m} 
\]

(4.1)

\[
\phi_j = \frac{\sum_{i=1}^{n} (\mu_{ij})^m \mu_{ij}}{\sum_{i=1}^{n} \mu_{ij}^m}, \forall j = 1, 2, ..., c
\]

(4.2)

where, ‘n’ is the number of data points, \( \phi_j \) represents the \( j^{th} \) cluster center. ‘m’ is
the fuzziness index \( m \in [1, \infty) \). ‘c’ represents the number of cluster center. ‘\( \mu_{ij} \)’
represents the membership of \( i^{th} \) data to \( j^{th} \) cluster center. ‘\( d_{ij} \)’ represents the
Euclidean distance between \( i^{th} \) data and \( j^{th} \) cluster center.

Main objective of fuzzy c-means algorithm is to minimize:
\[ J(U, V) = \sum_{i=1}^{n} \sum_{j=1}^{c} (u_{ij})^{m} \| x_i - v_j \|^2 \]  

(4.3)

where,

\[ \| x_i - v_j \| \] is the Euclidean distance between \( i \)th data and \( j \)th cluster center.

Let \( X = \{ x_1, x_2, x_3, \ldots, x_n \} \) be the set of data points and \( V = \{ v_1, v_2, v_3, \ldots, v_c \} \) be the set of centers.

1) Randomly select ‘c’ cluster centers.

2) Calculate the fuzzy membership \( \mu_{ij} \) using:

\[ \mu_{ij} = \frac{1}{\sum_{k=1}^{c} (\frac{d(x_i, v_k)}{d(x_i, v_j)})^{\frac{1}{m-1}}} \]  

(4.4)

3) Compute the fuzzy centers \( v_j \) using:

\[ v_j = \frac{\sum_{i=1}^{n} (u_{ij})^{m} x_i}{\sum_{i=1}^{n} (u_{ij})^{m}}, \forall j = 1, 2, \ldots, c \]  

(4.5)

4) Repeat step 2) and 3) until the minimum \( J \) value is achieved or \( \| U^{(k+1)} - U^{(k)} \| < \beta \).

where,

\( k \) is the iteration step, \( \beta \) is the termination criterion between \([0, 1] \), \( U = (u_{ij})_{n\times c} \) is the fuzzy membership matrix, \( J \) is the objective function.
The advantages are,

1) Gives best result for overlapped data set and comparatively better than k-means algorithm.

2) Unlike k-means where data point must exclusively belong to one cluster center here data point is assigned membership to each cluster center as a result of which data point may belong to more than one cluster center.

The disadvantages are listed below:

1) Apriori specification of the number of clusters.

2) With lower value of $\beta$ we get the better results but at the expense of more number of iterations.

3) Euclidean distance measures can unequally weigh underlying factors.
4.7 KERNEL-BASED FUZZY C-MEANS CLUSTERING

When applying the Kernel based FCM framework in image-segmentation problems, the multiresolution segmentation may end up with local optimization procedure. Global mutual fitting is the strongest constraint for the optimization problem and it reduces heterogeneity most over the scene following a pure quantitative criterion. Its main disadvantage is that it does not use the treatment order and builds first segments in regions with a low spectral variance leading to an uneven growth of the image objects over a scene. It also causes an unbalance between regions of high and regions of low spectral variance.

Comparison of global mutual fitting to local mutual fitting results show negligible quantitative differences, the former always performs the most homogeneous merge in the local vicinity following the gradient of the degree of fitting. The growth of image objects happens simultaneously as well in regions of low spectral variance as in regions of high spectral variance.

Now describe how kernels defined as an inner product in feature space. Kernel $K$ is corresponding to an inner product in a feature space $\mathcal{H}$ via a map $\Phi : K \rightarrow \mathcal{H}$. In order to compute inner products of the form $(\Phi(x), \Phi(y))$, we employ kernel representations of the following form:

$$k(x, y) = (\Phi(x), \Phi(y))$$  \hspace{1cm} (4.6)

Which allow us to compute the value of the inner product in $\mathcal{H}$ without having to explicitly compute the map $\Phi$. Note that kernels are usually complex-valued in the mathematical literature. However we only consider the real-valued kernels, which is the common practice in machine learning.
A function : \( \mathbf{X} \rightarrow \mathcal{H} \) is employed and \( x \) is transformed into \( \varphi(x) \). An explicit representation of \( \varphi(x) \) is not usable in general, however the inner product \( \langle \varphi(x), \varphi(y) \rangle \) is expressed by

\[
K(x, y) = \langle \varphi(x), \varphi(y) \rangle
\]

(4.7)

Where a function \( K \) is a positive-definite kernel.

Let us apply kernel functions to kernel-based fuzzy c-means clustering. In this section, we assume a transformation into a high-dimensional Euclidean space \( \varphi: \mathbb{R}^p \rightarrow \mathcal{H} \), where \( x_k \) mapped into \( \varphi(x_k) \).

Objective function of the kernel-based fuzzy c-means is the following:

\[
J(U, V) = \sum_{k=1}^{N} \sum_{c=1}^{C} u_{kc} \| \varphi(x_k) - \varphi_c \|^2_{\mathcal{H}}
\]

(4.8)

Where \( \| . \| \) is the norm of \( \mathcal{H} \).

The basic procedure FCM should be used for iterative minimization, but

\[
D_{\varphi}(x_k, \varphi_c) = \| \varphi(x_k) - \varphi_c \|^2_{\mathcal{H}}
\]

(4.9)

Should be used instead of \( D(x_k, \varphi_c) \). Cluster center is

\[
\varphi = (\varphi_1, \ldots, \varphi_c)
\]

(4.10)

However, optimal \( \varphi_c \) does not have an explicit representation in general. Hence, the dissimilarity for clustering is as follows

\[
D_{\varphi} = D_{\varphi}(x_k, \varphi_c) = \langle \varphi(x_k), - \varphi_c \rangle + \langle \varphi_c, \varphi_c \rangle - 2 \langle \varphi(x_k), \varphi_c \rangle
\]

\[
= \langle \varphi(x_k), \varphi(x_k) \rangle - 2 \langle \varphi(x_k), \varphi_c \rangle + \langle \varphi_c, \varphi_c \rangle
\]
\[
K(x, y) = \langle \phi(x), \phi(y) \rangle, K_{ij} = \langle \phi(x_j), \phi(x_i) \rangle.
\]

Hence, we use the updating equation \( D_{k;i} \) and \( u_{k;i} \) are as follows:

\[
D_{k;i} = K_{k;k} - \frac{2}{\sum_{\alpha=1}^{c} w_{\alpha k}} \sum_{j=1}^{c} u_{\alpha j} K_{j;k} + \frac{1}{\sum_{\alpha=1}^{c} w_{\alpha k}} \sum_{j=1}^{c} \sum_{l=1}^{c} u_{\alpha j} u_{\alpha l} K_{j;l}
\]

\[
u_{k;i} = \frac{(\frac{1}{D_{k;i}})^{m-1}}{\sum_{k=1}^{C} (\frac{1}{D_{k;i}})^{m-1}}
\]

On the other hand, the entropy-based fuzzy c-means using a kernel function is used for the following objective function:

\[
J_{\text{kfcm}}(U, V) = \sum_{k=1}^{C} \sum_{i=1}^{c} u_{k;i} || \phi(x_k) - \hat{\phi}_i ||^2_F + \lambda^{-1} \sum_{k=1}^{C} \sum_{j=1}^{c} u_{k;j} \log u_{k;i}
\]

The dissimilarity \( D_{k;i} \) is modified to:

\[
D_{k;i} = D_{k;i}(x_k, \phi_i)
\]

\[
= \langle \phi(x_k), \phi(x_k) - \phi_i \rangle
\]

\[
= \langle \phi(x_k), \phi(x_k) \rangle - 2 \langle \phi(x_k), \phi_i \rangle + \langle \phi_i, \phi_i \rangle
\]

\[
= \langle \phi(x_k), \phi(x_k) \rangle - \frac{2}{\sum_{\alpha=1}^{c} w_{\alpha k}} \sum_{j=1}^{c} u_{\alpha j} \langle \phi(x_k), \phi(x_j) \rangle + \frac{1}{\sum_{\alpha=1}^{c} w_{\alpha k}} \sum_{j=1}^{c} \sum_{l=1}^{c} u_{\alpha j} u_{\alpha l} \langle \phi(x_j), \phi(x_l) \rangle
\]
We thus obtain the updating equation $D_{k+1}$ and $u_{k+1}$ of entropy-based kernel fuzzy c-means:

$$D_{k+1} = \lambda_{k+1} - \frac{2}{\sum_{j=1}^{n} u_{j+1} K_{k+1}^{-1}} \sum_{j=1}^{n} u_{j+1} K_{k+1}^{-1} - \frac{1}{\sum_{j=1}^{n} u_{j+1}} \sum_{j=1}^{n} \sum_{l=1}^{n} u_{j+1} u_{l+1} K_{k+1}$$  \hspace{1cm} (4.16)

$$u_{k+1} = \frac{\exp(-\lambda_{k+1})}{\sum_{j=1}^{n} \exp(-\lambda_{k+1})}$$  \hspace{1cm} (4.17)

In Fuzzy K-C-Means the interest is on making the number of iterations equal to that of the fuzzy c-means, and still attain an optimum result. This implies that irrespective of the lower number of iteration, we will still get an accurate result. The algorithm has the following steps:

1. Read the image into the Matlab environment
2. Try to identify the number of iteration it might possibly do within a given period of time
3. Reduce number of iteration with distance check
4. Get the size of the image
5. Calculate the distance possible size using repeating structure
6. Concatenate the given dimension for the image size
7. Repeat the matrix to generate large data items in carrying out possibly distance calculation
8. Reduce repeating when possible distance has been attained
9. Iterations begin by identifying large component of data
10. Iteration stops when possible identification elapses
4.8 MULTIPLE KFCM (MKFCM)

The application of multiple or composite kernels in the FKCM has its advantages. In addition to the flexibility in selecting kernel functions, it also offers a new approach to combine different information from multiple heterogeneous or homogeneous sources in the kernel space. Specifically, in image-segmentation problems, the input data involve properties of image pixels sometimes derived from very different sources.

Therefore, the different kernel functions purposely for the intensity information and the texture information separately, and we then combine these kernel functions and apply the composite kernel in MKFCM to obtain better image-segmentation results. The pixel information in these images inherits from different temporal sensors. As a result, we can define different kernels for different temperature channels and apply the combined kernel in a multiple-kernel learning algorithm.

The general framework of MKFCM aims to minimize the objective function

\[ Q = \sum_{i=1}^{c} \sum_{j=1}^{N} u_{ijm} \| \phi_{com} (x_j) - \phi_{com} (x_i) \|^2 \]  

(4.18)

To enhance the Gaussian-kernel-based KFCM-F by adding a local information term in the objective function

\[ Q = \sum_{i=1}^{c} \sum_{j=1}^{N} u_{ijm} (1 - k(x_j, x_i)) + \alpha \sum_{i=1}^{c} \sum_{j=1}^{N} u_{ijm} (1 - k(x_j, x_i)) \]  

(4.19)

where \( x_i \) is the intensity of pixel \( j \). In the new objective function, the additional term is the weighted sum of differences between the filtered intensity \( x_i \) (the local spatial information) and the clustering prototypes. The
differences are also measured using the kernel-induced distances. Such kind of enhanced KFCM-based algorithm is denoted as AKFCM (with a standing for additional term).

The MKFCM algorithm evaluates the centroids so as to minimize the influence of outliers. Unlike FCM, it does not attempt fuzzification for elements having membership values above the calculated threshold. This reduces the computational burden compared to FCM, also there is an absence of external user-defined parameters. The removal of this initial trial and error factor makes MKFCM more robust, as well as insensitive to the fluctuations in the incoming data. The elevation and reduction of the membership values to 1 and 0, respectively, results in contrast enhancement in the observability of the incoming data. This helps in focusing on the ambiguous boundary region; thereby gaining in terms of the quality of segmentation.

To further improve the performance of segmentation, MKFCM that linearly combines three kernels, i.e., the first two kernels are the kernels for intensities and the local spatial information. To sum up, the merit of MKFCM-based image-segmentation algorithms is the flexibility in selections and combinations of the kernel functions in different shapes and for different pieces of information. After combining the different kernels in the kernel space, there is no need to change the computation procedures of MFKCM. This is another advantage to reflect and fuse the image information from multiple heterogeneous or homogeneous sources. MKFCM-based image-segmentation algorithms are inherently better than other KFCM-based image segmentation methods.

The MKFCM's significant flexibility in kernel selections and combinations and the great potential of this flexibility could bring to image segmentation problems. In the MKFCM framework, it is possible to easily fuse the texture information into segmentation algorithms by just adding a
kernel designed for the texture information in the composite kernel. As in the satellite image-segmentation and two-texture image-segmentation problems, simply adding a Gaussian kernel function of the texture descriptor in the composite kernel of MKFCM leads to better segmentation results.

The FCM clustering algorithm was first proposed by (Dunn 1974) and promoted as the general fuzzy c-means clustering algorithm by (Bezdek 1981). The main purpose of FCM algorithm is to make the vector space of a sample point be divided into a number of sub-spaces in accordance with distance measure. However, FCM algorithm fails to deal with significant properties of images, since neighbor pixels are strongly correlated, which leads to strong noise sensitivity. To overcome this weakness, (Krishnapuram and Keller 1993) proposed a new clustering algorithm named PCM. PCM relaxes the column sum constraint of fuzzy membership matrix in FCM and introduces a possibilistic partition matrix, so that possibilistic memberships may reflect the typical data points to their clusters. Compared with FCM, PCM can effectively eliminate the influence of noise and outliers on clustering results. However, firstly the price PCM pays for its freedom to ignore noisy points is that PCM is very sensitive to initializations, and often results in the coincident cluster problem.

To overcome the weaknesses of the original PCM algorithm, the objective functions of PCM and FCM were combined into a new objective function and it was presented to provide an improved version, called PFCM, which can be interpreted as PCM and FCM, respectively, in some special cases where some proper parameters were adopted. So, PFCM can inherit the merits of both clustering algorithms.

The algorithm divides the data set \( I = \{I_1, I_2, \ldots, I_n\} \) into \( c \) clusters and \( n \) is the number of all the pixels in the image. Let the membership function \( u_{ik}, u_{ik} \in [0, 1] \) show the degree of the pixel \( I_k, k=1, 2, \ldots, N \) belonging
to cluster \( i \) \((1 \leq i \leq c)\). Then the result can be denoted by a matrix of fuzzy membership function matrix \( U = [u_{ik}]_{n \times c} \). Its represent typicality by \( t_{ik} \) \([0, 1]\) and the typicality matrix by \( T = [t_{ik}]_{n \times c} \). According to the definition of the theory, we have \( c = 1, u_{ik} = 1 \) for every pixel in the image.

The objective function to be minimized is:

\[
J_{MFPCM} = \sum_{i=1}^{c} \sum_{k=1}^{N} \left( C_F u_{ik} + C_T t_{ik} \right) \| I_k - V_i \|^2 + \sum_{k=1}^{N} \gamma_i \sum_{r=1}^{c} (1 - t_{ir})
\]

where \( V = \{v_1, v_2, ..., v_6\} \) is the characterized intensity center. The parameters \( C_F > 0, C_T > 0, m > 1, \gamma > 1, \gamma_i > 0 \) are user defined constants. The constants \( C_F \) and \( C_T \) define the relative importance of fuzzy membership with typical values in the objective function. Note that \( u_{ik} \) has the same meaning of membership as that in FCM. Similarly, \( t_{ik} \) has the same interpretation of typicality as in PCM. Let, the objective function of PFCM can get the minimum by updating the membership \( U \), the typicality \( T \) and the cluster centers \( V \) as follows:

\[
U_{ik} = \left( \sum_{j=1}^{c} \left( \frac{D_{ij}}{D_{ij} + \delta} \right) \right)^{-1}
\]

\[
t_{ik} = \frac{1}{1 + (C_T/C_F) D_{ik}}
\]

\( \gamma_i \) is defined as

\[
\gamma_i = \frac{\gamma_i + \alpha}{\gamma_i + \alpha \sum_{k=1}^{N} u_{ik} R_k}
\]

The intensity \( I_k \) at location \( k \) far away from the neighborhood center should have less influence in the clustering criterion function than the locations close to the neighbourhood center.
4.9 ADAPTIVE FUZZY C-MEANS ALGORITHM

A new objective function for obtaining fuzzy segmentations of images with intensity in homogeneity and an iterative algorithm for minimizing this objective function. The objective function contains a multiplier field term that models the brightness variation caused by the in homogeneties. Based on nonparametric density estimation for automatically obtaining initial values for the centroids is needed in this algorithm. Finally, a multigrid technique is described for the solution for the multiplier field during each iteration of the algorithm.

The membership value in this method is calculated using expressions

$$\mu_f(x_i) = \frac{\left(\frac{1}{d_i}\right)^{\frac{1}{m-1}}}{\sum_{j=1}^{m} \left(\frac{d_j}{d_i}\right)^{m-1}} \quad (4.24)$$

The adaptive fuzzy clustering algorithm is efficient in handling data with outlier points. In comparison with c-means algorithm, it gives only very low membership for outlier points. Since the sum of distances of points in all the clusters involves in membership calculation this method tends to produce very less membership and which is the main limitation of it.

4.10 MODIFIED ADAPTIVE FUZZY C-MEANS (MAFCM)

The observed MRI signal is modeled as a product of the true signal generated by the underlying anatomy, and a spatially-varying factor called the gain field

$$X_k = X_k \cdot G_k \quad \text{for} \ k \in \{1,2,3\ldots,N\} \quad (4.25)$$

where $X_k$ and $Y_k$ are the corrected and observed intensities at the $k_{th}$ voxel,
respectively, $G_k$ is the gain field at the $k_{th}$ voxel, and $N$ is the total number of voxels in the MRI volume.

The application of a logarithmic transformation to the intensities allows the artifact to be modeled as an additive bias field

$$y_k = x_k + \beta_k \quad \text{for } k \in \{1,2,3,...,N\}$$

where $x_k$ and $y_k$ are the corrected and observed log-transformed intensities at the $k_{th}$ voxel, respectively, and $\beta_k$ is the bias field at the $k_{th}$ voxel. If the gain field is known, then it is relatively easy to estimate the tissue class by applying a conventional intensity-based segmented to the corrected data. Similarly, if the tissue classes are known, then we can estimate the gain field. It may be problematic to estimate either without the knowledge of the other. We will show that by using an iterative algorithm based on fuzzy logic.

### 4.10.1 Objective Function

The standard FCM objective function for partitioning $\{x_k\}_{k=1}^N$ into $c$ clusters is given by

$$J = \sum_{l=1}^c \sum_{k=1}^N \|x_k - \theta_l\|^p$$

(4.27)

Where $\{\theta_l\}_{l=1}^c$ are the prototypes of the clusters and $[u_{ik}] = U$ represents a partition matrix, namely $U_{ik}^p$

$$U_{ik} \in [0,1], \sum_{l=1}^c u_{ik} = 1 \text{ and } 0 < \sum_{k=1}^N u_{ik} < N$$

(4.28)

The parameter $p$ is a weighting exponent on each fuzzy membership and determines the amount of fuzziness of the resulting classification. The FCM objective function is minimized when high
membership values are assigned to voxels whose intensities are close to the centroid of its particular class, and low membership values are assigned when the voxel data is far from the centroid.

Introducing a term that allows the labeling of a pixel (voxel) to be influenced by the labels in its immediate neighborhood. As mentioned before, the neighborhood effect acts as a regularizer and biases the solution towards piecewise-homogeneous labelings. Such regularization is useful in segmenting scans corrupted by salt and pepper noise.

The modified objective function is given by

\[ J_m = \sum_{i=1}^{c} \sum_{k=1}^{N} u_{ik} \| y_k - \beta_i \|^2 + \frac{\alpha}{N_R} \sum_{i=1}^{c} \sum_{k=1}^{N} u_{ik} \left( \sum_{i=1}^{c} \| y_i - \beta_i \|^2 \right) \]  \hspace{1cm} (4.29)

where \( N_k \) stands for the set of neighbors of \( x_k \) and \( N_R \) is the cardinality of \( N_k \). The neighbors effect term is controlled by the parameter \( \alpha \).

The relative importance of the regularizing term is inversely proportional to the signal to noise ratio (SNR) of MRI signal. Lower SNR would require higher value of the parameter \( \alpha \).

Substituting equation (4.26) into equation (4.29), we have

\[ J_m = \sum_{i=1}^{c} \sum_{k=1}^{N} u_{ik} \| y_k - \beta_i - \theta_i \|^2 + \frac{\alpha}{N_R} \sum_{i=1}^{c} \sum_{k=1}^{N} \sum_{i=1}^{c} \left( \sum_{i=1}^{c} \| y_i - \beta_i - \theta_i \|^2 \right) \]

(4.30)

Formally, the optimization problem comes in the form

\[ \min_{U, \{\beta_i\}_{i=1}^{c}} J_m = U \in \mathcal{U} \]  \hspace{1cm} (4.31)
4.10.2 Parametric Estimation

The objective function $J_m$ can be minimized in a fashion similar to the standard FCM algorithm. Taking the first derivatives of $J_m$ with respect to $u_{hk}$, $v_h$, and $\beta_k$ and setting them to zero results in three conditions for $J_m$ to be at a minimum. In the following subsections, then derived these three conditions.

4.10.3 Membership Evaluation

The constrained optimization in equation (4.31) will be solved using Lagrange multipliers as given by

$$J_m = \sum_{i=1}^{c} \sum_{k=1}^{N} \left( u_{ik}^p D_{hk} + \frac{\alpha}{N_F} u_{ik}^p \gamma_i \right) + (1 - \sum_{i=1}^{c} u_{ik})$$  \hspace{1cm} (4.32)

Where

$$D_{hk} = \|y_h - \beta_k - \theta_i\|^2$$ \hspace{1cm} (4.33)

$$\gamma_i = \sum_{h \in N_R} \|y_h - \beta_k - \theta_i\|^2$$ \hspace{1cm} (4.34)

Taking the derivative of $J_m$ w.r.t $u_{hk}$ and setting the result to zero, we have

$$\left[ \frac{\partial J_m}{\partial u_{hk}} = p u_{hk}^{p-1} D_{hk} + \frac{\alpha}{N_R} u_{hk}^{p-1} \gamma_i - \lambda \right] = 0$$ \hspace{1cm} (4.35)

Solving for $u_{hk}^p$ we have

$$u_{hk}^p = \left( \frac{\lambda}{p(D_{hk} + \frac{\alpha}{N_R} \gamma_i)} \right)^{\frac{1}{p-1}}$$ \hspace{1cm} (4.36)
Since $\sum_{i=1}^{j} u_{ir} = 1$

Then

$$\sum_{j=1}^{j} 1 \left( \frac{\lambda}{y[r_i]} \right)^{\frac{1}{p-1}} = 1 \quad \text{(4.37)}$$

or

$$\lambda = \frac{y[r_i]}{\left( \sum_{j=1}^{j} 1 \left( \frac{\lambda}{y[r_i]} \right)^{\frac{1}{p-1}} \right)^{p-1}} \quad \text{(4.38)}$$

Substituting into equation (5.36), the zero-gradient condition for the membership estimator can be rewritten as,

$$u_{kr} = \frac{1}{\left( \sum_{j=1}^{j} 1 \left( \frac{\lambda}{y[r_i]} \right)^{\frac{1}{p-1}} \right)^{p-1}} \quad \text{(4.39)}$$

4.10.4 Cluster Prototype Updating

In the following derivation the standard Euclidean distance is used. Taking the derivative of $J_m$ w.r.t $v_i$ and setting the result to zero we have;

$$\sum_{k=1}^{K} u_{kr}^p (y_k - \beta_k - o_i) + \sum_{k=1}^{K} u_{kr}^p \frac{\lambda}{N_k} \sum_{r \neq R_k} (y_k - \beta_k - o_i) = 0 \quad \text{(4.40)}$$

Solving for $\theta_j$ we have;

$$\theta_{j}^s = \frac{\sum_{k=1}^{K} u_{kr}^p (y_k - \beta_k) \gamma_{kr} \sum_{r \neq R_k} (y_k - \beta_k)}{(1-\alpha) \sum_{k=1}^{K} u_{kr}^p} \quad \text{(4.41)}$$
4.10.5 Bias Field Estimation

In a similar fashion, taking the derivative of $J_m$ w.r.t $\beta_k$ and setting the result to zero we have

$$\left[\Sigma_{i=1}^{c} \frac{\partial}{\partial \beta_k} \Sigma_{h=1}^{H} u_{ih}^p (v_{ih} - \beta_k - \theta_i)^2\right]_{\beta_k = \beta_k^*} = 0 \quad (4.42)$$

Since only $k_{th}$ term in the second summation depends on $\beta_k$, we have

$$\left[\Sigma_{i=1}^{c} \frac{\partial}{\partial \beta_k} u_{ih}^p (v_{ih} - \beta_k - \theta_i)^2\right]_{\beta_k = \beta_k^*} = 0 \quad (4.43)$$

Differentiating the distance expression, we obtain

$$\left[j_{ih} \Sigma_{i=1}^{c} \frac{\partial}{\partial \beta_k} u_{ih}^p - \beta_k \Sigma_{i=1}^{c} u_{ih}^p \theta_i - \Sigma_{i=1}^{c} u_{ih}^p \theta_i\right]_{\beta_k = \beta_k^*} = 0 \quad (4.44)$$

Thus, the zero-gradient condition for the bias field estimator is expressed as

$$\beta_k^* = j_{ih} - \frac{\Sigma_{i=1}^{c} u_{ih}^p \theta_i}{\Sigma_{i=1}^{c} u_{ih}^p} \quad (4.45)$$

The MAFCM algorithm for correcting the bias field and segmenting the image into different clusters can be summarized in the following steps:

Step 1: Select initial class prototypes $\{\theta_i\}_{i=1}^{c}$. Set $\{\beta_k\}_{k=1}^{K}$ to zero.

Step 2: Update the partition matrix

$$u_{ih}^* = \frac{1}{\Sigma_{j=1}^{c} \left(\frac{\beta_j - \theta_i}{\sigma_i^2 + \beta_j - \theta_i}\right)^{p-1}} \quad (4.46)$$
Step 3: The prototypes of the clusters are obtained in the form of weighted averages of the patterns

\[
\hat{u}_k^i = \frac{\sum_{l=1}^L u_{kl} P(y_k - \hat{y}_k) \frac{1}{\sigma^2 + \sum_{l=1}^L u_{kl} P(y_k - \hat{y}_k)}}{(1 + \alpha) \sum_{l=1}^L u_{kl} P}
\]  

(4.47)

Step 4: Estimate the bias term as follows

\[
\theta_k^x = y_k - \frac{\sum_{l=1}^L u_{kl} \theta_l^x}{\sum_{l=1}^L u_{kl}}
\]  

(4.48)

Repeat steps 2-4 till convergence.

4.11 RESULTS AND DISCUSSION

We test and compare the proposed method (MAFCM) with some other reported algorithms on several synthetic images and synthetic brain MR images from two aspects. The performance of FCM-type algorithms depends on the initialization, this work does the initialization and iterations depend upon the input images and choose the one with the best objective function value. This increases the reliability of comparison results acquired in the simulations. Considering accuracy, the proposed method is concentrated on obtaining a robust segmentation for noisy images and a correct detection of small regions.

Generally, incorporating of spatial information into the segmentation process will dramatically increase the algorithm’s computational complexity. To compare the computational complexity of the FCM, KFCM, MKFCM and MAFCM segmentation algorithms. The segmentation complexity of each algorithm was measured in terms of the average iteration number and average running time.
The test images for lena image, cameraman image and MRI images as shown in the below figures. In this segmentation algorithm, the parameter $\alpha$ is a constant, which controls the influence of the global intensity force and local intensity force. When the intensity in homogeneity is severe, the bias estimation relies on the local intensity force. In such a case, we should choose small $\alpha$, as the weight of the global intensity force. Otherwise, the bias field estimation may perform poorly. For images with minor in homogeneity, the accuracy of segmentation relies on the global intensity force. In this case, we can use relatively larger, weight for global intensity. Thus, the global intensity reduces the misclassification for the pixels around the edges. The modified AFECM has been applied to segment the images shown in the below figures.

![Figure 4.8. Original Cameraman Image](image-url)
Figure 4.11 Segmented Lena Image

Figure 4.12. Original MRI 1 Image
Figure 4.13. Segmented MRI 1 Image

Figure 4.14 Original MRI 2 Image
Figure 4.15 Segmented MRI 2 Image

Table 4.1 No. of iterations and time complexity of the proposed algorithm

<table>
<thead>
<tr>
<th>Image</th>
<th>Initial Cluster Centre Value</th>
<th>Final Cluster Centre Value</th>
<th>No. of Iterations</th>
<th>Time Consumption(sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lena</td>
<td>74.21</td>
<td>159.04</td>
<td>15</td>
<td>32</td>
</tr>
<tr>
<td>Cameraman</td>
<td>21.48</td>
<td>141.83</td>
<td>12</td>
<td>21</td>
</tr>
<tr>
<td>MRI 1</td>
<td>51.24</td>
<td>223.14</td>
<td>9</td>
<td>19</td>
</tr>
<tr>
<td>MRI 2</td>
<td>51.71</td>
<td>221.64</td>
<td>9</td>
<td>18</td>
</tr>
</tbody>
</table>

The quantitative comparison of the accuracy of those segmentation results are reported in Table 4.1. It reveals that our MAFCM algorithms achieves not only the highest accuracy in all three cases, but also the best
robustness to noise. This experiment demonstrates again that the proposed algorithm has a better ability to resist the influence of noise.

Table 4.2  Comparison between the proposed algorithms with other FCM algorithms

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Final Cluster Centre Value</th>
<th>No. of Iterations</th>
<th>Time Consumption(sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FCM</td>
<td>234.65</td>
<td>40</td>
<td>45</td>
</tr>
<tr>
<td>KFCM</td>
<td>219.76</td>
<td>28</td>
<td>42</td>
</tr>
<tr>
<td>MKFCM</td>
<td>225.94</td>
<td>25</td>
<td>30</td>
</tr>
<tr>
<td>Proposed</td>
<td>221.64</td>
<td>12</td>
<td>15</td>
</tr>
</tbody>
</table>

The size of image patches is an important parameter in our MAFCM algorithm. It determines how much spatial information will be used, and hence represents a trade-off between the image information and the spatial smoothness constraint. Table 4.2 shows the segmentation accuracy of the MAFCM algorithm with image patches of different size.

Now we draw the graph between Cluster value of the text image between numbers of iteration produced during the segmentation process for the algorithms fuzzy c-means, kernel fuzzy c-means, multiple kernel fuzzy c-means and multiple adaptive fuzzy c-means algorithm. The test images are generated by adding zero-mean Gaussian noise with different STD to the synthetic image shown in Figure 4.16. It reveals that the accuracy of the algorithm decreases with the increase in the level of noise for all size of image patches.
Next we draw the graph between Time consumption between numbers of iteration produced during the segmentation process for the algorithms fuzzy c-means, kernel fuzzy c-means, multiple kernel fuzzy c-means and multiple adaptive fuzzy c-means algorithm as shown in the Figure 4.17.
The above Figure 4.16 and 4.17 shows that the time consumption for segmentation and number of iterations is very less for our proposed method compared with other algorithms. Generally, incorporating of spatial information into the segmentation process will dramatically increase the algorithm’s computational complexity. To compare the computational complexity of the FCM, KFCM, MKFCM and our MAFCM algorithms, we applied each of these four segmentation algorithms to the 512 × 512 Lena image. Each segmentation was performed 20 times, and the computational complexity of each algorithm was measured in terms of the average iteration number and average running time.

4.12 CONCLUSION

A modified adaptive fuzzy c-means clustering algorithm is presented for fuzzy segmentation of MR images that have been corrupted by intensity inhomogeneities and noise. An adaptive method to compute the weights for the neighbourhood of each pixel in the image is proposed. The proposed adaptive method can not only overcome the effect of the noise effectively, but also prevent the edge from blurring. To address intensity inhomogeneity, the proposed algorithm introduces the global intensity into the algorithm and combines the local and global intensity information into account to ensure the smoothness of the derived optimal bias field and improve the accuracy of the segmentations. The proposed model can segment a brain MR mage in 9-10 iterations within 20 seconds. With good initialization, the model may need less iteration and can obtain results in less time. A variety of images, including synthetic images, synthetic brain MR images and real brain MR images are used to compare the performance of the proposed algorithm.