Development of Unsupervised Image Segmentation Schemes for Brain MRI using HMRF model

Master of Technology

(Research)

by

Smita Pradhan

(Roll: 607EE003)

Department of Electrical Engineering

National Institute of Technology

Rourkela, Orissa- 769008, India

2010

Development of Unsupervised Image Segmentation Schemes for Brain MRI using HMRF model

Thesis

Submitted in Partial Fulfillment of the Requirements

for the degree of

Master of Technology

(Research)

by

Smita Pradhan

Under the guidance

of

Dr. Dipti Patra

Department of Electrical Engineering

National Institute of Technology

Rourkela, Orissa- 769008, India

2010

Department of Electrical Engineering
National Institute of Technology
Rourkela-769008, Orissa, India.

Certificate

This is to certify that the work in this thesis entitled *Development of Unsupervised Image Segmentation Schemes for Brain MRI using HMRF model* by *Smita Pradhan, Roll No. 607EE003* is a record of an original research work carried out by her under my supervision and guidance in partial fulfillment of the requirements for the award of the degree of Master of Technology (Research) during the session 2007-2010 in the department of Electrical Engineering, National Institute of Technology, Rourkela. Neither this thesis nor any part of it has been submitted for any degree or academic award elsewhere.

Dr. Dipti Patra
Associate Professor
Department of EE, NIT, Rourkela

Place: NIT, Rourkela
Date: 25 Mar 2010

Acknowledgment

I would like to acknowledge my gratitude to a number of people who have helped me in different ways for the successful completion of my thesis. I would like to express my sincerest gratitude to my guide Dr. Dipti Patra, Associate Professor, Department of Electrical Engineering, National Institute of Technology, Rourkela but for whose deft guidance this thesis would not have seen the light of the day. Her erudite scholarship, prudent observations, perceptive critical comments and painstaking efforts to improve the quality of my work, have been steering me in the proper direction of research, from beginning to end. Moreover, her deep patience and kind understanding have helped me to overcome same particularly difficult times, when the odd seemed invincible and insurmountable.

I am also grateful to Prof. Sunil Kumar Sarangi, Director, N.I.T. Rourkela, for his inspiring words and supports. I humbly acknowledge the creative criticism and constructive suggestions of Prof. B. Majhi, Prof. S. Meher, committee members, while scrutinizing my research work. I wish to place on record my thanks to Prof. B. D. Subudhi, for his valuable comments, which helped me to complete my research work.

I ventilate my deep sense of gratitude to Prof. Susmita Das, Prof. and K.R.Subhashini for their learned advice and constant encouragement. I appreciate immensely the invaluable time lent to me by Mrs. Sucheta Panda and sonologist Dr. Hemalata Satapathy for their long discussion about my research. I am also grateful to all the staff members of the Department of Electrical Engineering for their co-operations and support throughout this period.

This work was made thoroughly enjoyable by the friendly and congenial atmosphere of the Image Processing and Computer Vision (IPCV) Laboratory of Electrical Engineering Department. I express my sincere thanks to all the past and present members of this Lab.

I am highly indebted to N.I.T. Rourkela for providing me the facilities like library, computational facility and Internet access, without which it could not have been completed in time.

I am thankful to all the staff members of the Department of Electrical Engineering,

N.I.T. Rourkela for their co-operations throughout this period.

Finally, I owe my loving thanks to my parents, brother and sisters. Without their encouragement and understanding it would have been impossible for me to finish this work.

Abstract

Image segmentation is a classical problem in computer vision and is of paramount importance to medical imaging. Medical image segmentation is an essential step for most subsequent image analysis task. The segmentation of anatomic structure in the brain plays a crucial role in neuro imaging analysis. The study of many brain disorders involves accurate tissue segmentation of brain magnetic resonance (MR) images. Manual segmentation of the brain tissues, namely white matter (WM), gray matter (GM) and cerebrospinal fluid (CSF) in MR images by an human expert is tedious for studies involving larger database. In addition, the lack of clearly defined edges between adjacent tissue classes deteriorates the significance of the analysis of the resulting segmentation. The segmentation is further complicated by the overlap of MR intensities of different tissue classes and by the presence of a spatially and smoothly varying intensity in-homogeneity. The prime objective of this dissertation is to develop strategies and methodologies for an automated brain MR image segmentation scheme.

As an initial attempt in this direction, the brain MR image segmentation problem is addressed in an unsupervised framework and is formulated as pixel labeling problem. Stochastic model based approach has been considered for the same. Hidden Markov Random Field (HMRF) models have been used to model the tissue classes of the observed degraded image. The *a priori* class labels are modeled as Markov Random Field (MRF) model. As the problem is addressed in an unsupervised framework, HMRF model parameters are assumed to be unknown. It is assumed to have the *a priori* knowledge of MRF model parameters which are used to model the unknown class labels, but no knowledge of number of classes and image labels. The problem becomes an incomplete data problem. To handle this problem, Expectation-Maximization algorithm is used. In order to incorporate a variable spatial characteristics which varies with internal part of the brain, the energy function associated with the *a priori* model is modified by an biased factor. This factor controls the effect of spatial information to avoid identical spatial information throughout the brain. The proposed modified model is named as Biased HMRF

(BHMRF) model. Intensity inhomogeneity or multiplicative bias field in brain MR image is also corrected in the proposed scheme. The results obtained by the proposed BHMRF-EM framework are compared with that of HMRF-EM scheme. The proposed scheme is found to be outperforming the later one and is observed to be an efficient method for brain MR image segmentation corrupted by biasfield.

In order to address the problem from practical stand point, a new notion of image segmentation is introduced by incorporating the fuzzy clustering approach in HMRF framework. The proposed approach is formulated using fuzzy c-means (FCM) algorithm which is facilitated by *a priori* MRF distribution. In this regard, HMRF oriented modification of the fuzzy objective function is incorporated. HMRF-EM scheme is found to be sensitive to the initial set of parameters. This has been overcome by proposing fuzzy clustering -EM (FCEM) algorithm that does not require to have a proper choice of initial parameters. In the proposed HMRF-FCEM scheme, combined strength of fuzzy clustering approach as well as HMRF model are incorporated. The result obtained by the proposed FCEM algorithm in HMRF-FCEM scheme are compared with that of existing schemes and the results are quite comparable to the later ones. The performance of proposed algorithm could be successfully tested with an arbitrary set of initial model parameters.

Both BHMRF-EM and HMRF-FCEM schemes could be validated for healthy as well as diseased brain MR images.

Contents

List of Acronyms

MRF	Markov Random Field
HMM	Hidden Markov Model
HMRF	Hidden Markov Random Field
BHMRF	Biased Hidden Markov Random Field
EM	Expectation-Maximization
ICM	Iterated Conditional Mode
SA	Simulated Annealing
GD	Gibbs Distribution
GRF	Gibbs Random Field
FCM	Fuzzy C-Means
FCEM	Fuzzy Clustering Expectation Maximization
MAP	Maximum a Posteriori
ML	Maximum Likelihood
MR	Magetic Resonance
MRI	Magetic Resonance Imaging
GM	Gray Matter
WM	White Matter
CSF	Cerebro Spinal Fluid
CT	Computed Tomography
PD	Photon Density
RF	Radio Frequency
Algo.	Algorithm

Nomenclatures

X	Random field associated with the labels of the original image.
X_i	Random variable of i^{th} site of the original image.
Y	Observed random field.
Y_i	Random variable of the i^{th} site of the observed image.
x	Realization of X.
y	Realization of Y.
f	Random field associated with line field.
θ	Parameter associated with image model.
θ_l	Parameter of each class with Gaussian distribution.
μ	Mean value of Gaussian distribution.
σ	Standard deviation of Gaussian distribution.
μ_l	Mean value of each class label with Gaussian distribution.
σ_l	Standard deviation of each class label with Gaussian distribution.
L	Set of class labels.
l	Class label variable.
$P(.)$	Probability.
$P(.\|.)$	Conditional probability.
S	A rectangular image lattice.
(i, j)	A pixel of an image lattice.
$(M \times N)$	Size of the image lattice.
$\eta_{(i,j)}$	Neighbourhood of pixel (i, j).
η^m	m^{th}-order neighbourhood.
$U(.)$	Energy function of Gibbs distribution.
Z	Partition function.
T	Temperature parameter.
T_{in}	Initial temperature.
c	A clique.
C	Collection of all cliques.

ϕ	Clique parameter vector.
dist(A, B)	Euclidian distance between A and B.
$V_c(.)$	Clique potential of MRF model.
x^*	True and unknown labeling configuration.
$\hat{x}$	Estimate for x^*.
$U_p(.)$	A posteriori energy function.
$T1$	Longitudinal relaxation time.
$T2$	Transverse relaxation time.
ρ	Biased parameter.
B	Bias field
ψ_B	Covariance matrix
R	Mean residual
F	Lowpass filter
W_{ij}	Posterior probability
$J_\lambda(\psi)$	Fuzzy objective function
r_{li}	Fuzzy membership function
d_{li}	Dissimilarity function
π_l	Prior probability
λ	Degree of fuzziness
G_i	Lagrange multiplier
ε	Convergence threshold

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Image segmentation

Segmentation is the process of splitting an observed image into its homogeneous or constituent regions. The goal of segmentation is to simplify or change the representation of an image into something that is more meaningful and easier to analyze. It is important in many computer vision and image processing application. In computer vision, segmentation refers to the process of partitioning a digital image into multiple regions. There are basically three different approaches to image segmentation. First is region based, which relies on the homogeneity of spatially localized features and other pixel statistics, the second one is based on the methods of boundary finding relying on the gradient features at a subset of the spatial positions of an image (near an object boundary), whereas the third one is pixel classification approach. Additionally, image segmentation has applications separate from computer vision; it is frequently used to aid in isolating or removing specific portions of an image. Image segmentation is typically used to locate objects and boundaries in images. The result of image segmentation is a set of regions that collectively cover the entire image, or a set of contours extracted from the image. It provides additional information about the contents of an image by identifying edges and regions of similar color, intensity and texture, while simplifying the image from thousands of pixels to less than a few hundred segments. Each of the pixels in a region are similar with respect to some characteristic or computed property. The segmentation of 2D and 3D images is an important first step for a variety of image analysis and visualization tasks. Hence, image segmentation is one of the early vision problems and has a wide application domain. The

problem becomes more compound while segmenting noisy images. The segmentation problem can be categorized as (i) supervised and (ii) unsupervised approach.

1.1.1 Supervised Image Segmentation

In supervised framework, the model parameters are assumed to be known *a priori*. These model parameters are used for estimating the pixel labels in segmentation framework. The pixel labelling problem, using MRF model has been formulated using Maximum a Posteriori (MAP) criterion and Bayesian framework [1, 2]. Segmentation of noisy images including textured images using MRF model could be formulated in supervised manner successfully. Nanda *et al.* have proposed a supervised image segmentation method where the MRF model parameters are estimated using homotopy continuation method and MAP estimate of image labels are obtained by SA algorithm [8].

1.1.2 Unsupervised Image Segmentation

In unsupervised framework, the number of class labels and the model parameters are unknown. Estimation of image labels as well as model parameters is required simultaneously. Since the image label estimation depends upon the optimal set of parameters, the unsupervised image segmentation is viewed as an incomplete data problem. To handle such problem, an iterative scheme namely expectation maximization (EM) algorithm was suggested [11, 16]. Besag *et al.* estimated the parameters using iterated conditional mode (ICM) algorithm for restoration [2]. Zhang *et al.* has suggested an unsupervised scheme which alleviates the difficulty in computing expectation in EM algorithm for general models. In order to accomplish this objective, he has proposed a Monte Carlo averaging scheme and a scheme related to Besag's ICM algorithm [9].

1.2 Application of Segmentation

1. Medical Imaging:
    - Locate tumors
    - Measure tissue volumes

- Computer-guided surgery
- Diagnostic Treatment planning
- Study of anatomical structure

2. Locate objects in satellite images
3. Face recognition
4. Automatic traffic controlling systems
5. Machine vision

1.3 Image models

In recent years, stochastic models have become more popular in image processing. Out of the various stochastic models, Markov Random Field (MRF) model provides a better framework for many complex problems in image segmentation. This is due to the fact that, MRF model is based on the notion of neighborhood structure and therefore, helps in understanding global interaction through local spatial interactions. Moreover, the global interaction is governed by Gibbs distribution. Markov Random Field (MRF) based methods have been widely used by researchers [2, 21, 7, 8, 10].

The extension of an observable Markov Model is the Hidden Markov Model (HMM). Here the observation is a probabilistic function (discrete or continuous) of a state. All observations are dependent on the state that generated them, not on the neighboring observations. HMM is a finite set of states, each of which is associated with a probability distribution. In a particular state an outcome or observation can be generated, according to the associated probability distribution. It is only the outcome, not the state visible to an external observer and therefore states are "hidden" to outside; hence the name Hidden Markov Model [12]. This model is specifically useful where the data is hidden. A special case of HMM is that, the underlying stochastic process is considered as MRF instead of a Markov chain and therefore not restricted to one dimension. This special case is referred to as Hidden Markov Random Field (HMRF) model.

The segmentation only relies on the histogram of the data and therefore is sensitive to noise and other artifacts or variations. To overcome this limitation, a hidden Markov random field (HMRF) is derived. The HMRF model is based on the Markov random field theory, in which the spatial information is encoded through a neighborhood system. Hidden Markov random field (HMRF) model is a stochastic process generated by a Markov random field whose state sequence cannot be observed directly but can be observed through observations. Mathematically, it can be shown that the FM model is a degenerate version of the HMRF model. Each observation is assumed to be a stochastic function of state sequence. By Markov Random Field the segmentation algorithm captures three features that are of special importance for MR images, i.e nonparametric distributions of tissue intensities, neighborhood correlation and signal inhomogeneities. The advantage of the HMRF model derives from the way in which the spatial information is encoded through the mutual influences of neighboring pixels [13, 25, 17].

1.4 Brain MR Images

MRI is an advanced medical imaging technique providing rich information about the human soft tissue anatomy. It has several advantages over other imaging techniques enabling it to provide 3-dimensional data with high contrast between soft tissues. However, the amount of data is far too much for manual analysis/interpretation, and this has been one of the biggest obstacles in the effective use of MRI. For this reason, automatic or semi-automatic techniques of computer-aided image analysis are necessary. Segmentation of MR images into different tissue classes, especially gray matter (GM), white matter (WM) and cerebrospinal fluid (CSF), is an important task. Brain MR images have a number of features, especially the following: Firstly, they are statistically simple; MR Images are theoretically piecewise constant with a small number of classes. Secondly, they have relatively high contrast between different tissues. The contrast in an MR image depends upon the way the image is acquired. By altering radio frequency and gradient pulses and by carefully choosing relaxation timing, it is possible to highlight different component in the object being imaged and produce high contrast images. These two features facilitate segmentation.

1.4.1 Weighting

MR images can be acquired using different techniques. The resulting images highlight different properties of the depicted materials. The most common weightings are T1 and T2, which highlight the properties T1-relaxation and T2-relaxation respectively. Selection of the most appropriate weighting is important for a successful segmentation. According to Pham *et al.* the properties of the tissues that are to be segmented have to be known to make a well-founded decision [42].

T1-weighted Images

T1-images show high contrast between tissues having different T1-relaxation times. Tissues with long T1-relaxation time emit little signal and thus they will be dark in the resulting image. In T1-images air, bone and CSF have low intensity, gray matter is dark gray, white matter is light gray, and adipose tissue has high intensity. T1-images have high contrast between white matter and gray matter.

T2-weighted Images

In T2-images, white matter and gray matter are gray and have similar intensities. CSF is bright, while bone, air, and fat appear dark. As opposed to T1-images, T2-images have high contrast between CSF and bone. The contrast between white matter and gray matter is not as good as in T1-images.

Spin Density

Spin density or Photon Density (PD) is the most like Computed Tomography (CT) of all the MR contrast parameters. The spin density is simply the number of spins in the sample that can be detected. The observed spin density in medical imaging is always less than the actual spin density due to the fact that many spins are bound and lose signal before they can be observed.

1.4.2 Artifacts

A variety of artifacts may appear in MR images. Since the artifacts change the appearance of the image they may also affect the performance of a segmentation algorithm. The most important artifacts in image segmentation are intensity inhomogeneities and the partial volume effect.

Intensity Inhomogeneities

Intensity inhomogeneities are not always visible to the human eye, but can nonetheless have negative influence on automatic segmentation. This may manifest itself by making intensities in one part of the image brighter or darker than another part. It is often caused by the radio frequency (RF) coils. Different methods exist to compensate for the inhomogeneities. The inhomogeneity is often modeled as a field that varies smoothly over the image. The inhomogeneity field is often thought to be a multiplicative field, which means that the true pixel intensity is multiplied by the value of the field in that pixel. There are methods which extracts the inhomogeneities during segmentation. Wells *et. al* and van Li *et. al* alternate estimation of the inhomogeneity field with classification to obtain inhomogeneity corrections [19, 41].

The Partial Volume Effect

The partial volume effect occurs when a pixel cannot be accurately assigned to one tissue type. This is because the intensity in the pixel originates from more than one tissue. It occurs because one pixel contains many body cells and the signal emitted from these cells make up the detected intensity in this pixel. The partial volume effect is most apparent at edges between different tissues. It may deteriorate the sharpness of the edges between tissues. The partial volume effect can be a significant problem in brain segmentation since the brain has a complex folded surface [13]. The partial volume effect is caused by the fact that of limited resolution in the images. Smaller pixel sizes reduce the partial volume effect since the probability that more than one tissue type is contained in the same pixel is reduced.

1.5 Brain MR Image segmentation

Segmentation of medical imagery is a challenging task due to the complexity of the images, as well as to the absence of models of the anatomy that fully capture the possible deformations in each structure. Brain tissue is a particularly complex structure, and its segmentation is an important step for derivation of computerized anatomical atlases, as well as pre- and intra-operative guidance for therapeutic intervention.

MRI segmentation has been proposed for a number of clinical investigations of varying complexity. Measurements of tumor volume and its response to therapy have used image gray scale methods as applied to X-ray, Computerized Tomography (CT) or simple MRI datasets. However, the differentiation of tissues within tumors that have similar MRI characteristics, such as edema, necrotic, or scar tissue, has proven to be important in the evaluation of response to therapy. Other applications of MRI segmentation include the diagnosis of brain trauma where white matter lesions, a signature of traumatic brain injury, may potentially be identified in moderate and possibly mild cases. These methods, in turn, may require correlation of anatomical images with functional metrics to provide sensitive measurements of brain trauma. MRI segmentation methods have also been useful in the diagnostic imaging of multiple sclerosis, including the detection of lesions.

1.6 Literature Survey

The image segmentation is a challenging problem that has received an enormous amount of attention by many researchers [1, 2, 3, 4]. Pham *et al.* and James *et al.* have presented various techniques used in medical image segmentation and analysis [5, 6]. The segmentation problem can be categorized as supervised and unsupervised problem. For appropriate analysis, different image models have been proposed for taking care of spatial intrinsic characteristics. The popular stochastic model, provides the better framework for many complex problem in image segmentation is Markov Random Field (MRF) model [7, 10]. MRF model and its variants have been successfully used for brain MR image segmentation [12, 13]. Ruan *et al.* proposed a fuzzy Markovian method for brain tissue segmentation from magnetic resonance images that calculates a fuzzy membership in

each pixel to indicate the partial volume degree, which is statistically modeled [14].

In unsupervised framework, the number of class labels and the model parameters are assumed to be unknown. Hence, estimation of image labels and model parameters are required simultaneously. Since, the image label estimation depends upon the optimal set of parameters, the segmentation problem can be viewed as incomplete data problem. To handle this problem, an iterative scheme, named Expectation-Maximization algorithm has been proposed [16]. Zhang *et al.* proposed Hidden Markov Random Field (HMRF) model to achieve brain MR image segmentation in unsupervised framework [17]. The segmentation obtained by Zhang's approach greatly depends upon the proper choice of initial model parameters. As Expectation-Maximization algorithm yields solutions at the cost of high computational burden, in order to overcome this Marroquin *et al.* have proposed a new class of probabilistic model, called Hidden Markov Measure Field model, that solved the complex segmentation problem by minimization of differentiable energy function [18]. Wells *et al.* and Brady *et al.* have proposed an adaptive brain MR image segmentation scheme in EM framework [19, 20]. They have also taken spatial intensity in-homogeneity into account and have estimated the bias field. Recently Hung *et al.* proposed an automatic segmentation method based on a decision tree to classify the brain tissues in magnetic resonance (MR) images [22]. Guan *et al.* have proposed an automatic hot spot detection and segmentation of whole body PET images using threshold and the Hidden Markov model (HMM). They compare the fixed PET pixel data threshold and the fixed standard uptake values (SUV) threshold for segmenting hot spots [24]. Nanda *et al* proposed a Tabu search based unsupervised scheme using HMRF-EM framework which could segment the images properly taking arbitrary initial parameter [25]. Anand *et al.* transformed an original image in to a multi scale wavelet domain and the wavelet coefficients are processed by a soft thresholding method. Various wavelet filter based denoising methods are studied according to different thresholding values and applied to ultrasound images [23]. Joshi *et al.* modeled the fused multi spectral (MS) image using a low spatial resolution MS images as the aliased and corresponding noisy versions as high spatial resolution. The fused image is obtained for each of the MS bands by estimating the high spatial resolution and then modeling as separate inhomogeneous Gaussian markov

random fields (IGMRF) and a maximum a posteriori (MAP) estimation [26].

Now a days, fuzzy image segmentation is increasing popularity because of rapid extension of fuzzy set theory, the development of various fuzzy set based mathematical modeling, and its successful application in computer vision system [27]. Ichihashi *et al.* showed that the EM algorithm for GMM can be derived from the FCM type clustering, when considering a regularization by KL information fuzzy objective function, for selection of the distance metric [28]. Ahmed *et al.* proposed a bias correction fuzzy c-means algorithm in which they incorporated a neighborhood regularizer into the FCM objective function to allow labeling of a pixel to be influenced by the labels in its immediate neighborhood [29]. The algorithm is realized by incorporating the spatial neighborhood information into the standard FCM algorithm and modifying the membership weighting of each cluster. Parui *et al.* approached a mixture model suitable for segmentation of the color images. The certain color space in a pixel is clustered by employing the K-Means algorithm [30]. A General Reflex Fuzzy Min-Max Neural Network (GRFMN) is proposed to extract the underlying structure of the data by means of supervised, unsupervised and partially supervised learning by Biswas *et al.* [31]. Chen *et al.* proposed an adaptive FCM algorithm which is found to be robust in convergence. The objective function to be minimized has regularization terms that ensure the estimated bias field is smooth and slowly varying [32]. Siyal *et al.* presented a modified FCM algorithm formulated by modifying the objective function of the standard FCM and uses a special spread method for classification of tissues [33]. Wang *et al.* proposed a modified FCM algorithm, called mFCM for brain MR image segmentation [34]. Aboulella *et al.* proposed a statistical feature extraction technique for diagnosis of breast cancer mammograms by combining the fuzzy image processing with rough set theory [35]. Martin *et al.* described a way to segment the medical images using an appropriately defined fuzzy clustering based on a fuzzy relation. The considered relation is defined in terms of Euclidian distance [36]. Kang *et al.* presented a novel method for segmentation by incorporating spatial neighborhood information in to the standard FCM. An adaptive weighted averaging filter is given to indicate the spatial influence of the center pixel [37]. Panas *et al.* presented the adaptive fuzzy clustering/ segmentation (AFCS). In AFCS, the nonstationary nature of the image

taken into account by modifying the prototype vectors as function of sample location in the image. A multiresolution model is utilized for estimating the spatially varying prototype vectors for different window sizes. The segmentation of different resolutions is combined using a data fusion process in order to compute the final fuzzy partition matrix. The results provide segmentation having lower entropy [38]. Mohamed *et al.* described the application of fuzzy set theory in medical imaging. A fully automatic technique to obtain cluster is proposed. A modified fuzzy c-means classification algorithm is used to provide a fuzzy partition. The method is inspried by Markov random field (MRF) and is found to be less sensitive to noise as it filters the image while clustering [39]. Kannan *et al.* presented a new method called fuzzy membership c-means (FMCM) for segmentation of Magnetic Resonance Images (MRI). This work develops a specific method to construct the initial membership matrix to clusters in order to improve the strength of the clusters [40].

1.7 Motivation

Image segmentation is an essential step in medical images for subsequent image analysis tasks. Some of the issues that make medical image segmentation difficult, particularly in brain magnetic resonance images (MRI) are intensity in-homogeneity or bias field and partial volume problem. Many segmentation techniques have been developed by the researchers which help the physicians and neurosurgeons to investigate and diagnose the structure and function of the brain. Brain consists of three soft tissues such as gray matter (GM), white matter (WM) and cerebrospinal fluid (CSF). Any other soft tissue like brain tumor along with above soft tissues can be imaged using magnetic resonance imaging (MRI). Ideally, for any given set of MR imaging parameters, the intensity values of the pixels of any given tissue class should be constant or correspond to a Gaussian distribution with small standard deviation. In practice, spatial intensity in-homogeneities are often of sufficient magnitude to cause the distributions of signal intensities associated with the tissue classes to overlap significantly. In addition, the lack of clearly defined edges between adjacent tissue classes deteriorates the significance of the analysis of the resulting segmentation. The segmentation is complicated due to these factors. This has motivated

the need for automatic segmentation techniques that are robust in application involving a broad range of brain MR images. Hence, the main objective of this thesis is to address the unsupervised image segmentation schemes for brain MR images.

1.8 Problem Addressed

In this thesis, attempts are made to address the problem of brain MR image segmentation in unsupervised framework. The observed degraded image is assumed to be corrupted with white Gaussian noise and bias field. The unsupervised schemes have been proposed using HMRF model. The research work of this thesis can be broadly categorized as;

(i) unsupervised image segmentation using biased HMRF model and

(ii) fuzzy clustering based image segmentation using HMRF model.

1.9 Summary of the thesis

In this piece of work, attempts are made to address the problem of image segmentation of brain MRI in unsupervised framework. The observed degraded brain MR image is assumed to be corrupted with additive white Gaussian noise and multiplicative bias field. The segmentation problem is casting as pixel labelling problem and the model based approach is adhered for the same. The unknown *a priori* class labels for different tissue classes of brain MR image are modelled as MRF model while the observed degraded images are modelled as HMRF model. As the problem is formulated in unsupervised mode, both the model parameters and image labels are assumed to be unknown and are estimated together. This problem can be viewed as an incomplete data problem and hence the problem is formulated in EM framework inspired by the work of Zhang *et al.* [17]. In this thesis, the spatial interaction of pixel labels are encoded through HMRF model. HMRF model is modified as biased HMRF (BHMRF) model by incorporating the biased neighborhood in the energy function such that the anatomy of brain is encoded through the proposed scheme. The BHMRF-EM scheme yielded better performance than HMRF-EM scheme. The scheme is modified to estimate intensity in-homogeneity or bias field along with model parameters and image class labels. Further it is found that the performance of

HMRF-EM scheme proposed by Zhang *et al.* greatly depends upon the selection of initial model parameters [17]. In order to circumvent that problem fuzzy clustering based HMRF model is proposed. In this regard HMRF-FCEM algorithm is proposed that does not require to have proper choice of initial model parameters. The proposed algorithm yielded satisfactory results with arbitrary initial model parameters. The proposed algorithms are validated with synthetic images, simulated as well as real brain MR images.

1.10 Thesis Organization

The thesis is organized into the following chapters.

Chapter 1: Introduction

It deals with the formal description of image segmentation, brain MR image and its segmentation, literature review and a brief on thesis contribution.

Chapter 2: HMRF model and Fuzzy clustering based image segmentation

Background on Markov Random field model, Hidden Markov Random Field model and segmentation methods based on above models are focused in this chapter. Basic notion of fuzzy c-means (FCM) clustering technique is also included here.

Chapter 3: Unsupervised image segmentation and intensity in-homogeneity correction using Biased HMRF model

In this chapter, the segmentation of brain MR image is addressed in an unsupervised framework. The Hidden Markov Random Field (HMRF) model is employed for the observed degraded image. MRF model is employed for the noise free class labels. The tissue class of brain MR image is modeled as HMRF model with Gaussian emission distribution and the associated model parameters such as μ and σ for each class are assumed to be unknown. In order to incorporate a variable spatial characteristics which varies with internal part of the brain, the energy function associated with the *a priori* model is modified by a biased factor. This factor controls the effect of spatial information to avoid identical spatial information throughout the brain. The proposed modified model is named as Biased HMRF (BHMRF) model. Expectation-Maximization (EM) algorithm is used to estimate the model parameters as well as image labels jointly. This leads to the development of BHMRF-EM algorithm for unsupervised brain MR image segmentation.

Intensity in-homogeneity or bias field in brain MR images is inherent due to the presence of radio-frequency coil during MRI. The slowly varying multiplicative bias field is estimated in various tissue classes of brain MR images based on Expectation-Maximization (EM) algorithm. This leads to the joint estimation of intensity in-homogeneity, model parameters and image labels using the proposed modified BHMRF-EM algorithm.

Chapter 4: Unsupervised Image segmentation using HMRF-FCEM algorithm
In this chapter, a new notion of unsupervised brain MR image segmentation scheme is proposed by hybridizing the benefits of fuzzy clustering technique and HMRF model. In the proposed approach, HMRF model is incorporated into fuzzy clustering scheme by modifying the fuzzy objective function with HMRF orientation. In this regard, HMRF model is regarded as defining a number of fuzzy partitions which is same as number of class labels. HMRF oriented fuzzy objective function is proposed by considering the mean field approximation of the MRF probability. This hybridizes the benefits of the spatial interaction of the HMRF model, and the enhanced flexibility obtained by the fuzzy clustering algorithm. HMRF-FCEM algorithm is proposed to estimate the image labels as well as fuzzy membership function jointly. Eventually, the fuzzy objective function is minimized using the proposed algorithm and an estimate of the HMRF model parameters are obtained. The proposed algorithm does not depend upon the proper choice of initial model parameters.

Chapter 5: Conclusions
This chapter presents the concluding remark on unsupervised image segmentation schemes for brain MR images, with scope for further research work on the related problems.

1.11 Image Metrics

The quality of an image is examined by objective as well as subjective evaluation. The metrics used for performance comparison of different segmentation schemes are defined below.

Misclassification error (MCE) is a measure of percentage of misclassified pixels changes their gray scale values in the segmented image. It measures the difference between two images. In other words, it measures the efficiency of the proposed schemes with the for-

mer existed schemes. Hence, the lower value of MCE, better is the segmentation. The MCE can be calculated as

$$MCE = \frac{Number\ of\ misclassified\ pixels\ in\ a\ cluster}{Original\ number\ of\ pixels\ in\ the\ cluster} * 100$$

Another image metric used for comparison of different methods is the execution time. Execution time is defined as the time taken for the simulation of an algorithm. The less time an algorithm takes for execution, the more efficient it is considered.

Subjective or Qualitative measure:

Subjective assessment is required to measure the image quality. Unavailability of quantitative performance measure in case of image segmentation, subjective or qualitative measure is another option for comparison. In a subjective assessment measures characteristics of human perception become paramount, and the image quality is correlated with the preference of an observer or the performance of an operator for some specific task. Hence, as an usual case of image segmentation there is no quantitative performance evaluation measure because no ideal image can be used as reference. Any reasonable measure should be tuned to the human visual system. However perceptual quality evaluation is not a deterministic process. So, subjective evaluation is the way to prove the performance. Hence, human observer is the only way by which segmented image quality can be observed.

The processor used for simulation of the segmentation problem is **Pentium IV Intel core 2 Duo processor, 1.8 Ghz, 1 GB RAM, Fedora-6 version in Linux operating system**.

Chapter 2

Fundamentals of HMRF model Fuzzy clustering methods

2.1 Introduction

Image Segmentation techniques using spatial interaction models like Markov Random Field (MRF) to model the image have become very popular. The use of contextual information is indispensable in low level as well as high level Image Processing . Markov Random Field theory provides a convenient and consistent way of modeling the entities with contextual constraints. This is achieved through characterizing mutual relationship among such entities such as pixels of an image and other spatially correlated features using MRF probabilities. MRF forms a probabilistic model for a set of variables that interact on a lattice structure. This started with the influential work of Geman $\&$ Geman [1] who linked via statistical mechanics between mechanical systems and probability theory. The distribution for a single variable at a particular site is conditioned on the configuration of a predefined neighborhood surrounding that site.

Hidden Markov field (HMF) models are widely applied to different problems concerned with image processing. The use of hidden Markov model (HMM) is a powerful modern statistical technique that has been found to be extremely useful for a wide spectrum of applications in ecology, crypt analysis, image understanding, speech and handwriting recognition. Formally, a hidden Markov model, is a doubly embedded stochastic process with an underlying process that is not observable but can only be observed through another set of stochastic process that produce the sequence of observations. It

is a statistical model where the system being modeled is assumed to be Markov process with unknown parameters, and the challenge is to determine the hidden parameters, from the observable parameters, based on the assumption. In these models, the hidden process is a Markov field and estimated from its observable noisy image. This models are popular mainly due to the fact that the conditional probability distribution of the hidden layer with respect to the observed layer remains Markov.

2.2 Markov Random Field

Let consider a collection of random variables $\{X_{ij}\}$, that is a random field defined over a finite discrete rectangular lattice S of size $(M \times N)$. The lattice S is defined as $S = \{(i,j) : 1 \leq i \leq M, 1 \leq j \leq N\}$, where site (i,j) corresponds to each pixel of the discrete image lattice structure. A neighborhood system η on this rectangular lattice S can be defined as follows.

Definition 1 *A collection of subsets of S described as $\eta = \{\eta_{i,j} : (i,j) \in S, \eta_{i,j} \subset S\}$ is a neighborhood system on S if and only if $\eta_{i,j}$, the neighborhood of pixel (i,j), is such that*

1. a site is not neighboring to itself : $(i,j) \notin \eta_{ij}$
2. the neighboring relationship is mutual : If $(k,l) \in \eta_{ij}$, then $(i,j) \in \eta_{kl}$ for any $(i,j) \in S$

The neighbor set of η_{ij} is defined as the set of nearby sites within a radius r such that $\eta_{ij} = \{(k,l) \in S \mid \{dist((i,j),(k,l))\}^2 \leq r, (i,j) \neq (k,l)\}$, where dist$(A,B)$ denotes the Euclidean distance between A and B, r takes an integer value. A hierarchically ordered sequence of neighborhood systems is shown in Figure 3.1. where η^1, η^2, η^3.... are the “first-order”, “second-order”, “third-order”..... neighborhood systems respectively and are denoted by numbers 1,2,3....as shown in Figure 2.1. Due to the finite lattice used, the neighborhood of pixels on the boundaries are necessarily smaller unless a toroidal (periodic) lattice structure is assumed. A nearest neighborhood dependence of pixels on an image lattice is obtained by going beyond the assumption of statistical independence. The neighborhood systems that can be defined over S are neither limited to the hierarchically

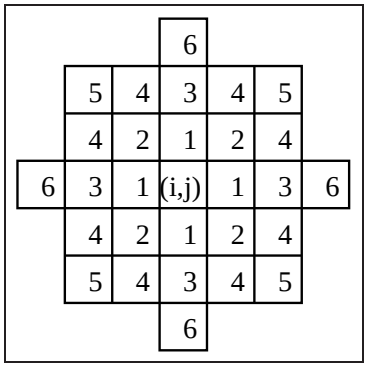


Figure 2.1: *Hierarchically arranged neighborhood system of Markov Random Field*

ordered sequence of neighborhood systems, nor they have to be isotropic or homogeneous.

Definition 2 *Let η be a neighborhood system defined over lattice S. A random field $X = \{X_{i,j}\}$ defined over lattice S is a Markov Random Field (MRF) with respect to the neighborhood system η if and only if*

1. All of its realizations have nonzero probabilities:
   $P(X = x) > 0$ for all x (property of Positivity)

2. Its conditional distribution satisfies the following property:
   $P\{X_{ij} = x_{ij} \mid X_{kl} = x_{kl}, (k,l) \in S, (k,l) \neq (i,j)\}$
   $= P\{X_{ij} = x_{ij} \mid X_{kl} = x_{kl}, (k,l) \in \eta_{ij}\}$ for all $(i,j) \in S$ (property of Markovianity)

where x_{ij} is the configuration corresponding to the random variable X_{ij} and so on.When the positivity condition is satisfied, the joint probability $P(X)$ of any random field is uniquely determined by its local conditional probabilities [2]. The Markovianity depicts the local characteristics of X which is characterized by the conditional distributions. The *Definition 2* says that the image value at a pixel does not depend on the image data outside its neighborhood, when the image data on its neighborhood are given. Hence, the most attractive feature of MRF is that *"images tend to have a degree of cohesiveness: pixels located near to each other tend to have the same or similar colours"* [1]. It does not constitute a theoretical restriction either, because all random field satisfy *Definition 2*,

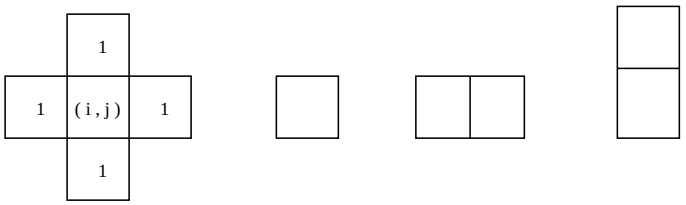


Figure 2.2: *Cliques associated with first-order neighborhood system*

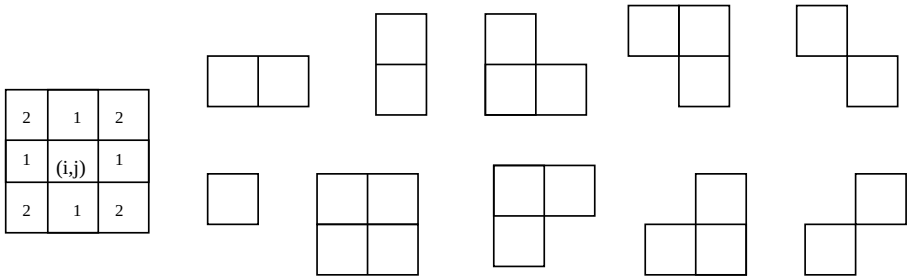


Figure 2.3: *Cliques associated with second-order neighborhood system*

with respect to a large enough neighborhood system, e.g. $\eta_{i,j} = S$ for all $\eta_{i,j} \in S$. On the other hand, MRF models, even with respect to small neighborhood systems such as η^2 prove to be very flexible and powerful. Let us define the *clique* associated with (S, η), a lattice neighborhood system pair:

Definition 3 A *clique of the pair* (S, η) *denoted by* c *is a subset of* S *such that*

1. *c consists of a single pixel, or*
2. *for* $(i, j) \neq (k, l),\ (i, j) \in c$ *and* $(k, l) \in c$ *implies that* $(i, j) \in \eta_{k,l}$

The collection of all cliques of (S, η) is defined by $C(S, \eta)$. The clique types associated with first-order and second-order neighborhood systems are shown in Figure 2.2 and Figure 2.3 respectively.

2.3 Gibbs Random Field

Gibbs Distribution (GD) or equivalently the Gibbs Random Field (GRF) can be defined as follows.

Definition 4 *Let* η *be a neighborhood system defined over a finite lattice* S*. A random field*

X is said to be a Gibbs Random Field (GRF) on lattice S with respect to a neighborhood system η if and only if its configuration obey a Gibbs distribution which has the following form

$$P(X = x) = \frac{1}{Z} e^{-\frac{1}{T}U(x)} \tag{2.1}$$

where,

$$Z = \sum_{x} e^{-\frac{1}{T}U(x)} \tag{2.2}$$

is the partition function. Z is simply a normalizing constant so that the sum of the probabilities of all realizations, x becomes one. T is a constant analogous to temperature which shall be assumed to be 1 unless otherwise stated and $U(x)$ is the energy function or Hamiltonian of a Gibbs distribution, which can be expressed as follows

$$U(x) = \sum_{c \in C} V_c(x) \tag{2.3}$$

Hence, energy is sum of clique potentials $V_c(x)$ over all possible cliques C. $V_c(x)$ are a set of potential functions depending on the values of x at the sites in the clique c. Thus, the key functions in determining the properties of the distribution are the potential functions $V_c(x)$. $P(x)$ measures the probability of the occurrence of a particular configuration x. The more probable is a particular configuration, has lesser energy. This is so because the energy is computed as a measure of the distance between the model and the raw image data. The potential functions are chosen to reflect the desired properties of the image so that the more likely images have a lower energy and are thus more probable. The temperature T controls the sharpness of the distribution. When the temperature is high, all configurations tend to be equally distributed and when it gradually decreases to zero, global energy minima is achieved. Gibbs energy formalism has the added advantage that if the likelihood term is given by an exponential, and the prior is obtained through a MRF model, the posterior probability continues to be a Gibbsian. This makes the MAP estimation problem equivalent to an energy minimization.

2.4 Markov-Gibbs Equivalence

Markov Random Field (MRF) is characterized by its local property (the Markovianity) whereas Gibbs Random Field (GRF) is characterized by its global property (the Gibbs

distribution). Hammersley-Clifford's famous theorem [2] (unpublished by the original authors) states that "*given the neighborhood structure η of the model, for any set of sites within the lattice S, their associated contribution to the Gibbs energy function should be non zero, if and only if the sites form a clique; a random field's having the Markov property is equivalent to its having a Gibbs distribution*". This theorem establishes the equivalence of these two types of properties and provides a very general basis for the specification of MRF joint distribution function. Many have been used throughout the literature [10]. The difficulties inherent in the MRF formulation are eliminated by use of this equivalence which are as follows:

1. Readily availability of joint distribution of random field
2. Obtaining local characteristics regardless of inconsistency
3. Characterizing the Gibbs Distribution model with few parameters

By use of this equivalence MRF theory provides a mathematical foundation for solving the problem of making a global inference using local information. It follows from the above equivalence that the local characteristics of the MRF are readily obtained from the joint distribution in 2.1 as

$$\begin{aligned} P(X_{i,j} = x_{i,j} \mid X_{k,l} = x_{k,l}\ (k,l) \in S,\ (k,l) \neq (i,j)) \\ = P(X_{i,j} = x_{i,j} \mid X_{k,l} = x_{k,l},\ (k,l) \in \eta_{i,j}) \\ = \frac{e^{-\sum_{c \in C} V_c(x)}}{\sum_{x_{i,j} \in S} e^{-\sum_{c \in C} V_c(x)}} \end{aligned} \tag{2.4}$$

2.5 Gibbs Sampler

To implement the Relaxation algorithm, Geman and Geman [1] developed the Gibbs Sampler to explore the energy surface. The interpretation of the Theorems derived by them are as follows.

- The interpretation of the **Theorem A** is *"At a constant temperature, if each site of an image lattice is visited infinite times, as time to infinity, the configuration X will*

be a sample from the Gibbs distribution and this distribution is independent of the initial configuration".

- The interpretation of **Theorem B** is *"To reach equilibrium state with lowest energy, the temperature is forced to decrease slowly. As time to infinity, X will be a sample from the Gibbs distribution at temperature absolute zero degree or the Gibbs distribution with minimum energy".*

The Gibbs sampler works by updating each random variable individually, but conditional on the states of the surrounding sites. The sequential implementation corresponding to a raster scan is used for Gibb's sampler. The state of image evolves by discrete changes. So for convenience time is discretized, say $t = 1, 2, 3.....$ At a given time, each site $x_{i,j}$ is represented by a random variable $X_{i,j}(t)$ with values in $G = 0, 1, 2...., n-1$. Hence the total configuration of the image is $X(t) = \{x_{i,j}(t)\};\ i, j \in S$. The starting configuration $X(0)$ is arbitrary and at any time t, the total configuration $X(t)$ evolves due to state change of individual site. At any instant of time only one site undergoes (possible) change. So the state at any two consecutive instant of time t and $t-1$ can differ by at most one coordinate. If $n_1, n_2, ...$ be the sequence in which the sites are visited for replacement; thus $n_t \in S$ and $X_{i,j}(t) = X_{i,j}(t-1), i \neq n_t$. For replacement at each site a sample is drawn from its local characteristics. In other words, a state $x \in Gn_t$ is chosen from the conditional distribution of X_{n_t}. Given the observed states of the neighboring sites. All other sites remaining unchanged, the change in total energy is the changes due to change at site n_t with respect to its neighborhood. Let $U(t-1)$ is the old energy and $U(t)$ be the new one. If $U(t)$ is found to be less than $U(t-1)$, then the change is accepted; otherwise it is accepted with a probability to avoid the sampling to stuck in a local minimum. When all the sites of the image are visited once, one iteration is said to be completed.

2.6 Hidden Markov Model (HMM)

HMM is a finite set of states, each of which is associated with a (generally multidimensional) probability distribution. Transitions among the states are governed by a set of probabilities called transition probabilities. In a particular state an outcome or observa-

tion can be generated, according to the associated probability distribution. It is only the outcome, not the state visible to an external observer and therefore states are "hidden" to the outside; hence the name Hidden Markov Model. Basically it is an extension of observable Markov Model. The observation is a probabilistic function (discrete or continuous) of a state. All observations are dependent on the state that generated them, not on the neighboring observations. Hidden Markov models exploit the "locality" of physical properties of a system. They are often used to construct models of physical systems when the information about the system is gathered using an apparatus that distorts the physical reality being observed in some manner. Markov Random Field (MRF) theory provides a basis for modeling contextual constraints. It is commonly accepted that the pixel intensities in an image exhibit high spatial statistical interdependence, i.e., background pixels have a high probability of occurring next to other background pixels. Likewise, pixels generally lie adjacent to other pixel. The key assumption is that a high spatial interdependence present in the image field can be easily incorporated into a MRF model. A description of MRF is given in section 2.2 before.

MRF is a multidimensional extension of Markov chain, but the generalization is complicated by the lack of a natural ordering of pixels in multidimensional space. Hidden Markov fields are a natural generalization of the HMM that have proved essential to the development of modern speech recognition, but again the multidimensional nature of the signals makes them inherently more complicated to handle. This added complexity contributed to the long time required for the development of successful methods and applications. Here, the output of the process is the set of states at each instant of time, when each state corresponds to an observable event and also the output in any given state is not random (deterministic). The above stochastic process could be called an observable Markov model since the output of the process is the states at each instant of time, where each state corresponds to a physical (observable) event. In this section we extend the concept of Markov models to include the case where the observation is probabilistic function of the state, i.e. stochastic process with an underlying stochastic process that is not observable (hidden), but can only be observed through another state of stochastic processes that produce the sequence of observations. Here, an example is produced for an

better insight into the theory of Hidden Markov model.

2.7 HMRF Image Model

The concept of a hidden Markov random field (HMRF) model is derived from hidden Markov models (HMM), which are defined as stochastic processes generated by a Markov chain whose state sequence cannot be observed directly, only through a sequence of observations. Each observation is assumed to be a stochastic function of the state sequence. The underlying Markov chain changes its state according to a $l \times l$ transition probability matrix, where l is the number of states. HMMs have been applied successfully to speech recognition and handwritten script recognition. Since original HMMs were designed as 1D Markov chains with first order neighborhood systems, it can not directly be used in 2D/3D problems such as image segmentation. Here, we consider a special case of a HMM, in which the underlying stochastic process is a Markov random field (MRF), instead of a Markov chain, therefore not restricted to 1D. We refer to this special case as a hidden Markov random field (HMRF) model [17]. Mathematically, an HMRF model is characterized by the following:

- **Hidden Random Field:** The Random field $X = X_i, i\epsilon S$ is an underlying MRF assuming values in a finite state space L with probability distribution. The state of X is unobservable.

- **Observable Random Field:** $Y = Y_i, i\epsilon S$ is a random field with a finite state space D. Given any particular configuration, $x\epsilon X$ every Y_i follows known conditional probability distribution $p(y_i|x_i)$ of the same functional form $f\left(y_i; \theta_{x_i}\right)$, where θ_{x_i} is the involved parameter. This distribution is called the emission probability function and Y is also referred to as the emitted random field.

- **Conditional Independence:** For any $x\epsilon X$, the random variables Y_i are conditional independent.

  Based on the above, we can write the joint probability of (X, Y) as

$$P(y, x) = P(y \mid x)P(x) = P(x) \prod_{i \in S} P(y_i \mid x_i) \tag{2.5}$$

According to the local characteristics of MRFs, the joint probability of any pair of (X_i, Y_i), given X_i's neighborhood configuration X_{N_i}, is

$$P(y_i, x_i \mid x_{N_i}) = P(y_i \mid x_i)P(x_i \mid x_{N_i}) \tag{2.6}$$

Thus, we can compute the marginal probability distribution of X_{N_i}, Y_i dependent on the parameter set (in this case, we treat as a random variable) and

$$\begin{aligned} p(y_i \mid x_{N_i}, \theta) &= \sum_{l \in L} p(y_i, l \mid x_{N_i}, \theta) \\ &= \sum_{l \in L} f(y_i, \theta_l) P(l \mid x_{N_i}) \end{aligned} \tag{2.7}$$

where $\theta = \{\theta_l, l \in L\}$. . This is the hidden Markov random field (HMRF) model. The concept of an HMRF is different from that of an MRF in the sense that the former is defined with respect to a pair of random variable families (X, Y) while the latter is only defined with respect to X. More precisely, the HMRF model can be described by the following:

1. $X = X_i, i \epsilon S$ - hidden MRF, with prior distribution p(x);
2. $Y = Y_i, i \epsilon S$ - observable random field, with emission probability distribution $p(y_i|x_i)$ for each y_i;
3. $\theta = \{\theta_l, l \in L\}$ - the set of parameters involved in the above distributions.

If we assume the random variables X_i are independent of each other, which means that for $\forall l \epsilon L$ and $i \epsilon S$, we have $p(l \mid x_{N_i}) = p(l) = w_l$, then equation reduces to

$$p(y \mid \theta) = \sum_{l \in L} w_l f(y; \theta_l) \tag{2.8}$$

This is the definition of the finite mixture model. Therefore a FM model is a degenerate special case of an HMRF model. It is obvious from the above that the fundamental difference between the FM model and the HMRF model lies in their different spatial properties. The FM model is spatially independent whereas the HMRF model may be spatially dependent. Therefore, the HMRF model is more flexible for image modeling in the sense that it has the ability to encode both the statistical and spatial properties of an image. With

a Gaussian emission distribution, the FM model is usually known as the finite Gaussian Mixture (FGM) or finite normal mixture (FNM) model. More specifically, the observable random variables have the following density function:

$$p(y \mid \phi) = \sum_{l \in L} w_l g(y; \theta_l) \tag{2.9}$$

where

$$g(y; \theta_l) = \frac{1}{\sqrt{2\pi\sigma_l^2}} exp\left(-\frac{(y_i - \mu_l)^2}{2\sigma_l^2}\right) and\ \theta_l = (\mu_l, \sigma_l)^T \tag{2.10}$$

Similarly, the HMRF model with a Gaussian emission distribution can be specified as:

$$p(y_i \mid x_{N_i}, \theta) = \sum_{l \in L} g(y_i, \theta_l) P(l \mid x_{N_i}) \tag{2.11}$$

where g and θ_l are defined as in . This type of HMRF as the Gaussian hidden Markov random field (GHMRF) model.

2.8 Fuzzy Clustering

2.8.1 Classical Sets

A classical set is a set that has a crisp boundary. For example, a classical set X of real numbers greater than 6 is expressed as

$A = \{x \mid x > 6\}$

In this set of real numbers there is a clear unambiguous boundary 6 such that if x is greater than this number. In this case x either belongs to this set A or it does not belong to this set. These types of sets are called Classical Sets. Classical sets are an important tool in mathematics and computer science but they do not reflect the nature of human concepts and thought. In contrast to a classical set, a fuzzy set is a set without crisp boundaries. That is, the process of an element belongs to a set to does not belong to a set is gradual. This transition is decided by the membership function of a fuzzy dataset. Real life problems have data which most of the time has a degree of trueness or falseness that is the data cannot be expressed in terms of classical set. A good example of this is; the same set A is a set of tall basketball players. According to the classical set logic a player 6.01 ft tall is considered to be tall whereas a player 5.99 ft tall is considered to be short.

2.8.2 Fuzzy Sets and Membership Function

Membership functions give the flexibility in modeling commonly used linguistic terms such as the water is hot or the temperature is high to fuzzy sets. Zadeh (1965) points out that, this imprecise data set information plays an important role in human approach to problem solving. It is important to note that fuzziness in a dataset comes does not come from the randomness of the elements of the set, but from the uncertain and imprecise nature of the abstract thoughts and concepts. If X is a collection of objects denoted by x, then a fuzzy set $A \epsilon X$ is defined as a set of ordered pairs $A = (x, \mu A(x)) \mid x \epsilon X$, Where $\mu A(x)$ is called the membership function (MF) for the fuzzy set A. The membership function maps each element of X to a membership grade between 0 and 1. If the value of the membership function is restricted to either 0 or 1, then A is reduced to a classical set and $\mu A(x)$ is the characteristic function of A. Usually X is referred to as the universe of discourse and may consist of discrete objects or continuous space.

2.8.3 Data Clustering Algorithms

Clustering of numerical data forms the basis of various classification and system modeling algorithms. The purpose of clustering is to identify natural groupings of data from a large data set to produce a concise representation of a system's behavior. Clustering algorithms are not only used to organize and categorize data, but are helpful in data compression and model construction. Clustering partitions the data set into several groups such that the similarity within a group is larger than among the groups. To achieve such partitions it is essential to have a similarity metrics that takes two input vectors and returns a value reflecting their similarity. As most of the similarity metrics are sensitive to the range of elements in the input vectors, each of the input variables must be normalized or scaled down. Clustering techniques are broadly classified as hard clustering and fuzzy clustering.

2.8.4 Fuzzy C-means Clustering Algorithm

Fuzzy C-means clustering (FCM) algorithm, also known as fuzzy Isodata, is a data clustering algorithm in which each data point belongs to a cluster to a degree specified by a membership grade. Bezdek *et al.* proposed this algorithm in 1973 as an improvement to

K-means algorithm also known as the hard C-means algorithm. Hard k-means algorithm executes a sharp classification, in which each object is either assigned to a class or not. The application of fuzzy clustering to the dataset function allows the class membership to have several classes at the same time but with different degrees of membership function ranging from 0 to 1. Fuzzy c-means (FCM) is a method of clustering which allows one piece of data to belong to two or more clusters. It is based on minimization of the following objective function

$$J_m = \sum_{i=1}^{N}\sum_{j=1}^{C} u_{ij}^{m} \parallel x_i - c_j \parallel^2 \tag{2.12}$$

where m the fuzzy factor, any real number greater than 1, j is the number of cluster decided by the user, u_{ij} is the degree of membership of x_i in the cluster j, x_i is the i^{th} of d-dimensional measured data namely throughput, storage level and volume, c_j is the d-dimension center of the cluster, and $\parallel x_i - c_j \parallel^2$ is any norm expressing the similarity between the measured data and the center. Fuzzy partitioning is carried out through an iterative optimization of the objective function shown above, with the update of membership matrix u_{ij} and the cluster centers c_j by,

$$u_{ij} = \frac{1}{\sum_{k=1}^{C}\left(\frac{\|x_i - c_j\|^2}{\|x_i - c_k\|^2}\right)^{\frac{2}{m-1}}} \quad and, \tag{2.13}$$

$$c_j = \frac{\sum_{i=1}^{N} u_{ij}^{m} x_i}{\sum_{i=1}^{N} u_{ij}^{m}} \tag{2.14}$$

This iteration will stop when max u_{ij} , $\mid u_{ij}^{(k+1)} - u_{ij}^{(k)} \mid \leq \varepsilon$ where ε is a termination criterion between 0 and 1 and usually set to 0.02 whereas k is the iteration steps. This procedure converges to a local minimum or a saddle point of J_m. The algorithm is composed of the following steps mentioned below.

The steps of the algorithm are as follows:

1. Fix c, $2 \leq c \leq N, m, 1 \leq m \leq \ infty$ initialize The class prototypes V

2. Compute the partition matrix
   $u_{ik}^{*} = \frac{1}{\sum_{j=1}^{c}(\frac{d_{ik}^2}{d_{ik}^2})^{\frac{1}{m-1}}}$

3. Compute the fuzzy cluster centers V using
   $V_i^* = \frac{\sum_{k=1}^{N} u_{ik}^m x_k}{\sum_{k=1}^{N} u_{ik}^m}$
4. Compute the change in the cluster centers values using a appropriate norm; if the change is small, stop. Else return to step 2.

2.8.5 Fuzzy Factor and Ideal Number of Clusters

The fuzzy factor m was is also known as fuzzifier. As the value of m approaches 1 the clusters formed tend to be hard and as the value of m tends to infinity the obtained clusters tend to go in a the fuzziest state. There is no theoretical justification on the value of m but is usually set to 2 and in a more generalized form tends to be between 1.5 and 3.

The number of clusters for a certain type of data will vary based on the data partition desired. The number of clusters can vary between 2 to infinity.

2.8.6 Significance of Membership Function in Cluster Analysis

As discussed in the earlier section, data are bound to each cluster by means of a membership function, which represents the fuzzy behavior of this algorithm. To do that, we build an appropriate matrix named U whose factors are numbers between 0 and 1, and represent the degree of membership between data and centers of clusters. In the FCM approach, instead, the same given datum does not belong exclusively to a well-defined cluster, but it can be placed in a middle way. In the case of FCM, the membership function follows a smoother line to indicate that every datum may belong to several clusters with different values of the membership coefficient.

Chapter 3

Unsupervised image segmentation and intensity in-homogeneity correction using Biased HMRF model

3.1 Introduction

Segmentation of brain MR images into different tissue classes, especially gray matter (GM), white matter (WM) and cerebrospinal fluid (CSF) is an important task for computer-aided image analysis. The MR image is degraded considerably by electronic noise, the bias field (intensity inhomogeneities in the Radio Frequency field) and partial volume effect during its acquisition. Often, model based approaches have been adhered to obtain proper segmentation of degraded MR images. In this regard, the problem is casting as a pixel labeling problem and the segmentation problem reduces to estimation of pixel labels. With an aim to develop the automatic segmentation method, the model based problem is viewed as unsupervised one. In this chapter, Hidden Markov Random field (HMRF) model proposed by Zhang *et al.* has been employed to formulate the unsupervised segmentation problem [17]. HMRF models have been used to model the tissue classes of the observed degraded image. The *a priori* class labels are modelled as Markov Random field (MRF) model. The model parameters, the number of class labels and the image labels are assumed to be unknown. The problem becomes an incomplete data problem and is formulated in Expectation-Maximization (EM) framework motivated by Dempster *et al.* [11]. Using this HMRF-EM framework proposed by Zhang *et al.*, the segmentation of brain MR images are obtained without any significant improvement in segmentation

accuracy and computational time for the cleaner data [17]. In order to overcome this bottleneck, we have proposed the Biased HMRF model with a biased neighborhood system. The energy function associated with the *a priori* model is modified by an biased amount of internal field. By this modification, the effect of internal structure is incorporated in the new model. With the BHMRF-EM framework, the image labels and the model parameters are estimated recursively in E- step and M-step respectively. MAP estimation of image are obtained using Iterated Conditional Mode (ICM) algorithm.

In this chapter, we have also extended the BHMRF-EM framework to incorporate the intensity inhomogeneity correction task. In this regard, we have developed a modified BHMRF-EM algorithm for estimation of bias field, bias field corrected image labels as well as model parameters.

3.2 Image model

Let X denotes the random field associated with the labels of the original image and x denotes the realization of that. The label process X is assumed to be MRF. We have already described about MRF model in chapter 2. Hence, the joint distribution can be expressed as

$$P(X = x \mid \phi) = \frac{1}{Z} e^{-U(x,\phi)} \tag{3.1}$$

where Z is the partition function, ϕ denotes the clique parameter vector. Let X is the unobservable and Y denotes the observed random field. Y_i denotes the pixel intensity and i is the individual site in S. It is assumed that for any realization x, the random variables are conditionally independent.

$$P(Y = y \mid X = x) = \prod_{i \in S} P(Y_i = y_i \mid X_i = x_i) \tag{3.2}$$

where S denote the set of all sites of the image. The joint probability of (X, Y) can be expressed as

$$\begin{aligned} P(Y = y, X = x) &= P(Y = y \mid X = x) P(X = x) \\ &= P(X = x) \prod_{i \in S} P(Y_i = y_i \mid X_i = x_i) \end{aligned} \tag{3.3}$$

Thus the marginal probability distribution of Y_i can be expressed by using the local characteristics of MRF as

$$P(Y_i = y_i \mid X_{\eta_i} = x_{\eta_i}, \theta) = \sum_{l \in L} P(Y_i = y_i \mid X_{\eta_i} = x_{\eta_i}, \theta)$$
$$= \sum_{l \in L} P(y_i, l, \theta_l)\ P(l \mid x_{\eta_i}) \tag{3.4}$$

where $\theta = \{\theta_l, l \in L\}$, L denotes the number of class labels. $\theta_l = [\mu_l, \sigma_l]$ is the model parameters for each class, μ_l, σ_l are the mean and covariance of each class of the image, l is corresponding class label. (3.4) is referred as the Hidden Markov Random Field model. With Gaussian distribution, (3.4) can be expressed as

$$p(Y_i = y_i \mid X_{\eta_i} = x_{\eta_i}, \theta) = \sum_{l \in L} g(y_i, \theta_l) P(l \mid x_{\eta_i}) \tag{3.5}$$

where

$$g(y_i; \theta_l) = \frac{1}{\sqrt{2\pi\sigma_l^2}} exp\left(-\frac{(y_i - \mu_l)^2}{2\sigma_l^2}\right) and\ \theta_l = (\mu_l, \sigma_l)^T \tag{3.6}$$

3.3 Proposed Biased HMRF model

As described in chapter 2, the spatial constraint encoded through the HMRF model aims to solve the data with noise and local variations. The segmentation of brain MR images using HMRF-EM scheme proposed by Zhang *et al.* is obtained without any significant improvement in segmentation accuracy and time complexity with cleaner data than a noisy data [17]. To overcome this limitation, a modified model is proposed to improve the neighborhood system of HMRF model by better characterizing the structure of human brain. While taking different weightings of MR images, several difficulties arises due to imaging artifacts. In human brain, WM and CSF are only adjacent to each other in the regions around the ventricles. In most of the T1 weighted brain regions it is a big task to differentiate between WM and CSF due to low contrast in intensity. Also in case of T2 weighted images, the contrast between GM and WM is negligible and CSF has high contrast. This information is encoded into segmentation framework and BHMRF model is proposed. In HMRF model, it is considered that a pixel is more likely to be a certain tissue type if the neighboring pixels are of that type. Based on this assumption, the energy

function is defined as $U_x = \sum_{c \in C} V_c(x)$ and the clique potential is normally defined as

$$V_c(x) = -\delta \quad if|x_i - x_j| = 0 \tag{3.7}$$
$$= \delta \quad if|x_i - x_j| \neq 0$$

In our proposed BHMRF model, the clique potential is modified as

$$V_c(x) = -\delta(x_i - x_j) - \rho * \delta(\mid x_i - x_j \mid -1) \tag{3.8}$$

where ρ is biased factor. In this way, a CSF pixel in the neighborhood system can also contribute to the central pixel being a gray matter but not being a white matter, and vicev-ersa. Therefore, this reduces the probability that WM and CSF are adjacent to each other.

3.4 MAP estimation of image labels

Let X be the random field associated with the noise free class label and x be the realization of the same. X is modeled as a MRF.

Let Y denote the observed image random field and y be the realization of it. Y is modeled as Hidden Markov Random Field (HMRF).

Let θ be the associated model parameters. In the pixel labeling problem, let x^* denote the true but unknown labeling configuration and $\hat{x}$ denote the estimate for x^*.

x^* is the realization of random field X, which is modeled as MRF. The observed image y is a realization of proposed BHMRF framework. The problem is to recover x^* from the observed image y. The following optimality criterion is adopted,

$$\hat{x} = arg\max_x P(X = x \mid Y = y, \theta) \tag{3.9}$$

where $P(X = x \mid Y = y, \theta)$, is the posterior probability distribution of X, the model parameters for each class $\theta_l = [\mu_l, \sigma_l]$ are taken from histogram analysis. Since X is unknown, the posteriori probability distribution $P(x \mid y, \theta)$ can not be evaluated. Hence, using Baye's rule, (3.9) can be expressed as

$$\hat{x} = arg\max_x \frac{P(y \mid x, \theta)P(x)}{P(y)} \tag{3.10}$$

Since Y is known, the denominator of (3.10) is a constant. Thus, (3.10) can be written as

$$\hat{x} = arg\max_x P(y \mid x, \theta)P(x) \tag{3.11}$$

Since,X is MRF, the prior probability distribution in (3.11) is given as $P(x) = \frac{1}{Z}e^{-U(x)}$. It is also assumed that the pixel intensity y_i follows a Gaussian distribution with parameters $\theta_l = \{\mu_l, \sigma_l\}$. Given the class label $x_i = l$,

$$P(y_i \mid x_i) = \frac{1}{\sqrt{2\pi\sigma_l^2}} exp\left(-\frac{(y_i - \mu_l)^2}{2\sigma_l^2}\right) \tag{3.12}$$

Using the assumption of conditional independence

$$P(y \mid x) = \prod_{i \in S} P(y_i \mid x_i) = \prod_{i \in S}\left[\frac{1}{\sqrt{2\pi}}\left(-\frac{(y_i - \mu_{x_i})^2}{2\sigma_{x_i}^2} - log(\sigma_{x_i})\right)\right] \tag{3.13}$$

(3.13) can be expressed as

$$P(y \mid x) = \frac{1}{Z'} exp(-U(y \mid x))) \tag{3.14}$$

$$U(y \mid x) = \sum_{i \in S} U(y_i \mid x_i) = \sum_{i \in S}\left[\frac{(y_i - \mu_{x_i})^2}{2\sigma_{x_i}^2} + log(\sigma_{x_i})\right]$$

and $Z' = (2\pi)^{N/2}$. Using the above, (3.11) can be expressed as

$$\hat{x} = arg\max_x \left[\frac{1}{Z} exp(-U(x)) \frac{1}{Z'} exp(-U(y \mid x))\right] \tag{3.15}$$

(3.15) is equivalent to minimizing the following

$$\hat{x} = arg\min_x [U(y \mid x) + U(x)]$$

where $U(y \mid x)$ and $U(x)$ are the energy functions corresponding to the conditional distribution $P(Y = y \mid X = x, \theta)$ and the *a priori* class distribution $P(X = x)$ respectively.

$$\hat{x} = arg\min_x \left[\sum_{i \in S}\left[\frac{(y_i - \mu_{x_i})^2}{2\sigma_{x_i}^2} + log(\sigma_{x_i})\right] + \sum_{c_i \in C} V_c(x, \phi)\right] \tag{3.16}$$

The MAP estimate of $\hat{x}$ in (3.16) is obtained by employing the ICM algorithm.

3.5 Iterated Conditional Mode Algorithm

Since it is difficult to maximize the joint probability of an MRF, Besag *et al.* proposed a deterministic algorithm called Iterated Conditional Modes (ICM) which maximizes local conditional probabilities sequentially. The ICM algorithm uses the greedy

strategy in the iterative local maximization. Given the data y and the other labels $x_{S-i}^{(k)}$, the algorithm sequentially updates each $x_i^{(k)}$ into $x_i^{(k+1)}$ by maximizing $P(x_i \mid y, x_{S-i})$, the conditional probability, with respect to x_i. Two assumptions are made in calculating $P(x_i \mid y, x_{S-i})$:

1. The observation components $y_1,\ y_2,\ y_3 ... \ y_m$ are conditionally independent given x and each y_i has the same known conditional density function $p(y_i \mid x_i)$ dependent only on x_i. Thus
$$p(y \mid x) = \prod_i p(y_i \mid x_i) \tag{3.17}$$
2. The second assumption is that x depends on the labels in the local neighborhood, which is the Markovianity.

From the two assumptions and the Bayes theorem, it follows that

$$P(x_i \mid y, x_{S-i}) \propto p(y_i \mid x_i) P(x_i \mid x_{N_i}) \tag{3.18}$$

Obviously, $P(x_i \mid y_i,\ x_{N_i}^k)$ is much easier to maximize than $P(x \mid y)$, which is the point of ICM. Maximizing 3.18 is equivalent to minimizing the corresponding posterior energy using the following rule.

$$x_i^{k+1} \leftarrow \underset{x_i}{arg\,max}\ U(x_i \mid y_i,\ f_{N_i}^{(k)}) \tag{3.19}$$

The result obtained by ICM depends very much on the initial estimator $x^{(0)}$ and the ICM is locally convergent [10].

3.6 Parameter estimation

The problem of parameter estimation is regarded as an incomplete-data problem in this scheme. Expectation-Maximization (EM) algorithm is a potential tool to handle the incomplete data problem. It solves the Maximum-Likelihood (ML) estimation of the model parameter.

E- step estimates the expected value of unknown variables, given the current parameter estimate. The MAP estimates of the image labels are obtained by the Iterated Conditional Model (ICM) algorithm.

M- step re-estimates the distribution parameters to maximize the likelihood of the data, given the expected estimates of the unknown variables.

The E and M steps are iterated recursively till the parameters converge to the optimal solution. Experiments on both real and synthetic brain MR images show that the segmentation results of proposed BHMRF-EM scheme outperforms the existing HMRF-EM scheme.

3.7 Expectation-Maximization algorithm

To perform unsupervised segmentation of image data, where both the class labels and model parameters are unknown, a method of concurrently estimating the underlying class labels of the image and associated model parameters is required. Alternately, the problem may be viewed as an incomplete data problem. Following this approach, the complete data comprises $Z = \{X, Y\}$, where Y is observed and X is the underlying or hidden component. Applying this to image segmentation, Y comprises the observed noisy image and X is a lattice on which the segmentation of the image is defined. With the complete data set $Z = \{X, Y\}$, a joint density function $P(Z \mid \theta) = P(X, Y \mid \theta) = \pounds(\theta \mid X, Y)$ is specified, which is the complete-data likelihood function and $\pounds(\theta \mid Y)$ is the incomplete-data likelihood function, θ is the set of parameters governing the observed data.

The Expectation-Maximization (EM) algorithm was first proposed by Dempster *et al.* as an iterative maximum likelihood procedure for parameter estimation from incomplete data [11]. The methodology has been extensively applied to the problem of image segmentation and specifically to brain MR image segmentation [17]. Since the EM algorithm yields maximum-likelihood estimate for the hidden data, the unsupervised segmentation problem is solved using the following steps.

- Obtain an initial parameter estimate $\theta^{t=0}$ using an initial guess.
- Use the EM algorithm to find the maximum-likelihood parameter estimate $\hat{\theta}$.
- Use a supervised algorithm to obtain the maximum-likelihood or maximum *a posteriori* estimate for the hidden data.

The EM algorithm is an iterative process, where each iteration consists of two steps. The first of these finds an expression for the expected value of the complete data log-likelihood $\log P(X, Y \mid \theta)$ with respect to the hidden data X, given the observed data Y and the current parameter estimates. The second step maximizes this expectation to estimate the parameters.

Expectation Step:

The *E-step* calculates the conditional expectation

$$Q(\theta \mid \theta^t) = E\{\log P(X, Y \mid \ \theta) \ \mid Y, \ \theta^t\} = \sum_{x \in X} P(X \mid Y, \ \theta^t) \log P(X, Y \mid \ \theta) \quad (3.20)$$

where the parameter set $\theta = \{\mu_l, \sigma_l \mid l \in L\}$ is the new parameter that is optimized to increase Q and θ^t is the current parameter estimate that is used to evaluate the expectation. For a given X, Q function is formulated as

$Q(\theta \mid \theta^t) = \sum_{i \in S} P(x_i \mid y_i, \ \theta^t) \ \log P(x_i, y_i \mid \theta)$

$= \sum_{i \in S} \sum_{l \in L} P^t(l \mid y_i) \ \log\{P(y_i \mid x_i) \ P(x_i \mid x_{\eta i}, \ \theta)\}$

$= \sum_{i \in S} \sum_{l \in L} P^t(l \mid y_i)\{ \ \log P(y_i \mid l_i) + \log P(l \mid x_{\eta i}, \ \theta)\}$

$= \sum_{i \in S} \sum_{l \in L} P^t(l \mid y_i)\{\log(\frac{1}{\sqrt{2\pi\sigma_l^2}} exp(-\frac{(y_i - \mu_l)^2}{2\sigma_l^2})) + \log P(l \mid x_{\eta i}, \ \theta)\}$

$= \sum_{i \in S} \sum_{l \in L} P^t(l \mid y_i)\{-\frac{(y_i - \mu_l)^2}{2\sigma_l^2} - \log \sigma_l - 0.5 \log 2\pi\} + \log P(l \mid x_{\eta i}, \ \theta)$

$= \sum_{i \in S} \sum_{l \in L} P^t(l \mid y_i)\{-\frac{(y_i - \mu_l)^2}{2\sigma_l^2} - \log \sigma_l + \log P(l \mid x_{\eta i}, \ \theta)\}$

$+P^t(l \mid y_i)(-0.5 \log 2\pi)$

$= \sum_{i \in S} \sum_{l \in L} P^t(l \mid y_i)(W + C)$

where $W = -\frac{(y_i - \mu_l)^2}{2\sigma_l^2} - \log \sigma_l + \log P(l \mid x_{\eta i}, \ \theta)\}$ and $C = -0.5 \log 2\pi$

Maximization Step:

The *M-step* maximizes $Q(\theta \mid \theta^t)$ to obtain the next estimate with respect to the first argument of the Q function which is the conditioner of the complete data likelihood.

$$\theta^{t+1} = \overset{arg\ max}{\theta}\ Q(\theta \mid \theta^t) \tag{3.21}$$

- Maximizing Q-function with respect to μ_l :

  $$\frac{\partial Q(\theta \mid \theta^t)}{\partial \mu_l} = 0 \text{ or } \frac{\partial}{\partial \mu_l}\{P^t(l \mid y_i)(-\frac{(y_i - \mu_l)^2}{2\sigma_l^2})\} = 0$$

  $$\text{or } P^t(l \mid y_i) y_i = P^t(l \mid y_i)\mu_l$$

  $$\text{or } \mu_l^{t+1} = \frac{\sum_{i \in S} P^t(l \mid y_i)\ y_i}{\sum_{i \in S} P^t(l \mid y_i)}$$

- Maximizing Q-function with respect to σ_l^2 :

  $$\frac{\partial Q(\theta \mid \theta^t)}{\partial \sigma_l^2} = 0$$

  $$\text{or } \frac{\partial}{\partial \sigma_l^2}\{P^t(l \mid y_i)(-\frac{(y_i - \mu_l)^2}{2\sigma_l^2}) - \log \sigma_l\} = 0$$

  $$\text{or } P^t(l \mid y_i)(y_i - \mu_l)^2 = P^t(l \mid y_i)\sigma_l^2$$

  $$\text{or } {\sigma_l^2}^{t+1} = \frac{\sum_{i \in S} P^t(l \mid y_i)(y_i - \mu_l)^2}{\sum_{i \in S} P^t(l \mid y_i)}$$

The EM algorithm consists of choosing an initial parameter $\theta^{(t=0)}$, then iterates the E-step and M-step successively until the parameter θ converges to $\hat{\theta}$.

3.8 Joint estimation of image labels and model parameters using Biased HMRF-EM framework

In the EM framework, image labels as well as the model parameters are estimated recursively until the model parameters converge to the optimal ones. The image label estimates

x^0 are obtained by computing the MAP estimation using the initial model parameters that are taken from histogram of the noisy images. The MAP estimate is obtained by ICM algorithm described in section 3.5. Using the label estimates x^0 and the degraded image Y, the model parameters θ^1 is estimated by maximizing the $Q(\theta, \theta^t)$. This recursive process is repeated till the parameters converge to optimal values. The optimal values of the parameters are used to obtain the desired segmentation result. The image labels and model parameters are obtained by the BHMRF-EM algorithm. The salient steps of the proposed BHMRF-EM algorithm is described below.

BHMRF-EM Algorithm:

1. Perform the initial parameter estimation and segmentation.

2. Calculate the likelihood distribution

$$p^{(t)}\left(y_i|x_i\right) = g^{(t)}\left(y_i; \theta\left(x_i\right)\right)$$

3. Estimate the class labels by MRF-MAP estimation

$$x^{(t)} = arg\max_{x \in X} P(y \mid x, \theta^t + P(x))$$

   ICM algorithm is used to estimate the class labels.

4. Calculate the posterior distribution

$$p^{(t)}\left(l \mid y_i\right) = \frac{g^{(t)}\left(y_i; \theta_l\right) p^{(t)}\left(l \mid x_{N_i}\right)}{p(y_i)}$$

5. Update parameters by

$$\mu_l^{(t+1)} = \frac{\sum_{i \in s} p^{(t)}\left(l \mid y_i\right) y_i}{\sum_{i \in s} p^{(t)}\left(l \mid y_i\right)} \qquad \sigma_l^{(t+1)^2} = \frac{\sum_{i \in s} p^{(t)}\left(l \mid y_i\right)\left(y_i - \mu_l\right)^2}{\sum_{i \in s} p^{(t)}\left(l \mid y_i\right)}$$

6. $t \leftarrow t + 1$ and repeat from 2 until enough iterations have been performed.

3.9 Bias Field estimation

In this chapter, we have taken the bias field model proposed by Wells *et al.* [19]. Let the observed intensity of the given image is $I = (I_1, ..., I_N)$ and
the true intensity of the given image is $I = (I_1^*, ..., I_N^*)$

The degradation effect of the bias field at pixel $i, 1 \leq i \leq N$ can be expressed as

$$I_i = I_i^* \times b_i \tag{3.22}$$

The bias field effect is treated as an additive artifact after logarithmic transformation.

Let the observed log transformed intensity is Y and the true log transformed intensity is Y^*. Then $Y = Y^* + B$. The bias field B is modelled with Gaussian prior probability density $p(B) = G_{\psi_B}(B)$ where ψ_B is the $N \times N$ covariance matrix. Assuming the true intensity value at pixel i following the Gaussian distribution with parameter $\theta(x_i) = (\mu_{x_i}, \sigma_{x_i})$ with given class labels x is

$$p(y_i^*|x_i) = g(y_i^*; \theta(x_i)) \tag{3.23}$$

With bias field b_i, the distribution can be written in terms of the observed intensity y_i as,

$$p(y_i|x_i, B) = g(y_i - b_i; \theta(x_i)) \tag{3.24}$$

Thus, the intensity distribution in terms of Gaussian mixture model is

$$p(y_i|B) = \sum_{j \in L} g(y_i - b_i; \theta(j) P(j)) \tag{3.25}$$

The optimal estimate of the bias field is obtained using the MAP principle as

$$\hat{B} = arg \max_B P(Y \mid B) p(B) \tag{3.26}$$

A zero gradient condition is to assess this maximum as

$$W_{ij} = \frac{p(y_i|x_i, \beta) p(x_i = j)}{p(y_i|\beta)} \tag{3.27}$$

$$b_i = \frac{[FR]_i}{[F\psi^{-1}1]_i}, \quad with \quad 1 = (1, 1, , 1)^T$$

where R is the mean residual for pixel i

$$R_i = \sum_{j \in L} \frac{W_{ij}\,(y_i - \mu_j)}{\sigma_j^2} \tag{3.28}$$

ψ is the mean inverse covariance

$$\psi_{ik}^{-1} = \sum_{j \in L} W_{ij}\sigma_j^2, \quad if \quad i = k$$
$$0 \quad otherwise \tag{3.29}$$

and F is the low-pass filter. W_{ij} is the posterior probability that pixel i belongs to class j given the bias field estimate. The E step assumes that the bias field is known and calculates the posterior tissue class probability W_{ij}. In the M step, the bias field B is estimated given the estimated W_{ij} in the E step. Once the bias field is obtained, the original intensity I^* is restored by dividing I by the inverse log of B. Initially, the bias field is assumed to be zero.

3.10 Joint estimation of bias field, image labels and model parameters using modified Biased HMRF-EM framework

In the EM framework, the bias field and class labels as well as the model parameters are estimated recursively until the model parameters converge to the optimal ones. The bias field B^0 and the image class label estimates x^0 are obtained by computing the MAP estimation using the initial model parameters. Using the label estimates x^0 and the bias field B^0, the model parameters θ^1 is estimated by maximizing the $Q(\theta, B^t)$. This recursive process is repeated till the parameters converge to optimal values. To estimate the class labels and model parameters along with bias field, a modified BHMRF-EM algorithm is proposed. The salient steps of the proposed BHMRF-EM algorithm is as follows.

Modified BHMRF-EM Algorithm:

1. Perform the initial parameter estimation and segmentation.

2. Estimate the bias field

$$b_i^{(t)} = \frac{[FR]_i}{[F\psi^{-1}1]_i}, with 1 = (1, 1, , 1)^T$$

3. Calculate the likelihood distribution

$$p^{(t)}\left(y_i^* | x_i, B\right) = g^{(t)}\left(y_i - b_i; \theta\left(x_i\right)\right)$$

4. Estimate the class labels by MRF-MAP estimation

$$x^{(t)} = arg \max_{x \in X} P(y \mid x, \theta^t + P(x))$$

5. Calculate the posterior distribution

$$p^{(t)}\left(l \mid y_i\right) = \frac{g^{(t)}\left(y_i; \theta_l\right) p^{(t)}\left(l \mid x_{N_i}\right)}{p(y_i)}$$

6. Update parameters by

$$\mu_l^{(t+1)} = \frac{\sum_{i \in s} p^{(t)}\left(l \mid y_i\right) y_i}{\sum_{i \in s} p^{(t)}\left(l \mid y_i\right)} \qquad \sigma_l^{(t+1)^2} = \frac{\sum_{i \in s} p^{(t)}\left(l \mid y_i\right)\left(y_i - \mu_l\right)^2}{\sum_{i \in s} p^{(t)}\left(l \mid y_i\right)}$$

7. $t \leftarrow t + 1$ and repeat from 2 until enough iterations have been performed.

3.11 Results and Discussions

In simulation, both synthetic as well as brain MR images are considered to validate the proposed algorithm. Synthetic images consisting of 3 and 5 classes are considered for the simulations. Besides, 5 simulated and 3 real brain MR images (healthy as well as diseased) are considered. The degraded images are obtained by adding white Gaussian noise of varying strength to the original image. The simulated brain MR images are obtained from "bicadmin@bic.mni.mcgill.ca" and "macampbell@davidson.edu" . The real brain MR images are obtained from "Department of Radiology, IGH, Rourkela" and "the whole brain Atlas".

Synthetic images:

Synthetic image of three class is considered as shown in Fig. 3.1 (a). The corresponding noisy version of SNRs 20 dB and 18 dB are shown in Fig. 3.1 (b) and 3.1 (e) respectively. The noisy images are modelled as Biased HMRF model and initial model parameters μ and σ considered for each class are selected from the histogram of respective noisy images. Proposed BHMRF-EM algorithm is used to obtain the segmented image of respective noisy images. The *a priori* MRF model parameter δ and the biased parameter are selected on the trial basis The performance of the algorithm is compared with Zhang's HMRF-EM algorithm [17]. The initial and converged parameters are tabulated in table 3.1. The results obtained by BHMRF-EM algorithm are shown in Fig. 3.1 (c) and (f) for SNR 20 dB and 18 dB respectively. Similarly the simulation and results obtained by HMRF-EM algorithm are shown in Fig 3.1 (d) and (g) respectively. As the number of classes is assumed to be unknown, the algorithm is run from higher number of initial classes and the algorithm converged to three classes. Here we have assumed 5 no. of initial classes. It is clear from the result that proper labelling of pixel could be obtained for noisy images of SNR 20 dB. The performance of the BHMRF-EM as well as HMRF-EM algorithm were with increase in noise strength which could be observed from the segmentation results of the noisy image of SNR 18 dB. It is evident from the image metrics that proposed BHMRF-EM algorithm converges faster than that of HMRF-EM algorithm. The $\%$ of MCE of the segmented image is also less in case of BHMRF-EM scheme. This is due to the incorporation of baised neighborhood in the *a priori* energy function. The identical spatial information used by HMRF model throughout the image is modified by assigning an biased parameter to the neighbourhood system. With increase in noise, i.e at 18 dB SNR, the performance of the algorithm degraded as seen from Fig. 3.1 (f), but the performance of HMRF-EM algorithm degraded more as in Fig. 3.1 (g).

A five class synthetic image is considered in our simulation as shown in Fig. 3.2 (a). The corresponding noisy versions of SNR 20 dB and 18 dB respectively. are shown in Fig. 3.2 (b) and (e). Fig. 3.2 (c) and (f) show the segmentation results obtained by BHMRF-EM algorithm and the model parameters are tabulated in table 3.2. The *a priori* model parameter δ and weighting parameter for each of the noisy images are also given in

table 3.2. It is observed from Fig. 3.2 (c) and (f) that BHMRF-EM algorithm yield proper segmented images. It is also observed from Fig. 3.2 (d) and (g) that the segmented results using HMRF-EM algorithm have more number of misclassified pixels. With increase in noise, there is more degradation in performance in both the schemes. There are many misclassified pixels five distinct visible classes. In both synthetic images, it is observed that the no. of misclassified pixels in the segmented images are less in our proposed scheme which signifies the quality of better accuracy. The convergence of the algorithm is also faster which is evident from the convergence time presented in table 3.1 and 3.2. This up gradation in performance of the BHMRF-EM scheme could be attributed to the incorporation of biased neighborhood structure.

Brain MR images:

After testing the developed algorithms with synthetic images successfully, 5 simulated and 3 real brain MR images are considered in our simulations. All simulated brain MR images are obtained from "Department of Radiology, IGH, Rourkela" and "the whole brain Atlas". A simulated brain MR image of size (128×128) degraded with 3% noise is shown in Fig. 3.3 (a). The "noise percentage" value represents the percent ratio of the standard deviation of the white Gaussian noise versus the signal for a reference class. The corresponding ground truth image is presented in Fig. 3.3 (b). The proposed BHMRF-EM algorithm is applied with varying weighting parameter . The algorithm starts with 6 number of initial classes. The *a priori* parameter δ is assumed to be 0.2. The initial parameters μ and σ for each class are assumed from the histogram of the original image. It is observed that the $\%$ of MCE changes with different weighting parameter . The similar study is also done with other brain MR images presented Fig. 3.4. The finding of $\%$ of MCE vercus are presented in table 3.3(a) for different brain MR images. It is observed that the $\%$ of MCE is less with the value of ρ between 0.3 and 0.5. So, for all brain MR images, the value of is taken as 0.5.

A diseased real brain MR images of size (175×215) degraded with 3% noise is shown in the Fig. 3.5 (a). The corresponding segmented image with different values of biased factor varies from 0.1 to 0.9 are shown in Figs. 3.5 (d)-(i) respectively. Six *a priori*

classes are assumed for this image and the corresponding BHMRF model parameters for four different classes are tabulated in Table 3.6. The *a priori* MRF model parameters for the two schemes is 0.25. It is observed from Figs. 3.(d)-(f) that four distinct classes could be obtained up to $\rho = 0.5$. At high biased value there are visibly four classes with a few misclassified pixels. Hence, the performance of the scheme gradually deteriorates with increase in biased value. The results are compared with the existing scheme HMRF-EM. This is evident from the Table 3.6 that the BHMRF-EM scheme converges in 23 secs, while the HMRF-EM found to be little faster than that of proposed scheme. But the $\%$ of MCE shows the accuracy in the segmentation of the proposed one.

Similarly, another two real sacroma diseased and multiple lessions of real brain MR images of size (175×215) with 3% noise are shown in Fig. 3.7 (a) and 3.8 (a). The corresponding ground truth images are shown in Fig. 3.7 (b) and 3.8 (b). The segmented image using BHMRF-EM framework are shown in Fig. 3.7 (c) and 3.8 (c). Fig. 3.7 (d) and 3.8 (d) shows the segmented image of HMRF-EM framework. All the initial and final model parameters, the a priori parameter and biased value for both the images are tabulated in Table 3.8 and Table 3.9 respectively. As the execution time for both the framework are not varying, from the misclassification error it can be observed that the proposed schemes outperfoms the former one using the biased value $\rho = 0.5$.

A simulated brain MR image of size (128×128) degraded with 3% and 5% noise are shown in the Fig. 3.6 (a) and (e). The ground truth image is presented in Fig. 3.6 (d). The segmented image using proposed BHMRF-EM framework are shown in Fig. 3.6 (b) and (f). Fig. 3.6 (c) and (g) shows the segmented image of HMRF-EM framework. All the initial and final model parameters are tabulated in Table 3.7. The $\%$ of MCE proves the efficiency of the proposed scheme.

At last, another two simulated brain MR image of size (241×181) and (256×256) are shown in in Fig. 3.9 (a) and 3.10 (a). The corresponding ground truth images are shown in Fig. 3.9 (b) and 3.10 (b). The segmented image using BHMRF-EM framework are shown in Fig. 3.9 (c) and 3.10 (c). Fig. 3.9 (d) and 3.10 (d) shows the segmented image of HMRF-EM framework. All the initial and final model parameters for both the images are tabulated in Table 3.10 and Table 3.11 respectively. From the table it can be

observed that the $\%$ of MCE is less in case of proposed BHMRF-EM framework rather than HMRF-EM framework.

3.11.1 Bias field estimation:

In our simulation for bias field correction using modified BHMRF-EM algorithm, one 3-class synthetic image and 3 brain MR images are considered. The bias field is ge

Synthetic images:

The 3- class synthetic image of size (128×128) is shown in Fig. 3.11 (a). The generated circular bias field is shown in Fig. 3.11 (b). The multiplicative bias field corrupted synthetic image is shown in Fig. 3.11 (c). After validation of the proposed modified BHMRF-EM algorithm on Fig. 3.11 (c), the extracted bias field is obtained properly and presented in Fig. 3.11 (d). The *a priori* model parameter along with the initial and converged sets of model parameters μ and σ are tabulated in Table 3.12. It is observed that the proposed modified BHMRF-EM algorithm could extract the bias field.

Brain MR images:

After testing the developed algorithms with synthetic image successfully, 3 simulated brain MR images are considered in our simulations. Fig. 3.12 (a) shows the simulated brain MR image of size (128×139). The generated circular bias field and the multiplicative bias field corrupted image are shown in Fig. 3.12 (b) and (c) respectively. The extracted bias field is shown in Fig. 3.12 (d). The initial and final model parameters are tabulated in Table 3.13.

The proposed algorithm is also validated with another 2 brain MR images of size (128×110) (128×147) are shown in Fig. 3.13 (a), Fig. 3.14 (a). The generated circular bias field are shown in Fig. 3.13 (b), Fig. 3.14 (b). Fig. 3.13 (c), Fig. 3.14 (c) are the multiplicative bias field corrupted images. Fig. 3.13 (d) and Fig.3.14 (d) indicates the extracted bias field of the brain MR image 3.13 (a), Fig. 3.14 (a) respectively. The initial and converged model parameters μ and σ for the 2 images are tabulated in Table 3.14 and Table 3.15. In all brain MR images, it is observed that the proposed algorithm could

successfully extracted the intensity in-homogeneity or bias field, hence could be able to segment the bias corrected images properly.

3.12 Conclusion

This chapter addressed the segmentation of brain MR images in unsupervised framework. Biased HMRF model is proposed to jointly estimate the model parameters and the image labels. The energy function of the *a priori* MRF model is modified in the proposed BHMRF model which could able to reduce the percentage of misclassification error and time complexity. The proposed BHMRF-EM scheme yielded the segmentation of synthetic as well as brain MR images with better performance than the result of HMRF-EM scheme. The proposed BHMRF-EM algorithm does not assume to have the knowledge of the number of classes. But, the MRF model parameter δ, is assumed to be known. The proposed algorithm is also exploited to take care of the intensity inhomogeneity of the brain MR images. In this regard, modified BHMRF-EM algorithm is proposed to estimate the bias field. The bias field, image labels and model parameters are estimated jointly. Although the proposed algorithm could yield better performance than the HMRF-EM scheme, but the algorithm is very much sensitive to the initial assumption of model parameters. Another limitation of the algorithm is to select δ of the MRF model. Our segmentation result may be further improved by developing a scheme which is insensitive to initial model parameters.

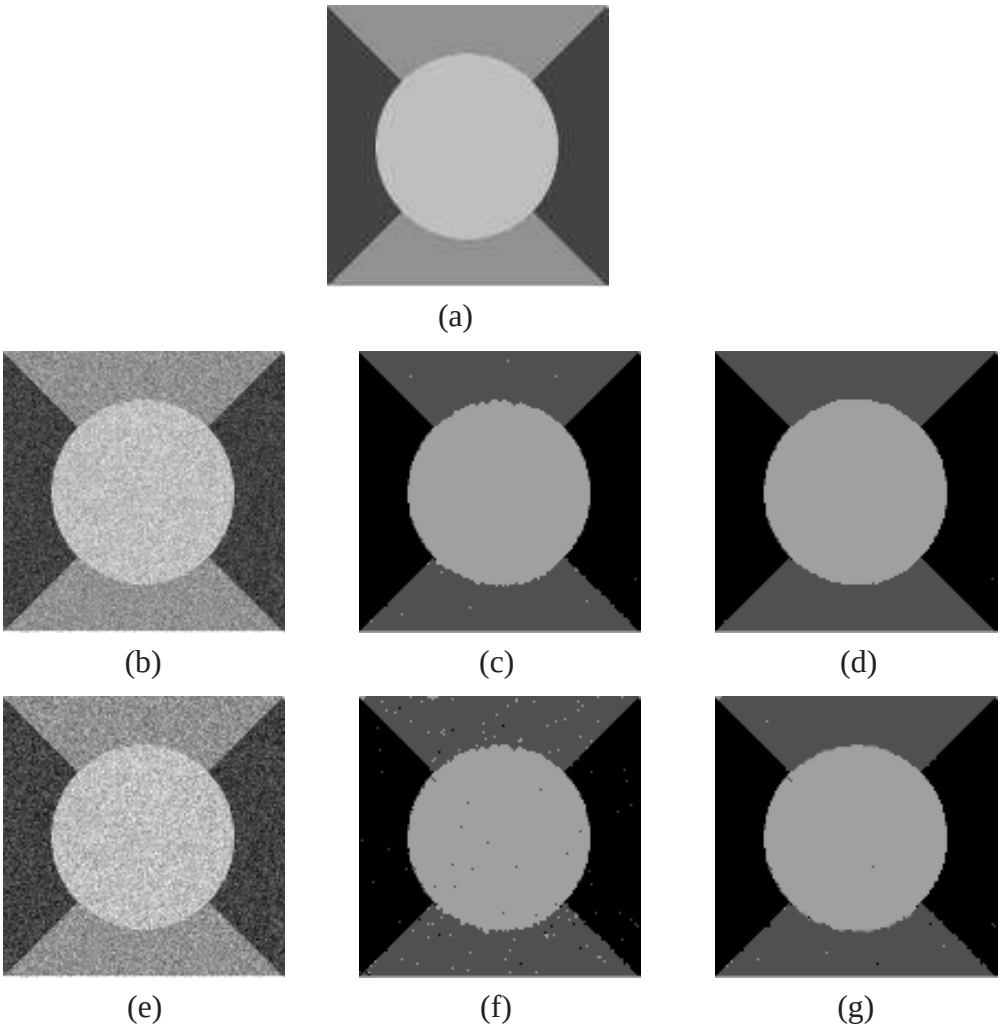


Figure 3.1: *Unsupervised image segmentation of synthetic 3-class image of size* (128×128)*: (a) Original image (b) and (e) noisy image with 20 db and 18 db SNR respectively (c) and (f) segmented image using HMRF-EM framework (d) and (g) segmented image using Biased HMRF-EM framework*

Schemes and Parameters		class → 1	2	3	% of MCE	Execu. time (sec.)
SNR= 20 dB	μ_i	0.61	1.48	1.97		
$\delta = 0.7$	σ_i	0.22	0.33	0.44		
BHMRF-EM	μ_f	0.83	1.81	2.38	1.82	28
$\rho = 0.26$	σ_f	0.16	0.17	0.16		
HMRF-EM	μ_f	0.82	1.84	2.28	2.03	34
	σ_f	0.16	0.21	0.27		
SNR= 18 dB	μ_i	0.42	1.13	1.80		
$\delta = 0.65$	σ_i	0.35	0.54	0.61		
BHMRF-EM	μ_f	0.82	1.78	2.32	2.45	28
$\rho = 0.23$	σ_f	0.20	0.21	0.23		
HMRF-EM	μ_f	0.82	1.81	2.38	2.86	28
	σ_f	0.20	0.22	0.25		

Table 3.1: *Image model parameters of synthetic 3-class image of size* (128×128) *with BHMRF-EM, HMRF-EM schemes of Fig. 3.1*

Schemes and Parameters		class → 1	2	3	4	5	% of MCE	Execu. time (sec.)
	μ_i	0.02	0.9	2.04	3.10	3.80		
$\delta = 0.62$	σ_i	0.37	0.68	0.54	0.60	0.43		
BHMRF-EM	μ_f	0.01	0.97	2.0	3.01	3.96	4.47	22
$\rho = 0.1$	σ_f	0.02	0.09	0.09	0.10	0.07		
HMRF-EM	μ_f	0.04	0.99	1.98	2.99	3.98	5.02	23
	σ_f	0.07	0.24	0.24	0.23	0.23		
SNR= 18 dB	μ_i	0.02	1.16	1.96	3.08	3.78		
$\delta = 0.6$	σ_i	0.37	0.45	0.60	0.59	51		
BHMRF-EM	μ_f	0.02	1.08	2.30	3.43	3.06	6.94	24
$\rho = 0.3$	σ_f	0.05	0.10	0.15	0.17	0.09		
HMRF-EM	μ_f	0.05	0.99	1.98	2.96	3.89	8.68	24
	σ_f	0.09	0.20	0.25	0.29	0.30		

Table 3.2: *Image model parameters of synthetic 5-class image of size* (128×128) *with BHMRF-EM, HMRF-EM schemes of Fig. 3.2*

Figure 3.3	ρ	0.2	0.3	0.4	0.5	0.7	0.9
	$\%$ of MCE	12.60	12.5	10.0	6.84	16.57	16.58
Figure 3.4	ρ	0.1	0.3	0.5	0.7	0.8	0.9
	$\%$ of MCE	20.17	18.23	18.07	19.38	19.45	22.24
Figure 3.5	ρ	0.1	0.5	0.6	0.7	0.9	
	$\%$ of MCE	13.70	13.47	15.76	18.54	19.82	

Table 3.3: $\%$ *of misclassification error with different biased parameter* ρ *using proposed BHMRF-EM framework of Fig. 3.3, 3.4, 3.5*

Schemes and Parameters		class $\rightarrow$ 1	2	3	4	$\%$ of MCE	Execu. time (sec.)
$\delta = 0.1$	μ_i	1.14	1.98	3.02	3.96		
	σ_i	0.55	0.57	0.67	0.23		
BHMRF-EM	μ_f	0.28	1.09	2.10	3.04		
$\rho = 0.5$	σ_f	0.69	0.21	0.12	0.09	6.84	20
HMRF-EM	μ_f	0.28	1.34	2.10	3.22		
	σ_f	0.30	0.20	0.14	0.09	11.35	18

Table 3.4: *Image model parameters of brain MR image of size* (128×128) *segmented image using BHMRF-EM framework with biased parameter* ρ*, segmented image using HMRF-EM framework schemes of Fig. 3.3*

Schemes and Parameters		class $\rightarrow$ 1	2	3	4	$\%$ of MCE	Execu. time (sec.)
$\delta = 0.2$	μ_i	0.74	1.86	3.0	3.96		
	σ_i	0.66	0.57	0.55	0.73		
BHMRF-EM	μ_f	0.96	1.23	1.59	4.03		
$\rho = 0.5$	σ_f	0.06	0.12	0.29	0.15	18.07	24
HMRF-EM	μ_f	0.98	1.51	3.07	4.16		
	σ_f	0.05	0.18	0.28	0.05	19.97	22

Table 3.5: *Image model parameters of brain MR image of size* (128×128) *segmented image using BHMRF-EM framework with biased parameter* ρ*, segmented image using HMRF-EM framework schemes of Fig. 3.4*

Schemes and Parameters		class → 1	2	3	4	% of MCE	Execu. time (sec.)
$\delta = 0.25$	μ_i	0.12	1.16	1.84	2.86		
	σ_i	0.22	0.51	0.61	0.56		
BHMRF-EM	μ_f	0.14	1.48	1.86	2.48		
$\rho = 0.5$	σ_f	0.08	0.10	0.19	0.14	13.47	23
HMRF-EM	μ_f	0.14	1.46	1.82	2.46		
	σ_f	0.05	0.09	0.19	0.14	19.29	20

Table 3.6: *Image model parameters of brain MR image of size* (175×215) *segmented image using BHMRF-EM framework , segmented image using HMRF-EM framework schemes of Fig. 3.5*

Schemes and Parameters		class → 1	2	3	4	% of MCE	Execu. time (sec.)
$\delta = 0.3$	μ_i	0.02	0.9	1.8	2.96		
	σ_i	0.12	0.57	0.58	0.68		
BHMRF-EM	μ_f	0.28	1.08	0.03	1.86		
$\rho = 0.5$	σ_f	0.04	0.09	0.01	0.21	5.19	24
HMRF-EM	μ_f	0.10	1.03	1.48	2.62		
	σ_f	0.02	0.06	0.14	0.22	14.65	22

Table 3.7: *Image model parameters of brain MR image of size* (128×128) *with BHMRF EM, HMRF-EM schemes of Fig. 3.6 respectively.*

Schemes and Parameters		class → 1	2	3	4	% of MCE	Execu. time (sec.)
$\delta = 0.3$	μ_i	1.18	1.84	3.0	3.86		
	σ_i	0.50	0.57	0.52	0.67		
BHMRF-EM	μ_f	0.14	1.62	1.93	2.36		
$\rho = 0.5$	σ_f	0.10	0.09	0.13	0.20	16.89	26
HMRF-EM	μ_f	0.12	1.25	1.72	2.24		
	σ_f	0.16	0.08	0.09	0.05	18.06	26

Table 3.8: *Image model parameters of brain MR image of size* (128×128) *segmented image using BHMRF-EM framework , segmented image using HMRF-EM framework schemes of Fig. 3.7*

Schemes and Parameters		class → 1	2	3	4	% of MCE	Execu. time (sec.)
$\delta = 0.1$	μ_i	1.34	2.14	3.16	4.08		
	σ_i	0.28	0.31	0.33	0.34		
BHMRF-EM	μ_f	1.36	1.83	2.74	3.52		
$\rho = 0.5$	σ_f	0.60	0.12	0.19	0.25	15.32	28
HMRF-EM	μ_f	1.35	1.79	2.64	3.44		
	σ_f	0.05	0.08	0.19	0.12	19.22	29

Table 3.9: *Image model parameters of brain MR image of size* (128×128) *segmented image using BHMRF-EM framework with different value of* ρ*, segmented image using HMRF-EM framework schemes of Fig. 3.8*

Schemes and Parameters		class → 1	2	3	4	% of MCE	Execu. time (sec.)
$\delta = 0.08$	μ_i	0.03	1.0	1.86	2.86		
	σ_i	0.32	0.65	0.52	0.48		
BHMRF-EM	μ_f	0.01	1.41	1.67	2.14		
$\rho = 0.5$	σ_f	0.09	0.21	0.08	0.08	19.01	25
HMRF-EM	μ_f	0.02	1.01	1.78	2.26		
	σ_f	0.10	0.22	0.09	0.08	21.09	26

Table 3.10: *Image model parameters of brain MR image of size* (241×181) *with BHMRF-EM, HMRF-EM schemes of Fig. 3.9 respectively.*

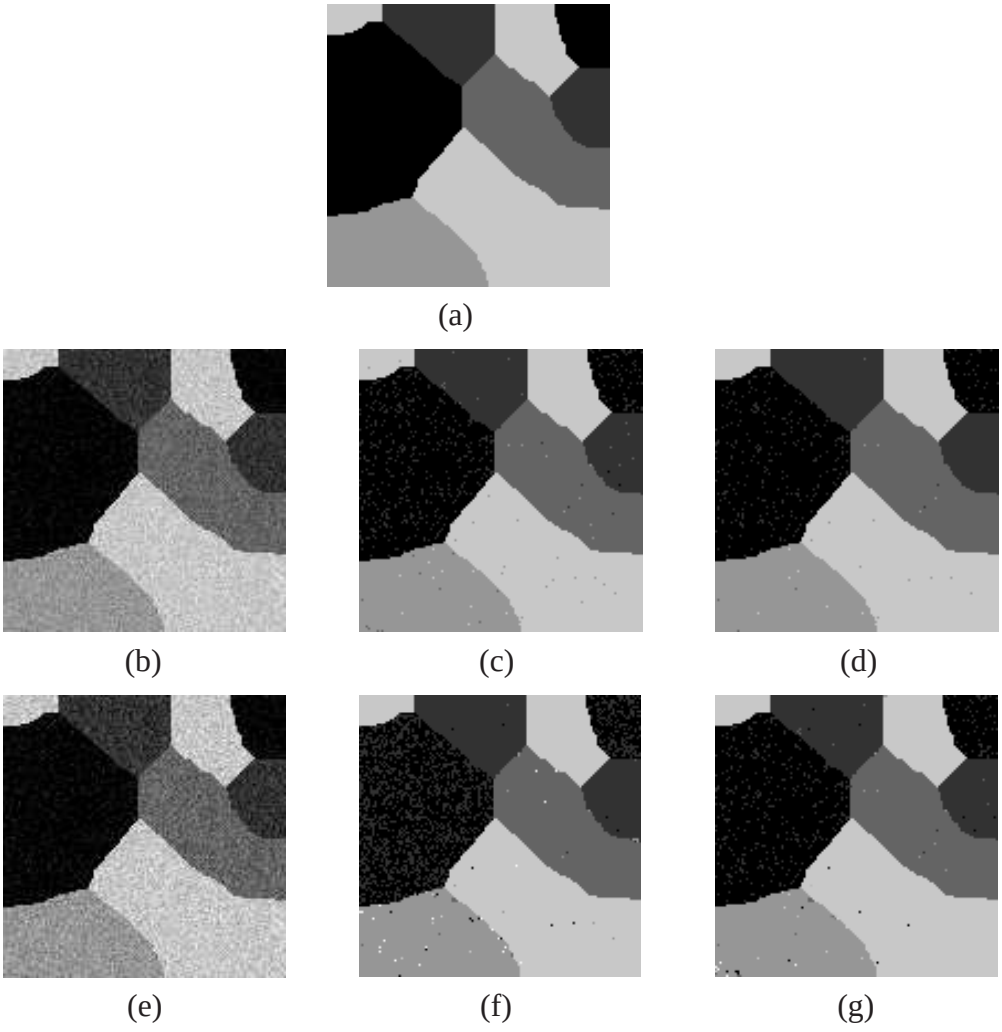


Figure 3.2: *Unsupervised image segmentation of synthetic 5-class image of size* (128×128)*: (a) Original image (b) and (e) noisy image with 20 db and 18 db SNR respectively (c) and (f) segmented image using HMRF-EM framework (d) and (g) segmented image using Biased HMRF-EM framework*

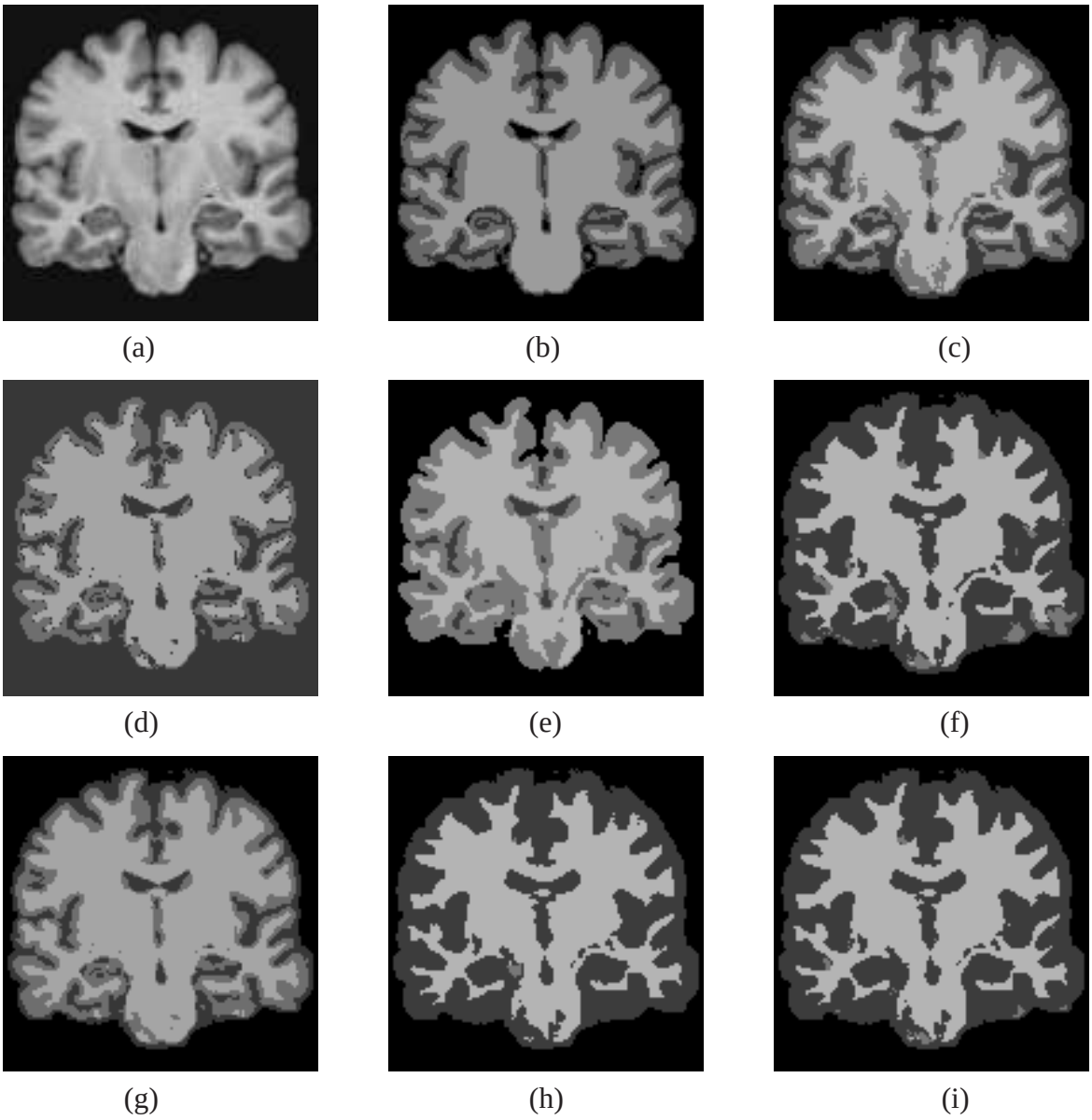


Figure 3.3: *Unsupervised image segmentation of Brain MR image of size* (128×128) *(a) Original image with 3* $\%$ *of noise (b) Ground Truth (c) segmented image using HMRF-EM framework (d)-(i) segmented image using Biased HMRF-EM framework with* ρ*= 0.2, 0.3, 0.4, 0.5, 0.7, 0.9*

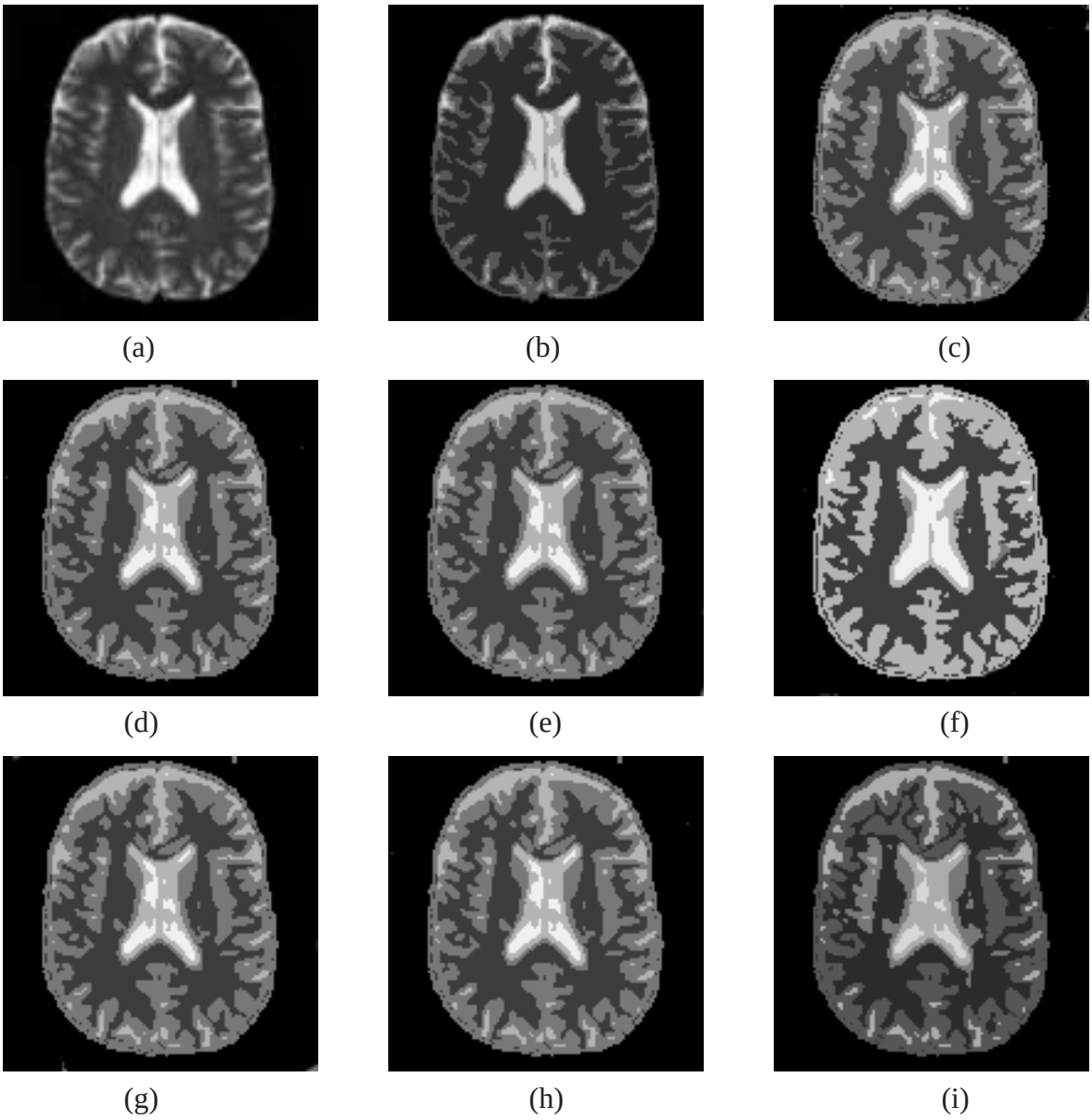


Figure 3.4: *Unsupervised image segmentation of Brain MR image of size* (128×128) *(a) Original image with 3* $\%$ *of noise (b) Ground Truth (c) segmented image using HMRF-EM framework (d)-(i) Biased segmented image using HMRF-EM framework with* ρ*= 0.1, 0.3, 0.5, 0.7, 0.8, 0.9*

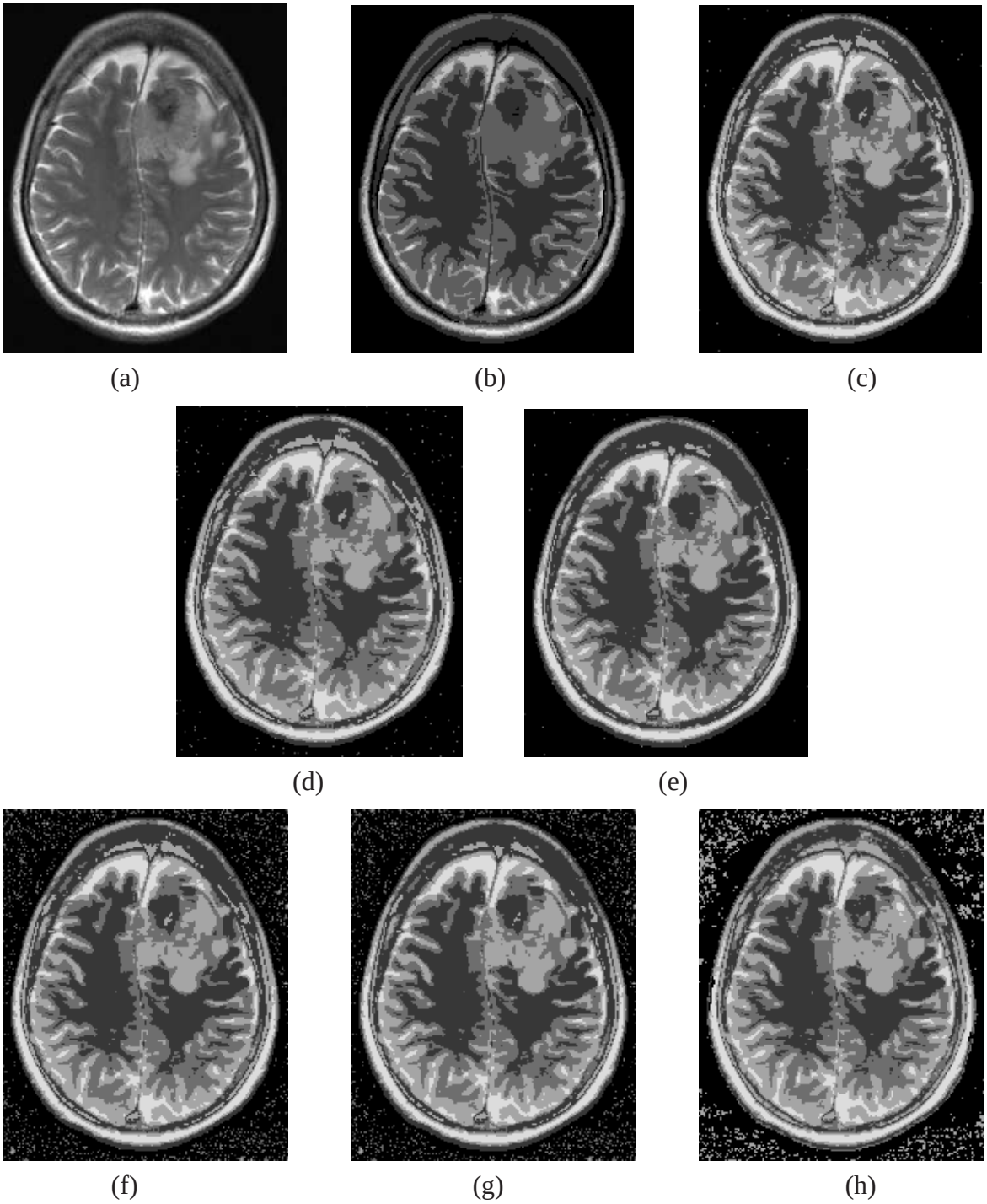


Figure 3.5: *Unsupervised image segmentation of tumor from a Brain MR image of size* (175×215) *(a) Real image with 3* $\%$ *of noise (b) Ground Truth (c) segmented image using HMRF-EM framework (d)-(h) segmented image using Biased HMRF-EM framework with* ρ*= 0.1, 0.5, 0.6, 0.7, 0.9*

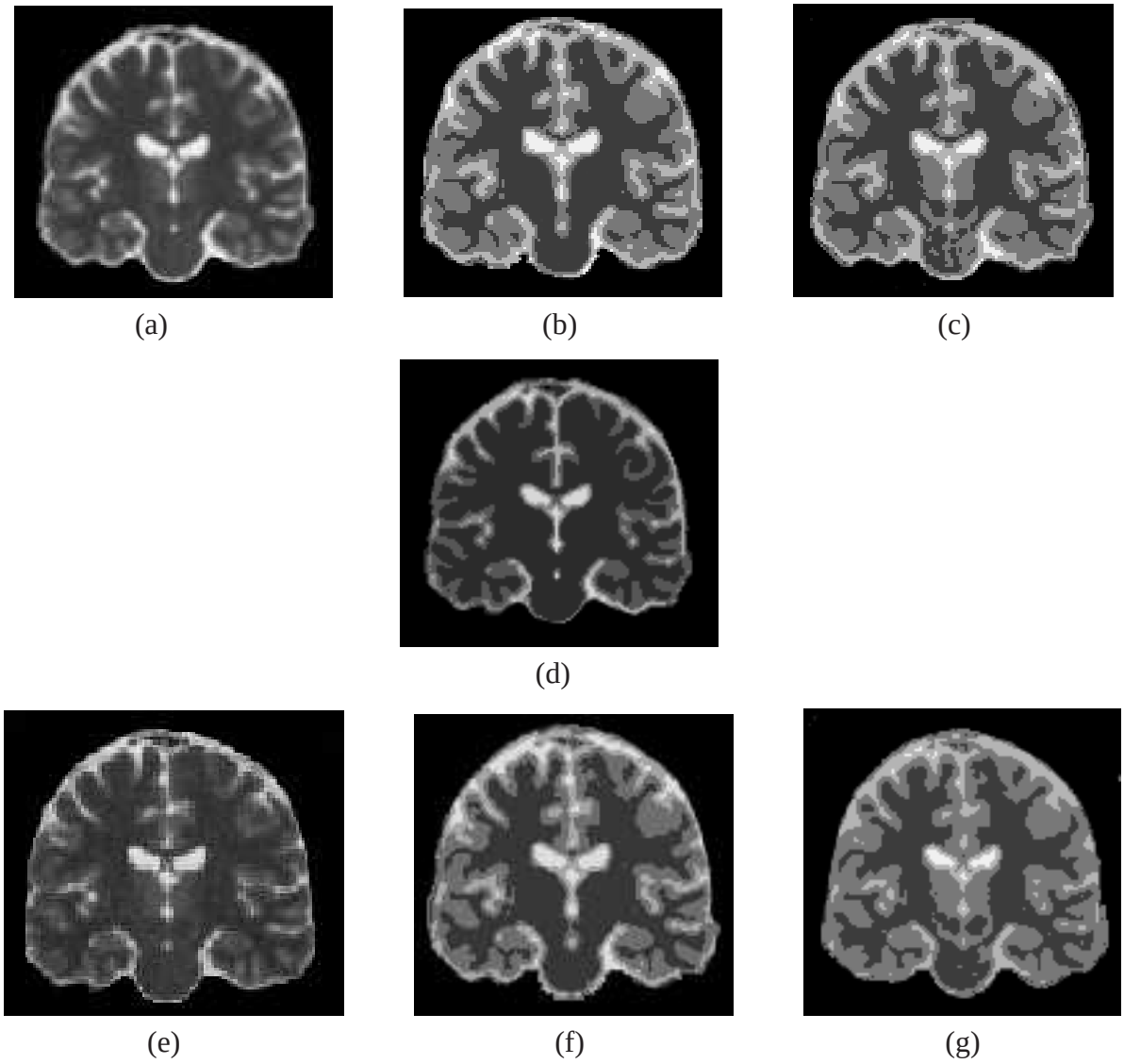
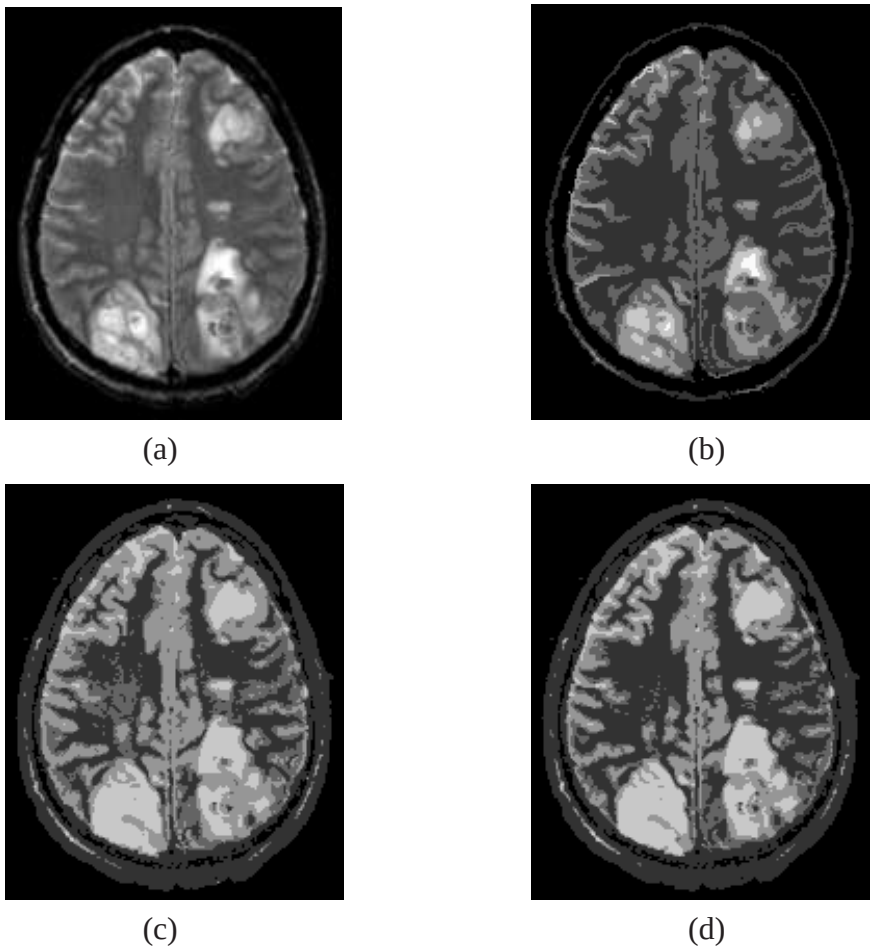


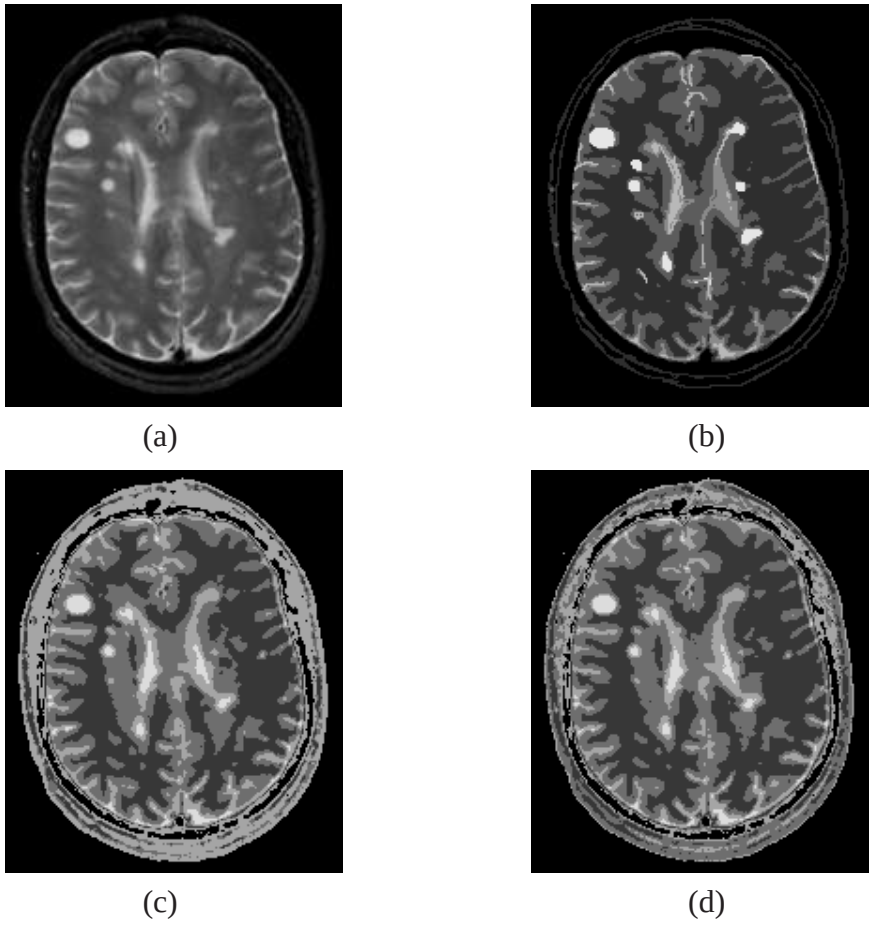
Figure 3.6: *Unsupervised image segmentation of Brain MR image of size* (128×128) *(a) and (e) Original image with 3* $\%$ *and 5* $\%$ *of noise (d) Ground Truth (b) and (f) segmented image using Biased HMRF-EM framework (c) and (g) segmented image using HMRF-EM framework*



(a) (b)

(c) (d)

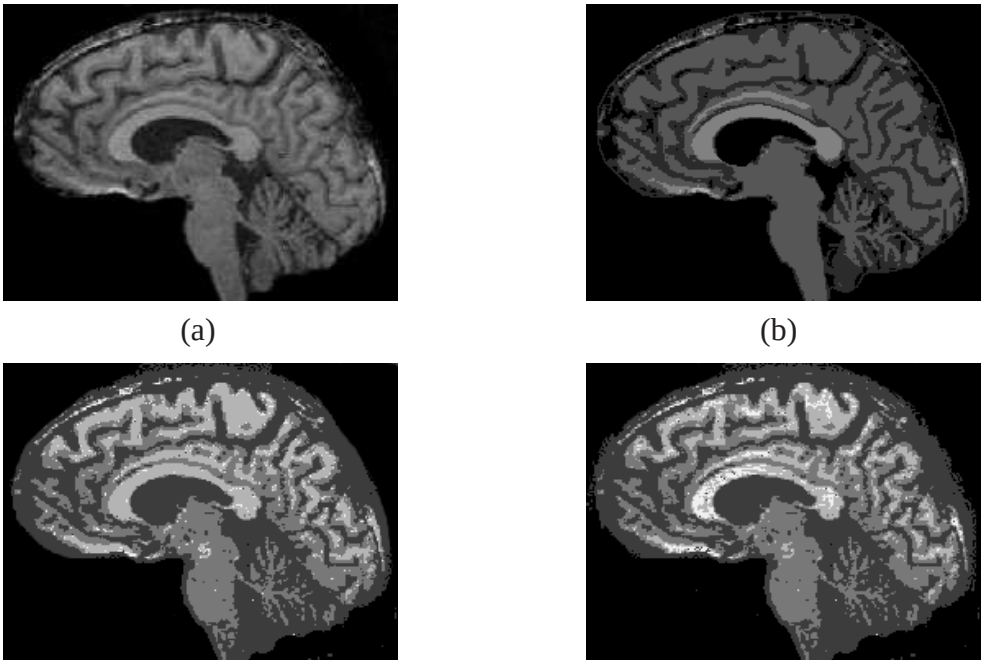
Figure 3.7: *Unsupervised image segmentation of Sarcoma diseased Brain MR image of size* (175×215) *(a) Real image with 3* $\%$ *of noise (b) Ground Truth (c) segmented image using HMRF-EM and framework (d) segmented image using Biased HMRF-EM framework*



(a) (b)

(c) (d)

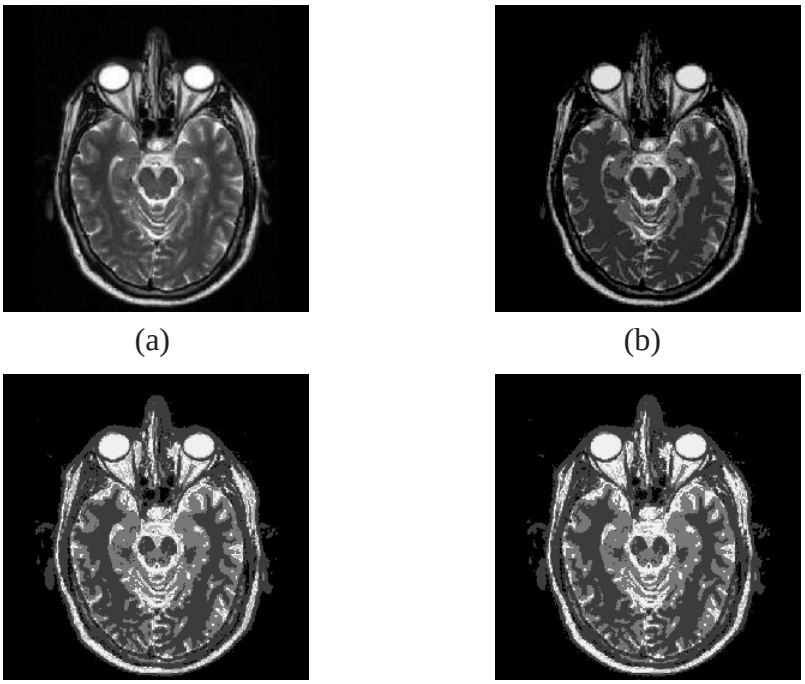
Figure 3.8: *Unsupervised image segmentation of Multiple sclerosis from a Brain MR image of size* (175×215) *(a) Real image with 3* $\%$ *of noise (b) Ground Truth (c) segmented image using HMRF-EM framework (d) segmented image using Biased HMRF-EM framework*



(a) (b)

(c) (d)

Figure 3.9: *Unsupervised image segmentation of Brain MR image of size* (241×181) *(a) Original image with 3* $\%$ *of noise (b) Ground Truth (c) segmented image using Biased HMRF-EM framework (d)segmented image using HMRF-EM framework*



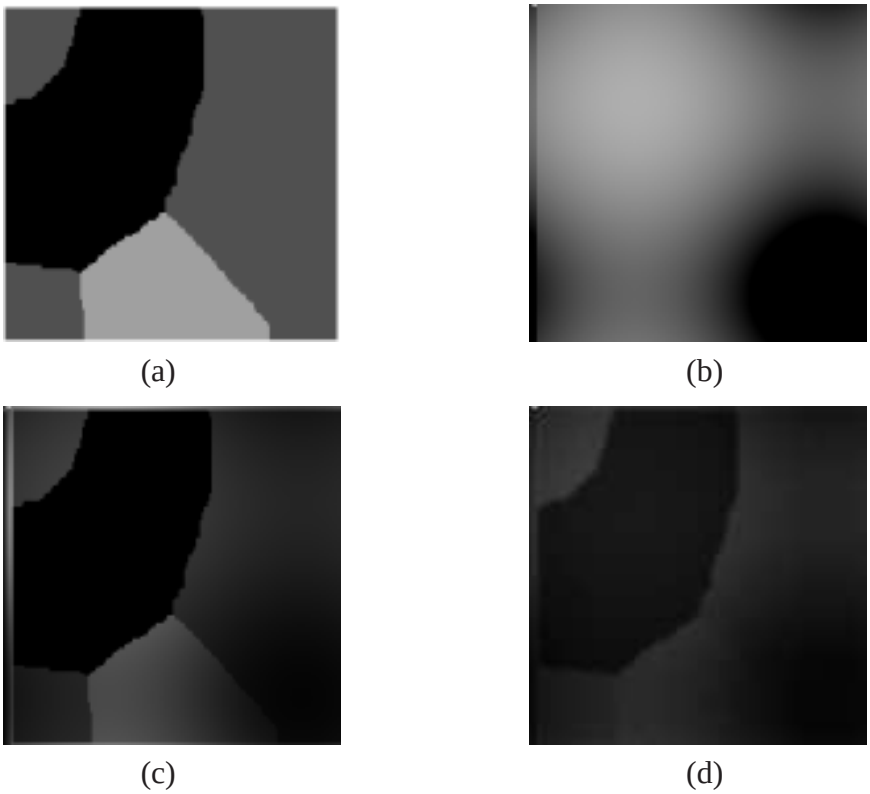
(a) (b)

(c) (d)

Figure 3.10: *Unsupervised image segmentation of Brain MR image of size* (255×255) *(a) Original image with 3* $\%$ *of noise (b) Ground Truth (c) segmented image using Biased HMRF-EM framework (d)segmented image using HMRF-EM framework*

Schemes and Parameters		class → 1	2	3	4	% of MCE	Execu. time (sec.)
$\delta = 0.01$	μ_i	0.08	0.78	1.96	2.96		
	σ_i	0.30	0.65	0.55	0.76		
BHMRF-EM	μ_f	0.01	1.05	1.55	2.83		
$\rho = 0.5$	σ_f	0.01	0.09	0.12	0.21	11.56	28
HMRF-EM	μ_f	0.10	1.09	1.48	2.46		
	σ_f	0.02	0.06	0.14	0.13	16.70	22

Table 3.11: *Image model parameters of brain MR image of size* (128×128) *with BHMRF EM, HMRF-EM schemes of Fig. 3.10 respectively.*

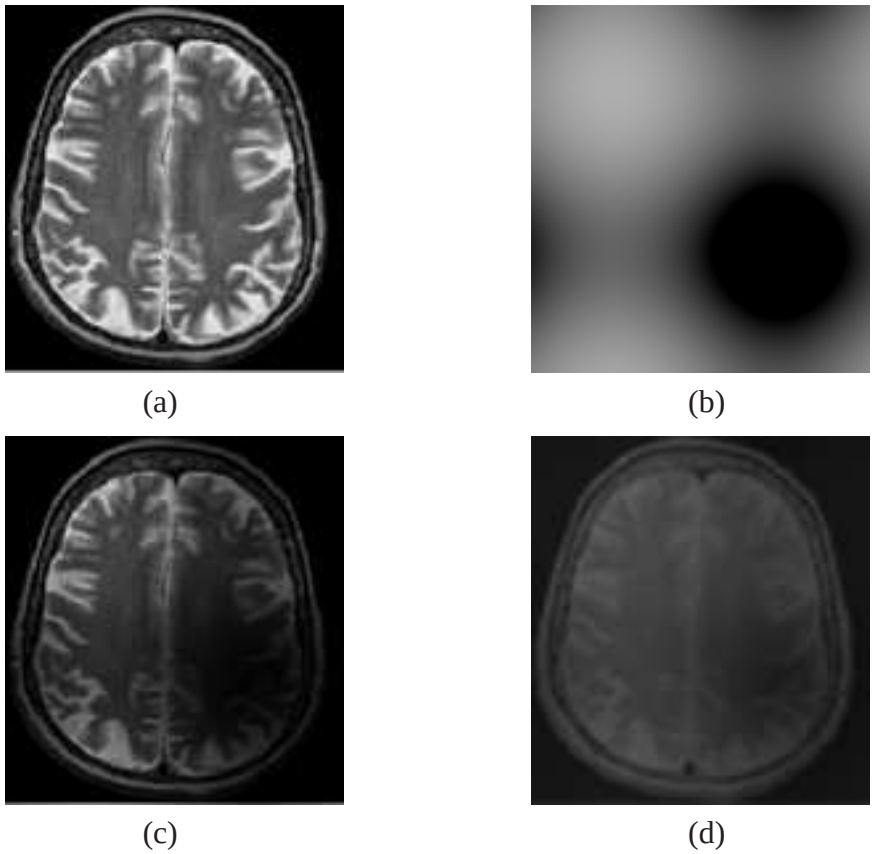


(a) (b) (c) (d)

Figure 3.11: *Synthetic 3-class image of size* (128×128) *with : (a) Original image (b) slowly varying circular bias field (c) multiplicative bias field corrupted image (d) extracted bias field using modified BHMRF-EM algorithm*

Parameters		class $\rightarrow$ 1	2	3
SNR = 25 dB $\delta = 1.2$	μ_i	0.875	1.85	2.00
	σ_i	0.55	0.38	0.4
	μ_f	0.99	1.99	2.06
	σ_f	0.24	0.23	0.12

Table 3.12: *Image model parameters of synthetic 3- class image of size* (128×128) *using modified BHMRF-EM algorithm of fig 3.11*

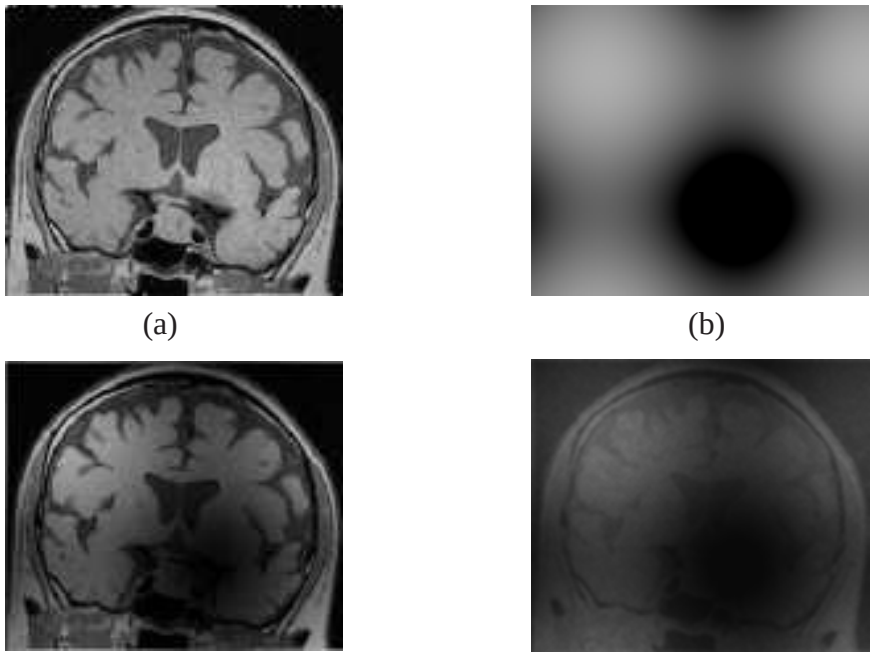


(a) (b) (c) (d)

Figure 3.12: *Brain MR image of size* (128×139) *with : (a) Original image (b) slowly varying circular bias field (c) multiplicative bias field corrupted image (d) extracted bias field using modified BHMRF-EM algorithm*

Parameters		class → 1	2	3	4
SNR = 25 dB $\delta = 1.2$	μ_i	0.013	1.43	2.83	2.00
	σ_i	0.32	0.60	0.66	0.50
	μ_f	0.0	3.11	3.12	2.475
	σ_f	0.0	0.198	0.197	0.054

Table 3.13: *Image model parameters of brain MR image of size* (128×139) *with SNR 25 dB using modified BHMRF-EM algorithm of fig 3.12*

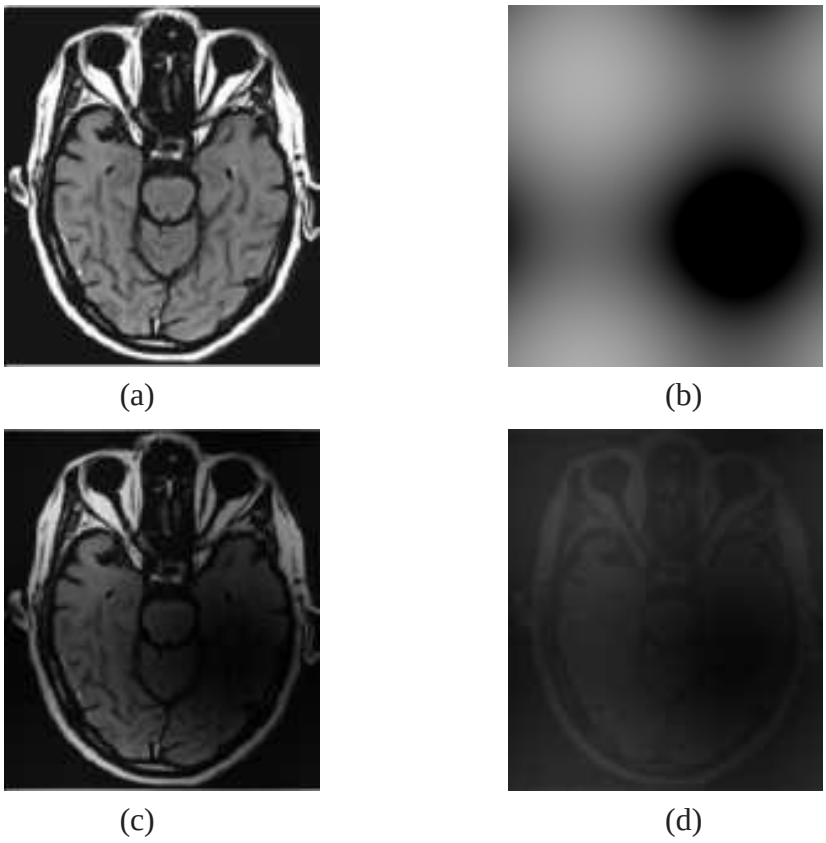


(a) (b) (c) (d)

Figure 3.13: *Brain MR image of size* (128×110) *with : (a) Original image (b) slowly varying circular bias field (c) multiplicative bias field corrupted image (d) extracted bias field using modified BHMRF-EM algorithm*

Parameters		class → 1	2	3	4
SNR = 25 dB $\delta = 1.2$	μ_i	0.96	2.13	2.88	0.1
	σ_i	0.57	0.55	0.6	1.50
	μ_f	1.2	2.06	3.07	3.65
	σ_f	0.12	0.2	0.89	0.57

Table 3.14: *Image model parameters of brain MR image of size* (128×110) *with SNR 25 dB using modified BHMRF-EM algorithm of fig 3.13*



(a) (b)

(c) (d)

Figure 3.14: *Brain MR image of size* (128×147) *with : (a) Original image (b) slowly varying circular bias field (c) multiplicative bias field corrupted image (d) extracted bias field using modified BHMRF-EM algorithm*

Parameters		class $\rightarrow$ 1	2	3	4
SNR = 25 dB $\delta = 0.5$	μ_i	0.2	1.01	1.39	3.28
	σ_i	0.34	0.60	0.62	0.65
	μ_f	0.35	0.36	1.92	4.19
	σ_f	0.06	0.08	0.19	0.06

Table 3.15: *Image model parameters of brain MR image of size* (128×147) *with SNR 25 dB using modified BHMRF-EM algorithm of fig 3.14*

Chapter 4

Unsupervised image segmentation using HMRF-FCEM algorithm

4.1 Introduction

In this chapter, a new concept of unsupervised brain MR image segmentation method is introduced by incorporating the HMRF model into fuzzy clustering procedure. As explained in chapter 3, the proposed BHMRF-EM scheme for brain MR image segmentation is sensitive to initial assumption of model parameters. This scheme leads to biased parameter estimates due to the conduction of the M-step of the EM algorithm considering the pixel labels are known quantities. In order to overcome this difficulty, an attempt has been made to incorporate fuzzy clustering approach and HMRF model together in one scheme. Fuzzy clustering methods are widely popular to conduct unsupervised image segmentation effectively [43, 44]. These methods are not very sensitive to initial assumption of cluster parameters, though obtain poor segmentation results with images corrupted by noise and other artifacts. The reason is that these methods do not take into account the spatial dependencies between the cluster data. To address the above issues, the assumptions of the HMRF model is incorporated effectively in to fuzzy clustering procedure. In the proposed scheme, HMRF model is used to model the degraded observed image and is regarded as defining fuzzy partition of the observed space motivated by the work of Celeux *et al.* [45]. The proposed approach is formulated by employing a mean field like approximation of the *a priori* MRF distribution. HMRF-FCEM algorithm is proposed to conduct the fuzzy clustering type treatment of the HMRF model using the fuzzy ob-

jective function. The use of the proposed algorithm eventually converge to the optimal set of clusters and model parameters. Hence proper segmentation of brain MR image is obtained.

4.2 Problem Statement

Let S is the set of sites. X denotes the random field associated with the labels of the original image and x denotes the realization of that. Let η denotes the neighborhood system on S. $p(x)$ is the probability distribution and can be modeled as a MRF with respect to the neighborhood system if

$$p(x_i \mid x_{s-i}) = p(x_i \mid x_{\eta_i})$$

Let Y denote the observed image random field and y be the realization of it. Y is modeled as Hidden Markov Random Field (HMRF). Depending on the conditional in dependency, the joint probability can be defined as

$$p(y, x) = p(y \mid x)p(x) \tag{4.1}$$

where,

$$p(y \mid x) = \prod_{i=1}^{s} p(y_i \mid x_i)$$

HMRF model associated with the computation of the posterior probabilities $p(x_i|y)$ and $p(x|y)$, which are obtained by means of Bayesian sampling. This requires a large amount of computations. In order to overcome this problem, mean field approximation of MRF is considered motivated by Celeux *et al.* [45]. Under this approximation, the joint prior of the Markov random field can be expressed as

$$p(x) = \prod_{i=1}^{s} p(x_i \mid x_{\eta_i})$$

Let each state of the L-state HMRF model be a cluster in the observable space Y. The *a posteriori* probability of observation y_i associated with the i^{th} site is

$$p(x_i = l \mid y_i) \equiv p(x_i = l \mid y_i, \hat{x}_{\eta_i}) \tag{4.2}$$

Considering each state of the model as a cluster in the observable space Y, it holds that

$$0 \leq p(x_i = l \mid y_i) \leq 1, \qquad \sum_{l=1}^{L} p(x_i = i \mid y_i) = 1$$

where $l = 1, ..., L$ and $i = 1, ..., s$. On this basis of observation, the considered HMRF model can be regarded as a fuzzy L-partition of the observation space Y. The fuzzy partition can be denoted as

$$R = \{r_{li}\} \tag{4.3}$$

where $r_{li}(l = 1, ..., L, i = 1, ..., s)$ represents the degree of observable vector y_i in the l^{th} state of the HMRF model. r_{li} is the fuzzy membership function having the properties:

$$0 \leq r_{li} \leq 1, \qquad \sum_{l=1}^{L} r_{li} = 1, \qquad 0 < \sum_{i=1}^{s} r_{li} < s \tag{4.4}$$

In this regard, HMRF model is treated as fuzzy clustering type under mean field like approximation of the MRF probability.

4.3 HMRF oriented fuzzy objective function

Motivated by the work of Ichihashi *et al.*, the fuzzy objective function is obtained by means of a regularization technique, where an FCM variant is regularized by KL information [28]. This new FCM variant is introduced into the fuzzy objective function. The modified fuzzy objective function becomes

$$J = \sum_{l=1}^{L} \sum_{i=1}^{s} r_{li}\, d_{li} + \lambda \sum_{l=1}^{L} \sum_{i=1}^{s} r_{li}\, log\left(\frac{r_{li}}{\pi_l}\right) \tag{4.5}$$

where

r_{li} : fuzzy membership function

d_{li} : dissimilarity function

π_l: prior probability of l^{th} cluster

λ : degree of fuzziness

$l = 1, ..., L$: cluster

$i = 1, ..., s$:site

In our proposed formulation, HMRF oriented modification of the fuzzy objective function is done by defining the dissimilarity function d_{li} as the negative log-likelihood of the l^{th} model state with respect to i^{th} site observation of the observed image.

$$d_{li}(\theta_l) \equiv -log \;\; p(y_i \mid x_i = l; \theta_l) \tag{4.6}$$

π_l can be considered as π_{li} which represents the pointwise prior probabilities of the HMRF model states. It is obtained on the basis of the mean-field approximation of the MRF as follows.

$$\pi_{li} = P\left(x_i = l \mid \hat{x}_{\eta_i}\right) = \frac{exp\left(-\sum_{c \ni i} V_c\left(x_{li}\right)\right)}{\sum_{h=1}^{L} exp\left(-\sum_{c \ni i} V_c\left(x_{hi}\right)\right)} \tag{4.7}$$

Eventually, the HMRF model is introduced into the fuzzy clustering procedure by the following modified fuzzy objective function.

$$J_\lambda(\psi) = -\sum_{l=1}^{L}\sum_{i=1}^{s} r_{li}\, logP\left(y_i \mid x_i = l; \theta_l\right) + \lambda \sum_{l=1}^{L}\sum_{i=1}^{s} r_{li}\, log\left(\frac{r_{li}}{\pi_{li}}\right) \tag{4.8}$$

where $\psi = \{R, \theta\}$ and

$$p(y_i \mid x_i = l, \theta_l) = \frac{1}{\sqrt{2\pi\sigma_l^2}} exp\left(-\frac{(y_i - \mu_l)^2}{2\sigma_l^2}\right) \tag{4.9}$$

Fuzzy clustering type treatment of the HMRF model is conducted by using this modified fuzzy objective function as given by (4.8).

4.4 Image label estimation

The image label estimation is formulated by defuzzification of the fuzzy membership function r_{li}. The following optimality criterion for each site is adopted,

$$\hat{x}_i = argmax_{l=1}^{L} r_{li} \tag{4.10}$$

4.5 Estimation of fuzzy membership function

The fuzzy membership function can be attained by minimizing the fuzzy objective function $J_\lambda(\psi)$ over r_{li} under the constraint

$\sum_{l=1}^{L} r_{li} = 1, \quad \forall i = 1, ..., s$

Introducing a Lagrange multiplier G_i for each data point to enforce the constraint, the minimization of fuzzy objective function becomes

$$\frac{\partial}{\partial r_{li}}\left[J_\lambda - \sum_{i=1}^{s} G_i \left(\sum_{h=1}^{L} r_{li} - 1\right)\right] = 0 \tag{4.11}$$

solving the (4.8), the fuzzy membership function becomes

$$r_{li} = \frac{\pi_{li}\, exp\left(-\frac{1}{\lambda} d_{li}\right)}{\sum_{h=1}^{L} \pi_{li}\, exp\left(-\frac{1}{\lambda} d_{li}\right)} \tag{4.12}$$

where, the dissimilarity function can be defined as

$$d_{li}^{(k)} = \frac{w}{2} log(2\pi) + \frac{1}{2} log|\sigma_l^{(k)}| + \frac{1}{2}\left(y_i - \mu_l^{(k)}\right)^T \sigma_l^{(k)^{-1}} \left(y_i - \mu_l^{(k)}\right) \tag{4.13}$$

4.6 Estimation of HMRF model parameters

The fuzzy objective function $J_\lambda(\psi)$ described in (4.8) is minimized to obtain an estimate of the HMRF model parameters given a data set. The minimization of $J_\lambda(\psi)$ is done iteratively using the proposed HMRF-fuzzy clustering EM (HMRF-FCEM) algorithm. Putting (4.9) in (4.8), the fuzzy objective function in terms of model parameters μ and σ becomes

$$J_\lambda(\psi) = -\sum_{l=1}^{L}\sum_{i=1}^{s} r_{li}\left[-\frac{1}{2}(y_i - \mu_l)^T \sigma_l^{-1}(y_i - \mu_l) - \frac{1}{2} log \mid \sigma_l \mid - \frac{w}{2} log(2\pi)\right] + \lambda \sum_{l=1}^{L}\sum_{i=1}^{s} r_{li} log\left(\frac{r_{li}}{\pi_{li}}\right) \tag{4.14}$$

minimizing the J_λ function with respect to μ_l and ignoring the terms not containing μ_l, is equal to

$$(y_i - \mu_l)^T \sigma_l^{-1}(y_j - \mu_l) = y_i^T \sigma_l^{-1} y_i \mu_l^T \sigma_l^{-1} \mu_l - 2 y_i^T \sigma_l^{-1} \mu_l$$

Since,

$$\frac{\partial \mu_l^T \sigma_l^{-1} \mu_l}{\partial \mu_l} = 2\sigma_l^{-1}\mu_l$$

$$\frac{\partial y_i^T \sigma_l^{-1} \mu_l}{\partial \mu_l} = 2\sigma_l^{-1} y_i \tag{4.15}$$

the solution of $\frac{\partial J_\lambda(\psi)}{\partial \mu_i} = 0$
Hence, the model parameter μ_l becomes,

$$\mu_l^{(k+1)} = \frac{\sum_{i=1}^{s} r_{li}^{(k)} y_i}{\sum_{i=1}^{s} r_{li}^{(k)}}$$

Similarly, minimizing (4.15) with respect to σ_l,

$$\frac{\partial log \mid \sigma_l \mid}{\partial \sigma_l^{-1}} = -\sigma_l$$

and

$$\frac{\partial (y_i - \mu_l)^T \sigma_i^{-1} (y_i - \mu_l)}{\partial \sigma_i^{-1}} = (y_i - \mu_l)(y_i - \mu_l)^T \tag{4.16}$$

As the solution of $\frac{\partial J_\lambda(\psi)}{\partial \sigma_i^{-1}} = 0$, the model parameter σ_l becomes,

$$\sigma_l^{(k+1)} = \frac{\sum_{i=1}^{s} r_{li}^{(k)} \left(y_i - \mu_l^{(k)}\right) \left(y_i - \mu_l^{(k)}\right)^T}{\sum_{i=1}^{s} r_{li}^{(k)}} \tag{4.17}$$

4.7 Joint estimation of image labels and model parameters using fuzzy clustering type EM framework

The image label estimates x^0 are obtained by defuzzification of the fuzzy membership function. The prior probabilities of the MRF is computed on the basis of mean field like approximation. The image labels as well as the fuzzy membership functions are estimated recursively until the model parameters converge to the optimal ones. The optimal values of the parameters are used to obtain the desired segmentation result. For joint estimation of the image labels as well as model parameters the HMRF-FCEM algorithm is proposed. The salient steps of the algorithm is as follows.

4.7.1 HMRF-FCEM Algorithm

1. Initialize the number of class labels to random values and select an arbitrary parameter set.

2. Estimate the image class labels $x^{(k)}$ by maximizing the fuzzy membership function

$$\hat{x}_i^{(k)} = argmax_{l=1}^{L} r_{li}^{(k)} \tag{4.18}$$

3. Compute the fuzzy membership functions $r_{li}^{(k+1)}$

$$r_{li}^{(k+1)} = \frac{\pi_{li}^{(k)}\ exp\left(-\frac{1}{\lambda}d_{li}^{(k)}\right)}{\sum_{h=1}^{L}\pi_{li}^{(k)}\ exp\left(-\frac{1}{\lambda}d_{li}^{(k)}\right)} \tag{4.19}$$

4. Update the parameters $\mu_l^{(k+1)}$ and $\sigma_l^{(k+1)}$

$$\mu_l^{(k+1)} = \frac{\sum_{i=1}^{s} r_{li}^{(k)} y_i}{\sum_{i=1}^{s} r_{li}^{(k)}}$$

$$\sigma_l^{(k+1)} = \frac{\sum_{i=1}^{s} r_{li}^{(k)} \left(y_i - \mu_l^{(k)}\right)\left(y_i - \mu_l^{(k)}\right)^T}{\sum_{i=1}^{s} r_{li}^{(k)}} \tag{4.20}$$

5. Step 2-4 are repeated until a stopping criterion is met. The stopping criterion for our algorithm is:

$$\mid J_\lambda(\psi^{(k+1)}) - J_\lambda(\psi^{(k)}) \mid \quad \leq \varepsilon \tag{4.21}$$

where ε is the convergence threshold.

4.8 HMRF-EM-SA algorithm

The performance of the proposed HMRF-FCEM algorithm is compared with that of existing HMRF-EM algorithm. As presented in Chapter 3, in E-step of HMRF-EM algorithm, Iterated Conditional Mode (ICM) algorithm is used. To avoid the difficulty of initial model parameter assumption from histograms of the degraded images, globally convergent Simulated Annealing (SA) algorithm is used to estimate the image labels in E- step.

4.8.1 Simulated Annealing Algorithm

Bayesian methods coupled with Markovian modelization usually result in a non-convex energy function. To find an estimate, one has to optimize this function. Unfortunately, this is a very hard computational problem known as combinatorial optimization. For example, considering an image 16 x 16 with only two possible labels at each pixel, we get a configuration space of 2^{256} elements But it is impossible to find the optimum by computing the possible values of the cost function. Here we can not use classical gradient descent methods because they stuck in a local minimum. It was realized in the early 80's

by Kirkpatrick *et al* [15] that there is an analogy between minimizing the cost function of a combinatorial optimization problem and finding energy minima of thermodynamical systems by slowly cooling a solid until equilibrium is reached. They have substituted the energy function of the solid by the cost function and executed the Metropolis algorithm at a sequence of slowly decreasing temperature. The so defined combinatorial optimization algorithm was named Simulated Annealing (SA).

SA algorithm is based on the analogy between the simulation of the annealing of solids and the solving of combinatorial optimization problems. It is inspired by an analogy between the physical annealing of solids (crystals) and combinatorial optimization problems. In the physical annealing process a solid if first melted and then cooled very slowly, spending a long time at low temperatures, to obtain a perfect lattice structure corresponding to a minimum energy state. SA transfers this process to local search algorithms for combinatorial optimization problem. It does so by associating the set of solutions of the problem attacked with the states of the physical system, the objective function with the physical energy of the solid, and the optimal solution with the minimum energy states.

Metropolis in the earliest days of scientific computing, introduced a simple algorithm that can be used to provide an efficient simulation of a collection of atoms in equilibrium at a given temperature. In each step of this algorithm, an atom is given a small random displacement and the resulting change ΔU, in the energy of the system is computed. If $\Delta U \leq 0$, the displacement is accepted, and the configuration with the displaced atom is used as the starting point of the next step. The case $\Delta U > 0$ is treated probabilistically: the probability that the configuration is accepted is $P(\Delta U) = exp(-\frac{\Delta U}{k_B T})$. Random numbers uniformly distributed in the interval $(0,\ 1)$ are a convenient means of implementing the random part of the algorithm. One such number is selected and compared with $P(\Delta U)$. If it is less than $P(\Delta U)$, the new configuration is retained , if not , the original configuration is used to start the next step. By repeating the basic step many times, one simulates the thermal motion of atoms in thermal contact with a heat bath at temperature T. The choice of $P(\Delta U)$ has the consequence that the system evolves into a Boltzman distribution [15].

Using the cost function in place of the energy and defining configuration by a set of parameters, it is straightforward with the Metropolis procedure to generate a population of configurations of a given optimization problem at some effective temperature. This temperature is simply a control parameter in the same units as the cost function. The simulated annealing process consists of first *melting* the system being optimized at a high effective temperature, then lowering the temperature at slow stages until the system freezes and no further changes occur. At each temperature, the simulation must proceed long enough for the system to reach a steady state. The sequence of temperature and the number of rearrangements of the parameters attempted to reach equilibrium at each temperature is known as an *annealing schedule* [15].

The SA algorithm used to obtain the MAP estimate is described below.

1. Initialize the temperature T_{in}.
2. Compute the energy U of the configuration.
3. Perturb the system slightly with suitable Gaussian disturbance.
4. Compute the new energy $U^{'}$ of the perturbed system and evaluate the change in energy $\Delta U = U^{'} - U$.
5. If $(\Delta U < 0)$, accept the perturbed system as the new configuration.
   Else accept the perturbed system as the new configuration with a probability $exp\ (-\Delta U)/k_B T$.
6. Decrease the temperature according to the cooling schedule.
7. Repeat steps 2-7 till the stopping criterion is met. The stopping criterion used here is the energy.

The steps of the HMRF-EM-SA algorithm are as follows.

1. Perform the initial parameter estimation and segmentation.

2. Calculate the likelihood distribution

$$p^{(t)}\left(y_i|x_i\right) = g^{(t)}\left(y_i;\theta\left(x_i\right)\right)$$

3. Estimate the class labels by MRF-MAP estimation

$$x^{(t)} = arg\max_{x\in X} P(y \mid x, \theta^t + P(x))$$

SA algorithm is used to estimate the class labels.

4. Calculate the posterior distribution

$$p^{(t)}\left(l \mid y_i\right) = \frac{g^{(t)}\left(y_i;\theta_l\right)p^{(t)}\left(l \mid x_{N_i}\right)}{p(y_i)}$$

5. Update parameters by

$$\mu_l^{(t+1)} = \frac{\sum_{i\in s} p^{(t)}\left(l \mid y_i\right)y_i}{\sum_{i\in s} p^{(t)}\left(l \mid y_i\right)} \qquad \sigma_l^{(t+1)^2} = \frac{\sum_{i\in s} p^{(t)}\left(l \mid y_i\right)\left(y_i-\mu_l\right)^2}{\sum_{i\in s} p^{(t)}\left(l \mid y_i\right)}$$

6. $t \leftarrow t+1$ and repeat from 2 until enough iterations have been performed.

4.9 Results and Discussions

The proposed HMRF-FCEM algorithm is validated with one 5-class synthetic image. Besides, 4 simulated and 3 real diseased brain MR images are considered. The degraded images are obtained by adding white Gaussian noise of varying strength to the original image. Brain MR images are obtained from the sources as described in chapter 3.

4.9.1 Synthetic Images:

The 5-class synthetic image considered for the simulation is shown in Fig. 4.1 (a). The corresponding noisy version of SNR 20 dB is shown in Fig. 4.1 (b). The initial model parameters μ and σ taken for each class are selected from the histogram of the noisy image and another set of parameter is on arbitrary basis. The parameter λ, which is the degree of fuzziness is selected on adhoc basis. The proposed HMRF-FCEM algorithm is used to obtain the segmented image of respective noisy image. The results obtained by

the proposed algorithm is shown in Fig. 4.1 (c) with parameters taken from histogram of the noisy image and Fig. 4.1 (d) with arbitrary set of parameters. As the number of clusters or classes are unknown, the algorithm starts from higher number of initial classes and eventually the algorithm converged to 5 classes. The performance of the proposed HMRF-FCEM algorithm is compared with that of HMRF-EM-SA algorithm. Fig. 4.1 (e) shows the segmented image using HMRF-EM-SA algorithm. The initial and converged parameters of the two schemes are tabulated in Table 4.1. It is observed that the segmented image obtained using HMRF-EM-SA algorithm has a number of misclassified pixels denoted by $\%$ of MCE to be 4.28. However, the segmented image obtained using proposed scheme with arbitrary model parameters has less number of misclassified pixels with $\%$ of MCE of 2.04 only. It is clear from the Table that even though the algorithm starts from two different initial conditions, the parameters converge to values that are vary close to each other. For example, μ of the third class starts from two different values and convergences to 2.51 and 2.27 which are very close to each other. Similar observations are also made for other parameters of other classes. Comparing the performance based on convergence time in sec., it is found that proposed scheme converges much faster than the HMRF-EM-SA scheme. Visually, the results obtained by proposed scheme are also found satisfactory.

4.9.2 Brain MR Images:

After successful implementation on the synthetic 5-class image, the proposed scheme is applied to 3 real and 4 simulated brain MR images. A simulated brain MR image of size (128×128) degraded with 3% noise is shown in Fig. 4.2 (a). The corresponding ground truth image is shown in Fig. 4.2 (b). After validation of the proposed algorithm with 2 different set of initial model parameters μ and σ, the segmented images are shown in Fig. 4.2 (c) and (d). The algorithm run with 6 number of initial class labels, but finally it converged to 4 number of class labels in both sets of initial parameters. The degree of fuzziness λ is taken as 0.5 for both the cases. Fig. 4.2 (e) shows the segmented image using HMRF-EM-SA algorithm. The *a priori* model parameter δ is considered on adhoc basis and for this image the value is taken as 0.1. The initial and final model

parameters along with the performance results consisting the $\%$ of misclassification error and convergence time are tabulated in Table. 4.2. From the table it is observed that the $\%$ of MCE of results obtained by proposed algorithm is 4.18 and 4.94 for Fig. 4.2 (c) and (d) respectively. The algorithm converges in 12 sec. and 15 sec. respectively for the same. In case of HMRF-EM-SA, the $\%$ of MCE is 5.09 with convergence time of 56 sec.

Similarly, another simulated MR image of size (128×128) degraded with 3% noise is shown in Fig. 4.3 (a). The corresponding ground truth image is shown in Fig. 4.3 (b). The segmented images using the proposed HMRF-FCEM algorithm with both sets of initial model parameters of μ and σ, are shown in Fig. 4.3 (c) and (d). The degree of fuzziness λ is considered as 0.5 for both the cases. The segmented image using HMRF-EM-SA algorithm is presented in Fig. 4.3 (e). The *a priori* model parameter δ is taken as 0.3. The initial and final model parameters along with the $\%$ of misclassification error and convergence time are tabulated in Table. 4.3. From the table it can be seen that the $\%$ of misclassification error of proposed algorithm is 2.92 with 20 secs. of convergence timing where as in case of HMRF-EM-SA it is 12.30 with 35 secs. of convergence time. It is observed from the table that the parameters converged to the values which are close to each other with two different set of initial parameters. This overcomes the difficulty of choice of initial model parameters in the proposed BHMRF-EM scheme presented in chapter 3.

The algorithm is validated with a sarcoma diseased real brain MR image of size (175×215) degraded with 3% noise is shown in Fig. 4.4 (a). The corresponding ground truth image is shown in Fig. 4.4 (b). Fig. 4.4 (c) and (d) are the segmented images of the proposed HMRF-FCEM algorithm with two defferent set of initial model parameters of μ and σ. The degree of fuzziness λ is considered as 0.3 for both the cases. The segmented image using HMRF-EM-SA algorithm is presented in Fig. 4.3 (e). The *a priori* model parameter δ is taken as 0.3. The less $\%$ of MCE is the accuracy in the segmentation, which can be observed from the table 4.4. The $\%$ of MCE of proposed algorithm is 3.80 with 28 sec. of convergence time but in case of HMRF-EM-SA it is 15.09 with 52 sec. of convergence time. Visually it is observed that the diseased area is clearly defined in the result obtained by the proposed algorithm rather than that of HMRF-EM-SA algorithm.

Another real brain MR image of size (175×215) degraded with 3% noise with multiple sclerosis is shown in Fig. 4.5 (a). The corresponding ground truth image is shown in Fig. 4.5 (b). The number of sclerosis present in the real image is successfully segmented to a different class label which can be visualised from the Fig. 4.5, where Fig. 4.5 (c) and (d) are the segmented images of the proposed HMRF-FCEM algorithm. The degree of fuzziness λ is considered as 1.0 for both the cases. The segmented image using HMRF-EM-SA algorithm is presented in Fig. 4.5 (e). The *a priori* model parameter δ is taken as 0.01. The $\%$ of MCE of proposed algorithm is 3.23 with 18 secs. of convergence timing. The misclassified pixels in case of HMRF-EM-SA is 14.14% and the convergence time is 36 secs.

Similarly, 2 more diseased brain MR images of size (175×215) and (128×128) degraded with 3% noise are shown in Fig. 4.6 (a) and 4.7 (a). The corresponding ground truth images are shown in Fig. 4.6 (b) and 4.7 (b). The segmented images using proposed HMRF-FCEM algorithm are shown in Fig. 4.6 (c) and 4.7 (c) and HMRF-EM-SA algorithm are shown in Fig. 4.6 (d) and 4.7 (d). The initial and final model parameters, the value of degree of fuzziness, the *a priori* MRF model parameter along with the $\%$ of MCE and convergence time are tabulated in Table 4.6 and Table 4.7 respectively. From Fig. 4.6 (c) the tumor along with the swelling area is segmented successfully assignig it as different class labels. In the same way, from Fig. 4.7 (c), the disease in the ventricle is viewed clearly but in Fig. 4.7 (d) it is not cleared.

The last brain MR image considered is shown in Fig. 4.8 (a). Fig. 4.8 (b) represents the ground truth image. The segmented image using HMRF-FCEM algorithm is shown in Fig.4.8 (c). From this it can be seen that the lateral and oxipetal horn of left and right ventricle are clearly segmented. In case of HMRF-EM-SA algorithm the above is not visible and the algorithm converges to less number of class labels. The segmented image using HMRF-EM-SA algorithm is shown in Fig.4.8 (d). The model parameters and other performance results are tabulated in Table 4.8.

All segmentation results of brain MR images using our proposed schemes are consulted with Dr. Hemalata Satapathy, sonologist and radiologist of CWS Hospital, Rourkela and found to be satisfactory.

4.10 Conclusion

In this chapter fuzzy clustering approach is effectively incorporated in HMRF model to hybridized the benefits of both for brain MR image segmentation. HMRF-FCEM algorithm is proposed to formulate the above problem where the fuzzy objective function is optimized by employing a mean-field-like approximation of the *a priori* MRF distribution. The proposed algorithm could estimate the image labels and model parameters recursively. The algorithm does not assume to have the knowledge of the number of classes. It could yield satisfactory results even with arbitrary initial condition. The phenomena is consistently observed in simulated as well as real brain MR images. The proposed algorithms performance is compared with that of HMRF-EM scheme with global convergent Simulated Annealing algorithm. It is observed that the results obtained with the proposed algorithm outperformed the other one. The potentiality of the algorithm is that it does not need to have *a priori* proper set of model parameters. Our results may be further improved by estimating the bias field to take care of the intensity inhomogeneity of the brain MR images.

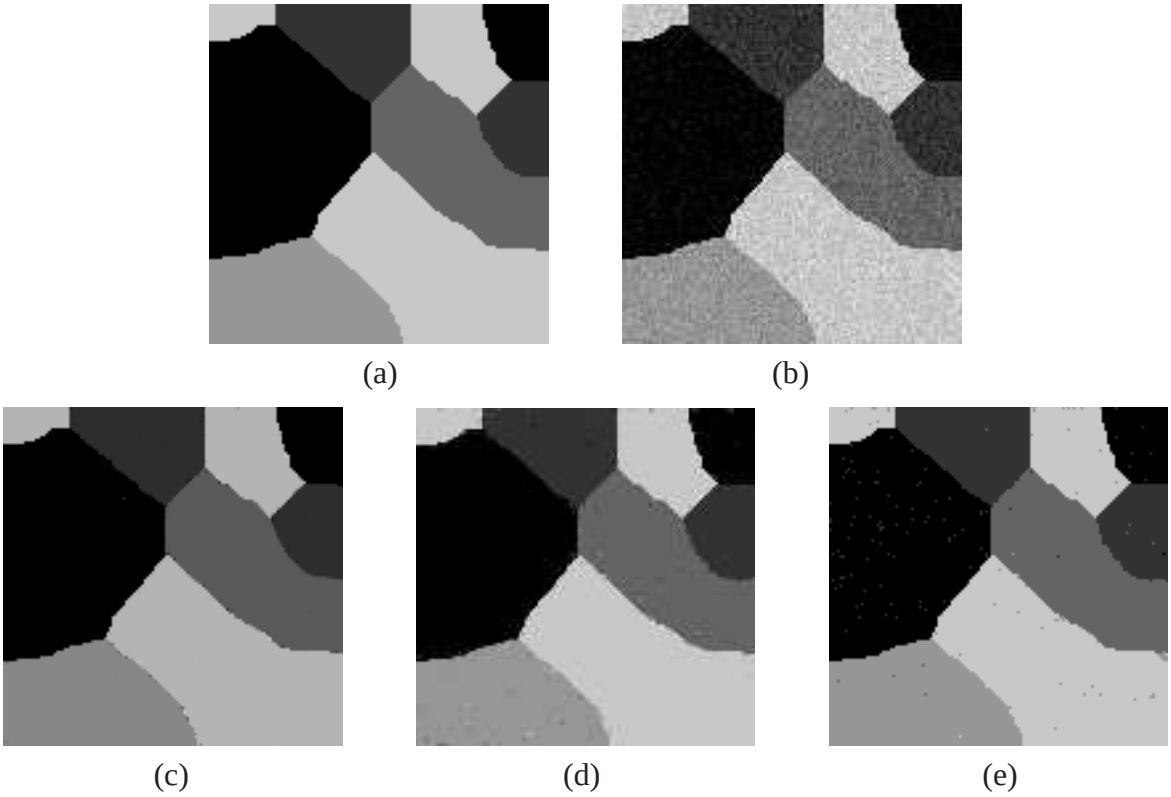


Figure 4.1: *Unsupervised image segmentation of synthetic 5-class image of size* (128×128)*: (a) Original image (b) noisy image with 20 db SNR (c) and (d) segmented image using HMRF-FCEM framework with histogram based initial parameters and arbitrary initial parameters (e) segmented image using HMRF-EM-SA framework*

Schemes and Parameters		class $\rightarrow 1$	2	3	4	5	$\%$ of MCE	Convrge. time (sec.)
Histogram based initial parameters	μ_i	0.0	0.9	2.04	3.10	3.80		
	σ_i	0.37	0.68	0.54	0.60	0.43		
HMRF-FCEM $\lambda = 1.0$	μ_f	0.25	1.54	2.51	3.35	4.07		
	σ_f	0.16	0.48	0.29	0.09	0.07	1.37	15
HMRF-EM-SA $\delta = 0.62$	μ_f	0.07	0.99	2.01	3.14	4.01		
	σ_f	0.10	0.26	0.28	0.39	0.25	4.28	48
Arbitrary initial parameters	μ_i	1.16	1.96	3.08	3.78	2.51		
	σ_i	0.37	0.46	0.62	0.52	0.57		
HMRF-FCEM $\lambda = 0.5$	μ_f	0.19	1.02	2.27	3.37	4.31		
	σ_f	0.67	0.84	0.59	0.08	0.12	2.04	18

Table 4.1: *Image model parameters of synthetic 5-class image of size* (128×128) *using HMRF-FCEM framework with histogram based initial parameters and arbitrary initial parameters, HMRF-EM-SA framework of Fig. 4.1*

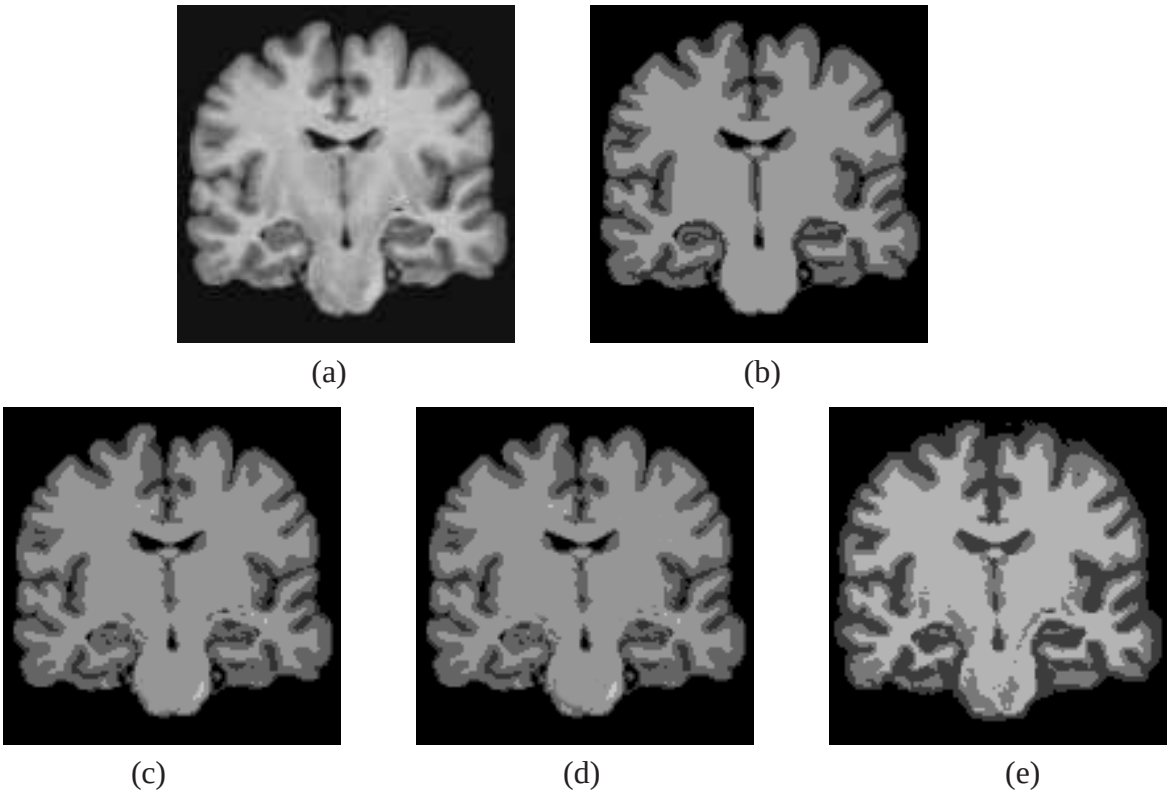


Figure 4.2: *Unsupervised image segmentation of Brain MR image of size* (128×128) *(a) Original image with 3* $\%$ *of noise (b) Ground Truth (c),(d) segmented image using HMRF-FCEM framework with histogram based initial parameters and arbitrary initial parameters, (e) segmented image using HMRF-EM-SA framework*

Schemes and Parameters		class → 1	2	3	4	% of MCE	Conver. time (sec.)
Histogram based initial parameters	μ_i	1.14	1.98	3.02	3.96		
	σ_i	0.55	0.57	0.67	0.23		
HMRF-FCEM $\lambda = 0.5$	μ_f	0.80	3.53	2.61	2.43		
	σ_f	0.48	3.32	1.52	0.88	4.18	12
HMRF-EM-SA $\delta = 0.1$	μ_f	0.09	1.26	1.89	2.99		
	σ_f	0.03	0.19	0.10	0.09	5.09	56
Arbitrary initial parameters	μ_i	0.09	1.80	2.96	4.08		
	σ_i	0.12	0.68	0.58	0.40		
HMRF-FCEM $\lambda = 0.5$	μ_f	2.35	2.93	2.71	4.04		
	σ_f	0.89	2.84	1.89	0.08	4.94	15

Table 4.2: *Image model parameters of brain MR image of size* (128×128) *using HMRF-FCEM framework with histogram based initial parameters and arbitrary initial parameters, HMRF-EM-SA framework of Fig. 4.2*

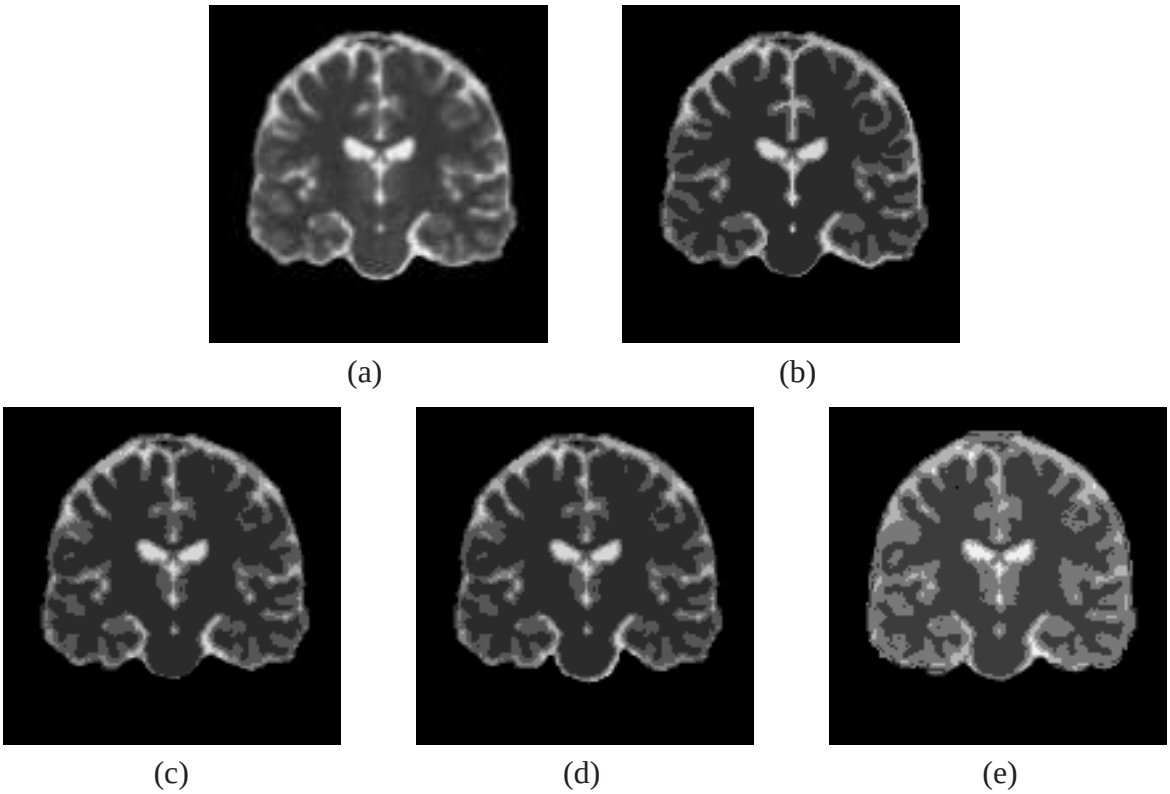


Figure 4.3: *Unsupervised image segmentation of Brain MR image of size* (128×128) *(a) simulated image with 3* $\%$ *of noise (b) Ground Truth (c)-(d) segmented image using HMRF-FCEM framework with histogram based initial parameters and arbitrary initial parameters (e) segmented image using HMRF-EM-SA framework*

Schemes and Parameters		class → 1	2	3	4	% of MCE	Conver. time (sec.)
Histogram based initial parameters	μ_i	0.0	0.9	1.8	2.96		
	σ_i	0.12	0.57	0.68	0.58		
HMRF-FCEM $\lambda = 0.5$	μ_f	0.62	2.67	2.90	3.03		
	σ_f	0.50	2.94	2.21	1.45	2.92	20
HMRF-EM-SA $\delta = 0.3$	μ_f	0.07	1.05	1.56	2.73		
	σ_f	0.01	0.07	0.16	0.22	12.30	35
Arbitrary initial parameters	μ_i	1.0	1.86	2.86	4.08		
	σ_i	0.65	0.52	0.48	0.10		
HMRF-FCEM $\lambda = 0.5$	μ_f	1.76	1.73	2.58	3.95		
	σ_f	0.35	1.35	2.19	1.38	3.45	24

Table 4.3: *Image model parameters of brain MR image of size* (128×128) *using HMRF-FCEM framework with histogram based initial parameters and arbitrary initial parameters, HMRF-EM-SA framework of Fig. 4.3*

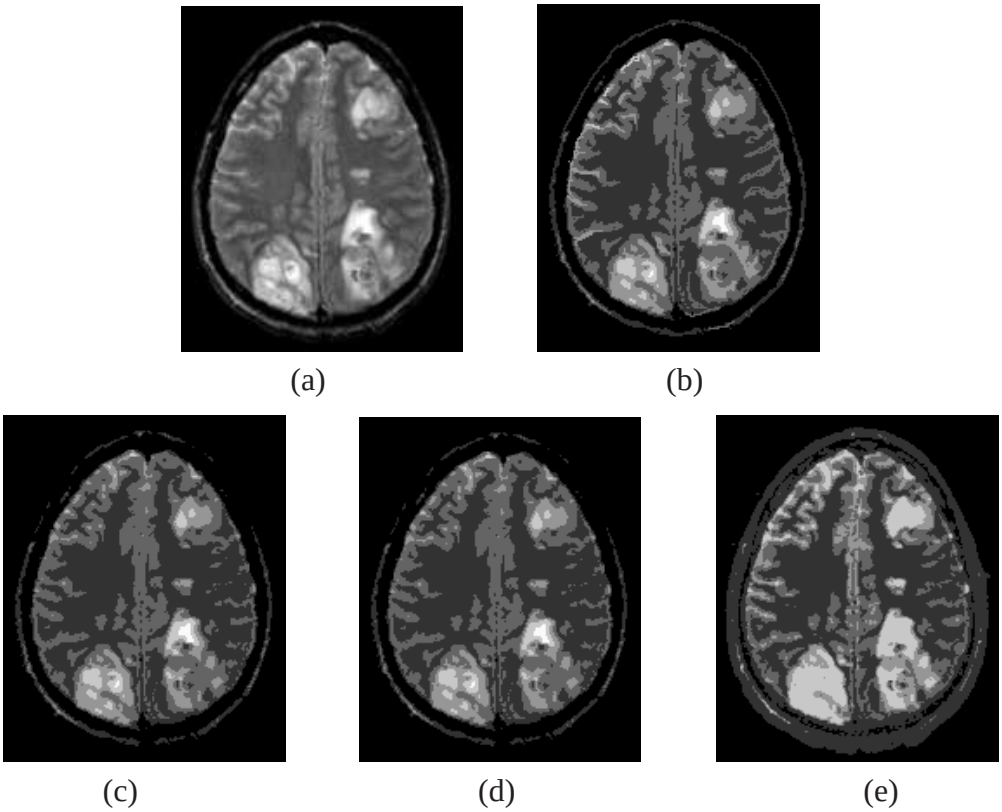


Figure 4.4: *Unsupervised image segmentation of Sarcoma diseased Brain MR image of size* (175×215) *(a) Real image with 3* $\%$ *of noise (b) Ground Truth (c)-(d) segmented image using HMRF-FCEM framework with histogram based initial parameters and arbitrary initial parameters (e) segmented image using HMRF-EM-SA framework*

Schemes and Parameters		class → 1	2	3	4	% of MCE	Conver. time (sec.)
Histogram based initial parameters	μ_i	0.50	0.57	0.52	0.67		
	σ_i	1.18	1.84	3.0	3.86		
HMRF-FCEM	μ_f	2.74	2.63	2.88	3.07		
$\lambda = 0.3$	σ_f	1.91	1.05	1.27	2.23	3.80	28
HMRF-EM-SA	μ_f	1.60	1.95	2.39	3.38		
$\delta = 0.3$	σ_f	0.08	0.10	0.12	0.23	15.09	52
Arbitrary initial parameters	μ_i	0.14	1.02	1.98	2.09		
	σ_i	0.38	0.54	0.58	0.69		
HMRF-FCEM	μ_f	0.78	2.90	2.86	3.09		
$\lambda = 0.3$	σ_f	0.23	1.29	0.83	1.91	4.04	25

Table 4.4: *Image model parameters of Real brain MR image of size* (175×215) *using HMRF-FCEM framework with histogram based initial parameters and arbitrary initial parameters, HMRF-EM-SA framework of Fig. 4.4*

Schemes and Parameters		class → 1	2	3	4	% of MCE	Conver. time (sec.)
Histogram based initial parameters	μ_i	1.34	2.14	3.16	4.08		
	σ_i	0.28	0.31	0.33	0.34		
HMRF-FCEM	μ_f	1.11	2.0	2.34	4.13		
$\lambda = 1.0$	σ_f	0.69	3.08	2.46	0.36	3.23	18
HMRF-EM-SA	μ_f	1.35	1.81	2.62	3.42		
$\delta = 0.01$	σ_f	0.05	0.11	0.20	0.12	14.14	36
Arbitrary initial parameters	μ_i	0.24	1.08	1.72	3.04		
	σ_i	0.29	0.60	0.54	0.67		
HMRF-FCEM	μ_f	2.18	2.70	3.22	2.64		
$\lambda = 1.0$	σ_f	0.39	1.32	1.68	3.28	5.39	25

Table 4.5: *Image model parameters of brain MR image of size* (175×215) *using HMRF-FCEM framework, HMRF-EM-SA framework of Fig. 4.5*

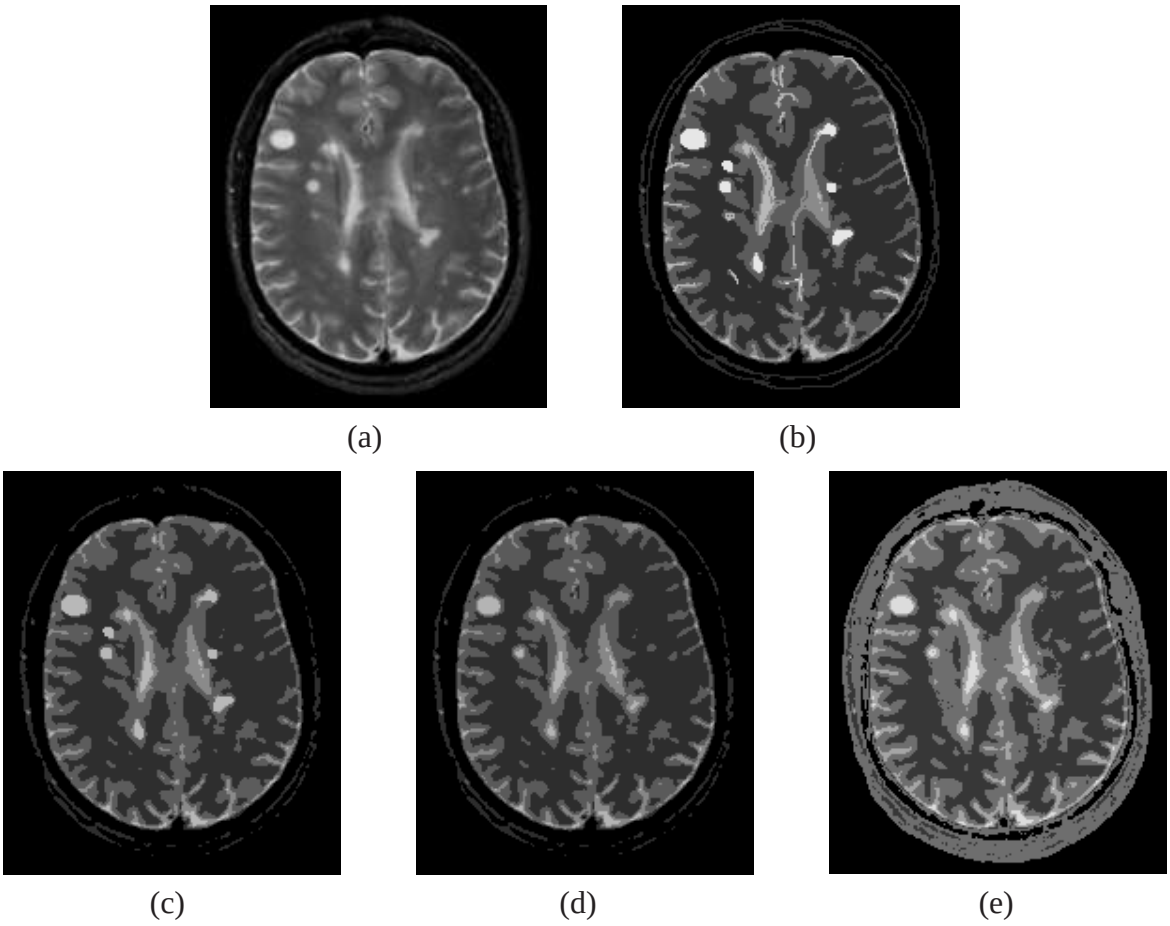
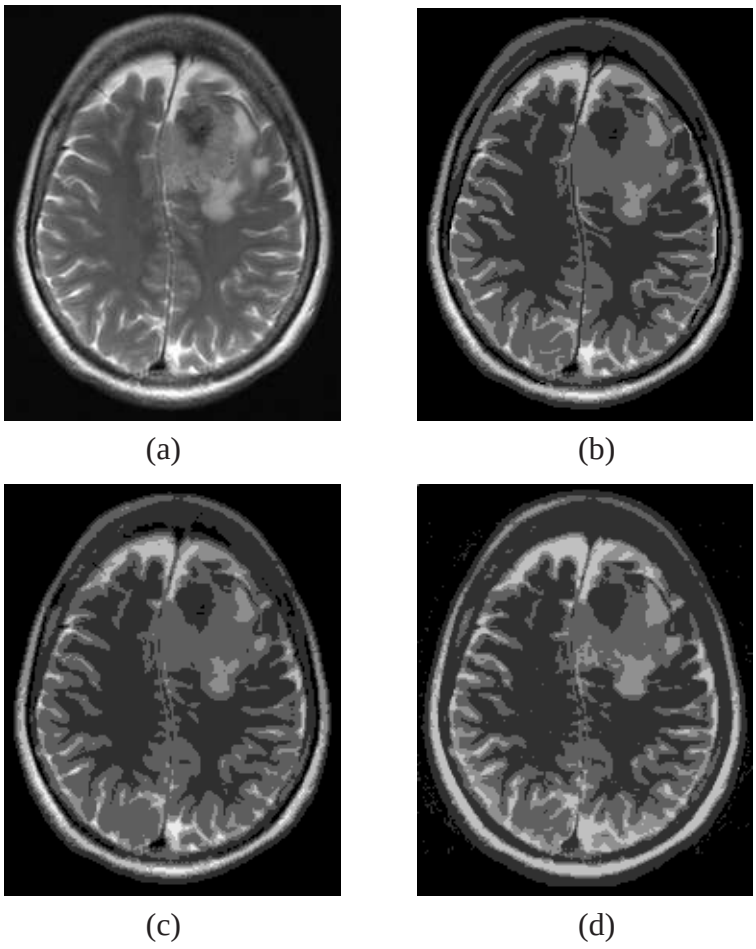


Figure 4.5: *Unsupervised image segmentation of Multiple sclerosis from a Brain MR image of size* (175×215) *(a) Real image with 3* $\%$ *of noise (b) Ground Truth (c) and (d) segmented image using HMRF-FCEM framework with histogram based initial parameters and arbitrary initial parameters (e) segmented image using HMRF-EM-SA framework*

Schemes and Parameters		class $\rightarrow$ 1	2	3	4	$\%$ of MCE	Conver. time (sec.)
Histogram based initial parameters	μ_i	0.12	1.16	1.84	2.86		
	σ_i	0.22	0.51	0.61	0.56		
HMRF-FCEM $\lambda = 1.0$	μ_f	0.14	2.01	2.93	3.15		
	σ_f	0.02	0.32	1.96	0.74	3.50	20
HMRF-EM-SA $\delta = 0.25$	μ_f	0.16	1.52	1.82	2.36		
	σ_f	0.01	0.09	0.07	0.19	18.09	42

Table 4.6: *Image model parameters of Real brain MR image of size* (175×215) *using HMRF-FCEM framework, HMRF-EM-SA framework of Fig. 4.6*

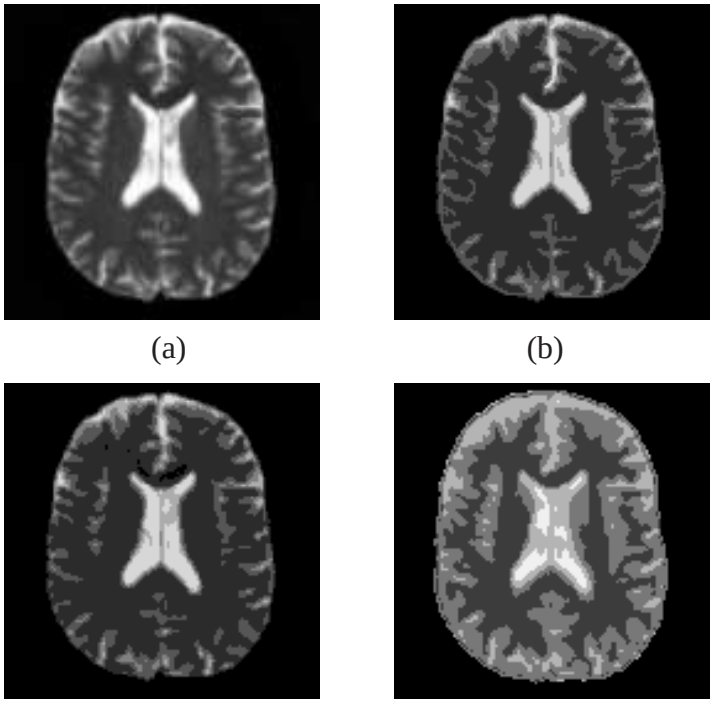


(a) (b) (c) (d)

Figure 4.6: *Unsupervised image segmentation of tumor from a Brain MR image of size* (175×215) *(a) Real image with 3* $\%$ *of noise (b) Ground Truth (c) segmented image using HMRF-FCEM framework (d) segmented image using HMRF-EM-SA framework*

Schemes and Parameters		class $\rightarrow$ 1	2	3	4	$\%$ of MCE	Conver. time (sec.)
Histogram based initial parameters	μ_i	0.74	1.86	3.0	3.96		
	σ_i	0.66	0.56	0.57	0.73		
HMRF-FCEM $\lambda = 2.0$	μ_f	0.64	2.29	2.92	3.17		
	σ_f	0.52	1.71	1.87	2.50	2.41	28
HMRF-EM-SA $\delta = 0.2$	μ_f	0.04	0.96	1.59	3.04		
	σ_f	0.09	0.05	0.18	0.31	10.97	40

Table 4.7: *Image model parameters of diseased brain MR image of size* (128×128) *using HMRF-FCEM framework, HMRF-EM-SA framework of Fig. 4.7*



(a) (b)

(c) (d)

Figure 4.7: *Unsupervised image segmentation of diseased Brain MR image of size* (128×128) *(a) Original image with 3* $\%$ *of noise (b) Ground Truth (c) segmented image using HMRF-FCEM framework (d) segmented image using HMRF-EM-SA framework*

Schemes and Parameters		class $\rightarrow$ 1	2	3	4	$\%$ of MCE	Conver. time (sec.)
Histogram based initial parameters	μ_i	0.01	1.10	1.73	3.0		
	σ_i	0.25	0.69	0.41	0.34		
HMRF-FCEM $\lambda = 1.0$	μ_f	1.24	2.15	3.29	3.72		
	σ_f	0.27	2.88	1.03	1.87	2.67	40
HMRF-EM-SA $\delta = 0.3$	μ_f	1.34	1.92	2.97	2.57		
	σ_f	0.17	0.08	0.07	0.02	14.40	84

Table 4.8: *Image model parameters of brain MR image of size* (128×128) *with HMRF-FCEM, HMRF-EM-SA framework of Fig. 4.8*

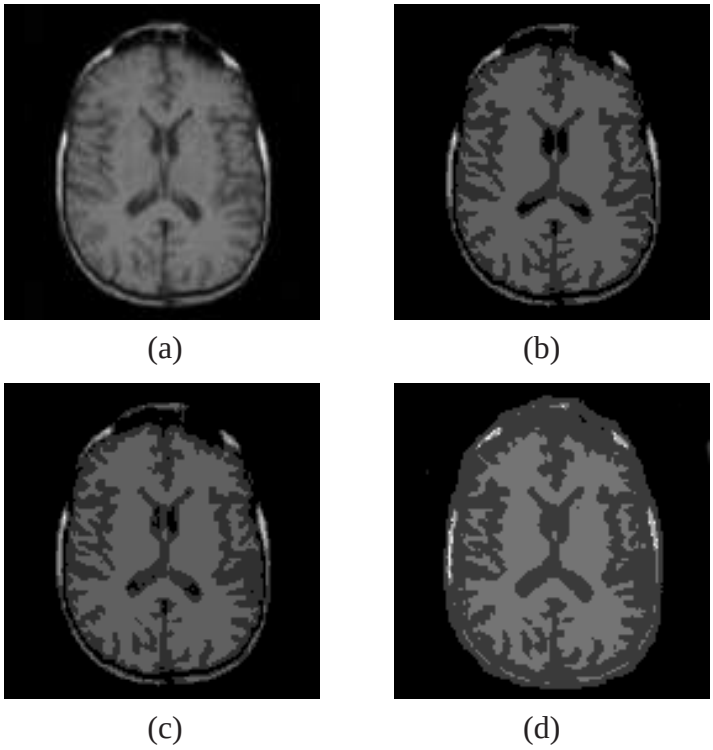


Figure 4.8: *Unsupervised image segmentation of Brain MR image of size* (128×128) *(a) Original image with 3* $\%$ *of noise (b) Ground Truth (c) segmented image using HMRF-FCEM framework (d)segmented image using HMRF-EM-SA framework*

Chapter 5

Conclusions

The objective of this dissertation is to devise methods and strategies for segmentation of brain MR images in unsupervised framework. This work attempt to develop unsupervised brain MR image segmentation schemes that would facilitate the clinical experts for an automatic segmentation and accurate diagnosis.

The initial portion of this thesis provides a background on HMRF, MRF models and fuzzy clustering methods for image segmentation. These are covered in Chapter 2. Standard fuzzy c-means algorithm is also included here.

The initial part of the research work is dedicated towards devising unsupervised brain MR image segmentation scheme using HMRF model which is included in Chapter 3. In this framework, the problem is cast as a pixel labeling problem, and MRF and HMRF models are employed to model the *a priori* unknown class labels and the observed degraded image respectively. HMRF model parameters (μ, σ) for each tissue class are assumed to be unknown. The MRF model parameters δ is assumed on an adhoc basis. HMRF model is modified as BHMRF model, where the energy function of the *a priori* model is modified by incorporating the biased neighborhood structure. This proposed modified BHMRF model is formulated in EM framework to jointly estimate the model parameters as well as image labels. Image label estimates are obtained by ICM algorithm in the E- step of the EM algorithm. It is observed that the proposed BHMRF-EM algorithm outperformed HMRF-EM algorithm. The algorithm yielded satisfactory results in both synthetic as well as brain MR images. Both algorithms are compared with respect to two performance measure, i.e. percentage of misclassification error and execution time

in sec. It is observed that due to incorporation of biased neighborhood interaction in the energy function of the *a priori* MRF model, natural structures of brain could be taken care of in the model. This could able to reduce the misclassification error in cleaner data which was the bottleneck in HMRF-EM scheme. The percentage of misclassification error is also found to be less in case of noisier data in our proposed scheme with comparable execution time.

This proposed BHMRF-EM algorithm is modified and exploited to take care of the intensity in-homogeneity are the bias field in the brain MR images. The proposed modified BHMRF-EM algorithm could jointly estimate the image labels, model parameters as well as bias field. This is indeed a step towards making an automated brain MR image segmentation scheme. Our algorithm starts with assuming more number of classes than that of actual ones and converges to the required number of classes. Since the actual number of classes are unknown, the selected number of initial classes are also on trial and error basis. The initial model parameters are assumed on histogram based strategy. It will be worth pursuing in future to develop strategies to overcome these problems.

Segmentation of brain MR images based on a new notion where HMRF model is incorporated in to fuzzy clustering scheme is introduced in chapter 4. In the proposed scheme, benefits of both HMRF model and fuzzy clustering procedure are combined together. Motivated by Celeux *et al.*, in our proposed approach, MRF model is used to model the *a priori* unknown class labels by employing a mean-field-like approximation. This assumption of MRF yields good estimates of the Markov posteriors with less computational costs. HMRF oriented fuzzy objective function is formulated by considering HMRF model to be defining as a fuzzy HMRF model to be defining as a fuzzy partitions of the observation space. The dissimilarity function become the negative log-likelyhood of the model state in the modified fuzzy objective function. This modified function is minimized by the proposed HMRF-fuzzy clustering EM (HMRF-FCEM) algorithm. The problem of selection of initial model parameters is overcome by our proposed algorithm. The image labels and fuzzy membership function is optimized. This proposed algorithm is an alternative to EM- type treatments of the HMRF model, based on the mean-field-like approximation of the MRF prior. The algorithm performed satisfactory when starting

from arbitrary initial model parameters. this algorithm does not need to have proper choice of initial model parameters. This is indeed a step towards making a complete automated brain MR image segmentation scheme. Since the actual number of classes are unknown, the selected number of initial classes are on trial and error basis. Large number of initial classes means estimation of large number of parameters which in turn increases computational burden. It will be worth persuing in future to develop strategies to overcome these problems. However, the proposed algorithm could be successfully tested for synthetic as well as real brain MR images with better performance than the estimating HMRF-EM algorithm. In E-step, the image labels are estimated using global convergent simulated annealing (SA) algorithm in stead of ICM algorithm. Though this HMRF-EM-SA algorithm overcomes the initial assumption of model parameters, but the computational burden is very high. In our proposed HMRF-FCEM algorithm, the performance is tested on percentage of misclassification error and execution time in sec. Both performance measures show better scales in HMRF-FCEM algorithm than HMRF-EM-SA algorithm. Joint estimation of bias field would enhance the worth of the work.

Bibliography

[1] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the bayesian restoration of images,"*IEEE Trans. Pattern Analysis and Machine Intelligence*, vol.6, no. 6, pp 721-741,1984.

[2] Julian Besag, "On the statistical Analysis of Dirty pictures,"*Royal Statistical Society*, 48, no. 3, pp. 259-302, 1986.

[3] Nikhil R. Pal and Sankar K. Pal, "A review on image segmentation techniques,"*Pattern Recognition*, vol. 26, no. 9, pp. 1277-1294, Sept. 1993.

[4] Todd R. Reed and J. M. Hans Du Buf, "A review of recent texture segmentation and feature extraction techniques," *CVGIP: Image Understanding*, vol.57, no.3, pp.359-372, 1993.

[5] D. L. Pham, C. Xu, and J. L. Prince. "Current methods in medical image segmentation," *Annual Review of Biomedical Engineering*,vol.2 pp. 315-337, 2000.

[6] James S. Duncan and Nicholas Ayache, "Medical Image Analysis: Progress over Two Decades and the challenges Ahead," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol.22, no.1, pp. 85-106, 2000.

[7] S.A.Barker and P.J.W.Rayner, "Unsupervised image segmentation using Markov random field models," *Pattern Recognition* vol. 33, no. 4, pp. 587-602, April. 2000.

[8] P. K. Nanda, "MRF model learning and application to image restoration and segmentation,"*Ph.D Dissertation, IIT Bombay* , 1995.

[9] J.Zhang, J.W.Modestino and D.A.Langan, "Maximum Likelihood parameter estimation for unsupervised stochastic model-based image segmentation,"*IEEE Trans. Image Processing*, vol.3, no. 4, pp. 404-420, 1994.

[10] Stan Z. Li, "Markov Random Field Modelling in Image Analysis," Springer-Verlag, 2001.

[11] A. P. Dempster, N. M. Laird and D. B. Rubin, "Maximum Likelihood from Incomplete Data via the EM algorithm,"*J.Royal Statistical Society*, vol. 39, pp. 1-38, 1977.

[12] Padhraic Smyth, "Belief networks, hidden Markov models and Markov random fields:A unifying view,"*Pattern Recognition Letters*, vol. 18, no. 11, pp. 1261-1268, Nov. 1997.

[13] Yongyue Zhang, Michael Brady and Stephen Smith, "A hidden markov random field model for partial volume classification,"*Neuro Image*, vol. 13, no. 6, pp.291, June 2001.

[14] S. Ruan, B. Moretti, J. Fadili and D. Bloyet,"Fuzzy Markovian segmentation in application of magnetic resonance images,"*Computer Vision and Image Understanding.* vol. 85, pp. 54-69, 2002.

[15] A. Kirkpatrick, "Optimization by Simulated Annealing: Quantitative studies," *Statistical physics*, vol. 34, pp. 975-986, 1984.

[16] Todd K.Moon, "The Expectation-Maximization Algorithm," *IEEE Signal Processing*, pp. 47-59, Nov. 1996.

[17] Youngyue Zhang, Michael Brady and Stephen Smith, "Segmentation of Brain MR Images Through a Hidden Markov random Field Model and the Expectation-Maximization Algorithm,"*IEEE Trans. Medical Imaging*, vol.20, no.1, pp. 45-57, 2001.

[18] J. L. Marroquin, E. A. Santana and S. Botello, "Hidden Markov Measure Field Models for Image segmentation," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol.25, no.11, 1380-1387, 2003.

[19] W. M. Wells, W. E. L. Grimson, R. Kikinis, F. A. Jolesz, "Adaptive Segmentation of MRI Data,"*IEEE Trans. Medical Imaging*, vol. 15, pp. 429-442, Aug. 1996.

[20] R.Guillemaud and Michael Brady, "Estimating the Bias Field of MR Images,"*IEEE Trans. Medical Imaging*, vol. 16, pp. 238-251, June. 1997.

[21] K. Held, E. R. Kops, B. J. Krause, W. M. Wells and R. Kikinis, H. W. M. Gartner, "Markov Random Field Segmentation of Brain MR Images,"*IEEE Trans. Medical Imaging*, vol. 16, pp. 878-886, Dec. 1997.

[22] Wen-Hung Chaoa,b, You-Yin Chena, Sheng-Huang Linc, Yen Yu I. Shihd, Siny Tsange , "Automatic segmentation of magnetic resonance images using a decision tree with spatial information," *Computerized Medical Imaging and Graphics* vol. 33, pp. 111-121, 2009.

[23] Thakur A. and R.S. Anand, "Image Quality Based Comparative Evaluation of Wavelet Filters in Ultrasound Speckle Reduction," *Digital Signal Processing*, Vol.15, No.5, pp 455-465, 2005.

[24] H. Guan, T. Kubota, X. Huang, X. S. Zhou, and M. Turk. "Automatic hot spot detection and segmentation in whole body fdg-pet images," *Proceedings of IEEE International Conference on Image Processing (ICIP)*, 2006.

[25] P. K. Nanda, D. Patra and A. Pradhan, "Brain MR Image Segmentation using Tabu Search and hidden Markov random field model," *Proc. of 2nd Indian International conference on Artificial Intelligence*, pp. 3143-3160, 2005.

[26] M. Joshi and A. Jalobeanu, "Multiresolution fusion in remotely sensed images using an IGMRF prior and MAP estimation," *IEEE Trans. on Geoscience and remote sensing* , vol. 48, no. 3, pp.1245-1255, Mar. 2010.

[27] J. Suckling, T. Sigmundsson, K. Greenwood and E.T.Bullmore, "A modified fuzzy clustering algorithm for operator independent brain tissue classification of Dual Echo MR images," *Magnetic Resonance Imaging*, vol. 17, No. 7, pp. 1065-1076, 1999.

[28] H.Ichihashi, k. Honda, and N. Tani,"Gaussian mixture pdf approximation and fuzzy c-means clustering with entropy regularization,"*Proceedings of 4th Symp. on Asian Fuzzy system* , pp.217-221, 2000.

[29] M.N.Ahmed, S.M.Yamany, Nevin Mohamed, A.A.Farag and T.Moriarty, "A modified fuzzy C-means algorithm for bias-field estimation and segmentation of MRI data,"*IEEE Trans. Medical Imaging*, vol. 21, No. 3, pp. 193-199, Mar. 2002.

[30] A. Roy, S. K. Parui, A. Paul, U. Roy, "A Color Based Image Segmentation and its Application to Text Segmentation," *IEEE Conference on Computer Vision, Graphics and Image Processing*, pp. 313-319, Mar. 2008.

[31] A.V. Nandedkar, P.K. Biswas, "A General Reflex Fuzzy Min-Max Neural Network," *International Conference on Pattern Recognition*, pp. 650-653 , Aug. 2006.

[32] W. Chen, M. L. Giger , "A Fuzzy C-means (FCM) based algorithm for intensity inhomogeneity correction and Segmentation,"*IEEE Trans. on Biomedical Imaging*, vol.2, pp.1307-1310, April 2004.

[33] M.Y. Siyal and Lin Yu, "An intelligent modified fuzzy c-means based algorithm for bias estimation and segmentation of brain MRI," *Pattern Recognition letters*, vol. 26, pp. 2052-2062, 2005.

[34] Ping Wang, HongLei Wang, "A modified FCM algorithm for MRI brain image segmentation," *IEEE Trans. on Future Bio-Medical Information Engineering*, pp.26-29, Dec.2008.

[35] AboulElla Hassanien, "Fuzzy rough sets hybrid scheme for breast cancer detection," *Image and Vision Computing*, vol.25, pp. 172-183, 2007.

[36] Tabakov Martin, "A fuzzy clustering technique for medical image segmentation," *Proceedings of internal symposium on evolving fuzzy systems*, pp. 118-122, sept. 2006.

[37] K. Jiayin, Min Lequan, "Novel modified fuzzy c-means algorithm with applications," *Digital signal processing*, vol. 19, no.2, pp. 309-319, Mar, 2009.

[38] Tolias Yannis A. and Panas Stavros M., "Image segmentation by a fuzzy clustering algorithm using adaptive spatially constrained functions," *IEEE Transactions on systems, Man, and Cybernetics* vol.28, issue.3, pp. 359-369, May, 1998.

[39] Mohamed N. A.,Ahmed M. N., "Modified fuzzy c-means in medical image segmentation," *Proceedings of IEEE intenational conference on Acoustics, speech, and signal processing*, vol. 6, pp. 3429-3432, 1999.

[40] Kannan S.R., "A new segmentation system for MR images based on fuzzy techniques," *Applied soft computing*, vol. 8, issue. 4, pp. 1599-1606, Sept. 2008.

[41] X. Li, L. Li, H. Lu, D. Chen, and Z. Liang, "Inhomogeneity correction for magnetic resonance images with fuzzy c-means algorithm," *Proc. SPIE* 2003.

[42] D.L. Pham, J. L. Prince, "Adaptive fuzzy segmentation of magnetic resonance images, *IEEE Trans. Medical Imaging*, vol. 18, 1999.

[43] Chuang Keh-shih, T. H. Long, "Fuzzy C-means clustering with spatial information for image segmentation," *Computerized Medical imaging and graphics*, 2006.

[44] Yang Zhang, Chung Fu-Lai, "Robust fuzzy clustering based image segmentation," *Applied soft computing*, pp.80-84, Jan, 09.

[45] G. Celeux, F. Forbes and N. Peyrard, "EM procedures using mean field like approximations for Markov model based image segmentation," *Pattern recognition*, vol. 36, no. 1, pp. 131-144, 2003.

List of Publications

Communicated:-

- D. Patra and S. Pradhan, "Development of Unsupervised brain MR Image Segmentation using fuzzy clustering based Hidden Markov Random Field Model," submitted to *SPCON-10*.

Published:-

1. T. Haribabu, Smita Pradhan and Dipti Patra, "Tissue Classification of Brain MR Images using Adaptive Fuzzy C-Means Algorithm," In Proceeding of National Conference on *Computational intelligence, control and computer vision in robotics and automation*, pp.162-166, Mar 2008, NIT, Rourkela.

2. Dipti Patra, and Smita Pradhan, "Intensity Inhomogeneity correction of Brain MRI using EM Algorithm," In Proceeding of National Conference on *Computational intelligence, control and computer vision in robotics and automation*, pp.194-198, Mar 2008, NIT, Rourkela.

3. Smita Pradhan and Dipti Patra, "Brain MR Image Segmentation using Biased Hidden Markov Random Field Model," In Proceeding of National Conference on *Advancement in Wireless Technologies and its Applications*, pp. 130-134, Dec 2008, NIT,Surat.

4. Smita Pradhan and Dipti Patra, "Unsupervised Brain MR Image Segmentation using Biased Hidden Markov Random Field Model," *Indian International Conference on Artificial Intelligence (IICAI)*, pp. 1519-1530, Dec 2009, SIT,Banglore.

5. Smita Pradhan and Dipti Patra, "Unsupervised Brain MR Image Segmentation using HMRF-FCM framework," *IEEE Indian International Conference*, pp. 93-96, Dec 2009, DAIICT,Gandhinagar.