Maximum Likelihood Parameter Estimation of Mixture Models and Its Application to Image Segmentation and Restoration

by

Mohammed Saeed

B.S.E.E, Purdue University (1995)

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

Science Master

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 1997

Author .....
Department of Electrical Engineering and Computer Science
May 30, 1997

Certified by .....
W. Clement Karl
Research Scientist
Thesis Supervisor

Accepted by .....
Arthur C. Smith
Chairman, Departmental Committee on Graduate Students

Maximum Likelihood Parameter Estimation of Mixture Models and Its Application to Image Segmentation and Restoration

by

Mohammed Saeed

Submitted to the Department of Electrical Engineering and Computer Science on May 30, 1997, in partial fulfillment of the requirements for the degree of Science Master

Abstract

Our contributions involve extending and improving Gaussian mixture model (GMM)-based approaches to segmentation and restoration. For segmentation, we extend the GMM paradigm by incorporating a multiscale correlation model of pixel dependance into the standard approach. In particular, we modify the standard GMM by introducing a multiscale neighborhood clique that incorporates the correlation between pixels in space and scale. We modify the likelihood function of the image field by a penalization term that is derived from our multiscale neighborhood clique. Maximum Likelihood (ML) estimation via the Expectation Maximization (EM) algorithm is used to estimate the parameters of our new model. Then, utilizing the parameter estimates, we segment the image field with a MAP classifier. We demonstrate that our new algorithm provides superior segmentations of synthetic images, as well as Magnetic Resonance (MR) images of the human brain, yet is computationally efficient.

The restoration of digital images that are corrupted with additive Gaussian and/or impulsive noise is a difficult task. A salient requirement for an effective algorithm is that details and edges in the image must be preserved. In this thesis, we demonstrate that a window-based Gaussian mixture model can be applied in the development of a robust nonlinear filter for image restoration. Via the EM algorithm, we utilize ML estimation of the spatially-varying model parameters to achieve the desired noise suppression and detail preservation. We demonstrate that this approach is a powerful tool which gives us information into the local statistics of noisy images. We demonstrate that the estimated local statistics can be efficiently utilized for outlier detection, edge detection, and suppression of quantization error in image coding.The advantages of our algorithm is that it can simultaneously suppress additive Gaussian and impulsive noise, while preserving fine details and edges.

Thesis Supervisor: W. Clement Karl
Title: Research Scientist

Acknowledgments

Thus far MIT graduate school has been the most challenging and rewarding experience in my life. I have had the opportunity to conduct interesting research, enroll in intriguing courses, and I have had the good fortune of teaching aspiring young students with brilliant futures ahead of them. Here, it is my pleasure to acknowledge all the people who have contributed to my great experience at MIT.

First, I am deeply indebted to Professor Clem Karl for serving as my thesis advisor. Clem always challenged me to strive for excellence. He always asked thought-provoking questions and gave me the room to freely think about how I should solve the problems addressed in this thesis. He also spent a great deal of effort in helping me clarify the writing of my thesis. The skills I learned in writing this thesis will prove invaluable in the future.

I would also like to thank my co-advisor, Professor Truong Nguyen. He was very gracious in offering his expertise and guidance. I very much appreciate his enthusiasm and confidence he placed in me. His suggestions greatly enhanced the research that I carried out. I very much look forward to future interactions with him.

Hamid Rabiee, of Oregon State University and Intel, provided many great ideas for my research. Hamid was always available to bounce ideas off of. The professional relationship I have with Hamid is dwarfed by our great friendship. In many ways, Hamid has always been like a big brother for me. His wife and two children are my family's dearest friends. The unending support and loyalty they have shown us is a great blessing, indeed. Hamid has always been a great role model for me. He is a very principled, humble, and honorable man whom I pray will always be there for me throughout my life.

I have had the rewarding experience of serving as a teaching assistant. I must thank Professors Jeffrey Lang and James Roberge for allowing me to TA the past four terms. The opportunity of teaching young students will be an experience I will always treasure. Thus, I must thank my students for the great times I had with them and for all that I learned from them. I wish them the best of luck in their careers at MIT and beyond. I had the pleasure to work with and learn from Professors Paul Gray, Frank Morgenthaler, Harry Lee, and James Chung.

The most pleasant aspect of MIT is the many great friends I have made. First I must acknowledge Gassan Al-Kibsi. I learned so much from him about life in general and I will always be indebted to him for our great friendship and the support he has shown me the past two years. I also thank Kashif Khan, Aiman Shabra, and Jalal Khan for the their kindness, support, and all the great times we shared. I thank all my other friends of the MIT MSA and wish them the best.

My dear friends outside of MIT will never be forgotten. I thank Naim, Jaffar, Adil, and Khalid for the many great weekend get togethers we had. I am constantly inspired and humbled by how hard they work and strive to further their education.

While I was at Purdue University for my bachelor's degree, I met many phenomenal people whose impact will be felt throughout my life. I thank my mentor, Dr. Leslie Geddes, for introducing me to the wonderful world of research. His brilliance and humility was truly inspiring. I also thank Professor Martha Chiscon who first allowed me to TA her biology class. She is a truly remarkable woman and her support was essential for the fur-

thering of my education.

The many friends I had at Purdue strongly enriched my life. I must thank Hassan Syed for the great friendship we have had over the years. His brothers, mother, and father are great family friends. We will always be indebted for their sincerity and kindness. My other great friends at Purdue include Mehdi, Salah, Murtazha, Nizzam, and Sadiquan. I thank all of them for the great times we have had. I will always cherish my friendship with Gautam Nayak of Tulane University. We had so many great late-night study sessions that I will never forget. His great sense of humor may never be matched by anybody (except myself).

All that I am today is due to the love and support of my family. My mother's love for me has been one of my greatest assets in life. She has always supported me in everything I have accomplished and I will always be grateful and truly indebted to her for the rest of my life. My father has always been my greatest role model. He truly struggled in completing his Ph.D. with a family of five children. He has always guided us and supported us with all his love. I can only hope to be as good a father, husband, and friend that he has been. Thus, it is to my parents that I dedicate this thesis.

I wish my younger brother, Ali, the best and I will always be proud of him. My sisters, Huda, Ziena, and Zainab, will always be close to my heart. I thank my brothers-in-law, Isa and Redha, for their kindness and caring.

My greatest strength is my relationship with God. All praise and glory to Him. I pray that He guides me to success in this life and the hereafter.

Table of Contents

List of Figures

Chapter 1

Introduction

1.1 Image Segmentation and Restoration

Among the most important and common tasks in image processing are image restoration and image segmentation [1]. Image segmentation is the process of separating an image into homogenous regions [1,2]. Image segmentation is of fundamental importance in such areas as video coding, computer vision, and medical imaging. The most widely-used image segmentation algorithms have focused on different strategies based upon: region growing and splitting, edge-detection techniques, Bayesian statistical models, and recently emerging multiresolution approaches [2].

The presence of noise in digital images is common and in many cases lowers the quality and usefulness of an image. Images can be corrupted by additive impulsive and white Gaussian noise due to a noisy sensor, data storage, and transmission over noisy channels [3]. In many cases, the noise cannot be completely characterized statistically, *a priori.* Image restoration is the process of recovering the visual information in a degraded image. Most of the traditional methods of image restoration are linear and assume an additive Gaussian noise model for the data [4]. While these procedures may be optimal when the noise model is exact, they are sensitive to small deviations from this model, and more robust techniques are required [3,4].

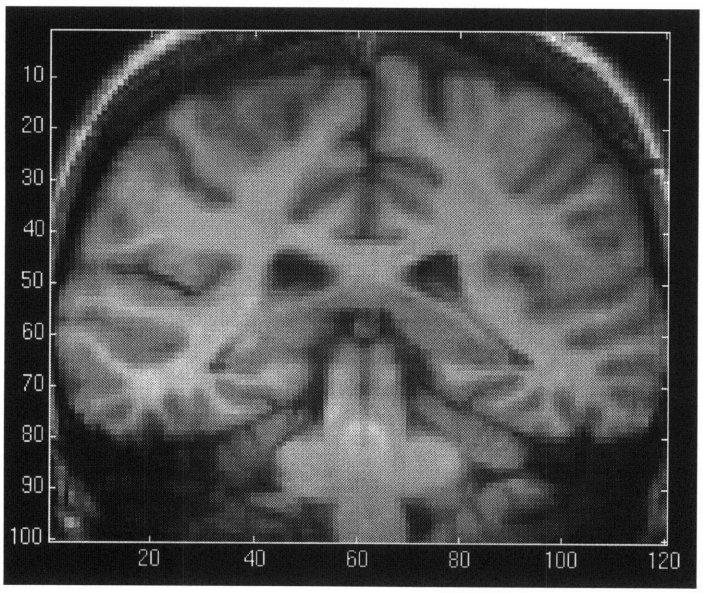
Thus, a common theme in image segmentation and restoration is that uncertainty plays a fundamental role [1]. Due to this uncertainty, we develop probabilistic models which must characterize the data of interest accurately, while at the same time be efficient in implementation. The aim of this thesis is in the development of robust methods for image restoration and segmentation through the estimation of spatially varying parameters of the

statistical model characterizing non-stationary images. The statistical model we develop will be applied to the challenging problem of automatic segmentation of Magnetic Resonance (MR) images of the human brain. Furthermore, a slight modification of the model will be used to develop a novel non-linear filter as a robust approach to image restoration.

1.2 Part 1: Background on MR Brain Image Segmentation

Magnetic Resonance Imaging (MRI) has been successfully used to produce two-dimensional images of the brain at high resolution (see figure 1.1). These high quality images have allowed physicians to readily diagnose various brain pathologies -- from mental disorders to the growth of cancerous tumors [5]. Furthermore, series of two dimensional MR images have been used to construct three-dimensional images of various brain structures. Brain tissues will have different intensities in MR images based upon their physical properties. Myelinated axons will appear white, clusters of cell bodies will appear gray, and cerebrospinal fluid (CSF) and air pockets will appear black [5]. A common first step in brain morphometrics is the segmentation of MRI data by human experts into gray and white matter regions. This process is quite meticulous and requires a great deal of time, and costly resources [6].

Figure 1.1: MR Image of Brain (T1-Weighted Coronal Section)



1.2.1 Background

Several researchers have proposed using automatic algorithms to segment brain MR images into the various major classes of brain tissue - white matter, grey matter, and cerebrospinal fluid (CSF). An effective algorithm must address the following challenges in yielding clinically acceptable segmentations: partial voluming which causes mixing of pure class intensities near the edges, impulsive and additive Gaussian noise which is an inherent problem in the acquisition of MR images, isolating the brain from surrounding bone and neck musculature, and the high intensity variance amongst pixels from the same tissue class. Furthermore, in many MRI systems, intrascan intensity inhomogeneities, due to radio frequency coils, may present a great deal of difficulty by producing images that are corrupted by a bias field. In images corrupted by a bias field, absolute intensity is no longer a good indicator of tissue type. Additionally, the proposed algorithm should require minimal user supervision.

A very popular approach is the use of Bayesian statistics and the development of an *a priori* model of the image field. Several researchers have proposed the use of a Markov

Random Field (MRF) model which utilizes the strong correlation between neighboring pixels [2]. One important disadvantage of the MRF-based approach is in its computational complexity in optimizing all the parameters, which require techniques such as simulated annealing [2]. For example, Ashton et al. [6] employed the use of a Gibbs Random Field (GRF) to model the brain image. Using the GRF model, the brain tissue is segmented. A contour model is then used to gain information about the general size, shape, and location of tissue structures. The authors also merged the segmentations of a series of MR images for reconstruction of 3-D models of brain structures.

Another approach which has been proposed is the use of neural networks [7]. However, the approach requires training data and is computationally inefficient. An important advantage of the neural network approach was that it allowed for a set of rules to be utilized in the segmentation scheme which incorporated the geometry of the brain in the field labeling process.

Another method which has been applied to MR brain image segmentation is the use of conventional edge detection and region growing [6]. However, many salient edges in MR images of the brain are "ramp edges." In images, a ramp edge can be defined as a gradual change in intensity over space. The aforementioned problem of partial voluming is manifested in MR images of the brain as ramp edges between homogenous regions of tissues. The detection of ramp edges of this type is a very challenging problem in signal processing [3].

Wells et. al have applied a mixture model to MR brain images, and utilized Maximum Likelihood estimates of the parameters of the mixture model in developing a segmentation scheme [5]. Wells "adaptive segmentation" algorithm uses knowledge of tissue intensity properties to correct MR images corrupted by bias fields and then segments the processed MR images. Wells models the pixel intensities of the brain tissue

as a Gaussian mixture model and utilizes a Maximum Likelihood (ML) based parameter estimation technique to correct and segment the image field. A salient short-coming of Wells' model of the image field is that in using the Gaussian mixture model, the spatial correlation of pixels is ignored.

Ambroise et al. [8] have proposed to modify the Gaussian mixture model by introducing spatial contiguity constraints. These constraints were defined by a neighborhood structure that they introduced into a likelihood function of the data. The resulting segmentations that their approach generated were more accurate than the traditional Gaussian mixture model-based approach.

In our research, we shall employ a multiresolution-based approach to image segmentation which extends the work of Ambroise et al. Thus, we include here a literature review of methods used for multiresolution image segmentation.

Within the past few years, wavelets have emerged as a powerful tool employed by the signal processing community to be used in the study of signals and images of interest. Mallat and Zhong have proposed the use of wavelets for the decomposition of a signal into different levels of resolution [9]. They have described a multiscale algorithm to characterize edges based upon wavelet transform extrema. At each scale, the wavelet transform extrema are detected based upon Lipschitz regularity. The Lipschitz regularity is simply a metric for local singularity in a signal. Thus, an abrupt change in the intensity of a signal would be classified as a local singularity. The decay of the maxima across scales is approximated using an exponential function. The exponential function provides a means of finding the Lipschitz regularity across scales. Using the Lipschitz criteria, the authors were able to define the location and strength of edges. Edges represent a partition of images into disjoint regions. A salient difference between edge detection and

segmentation is that in edge detection, the focus is on finding the boundaries between regions, not on the regions themselves.

Bongovanni et al. [10] have used a pyramid structure to detect bimodality in a population distribution. Briefly, an image is broken up into N 8x8 blocks (where N is power of 2). Then, within each block population, the bimodal pixel distribution parameters are estimated. Sets of four neighboring blocks are "merged." The information regarding the bimodal distributions of each of these "children" blocks are used to arrive at a bimodal distribution for the "parent" block, given that there is a bimodal distribution. The process is continued until we arrive at a bimodal segmentation of the entire image.

In a study by Park et al. [11], the authors employ the use of local variance within a window to indicate the presence of edges. The window size is varied, and local variance calculations are computed. Based upon the variance values, an optimal window size is found for each pixel. Depending on the window size and the local statistics, an edge detector of a various scale is applied locally. Thus, they propose that the local variance is a strong indicator of the resolution corresponding to an individual pixel. The resolution of a pixel is defined by its local neighborhood. For example, near edges, the local variance will be high, thus a smaller window size and high resolution edge detector is used.

Bello [2], uses both MRFs and wavelets to perform segmentation. Using the wavelet packet approach of Coifman, certain subbands are picked based upon their mean square intensity per pixel. These subbands are segmented using a MRF model. The correlation across channels is used in the refinement of MRF parameters using simulated annealing. Data is fused at specific resolution levels in order to obtain a texture label segmentation at that level. A distance measure is developed to judge the correlation across channels.

Based upon this distance measurement, a decision whether to merge data from different subbands is made.

Vincken et al. [12], developed a multiscale image segmentation technique which employs probabilistic linking across scales. An image is convolved with a Gaussian kernel to get coarser approximations of the image (lower resolutions). A model-directed linking scheme is used to establish a "child-parent" linkage of segments from adjacent scales.

Fosgate et al. [13] developed a multiscale segmentation algorithm for use in the classification of synthetic aperture radar (SAR) imagery. Essentially, the authors utilized a window-based multiscale stochastic process based upon an autoregressive model of SAR imagery at different resolutions.

1.3 Part 2: Background on Nonlinear Filters for Image Restoration

The restoration and enhancement of degraded images are of fundamental importance in image processing applications. Images can be corrupted by various noise processes such as additive Gaussian and impulsive noise due to a noisy sensor, data storage, very low-bit rate lossy compression, and transmission over noisy channels. An example of impulsive and Gaussian noise is shown in figure 1.2. Most of the traditional methods of image restoration and enhancement are linear and assume an additive Gaussian noise model for the data [3,4]. Linear filters are optimal under Gaussian models of the noise distribution, but are generally sensitive to small deviations from these models. The most widely used linear filter is the mean filter. It achieves noise reduction by averaging over a neighborhood of pixels. However, if the noise distribution is long-tailed or impulsive, the result is not satisfactory. Another disadvantage of the mean filter is that it tends to blur the edges, and often eliminates fine detail of the image. Thus, nonlinear filters have been developed which are

able to suppress the noise while preserving the integrity of edge and detail information [3,4].

Figure 1.2: Examples of Impulsive and Gaussian Noise

(A) - Original Image

(B) - Image corrupted with Impulsive and Gaussian Noise

1.3.1 Background

An important family of nonlinear filters is based on the theory of order statistics [4]. The median filter, as the most popular order statistics filter, has been widely used for its simplicity and robustness, despite its relative poor performance in preserving the details, enhancing the edges, and suppressing Gaussian noise [3,4]. The median filter was introduced by Tukey [14] as a method to smooth discrete-time signals. A median filter replaces the center pixel of a sliding processing window by the median of the pixel window population. We include a brief review of other well-known nonlinear filters.

Robust statistics have also been utilized in the development of nonlinear filters. In conjunction with maximum likelihood estimation, the relative ranks of the data within the window are utilized to arrive at an output. The L-estimators are defined as fixed linear combinations of order statistics. Some examples of these types of filters are the α-trimmed mean (α-TM) and the K-Nearest Neighbor (KNN) filters [15].

There has been recent research into the generalization of median filters by using a combination of different order-statistics. Bovik et al. [16], have used a weighted linear combination of order statistics of the input sequence which utilizes the properties of the mean and median filters. The outputs are chosen to minimize the resultant mean-square error. Furthermore, recent work into adaptive filters have yielded robust filters which are able to suppress signal-dependant noise as well as random noise [17].

Rabiee et al. [3] have developed a robust nonlinear filter which combines generalized maximum likelihood reasoning and order statistics (GMLOS). The GMLOS filter is able to suppress impulsive noise while preserving detail and edges. Furthermore, the GMLOS filter demonstrates edge-sharpening properties. However, the performance degrades in the presence of additive Gaussian noise with high variance.

Homomorphic filters are one of the oldest classes of what are termed digital nonlinear filters [18]. Nonlinear mean filters can be considered to be special cases of homomorphic systems. The center pixel of a moving processing window is replaced by a nonlinear weighting of the neighboring pixel intensities. Such a filter scheme is more effective than the mean filter in removing impulsive noise while still suppressing Gaussian noise, but the filter is not robust to simultaneous positive and negative spike noise (salt-and-pepper).

Morphological filters belong to the class of nonlinear filters which have their origins in shape analysis and set theory in mathematics [18]. The erosion and dilation operators have proven to be robust to additive Gaussian noise and in preserving edges and details. However, a short-coming of morphological filters is in failure to suppress impulsive noise when it occurs at high incidence in digital images [3].

As a departure from order statistics filters, Park et al. [19] utilized a MRF to model the pixel population within a sliding processing window. While their filter was able to suppress impulsive noise, the suppression of additive Gaussian noise was not demonstrated.

Furthermore, the MRF model substantially increases the computational complexity of the filter.

Thus, there is still active research in the development of nonlinear filters which are able to suppress noise while preserving the salient features of an image such as fine details and edges. The challenge is in developing an algorithm which is robust to both the impulsive and additive Gaussian noise models, while at the same time is computationally efficient to implement.

1.4 Thesis Contributions

Our contributions involve extending and improving Gaussian mixture model-based approaches to segmentation and restoration. For segmentation, we extend the GMM paradigm by incorporating a multiscale correlation model of pixel dependance into the standard approach. As will be shown, this multiscale correlation model provides superior segmentations, yet is computationally efficient.

We also extend the GMM methodology applied to image restoration by exploiting the identified GMM parameters for outlier detection and edge preservation. We will show that such knowledge allows us to achieve improved restorations. Furthermore, we apply our novel algorithm to suppress quantization error in image coding.

In the following paragraphs, we summarize the major contributions of our thesis.

1.4.1 Image Segmentation Using Mixture Models and Multiresolution Analysis

A popular approach to image segmentation involves the use of the Gaussian mixture model and maximum likelihood (ML) parameter estimation via the Expectation Maximization (EM) algorithm. Most such approaches assume the pixels are statistically independent of one another. The independence assumption allows the model to utilize the well-characterized Gaussian density function and greatly simplifies the computational complexity of parameter estimation algorithms. In reality, natural images are nonstationary

and the pixels are statistically correlated with one another. Such correlation motivates the use of correlation models such as MRF-based models. While these MRF-based models incorporate correlation into the segmentation schemes, they are computationally expensive.

In this thesis, we model the statistical dependance of neighboring pixels to achieve more accurate segmentation than can be obtained under the standard GMM i.i.d assumption, yet do so at a similar computational cost. In particular, we extend the work of Ambroise et al. by developing a multiresolution neighborhood structure. We assume that pixels are statistically dependant upon their local neighborhoods as well as pixels of coarser resolutions of the image. We have developed an image field model that captures this statistical dependance across both scale and space. By assuming this statistical dependance, our segmentation algorithm is robust to noise and the segmentation maps maintain the contiguous nature of regions, producing high quality segmentations of images in general, and MR brain images in particular. Moreover, our novel multiscale algorithm is computationally efficient.

1.4.2 Novel Nonlinear Filter Via ML Estimation of Spatially Varying Parameters

The restoration of digital images that are corrupted with additive Gaussian and/or impulsive noise is a difficult task. A salient requirement for an effective algorithm is that details and edges in the image must be preserved. Fine local details and edges can be modeled as significant deviations of local statistics which convey important visual information. Such deviations in local statistics implies that images are nonstationary. While most previous approaches have utilized a sliding processing window due to the nonstationarity of natural images, the methods utilized were sensitive to deviations in the noise model. In this thesis, we demonstrate that a window-based Gaussian mixture model can be applied in the development of a robust nonlinear filter for image restoration. Via the Expectation

Maximization (EM) algorithm, we utilize ML estimation of the spatially-varying model parameters to achieve the desired noise suppression and detail preservation. We demonstrate that this approach is a powerful tool which gives us information into the local statistics of noisy images. The advantages of our algorithm is that it can simultaneously suppress additive Gaussian and impulsive noise, while preserving fine details and edges. In contrast to our algorithm, most existing techniques use simple methods by just replacing the center pixel of window with a weighted mean. Such techniques do not exploit the information contained in the local statistics in restoring a center pixel of the processing window. We demonstrate how the local statistics can be utilized in the challenging tasks of edge detection as well as suppression of quantization noise in transform-based image coding.

1.5 Thesis Organization

Chapter 1 includes a literature review on relevant research of various techniques of image segmentation in general and MR image segmentation specifically. Furthermore, a review of published studies on nonlinear filters for image restoration is included.

Chapter 2 includes a review of Maximum Likelihood (ML) parameter estimation of Gaussian mixture models via the EM algorithm. Chapter 3 presents our research in the development of a robust segmentation algorithm which combines multiresolution analysis and Gaussian mixture models of images. The algorithm is applied to synthetic images as well as MR images of the brain. Chapter 4 presents an application of Gaussian mixture models to the development of a novel nonlinear filter for image restoration. Chapter 5 includes the application of our new nonlinear filter to the suppression of quantization artifacts in very low-bit rate image compression. Finally, Chapter 6 concludes with a discussion of the salient results of our research as well as suggests some possible future research directions.

Chapter 2

Review of Maximum Likelihood Parameter Estimation of Mixture Models

Before we begin with a detailed description of our novel image segmentation algorithm, we include here a review of preliminary concepts and background material related to applications of Gaussian mixture models to image segmentation. After introduction of this review material, we will in subsequent chapters show our new developments based upon the fundamental work presented in this chapter.

Our review proceeds in the following way. First, for the task of image segmentation, we discuss the model of the image field. Second, we explicitly define the generalized Gaussian mixture model (GMM) and its applicability to modelling image fields. Then we include the framework by which the GMM model parameters are estimated. In particular, we motivate the ML parameter estimation approach. We include a description of how ML parameter estimation is achieved using the EM algorithm. Finally, we have performed simulations to study the performance of previous GMM-based image segmentation algorithms. We discuss their drawbacks and motivate our developments in subsequent chapters.

Thus, the material presented in this chapter presents no new results or algorithms. The chapter's intended purpose is to assemble together much of the theory and background that we will refer to throughout the thesis. Also, this will allow us to compare the results presented here with our new results that are presented in subsequent chapters.

2.1 Model-based Image Segmentation

The overall goal of image segmentation is to identify regions that display similar characteristics in some sense. We accomplish this by assigning each pixel to be a member of one of K classes or homogeneous regions. Robust segmentation algorithms often utilize a model of the image field [20]. A statistical model of the image field is in the form of a probability density function (*pdf*) of the pixel intensities. Often the parameters of the *pdf* are not known *a priori*, thus we can utilize parameter estimation theory to achieve an efficient and consistent estimate of the model parameters [21]. To be an efficient estimator of a nonrandom parameter, the estimator must be unbiased, and its error variance must be minimized. The estimator is consistent when its mean-square error converges to zero as the data size increases to infinity [21].

2.2 Generalized Independent Gaussian Mixture Model

In this section, we formally describe the Gaussian Mixture Model (GMM). Our objective is to provide a framework for modelling the image field. In subsequent sections, we show how images can be segmented using ML parameter estimation of the GMM.

The GMM, we assumes the image field, Y(i,j), consists of intensities from K different classes. As an example, in MR brain images these classes represent different tissues (i.e. White Matter, Gray Matter, and CSF). In the GMM, pixel intensities are assumed to be independent and identically distributed (i.i.d). The i.i.d assumption allows for simple computation with the well-known Gaussian density functions. The intensity of each class is characterized by one of K different Gaussian density functions. Therefore, the model for the data is given by:

$$p(y_i) = \sum_{j=1}^{K} p(y_i = Y|k_i = j)p(k_i = j) \tag{2.1}$$

where $p(y_i = Y|k_i = j)$ is the conditional probability of each pixel and is defined by the Gaussian:

$$p(y_i = Y|k_i = j) = N\left(\mu_j, \sigma_j^2\right) \tag{2.2}$$

and $p(k_i = j)$ is the prior probability that the class of pixel i is class j. We point out that the notation emphasizes individual pixel statistics rather than the entire image. Thus (2.1) defines a Gaussian mixture.

To characterize the GMM, we define the parameter vector $\Phi = [\mu_1\ \sigma_1^2\ \mu_2\ \sigma_2^2\ ..\ \mu_k\ \sigma_k^2]^T$. Given the data and with knowledge of Φ, we can easily compute the Maximum a posteriori (MAP) estimate of the class, $\hat{k}_i$ at pixel i. We define the MAP estimate, $\hat{k}_i$, of the class of pixel i as:

$$\hat{k}_i = \underset{j}{\operatorname{argmax}}\, p(k_i = j|y_i = Y) \tag{2.3}$$

We can proceed to segment an image by assigning class memberships to each pixel individually using the above MAP estimate of the pixel class.

In practice, the conditional density parameters, Φ_i (e.g. μ_i and σ_i^2), and prior probabilities, $p(k_i = j)$, are not known *a priori*. For these reasons, the Maximum Likelihood (ML) estimation technique is used to find the estimated value of Φ_i based upon data in the image field. Since the class correspondence of each pixel in the image field is not known *a priori*, ML estimation for the conditional density parameters is a challenging nonlinear optimization problem. An attractive iterative technique to solve this problem is the Expectation Maximization (EM) algorithm [20].

2.3 Maximum Likelihood Parameter Estimation of Gaussian Mixture

Models

We shall describe the applications of the EM algorithm for ML-based parameter estimation. Let Y be a set of data of interest. Let $p(Y|\Phi)$ be defined as the probability density function (pdf) which describes the intensity distribution of Y. However, the challenge is to estimate Φ without knowledge of the class of each pixel. ML estimation is defined as finding $\hat{\Phi}_{ML}$ such that:

$$\hat{\Phi}_{ML} = \arg\max_{\Phi} \; p\langle Y|\Phi\rangle \tag{2.4}$$

We have defined y_i to be the intensity of pixel i. Furthermore, we let k_i indicate the class of pixel i. In general k_i is not observable and is referred to as the missing data. In keeping with the terminology used for k_i, y_i by itself is referred to as the incomplete data. A very popular approach to solving this incomplete data problem is by use of the EM algorithm, which alternately estimates the missing data and then takes this estimate as a given to estimate the parameters [21].

2.3.1 Expectation Maximization Algorithm

The purpose of this section is to include a description of the EM algorithm. We utilize the EM algorithm throughout most of this thesis find ML estimates of our model parameters.

The EM algorithm is an iterative technique, which converges to ML estimates when there is only one local maxima of the log-likelihood function [20]. For a rigorous proof of the convergence properties of the EM algorithm, please see the references of [20, 22].

Functionally, the EM algorithm is described by two steps at each iteration: the expectation step (E-step) is described by

$$R\langle\Phi|\hat{\Phi}^{\langle p\rangle}\rangle = E[\log p\langle Y|\Phi\rangle|y,\hat{\Phi}^{(p)}] \tag{2.5}$$

which represents the expectation of the log-likelihood function given the parameter estimates at the p^{th} iteration. Furthermore, the M-step is described by the following:

$$\hat{\Phi}^{\langle p+1 \rangle} = \arg\max_{\Phi} R\langle \Phi | \hat{\Phi}^{\langle p \rangle} \rangle \tag{2.6}$$

which estimates the parameters that maximize the expectation of the log-likelihood function. Thus, we see that at each iteration, the M-step produces a "new" estimate of the parameters.

The E and M-step for the GMM-based EM algorithm have been derived in many past studies on the generalized Gaussian mixture model [20]. We shall include the results of this derivation. The E Step can be deduced to the following:

$$z_{ik}^{(p)} = \frac{p\langle k_i = k | \Phi^{(p)} \rangle p\langle y_i | k_i = k, \Phi^{(p)} \rangle}{\sum_{j=1}^{K} p\langle k_i = j | \Phi^{(p)} \rangle p\langle y_i | k_i = j, \Phi^{(p)} \rangle} \tag{2.7}$$

In simplified form, this expression is based on Bayes' Law. Simply, z_{ik} is the probability that pixel i belongs to class k, given its intensity, y_i, and current model parameters at the p^{th} iteration of the process, $\Phi^{(p)}$.

Furthermore, $p\langle y_i | k_i = k, \Phi^{(p)} \rangle$ is the conditional probability of pixel i given it is a member of class k. The summation in the denominator is a normalization term.

$$n_k^{(p)} = \sum_{i=1}^{N} z_{ik}^{(p)} \tag{2.8}$$

$$p\langle k_i = k | \Phi^{(p+1)} \rangle = \frac{n_k^{(p)}}{N} \tag{2.9}$$

Equations (2.8) and (2.9) are used to estimate the class prior probabilities at step p. Here, N is the total number of pixels in the image field. After the E-step, the following two equations will form the M-step

$$m_k^{(p+1)} = \frac{1}{n_k^{(p)}} \sum_{i=1}^{n} y_i z_{ik}^{(p)} \tag{2.10}$$

$$\sigma_k^{2(p+1)} = \frac{1}{n_k^{(p)}} \cdot \sum_{i=1}^{n} z_{ik}^{(p)} \cdot \left(y_i - m_k^{(p+1)} \right)^2 \tag{2.11}$$

The M-step produces the new parameter estimates to be used in the next iteration of the process. For the Gaussian density function, the mean and variance are sufficient to fully characterize the pdf. The iterations are continued until the parameters converge to local maxima of the log-likelihood function. We see that the estimator is only a function of the data, and thus is valid. Furthermore, since these are ML estimates, they are efficient estimates given the efficient estimator exists. From estimation theory it is known that when an efficient estimator exists, it is the ML estimator [21].

2.4 Overview of GMM-based Algorithm for Image Segmentation

In image processing, a practical application of the EM algorithm is in estimating the parameters of the *pdf*'s which characterize the distributions of the pixel intensities. The parameter vector, $\Phi = [m_1\ \sigma_1^2\ ..\ m_K\ \sigma_K^2]^T$, was defined in section 2.2. Further let Y be a (N x 1) vector of all the pixel intensities in the given image. Then assuming the pixels are independent, we define $p(Y|\Phi)$ as

$$p\langle Y|\Phi\rangle = \prod_{i=1}^{N} \sum_{k=1}^{K} p\langle y_i \in k\rangle p\langle y_i|\Phi, y_i \in k\rangle \tag{2.12}$$

Then we estimate $\hat{\Phi}_{ML}$ using equation (2.4). We resort to the EM algorithm for its solution. The implementation of the GMM-based image segmentation algorithm is partitioned into the following set of rules:

1) Find an initial estimate of the parameters, $\Phi^{(0)}$. For example, in the case of MRI segmentation, this requires a general guess for the means and variances for the three different classes (white matter, gray matter, and CSF). White matter will have the highest mean and CSF will have the lowest mean. Because of the variation between different MRI acquisition systems, the histogram of the image should be used to generate the initial parameters. The prior class probabilities, $p(k_i = k)$, are assumed to be equally likely.

2) The EM algorithm is used to find the ML estimates of parameters which characterize the *pdf* of the image field.

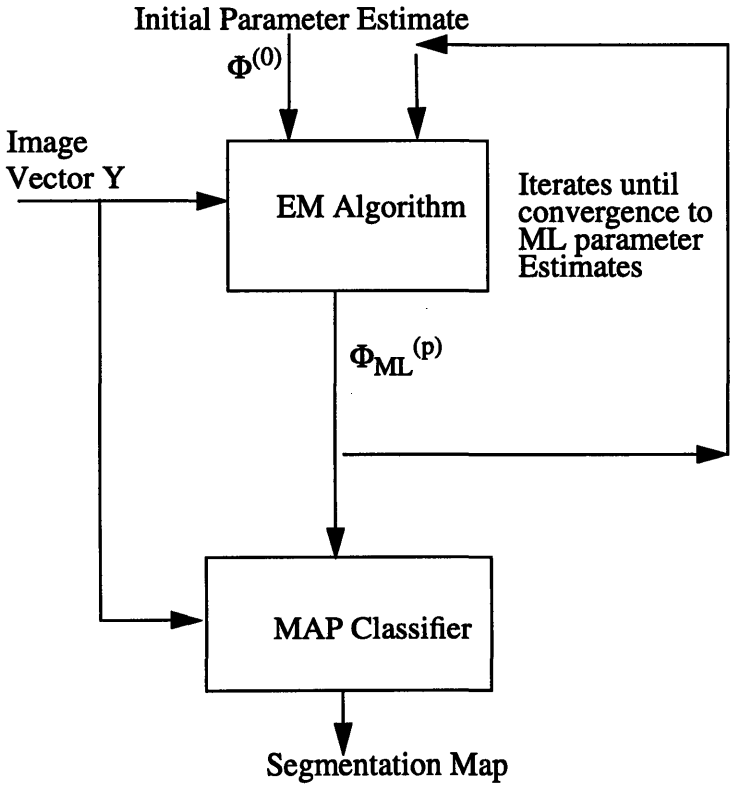
3) Using the parameters generated by the EM, a *maximum a posteriori* (MAP) estimator is used to decide which class each pixel belongs to based on an *a prior* model of the image field.

The Gaussian mixture model has been used in previous studies in automated MRI brain segmentation as well as image segmentation in general [5]. The GMM assumes the pixel distribution is independent and identically distributed (i.i.d). The advantage this assumption has is that it reduces the computational complexity of the segmentation task by allowing the use of the simple, and well-characterized Gaussian density functions. However, a short coming of this model is that it fails to utilize the strong spatial correlation between neighboring pixels which is a prominent characteristic of natural images in general and MR brain images in particular [2]. Let us define here the term, "spatial correlation," as we refer to it in this thesis. We assume that neighboring pixels are spatially correlated because they have a high probability of belonging to the same class. For example, in natural images, one can make a reasonable assumption that neighboring pixels, located

in the interior of a homogenous region, belong to the same class, k_a. If a pixel, i, of this neighborhood has an intensity closer to the mean of another class, k_b, a classifier such as the Minimum Distance Classifier (MDC) would incorrectly classify pixel i to belong class k_b while the surrounding pixels were classified to be members of k_a. Clearly, the prior knowledge regarding pixel correlation is important in being able to correctly classify pixels.

The Gaussian mixture model was described in section (2.2). After obtaining ML estimates of the GMM parameter via the EM algorithm, we can segment the image field using the MAP classifier which was described by equation (2.3). Figure (2.1) illustrates the overall segmentation process.

Figure 2.1: Schematic of Segmentation via the EM algorithm



2.5 Experimental Results and Conclusions

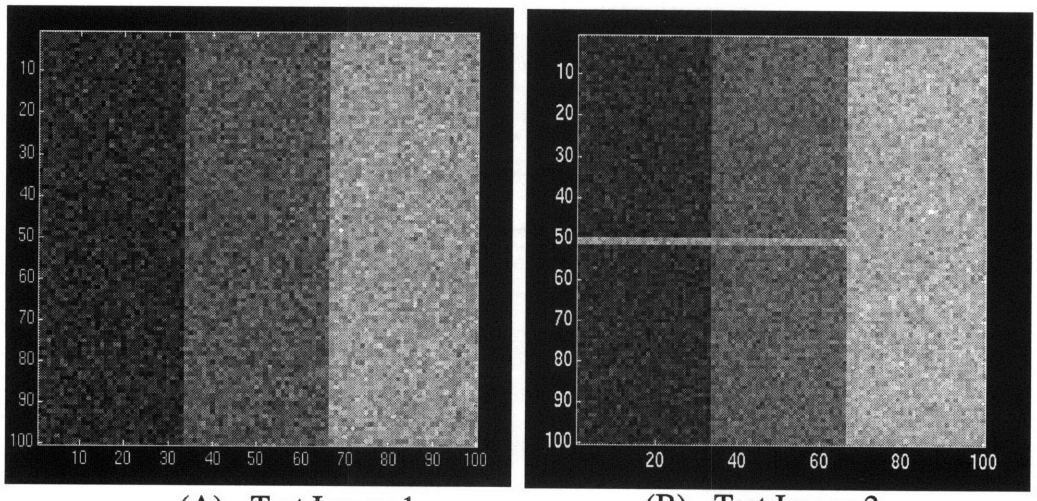
In this section, we present examples of the GMM-based segmentation algorithm. Our objective is to evaluate the performance that can be achieved using the standard GMM. We show the drawbacks of standard GMM-based segmentation algorithms using both synthetic and real images. This will allow us to motivate our new developments in the forthcoming chapter and compare with the results of this chapter.

2.5.1 Segmentation of Synthetic Test Image

We use synthetic images where we know what the "correct" segmentation should be. In the first set of experiments we applied the EM algorithm for MAP segmentation of synthetic images shown in figure (2.2). The synthetic images were formed by three vertical blocks. Each block consists of pixels from one of three Gaussian processes of a given mean and variance.

Pixels within the same block come from the same process. Pixels are also i.i.d. Classes 1, 2, and 3 have means of 50, 100, and 150, respectively. All classes have a variance of 225. Note, an equivalent description of the synthetic image is to state that the image consists of three blocks of constant intensities of 50, 100, and 150, and the image is then corrupted with an additive white Gaussian noise with a variance of 225. As figure (2.3) illustrates, the histogram distributions of the processes overlap due to the high variance. The image of figure (2.2b), is similar to (2.2a), except for the two-pixel wide horizontal line across the image field. The purpose of this line is that we would like to study the performance of segmentation algorithms in segmenting the fine features of an image.

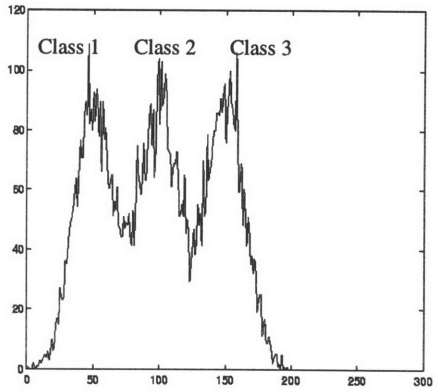
Figure 2.2: Synthetic Test Images



(A) - Test Image 1 (B) - Test Image 2

(A) Classes 1,2, and 3 and have means of 50, 100, and 150, respectively. All classes have a variance of 225. (B) Same as (A) with a 2-pixel wide horizontal strip from class 3 running through the other classes. The correct segmentation for these two images is obvious.

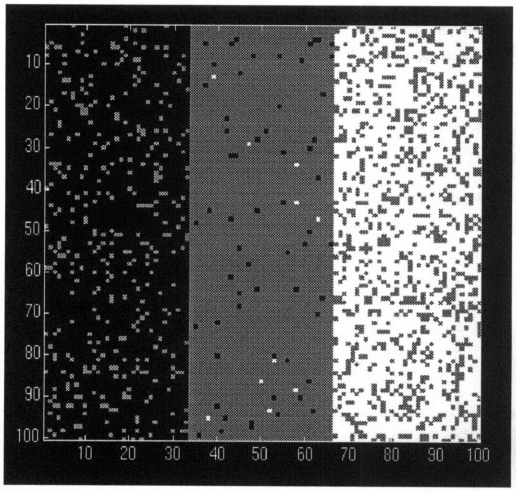
Figure 2.3: Histogram of synthetic image in Figure (2.2A).



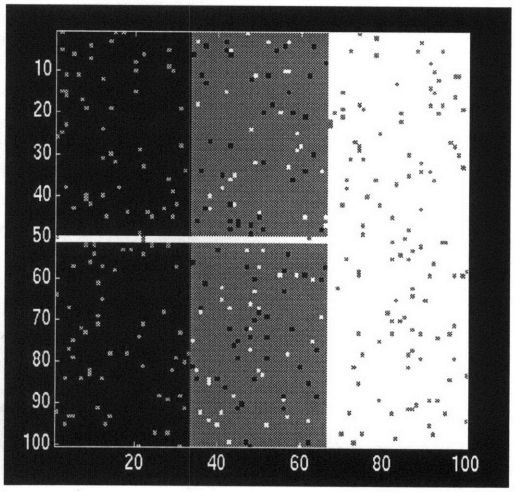
The segmentations of the test images are shown in figure (2.4). Clearly, the segmentation failed to reflect the spatial correlation existing between pixels of the same region. The

spottiness of these segmentation maps indicate classification errors. Because of i.i.d assumption, those pixels which have intensities at the tails of the Gaussian pdf's will be erroneously classified.

Figure 2.4: Segmentation of the test Images.



(A) - MAP Segmentation via EM Algorithm.



(B) - MAP Segmentation via EM Algorithm

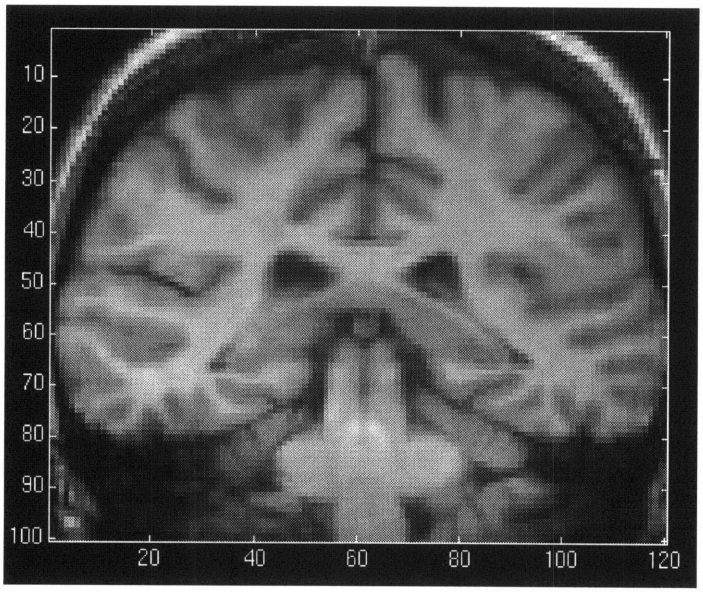
2.5.2 Segmentation of MR Image of Brain

An MR image was made available from the Center for Morphometric Analysis (CMA) at the Massachusetts General Hospital in Charlestown, MA. Figure (2.5) shows a T1-weighted coronal brain scan. Also, the image histogram is calculated.

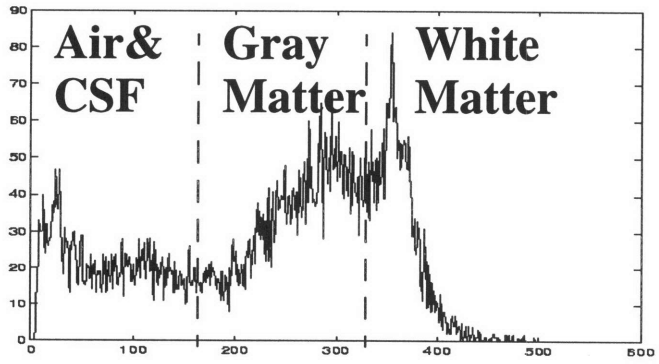
Often, Gaussian noise corrupts MR images of the brain in the acquisition process. We accentuate this corruption process by adding white Gaussian noise to an MR image of the brain. Thus, figure (2.6) is the noisy version of the MR brain image. Specifically, additive white Gaussian noise (AWGN) with zero mean and variance of 225 was added to the image. The segmentation map of the noisy image was generated using the EM algorithm. The result is shown in figure (2.7). A "good" segmentation should have contiguous regions [23]. Again since the GMM-based segmentation has failed to exploit the spatial

correlations, the resulting segmentation is poor. This failure is due to the GMM-based assumption that pixels are i.i.d.

Figure 2.5: MR Image of the brain with corresponding image histogram



(A) - MR Image of Brain (T1-Weighted Coronal Section)



(B) - Image histogram. White matter corresponds to higher intensities, whereas gray matter is intermediate, and CSF and air have low intensity.

Figure 2.6: Brain MRI with Additive Gaussian Noise

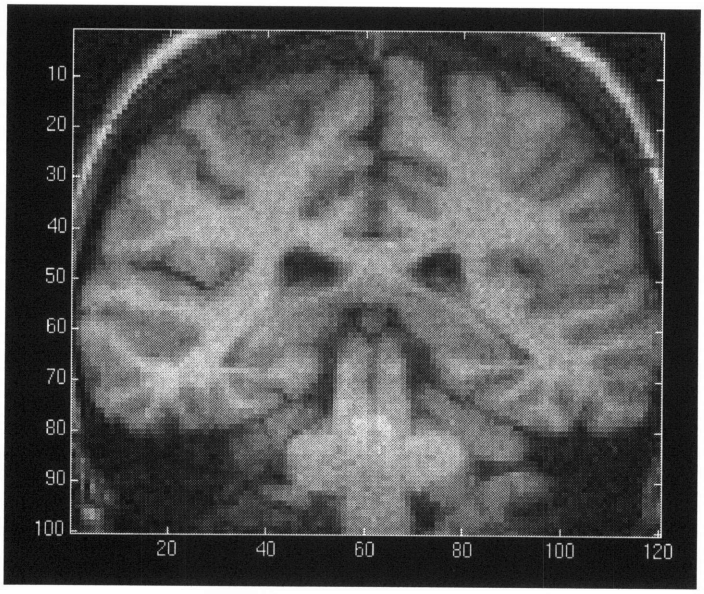
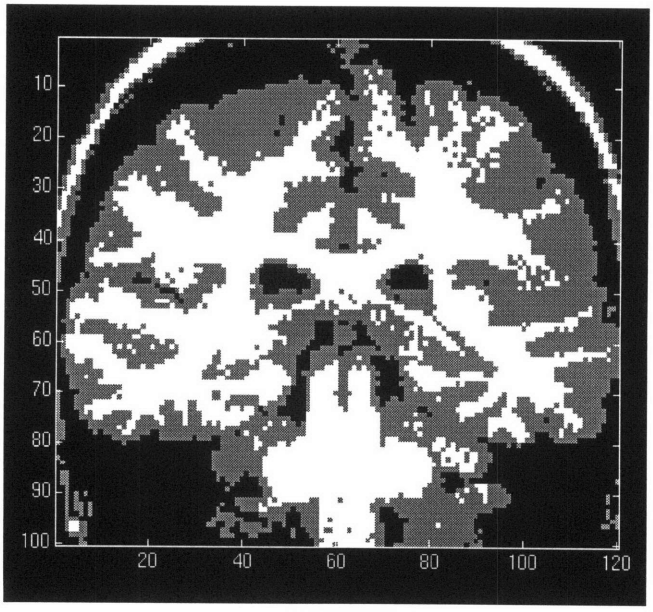


Figure 2.7: Segmentation map of noisy image with GMM-based segmentation algorithm.



2.5.3 Conclusions

In this chapter, we have reviewed the GMM. Using ML estimation of the GMM parameters, we demonstrated how images can be segmented. ML estimation was achieved using the EM algorithm. We performed simulations to study the performance of GMM-base d segmentation algorithms. From the results of these simulations, we have illustrated the drawbacks of the GMM-based segmentation. While this segmentation algorithm is computationally efficient to implement, it failed to reflect the spatial correlation of images because the GMM assumption is that the data is i.i.d. This failure was evident in the classification errors on a pixel-by-pixel basis.

Recently, Ambroise et al, have proposed to modify the likelihood function of the GMM. Specifically, they incorporated a neighborhood penalization term which had the desired effect of biasing their modified algorithm to classify pixels to the same class as their neighbors. In the forthcoming chapter, we extend their work by utilizing multiresolution analysis of the image field. We propose a multiresolution neighborhood which is defined in both scale and space. Moreover, our novel approach is computationally tractable.

Chapter 3

A New Statistical Model for Image Segmentation

As previously mentioned, in most natural images, there is strong correlation between neighboring pixels. For example, if the neighbors of a given pixel are classified to belong to a certain region class the probability that the pixel itself also belongs to the same class is more likely. Since the GMM-based segmentation is premised on the assumption that pixels are i.i.d, the spatial correlation is ignored in the model, and hence results in poor segmentations. In images, spatial correlation can be extended across resolutions as demonstrated in recently published multiresolution-based statistical models of images [12,13]. In this chapter we introduce a new multiresolution model of the image field that efficiently incorporates the spatial correlation into the EM algorithm across different scales, and thus into the resulting segmentation.

The chapter proceeds in the following way. First, we motivate the multiresolution-based approach our algorithm utilizes. Second, we develop a neighborhood clique which will be defined across scale and space. Then we unite multiresolution analysis with our neighborhood clique and formally describe our novel algorithm. Following that, we present applications of our algorithm in segmenting synthetic test images and MR brain images. We discuss the results of our simulations and compare our methods with the GMM-based segmentation algorithm of the previous chapter.

3.1 Multiresolution analysis of images

Here, we motivate the multiresolution approach to image segmentation. Within the image processing community, multiresolution-based image segmentation has emerged as a powerful method for producing high-quality segmentations of images [13]. In fact, the

human visual system (HVS) is known to process images using information at different resolutions [24]. Thus, we see that image segmentation is a multiresolution problem. Burt and Adelson proposed the Laplacian Pyramid as a compact representation of images [25]. Each resolution of the pyramid will contain information which emphasizes different features of the image. At increasingly finer resolutions the detail of an image is more prominent. At the lower resolutions, only the largest features of an image are detectable. Our goal is to efficiently combine the information at various scales and produce accurate segmentations of the image field.

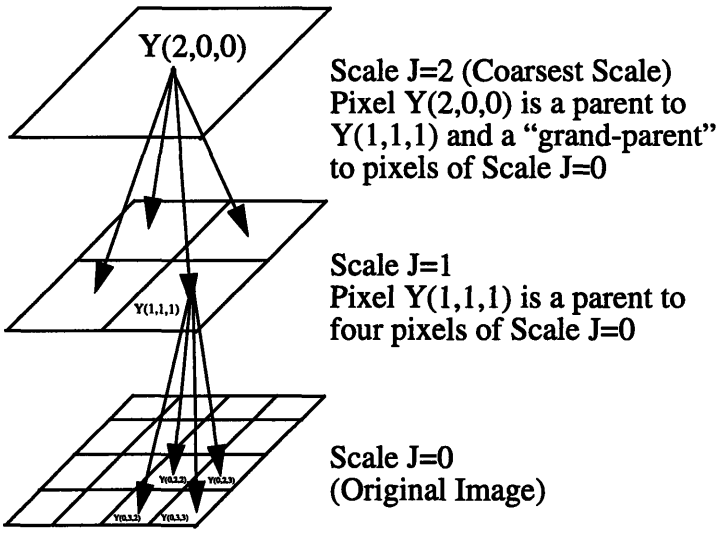
3.1.1 Wavelet-based decomposition of images

Using the Discrete Wavelet Transform (DWT), an image can be resolved into different frequency subbands using the pyramid algorithm [26]. By choosing an orthonormal basis to represent the image, one can construct a set of two-dimensional filters to isolate the subbands of interest. For many of the popular basis functions, their corresponding filters have been studied in great depth [26]. In the current study, we will mainly, except where otherwise stated, implement the DWT using the Haar basis because of its simplicity. The DWT can be implemented using other basis functions. We shall also briefly discuss the use of higher order basis functions of the Daubechies family in the implementation of our algorithm. From our experimental results, the majority of the energy in brain MR images is captured in the low frequency subbands. Thus, we will only use the low-frequency subbands isolated by the DWT. Starting at the original scale, S_0, of an NxN image, one can produce a collection of successively coarser low-pass filtered images $\{S_0, S_1,.., S_J\}$, produced by the DWT. Via the downsampling involved in the pyramid algorithm, the number of data points decreases by a factor of 4 at each coarser scale without aliasing. Thus, the image obtained from the coarsest scale, S_J, will be of size $\frac{N}{2^J} \times \frac{N}{2^J}$. By operating on the coarse resolutions, we can estimate parameters using less data points.

As figure (3.1) demonstrates, the low pass sequences of images can be mapped onto a quadtree structure. Each pixel at resolution J corresponds to four "child" pixels at resolution J-1. Thus the pixel at resolution J is defined as a "parent" pixel to its respective "child" pixels at resolution J-1. In turn, the pixels at resolution J-1 are parents to their corresponding child pixels at resolution J-2. In this study, we take advantage of this quadtree structure to develop an algorithm for image segmentation. We will make the following assumptions: the pdf of pixel i at resolution J-2 is dependant upon its neighbors, furthermore, the pdf is dependant upon the parent pixel at resolution J-1 and grandparent at resolution J, as well as the neighbors of the parent and grandparent. Thus, our model will attempt to utilize the dependance of the pdf across both scale and space towards the goal of a more robust segmentation algorithm. Hence, our aim is to modify the Gaussian mixture density pdf such that we penalize the likelihood of pixel membership to a certain class when its neighbors, parent, and parent's neighbors have a low probability of belonging to this same class.

From chapter one, it was pointed out that other researchers have implemented multiscale Markov random field models which also are directed at penalization of this "scale-space" likelihood. However, implementation of Markov random field models are computationally intensive and optimization of the model parameters requires the use of techniques such as simulated annealing and other gradient search methods [2].

Figure 3.1: Multiscale Quadtree illustration.



Pixels are indicated by scale and space. Y(X,Y,Z) indicates a pixel at scale X and location within that scale is denoted by (Y,Z). A pixel at scale J=2 is shown with its 4 children at J=1, and 16 grand-children at scale J=0

3.1.2 Multiresolution EM Algorithm

The EM algorithm has been shown to be a method of maximizing the following equivalent log-likelihood function of the image field []

$$L(Z,\Phi) = \sum_{k=1}^{K}\sum_{i=1}^{n} z_{ik}\log\left(p\left(k_i = k\right) p\langle y_i|\Phi_k, (y_i \in k)\rangle\right) - \sum_{k=1}^{K}\sum_{i=1}^{n} z_{ik}\log\left(z_{ik}\right) \tag{3.1}$$

We have kept consistent with the notation used throughout this chapter. Z is a matrix whose elements are all the z_{ik} of the image. Recall, that the EM algorithm iterates until the parameter matrix, Φ, converges to a local maxima of the log-likelihood function. In the E-step of the EM algorithm, we have defined an explicit expression for z_{ik} as shown in equation (2.7). Specifically, given the intensity y_i and Φ, z_{ik} is the probability that pixel i

belongs to class k. Thus, the outputs of the EM algorithm, Φ and Z, also maximize $L(Z,\Phi)$. The proof that the EM algorithm maximizes $L(Z,\Phi)$ can be found in [27].

As is apparent, $L(Z,\Phi)$ does not account for the spatial correlation of the data. We propose to modify the likelihood equation (3.1), by the addition of a penalization term, $V(Z)$. Our penalization term will bias the likelihood of a pixel, i, belonging to the same class, k_a, of its neighbors. Thus, we can view $V(Z)$ as modifying the pdf to incorporate desirable correlation properties. This prior probability on the pixel class probability is of a Gibbs form and thus like an MRF on the class probabilities. The new likelihood expression is given by the following equation:

$$U(Z,\Phi) = L(Z,\Phi) + V(Z) \tag{3.2}$$

The penalization term, $V(Z)$, will incorporate the quadtree data structure illustrated in figure (3.1) as well as a simple "clique" or pixel neighborhood system. Within the same resolution, we define the neighborhood of a pixel, i, to be all pixels which are adjacent to pixel i (top, down, right, left, and diagonal). Furthermore, a pixel at resolution J-2, will be defined to have a neighborhood at resolution J-1 which consists of the parent of i as well as the parent's neighborhood. This neighborhood can be extended further across resolution to include the "grand-parents" of i at resolution J. In practice, we only incorporate the information from scales J=0, 1, and 2. The neighborhoods at each resolution will have different weightings in the neighborhood interaction weights of the penalization term. Let us define the following neighborhood interaction weight (NIW):

$$v_{irj} = \begin{cases} \alpha & \text{if pixel i and r are neighbors at resolution 0} \\ \beta & \text{if pixel i and r are neighbors at resolution 0 and 1} \\ \gamma & \text{if pixel i and r are neighbors at resolution 0 and 2} \\ 0 & \text{else} \end{cases} \tag{3.3}$$

Using the NIW of (3.3), we propose the following penalization term

$$V(Z) = \sum_{j=0}^{J}\sum_{k=1}^{K}\sum_{i=1}^{N}\sum_{r=1}^{N} z_{jik} z_{jrk} v_{irj} \tag{3.4}$$

Where z_{jik} is the probability of pixel, i, from resolution j being a member of class k. $V(Z)$ weights neighborhoods which have pixels that are members of the same class more than heterogenous neighborhoods. Furthermore, we see that $V(Z)$ is only dependant on the probability matrix, Z, whose elements are the individual pixel probabilities, z_{jik}.

A modified version of the EM algorithm can be used to maximize the new penalized likelihood equation, $U(Z,\Phi)$. We shall call the modified EM algorithm the *Multiresolution EM (MEM) algorithm*. The attractiveness of the MEM algorithm is in the approach of utilizing a multiresolution neighborhood. The coarser resolutions will allow for the segmentation of the more prominent features in the image. However, the information at the finer levels is important for accurate segmentation along boundaries and for segmenting highly detailed regions. Thus this has two advantages: 1) the MEM algorithm has desirable correlation properties and avoids blurring, and 2) misclassifications are reduced. Here, we present an overview of the MEM algorithm:

3.2 Overview of MEM algorithm

1) First, we must generate a multiresolution sequence of images using the DWT. In this thesis, we utilize scales J=0, 1, and 2 in the MEM algorithm. Note, J=0 is the original image, and J=2 is the coarsest scale.

2) Then, using the standard GMM-based segmentation algorithm of the previous chapter, we segment scales J=1 and J=2. Recall, that the segmentation of an image is based on a MAP classifier. Thus, this step also produces the probability matrices Z_1 and Z_2, where Z_J is a matrix whose elements are the individual z_{jik}. We hold Z_1 and Z_2 fixed, and we uti-

lize these matrices in our multiresolution segmentation algorithm. We emphasize that thus far, the segmentation of one scale does not affect another scale. It is possible to have alternative implementations in which all scales affect one another in a scale-recursive fashion, and this approach may be investigated as an extension of this thesis.

3) Now, we are able to modify the EM algorithm to segment the finest level image. Since the additional term, $V(Z)$, does not contain any parameters, Φ, of the image model, the M-step will remain the same. However, the E-step will change since it is dependant upon z_{jik}. In particular, equation (2.7) must be modified to incorporate this modified multiresolution likelihood. In the multiresolution framework, z_{ik} will be expressed as

$$z_{0ik} = z_{ik}^{(p)} = \frac{p\langle k_i = k|\Phi^{(p)}\rangle p\langle y_i|k_i = k, \Phi^{(p)}\rangle \exp\left(\sum_{j=0}^{J}\sum_{r=1}^{n} z_{jrk}v_{irj}\right)}{\sum_{s=1}^{K} p\langle k_i = s|\Phi^{(p)}\rangle p\langle y_i|k_i = s, \Phi^{(p)}\rangle \exp\left(\sum_{j=0}^{J}\sum_{r=1}^{n} z_{jrs}v_{irs}\right)} \tag{3.5}$$

The above expression was derived using Lagrange multiplier optimization and is the equivalent of equation (2.7). The details of our derivation are quite similar to the derivation of Ambroise et al. [8]. We use the terms z_{0ik} and z_{ik} interchangeably. We can interpret equation (3.5) as a modified conditional probability of pixel i belonging to class k. The probability is conditioned on the current parameters, Φ, and the likelihoods of the neighbors of pixel i belonging to class k. The rest of the E-step is the same as the monoresolution EM algorithm. The M-step is not changed. Note, we are interested only in calculating the new z_{ik} at the original resolution, J=0, only. From step 2 of our algorithm, recall that we hold z_{1ik} and z_{2ik} fixed.

From equation (3.5), we see that z_{0ik} appears on both sides of the equation. Specifically, the summations in the exponential contain z_{0ik} if the summation index, j, equals

zero. This suggests an iterative algorithm to compute z_{0ik}. In equation (3.5), we point out that the index, p, refers to the iteration of the EM algorithm. Thus, at every iteration of the EM algorithm, we solve for z_{0ik}. Experimentally, we have found that a few iterations of equation (3.5) produces a reasonable probability matrix, Z.

4)After obtaining Z, we proceed to the M-step. We stress that our new developments have not modified the M-step of the EM algorithm. Thus, given Z, we estimate the parameters, $\Phi^{(p)}$, of our model using the equations (2.8) through (2.11).

5) With the overall updated EM model parameters, we go back to step 3. Experimentally, we have found that five to six iterations are required for the elements of Φ and Z to converge to their final values.

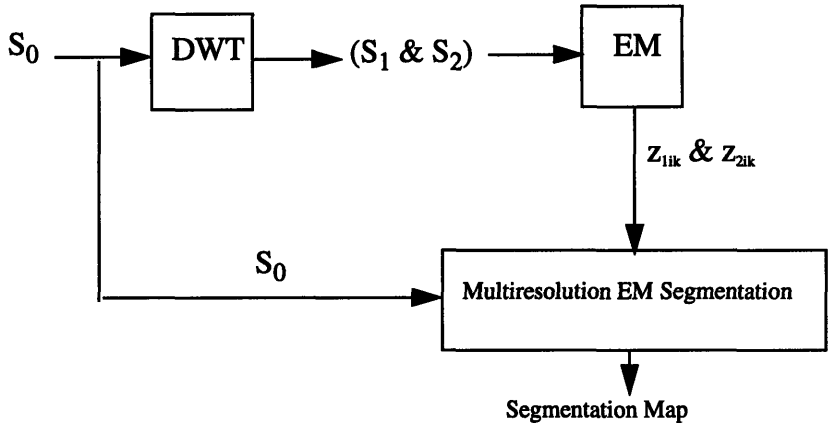
6) We finally segment the original image by assigning class memberships to each pixel, i, by choosing the class for which z_{ik} is maximum:

$$\hat{k}_i = \arg\max_k z_{ik} \tag{3.6}$$

3.2.1 Applications of the MEM algorithm towards image segmentation

Given an image Y, we compute a two-level Discrete Wavelet Transform (DWT) using any desired bases. In this thesis, we have used the Haar basis and also include an example using a seventh-order Daubechies Wavelet (DB7). The DWT will provide a collection of low-pass filtered images, $\{S_0..S_2\}$, where S_0 is the original image and S_2 is the coarsest image. We derive z_{ik1} and z_{ik2} using the conventional EM algorithm via the monoresolution Gaussian mixture model, and segment S_1 and S_2. As previously noted, we hold z_{ik1} and z_{ik2} fixed. Then after this information is provided, we can apply the MEM algorithm to S_0. Figure (3.2) provides an illustration of this process.

Figure 3.2: Illustration of Multiresolution segmentation algorithm. .



The Discrete Wavelet Transform (DWT) is used to generate a collection of images at different resolutions. The conventional EM algorithm is used for resolutions S_1 and S_2, and the output values are used to segment the original image, S_0

3.2.2 Evolution of Model Parameters Across Scales

An attractive feature of this multiresolution implementation of the EM algorithm is that the output parameters of one resolution can be used as the initial estimates of the next finer resolution. Starting at the coarsest scale S_J, (experimental results indicate that J = 2, or 3 give best results), we implement the ML-based segmentation using the EM algorithm as described above. After arriving at the ML estimate of the parameter vector Φ, the afore-mentioned MAP detector is used to segment the image at the current scale. Thus, after segmentation, the pdf's of each tissue class will be characterized by the parameter matrix. Using the estimated parameter matrix, Φ_J, of the previous scale, we can derive the initial parameter estimate, Φ_{J-1}, needed for the current image field. The evolution across scales of the parameters of the tissue class is defined by the following derivation.

Given data points $\{X_{1,2,..,N}\}$ in a homogenous region of a specific tissue class from scale S_{J-1} with mean, μ_1, and variance σ_1^2, the DWT performs a discrete-time convolution with impulse response of the low-pass filter. In the case of the Haar filter, the low-pass coefficients are defined to be: $\{h(0) = 1/\sqrt{2}, h(1) = 1/\sqrt{2}\}$. Let $Y_{1,2,..,N}$ be the output points of the convolution. We see in general that:

$$Y_k = h(0) \cdot X_{k-1} + h(1) \cdot X_k \tag{3.7}$$

$$E[Y_k] = m_1 \cdot (h(0) + h(1)) \tag{3.8}$$

$$Var[Y_k] = \sigma_1^2 \cdot (h^2(0) + h^2(1)) \tag{3.9}$$

Thus we establish that given the parameter, Φ_J, from scale J, we can derive the initial estimates of the parameter matrix at scale J-1. We employ this analysis for the initial estimates used in segmenting S_2 and S_1. The MEM algorithm's initial estimate needed for segmenting S_0 is provided by using the final estimates of the parameters of S_1.

3.2.3 Summary of MEM algorithm

The aim of the MEM algorithm is modify the pdf of the GMM such that we penalize the likelihood of a pixel belonging to a class other than the class its neighbors belong to. Thus, our procedure is:

1. Generate a multiresolution sequence of images using the DWT that is three levels deep.

2. Segment the coarser images using the well-known GMM-based segmentation algorithm.

3. Using the information provided by the coarser segmentations, we incorporate the penalization term from Equation (3.2) into the likelihood function of the image field. We point out that we do initialize each scale.

4. We modify the E-step of the EM algorithm to account for the penalization term as shown in Equation (3.5).

5. Using the probability matrix, Z, that results from the EM algorithm, we segment the finest level image.

3.3 Experimental Results and Conclusions

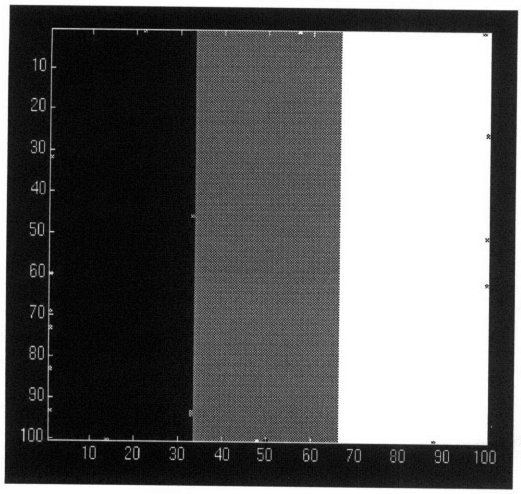
3.3.1 Segmentation of Synthetic Test Image

To demonstrate the robustness of the MEM algorithm and compare its performance against the traditional GMM-based segmentation algorithm, we applied the MEM to a series of test images. Specifically, our goal is to demonstrate that a multiresolution segmentation will result in a more accurate segmentation of the image field. We also will demonstrate that while the NEM algorithm of Ambroise et al. performs better than the GMM-based segmentation algorithm, the NEM's segmentations are not as accurate as our algorithm's segmentations. Moreover, we also illustrate that an ad-hoc algorithm based upon low-pass filtering, degrades the edges of an image field and thus results in a poor segmentation.

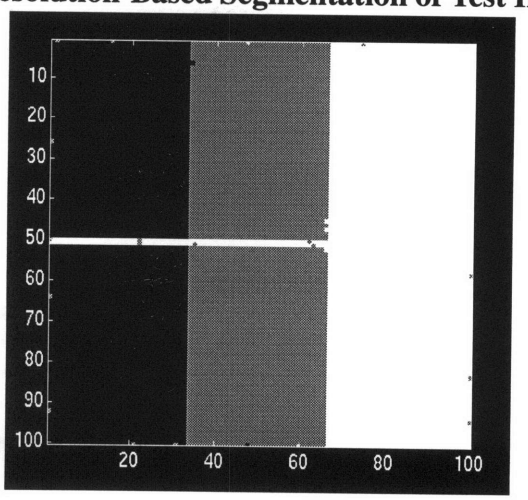
In our first set of experiments, we segment the same two test images, (2.2a) and (2.2b), of section (2.5.1) in chapter 2. The segmentations of the test images are shown in figures (3.3a-f). Clearly, the GMM-based segmentation failed to demonstrate the spatial correlation existing between pixels of the same region. The MEM algorithm provides a subjectively and quantitatively superior segmentation. We also compared our MEM algorithm to the Neighborhood EM (NEM) algorithm of Ambroise et al. While the NEM algorithm

performs better than the conventional EM algorithm, our new MEM algorithm yields a more accurate segmentation map.

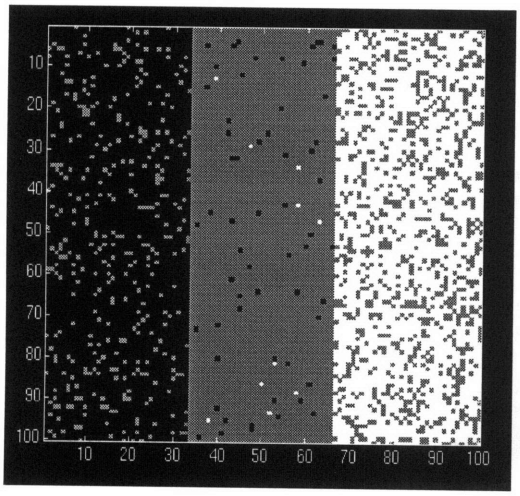
Figure 3.3: Monoresolution and Multiresolution-Based Segmentation of Test Images



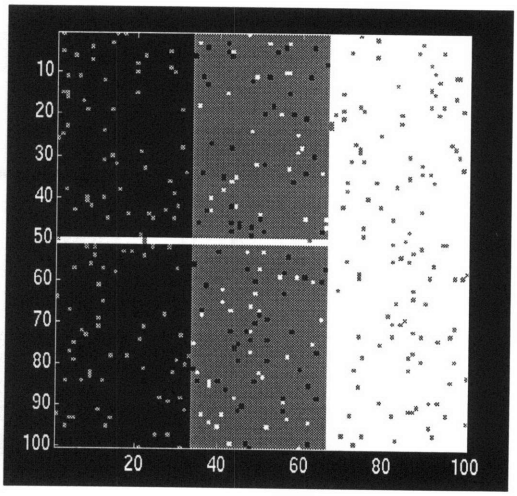
A- Segmentation of 2.2a via NEM



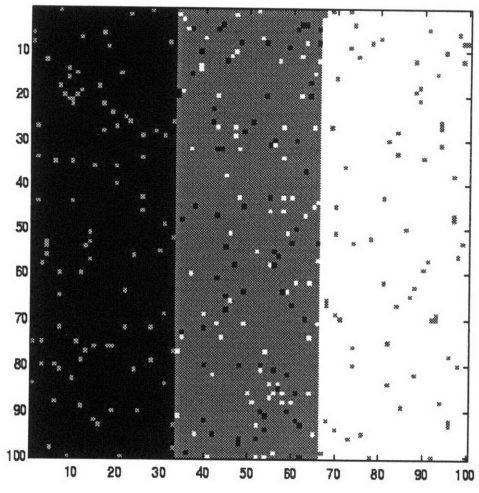
B- Segmentation of 2.2b via NEM



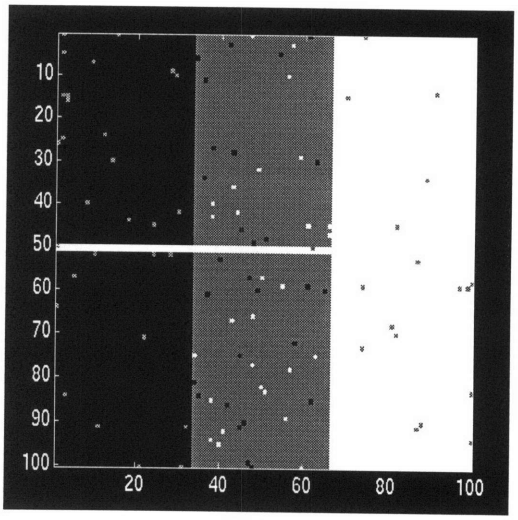
C- Segmentation of 2.2a via GMM



D- Segmentation of 2.2b via GMM



E- Segmentation of 2.2a via NEM

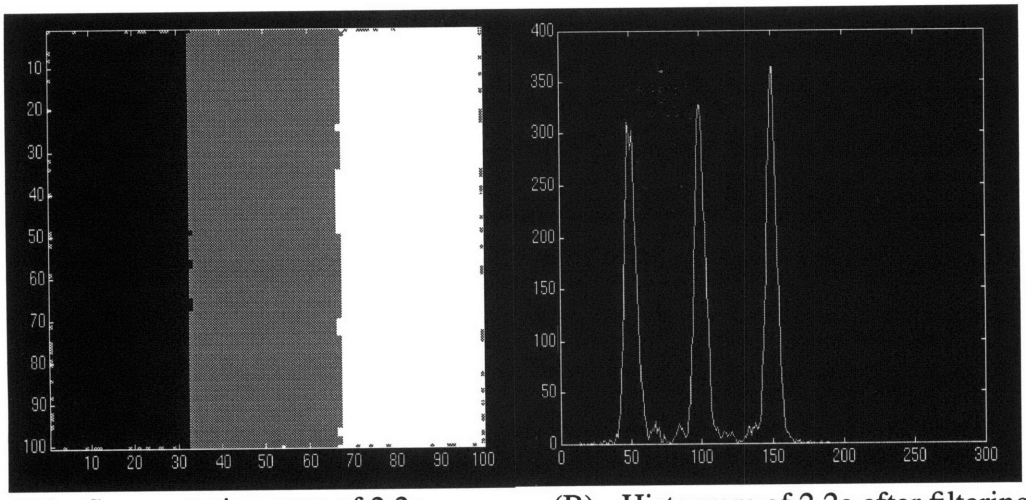


F- Segmentation of 2.2b via NEM

An ad-hoc algorithm was also used to segment the test image. In many standard approaches to image segmentation, the image is first low-pass filtered to suppress the noise. In this experiment, the test image was convolved with a low pass filter either once or twice and then segmented with the conventional EM-based segmentation algorithm. As figure (3.4) demonstrates, the low-pass filtering succeeded in reducing the variances of the three classes, and thus improved the segmentation. However, as a result of the low-pass filtering, the borders of the regions are blurred, and the resultant segmentation near the edges is poor. This effectively is creating the partial voluming which causes class confusion. In figure (3.5), we illustrate this occurrence by including an example of a one dimensional signal being processed by a set of low-pass filters. The filtering has severely degraded the edges and this results in a "mixed" set of data which will be erroneously classified. The extension of this example to two dimensions is straightforward.

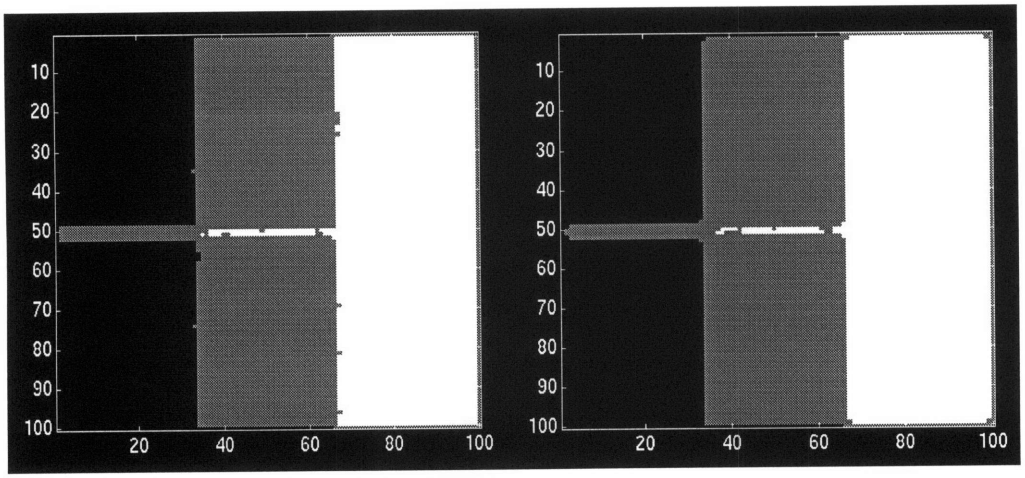
This blurring effect is most prominent in the segmentation of Test Image 2 of figure (2.2b). The horizontal line is not detected using this ad-hoc algorithm. However, because the MEM algorithm utilizes the information available in all the scales, we succeed in segmenting the horizontal line. Had we only utilized the information in the coarser resolutions, the horizontal line would not have been detected. Thus, this example demonstrates that the MEM algorithm is able to reliably segment structured images with fine details, in the presence of noise.

Figure 3.4: Ad-Hoc Segmentation of Test Image after Low-Pass Filtering



(A) - Segmentation map of 2.2a. Note the poor performance near the edges of blocks.

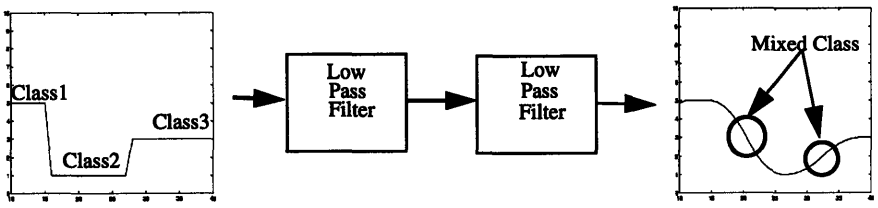
(B) - Histogram of 2.2a after filtering. Note the reduction of variance of the classes.



(C) - Segmentation map of 2.2b after low pass filtering once.

(D) - Segmentation map of 2.2b after low pass filtering twice

Figure 3.5: 1-D example of Partial Voluming.

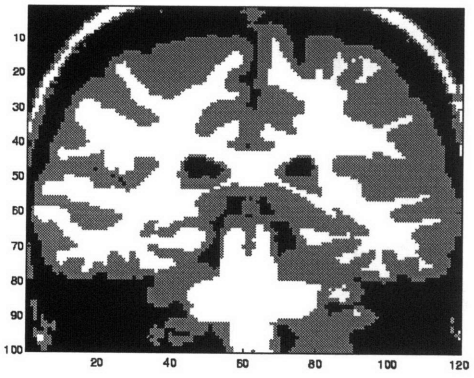


1-D example of Partial Voluming due to low-pass filtering operations. We see that after filtering, it becomes challenging to classify the "mixed" data points along the edges. The extension of this example to 2-D images is straightforward.

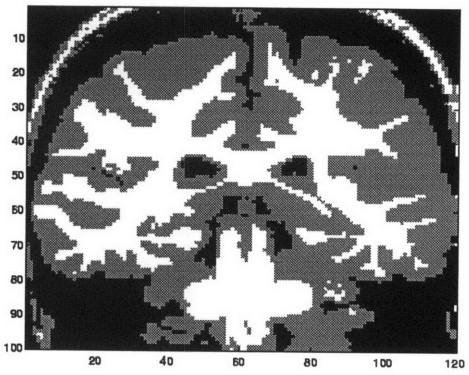
3.3.2 Segmentation of MR Image of Brain

Here, we segment the same brain image, figure (2.5a), shown in the previous chapter. Using the MEM algorithm, figure (3.6) shows a MEM-based segmentation map of the brain image. We also include the segmentations generated by the GMM-based algorithm as well as the NEM algorithm. Furthermore, the probability matrix, Z, can be visualized as a "probabilistic segmentation map" as shown in figure (3.7). Z can be used to generate probabilistic segmentations for the corresponding three tissue classes. Each subimage indicates which pixels have the highest probability of belonging to a specified class. Probabilistic segmentation maps can be generated using the other segmentation approaches that we have studied, although we do not show them here.

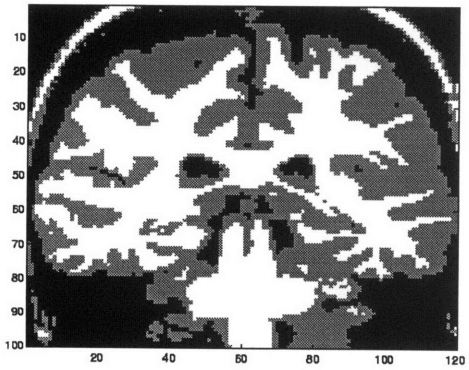
Figure 3.6: Comparison of segmentation maps of MR image of human brain



A- MEM Segmentation

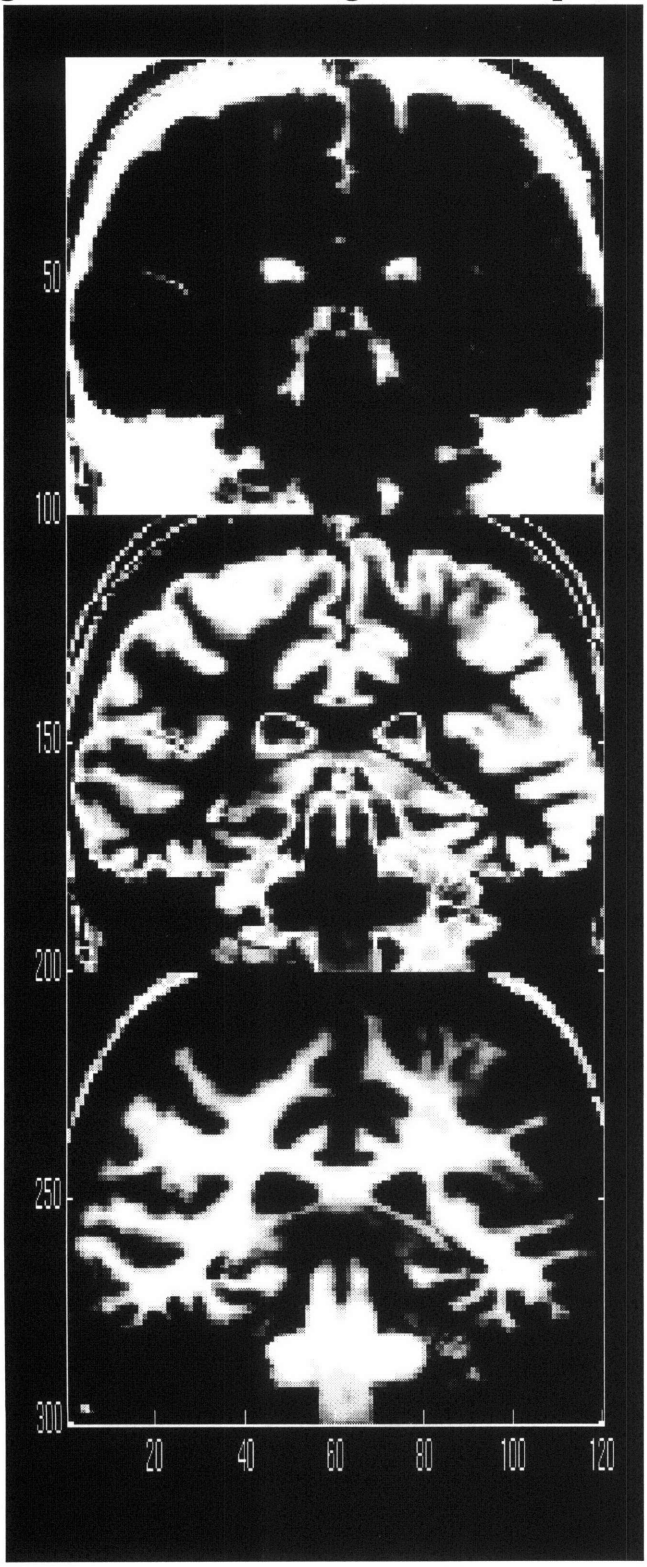


B- GMM-based Segmentation



C- NEM Segmentation

Figure 3.7: Probabilistic Segmentation Maps using MEM algorithm



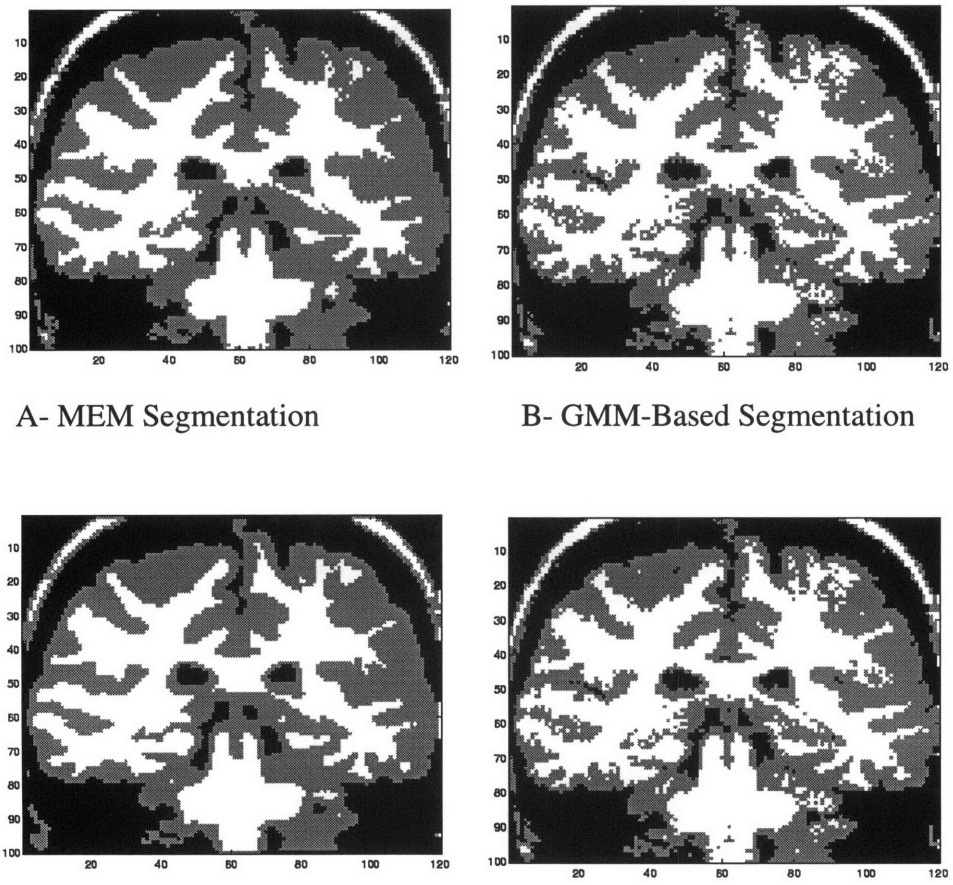
(A) - Pixels with a high probability of belonging to CSF or air are given higher intensities in this subimage.

(B) - Pixels with a high probability of belonging to Gray Matter are given higher intensities in this subimage.

(C) - Pixels with a high probability of belonging to White Matter are given high intensities in this subimage.

We also attempt to segment the noisy MR brain image, figure (2.6), of the previous chapter. A segmentation map of the noisy image was generated using the MEM algorithm. The segmentation is compared with a the GMM-based segmentation, the NEM algorithm, and low-pass filter ad-hoc algorithm. Figure (3.8) shows the segmentation maps of all these aforementioned schemes. We see that even in the presence of excessive noise, the MEM algorithm is able to derive a smooth segmentation of the image field. However, both the GMM-based algorithm and the NEM algorithm had apparent errors in classification of the pixels. The low-pass filter ad-hoc algorithm did result in a smooth segmentation of the image field. This can be explained because the filtering reduced the noise and thus, allowed for a smooth segmentation. However, as we have pointed out, in images with fine details and sharp edges, the low-pass filtering can have the adverse affect of partial voluming. Because of the multiresolution approach of our MEM algorithm, we avoid the partial voluming.

Figure 3.8: Segmentation maps of noisy brain image



A- MEM Segmentation B- GMM-Based Segmentation

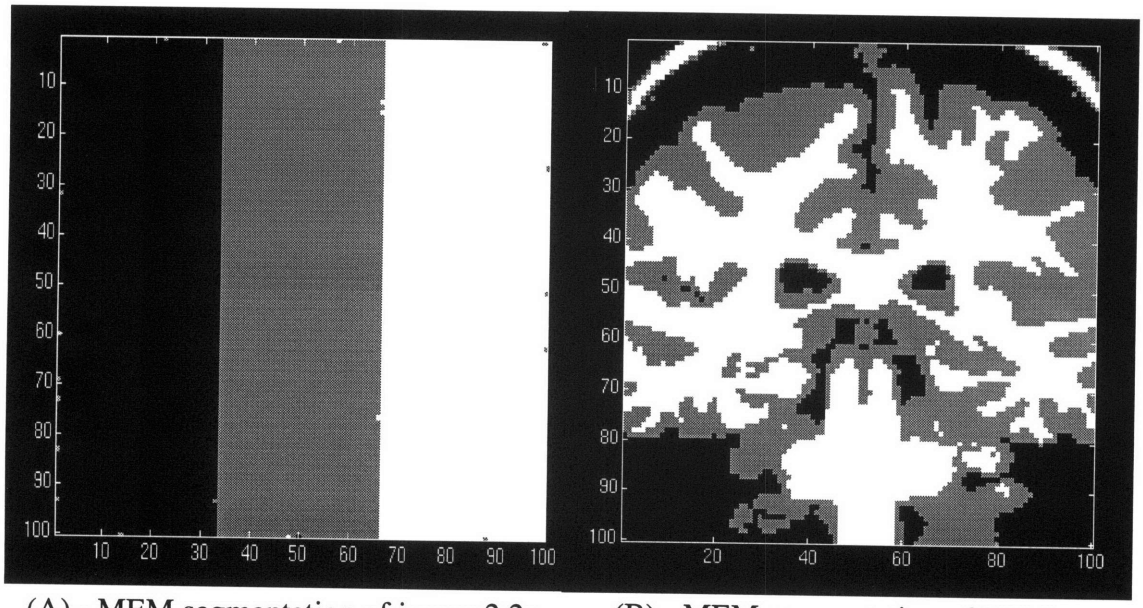
C- Low-Pass Ad-Hoc Segmentation D- NEM Segmentation

3.3.3 Higher Order Basis Functions

The MEM algorithm is modular and can be implemented efficiently with other basis functions. Here, using a seventh order Daubechies wavelet (DB7), we include results of segmentation of both the "synthetic" test image as well as an MR image of the brain. The "Wavelets Toolbox" of MATLAB was used for implementation of the DWT. Figure (3.9) includes segmentation maps of the images of figures (2.2a) and (2.5). While the brain segmentation is comparable to the Haar-based implementation, there is a noticeable difference in the segmentation of the "synthetic" image, (2.2a). There are a few more errors in

classification near edges when using DB7. This can be understood from knowledge of the frequency response of the DB7 filter. Because DB7 has more energy at the lower frequencies, the edges will be more blurred because edge information is typically at higher frequencies.

Figure 3.9: Segmentation Using Higher Order Wavelets



(A) - MEM segmentation of image 2.2a using Seventh Order Daubechies Wavelet

(B) - MEM segmentation of MR Image using Seventh Order Daubechies Wavelet.

3.3.4 Conclusions

In this chapter we introduced a novel multiresolution algorithm for image segmentation. Our algorithm incorporates a penalty based upon a multiresolution neighborhood clique into the likelihood function of the image field. This penalty captures the correlation between neighboring pixels. The conventional GMM-based algorithm assumes pixels are i.i.d. We showed that the MEM algorithm is robust with respect to additive Gaussian noise and produces better segmentation maps than the conventional GMM-based segmentation algorithm. Specifically, the MEM algorithm avoids the partial voluming which usually occurs when images are filtered. The simulations indicated that the MEM algorithm is able to segment fine details in images. The MEM algorithm is modular and can be implemented with different basis function for multiresolution analysis. Moreover, our scheme is

computationally tractable. In contrast, MRF-based models require much more computational processing.

Chapter 4

Image Restoration Via ML Parameter Estimation of Spatially Varying Models

Images can be corrupted by various noise processes such as impulsive and additive Gaussian noise due to a noisy sensor, lossy compression, and transmission over noisy channels. In modern communication systems, digital filters are often utilized to process data, voice, image, and video streams. In general, a digital filter used in image processing must satisfy one or more of the following requirements:

1. Restore the original image from its noisy version (smoothing)
2. Enhance certain features (edges) of the degraded image (sharpening)
3. Preserve fine detail in images (detail preservation)
4. Be implementable in real time (computational efficiency)

In restoring digital images, we attempt to simultaneously suppress noise and preserve the essential visual information in images. To accomplish this, an accurate model of the image field is important. The model parameters are not known a priori. Thus, given the data of the image field, we can utilize robust estimation techniques to find the estimates of the model parameters.

In this chapter, we shall develop a nonlinear filter to restore digital images corrupted by noise. Most nonlinear filters used for image restoration apply a "sliding" processing window. The windowing is used because most natural images are nonstationary. We also apply this windowing technique and model the data within the window using the generalized Gaussian mixture model. As we move the window about the image field, we use MAP estimation to restore the center pixel of the window. We shall utilize ML

parameter estimation via the EM algorithm to estimate the model parameters at each window location.

The rest of the chapter is organized as follows: the next section describes in more detail the theory, implementation, and applications of our algorithm. Then we include a section which analyzes the performance of our filter on the benchmark test image, "Lena." Moreover, we include a comparative study of our algorithm and the mean and median filters. We conclude the chapter with a general discussion of the results and implications of our study, and we consider future extensions of this work.

4.1 Bayesian EM Filter

4.1.1 Noise Model

We assume that each pixel of the degraded input data set from the image field obeys an additive model

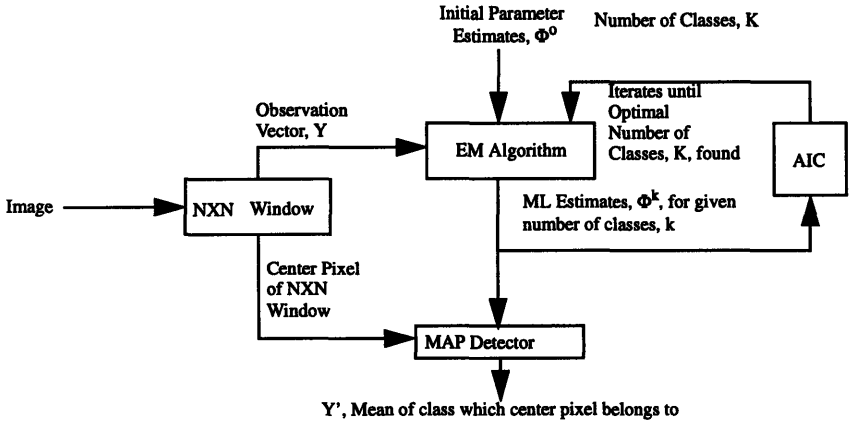
$$y_i = x_i + w_i \tag{4.1}$$

where the original pixel x_i and the noise process w_i are statistically independent. Furthermore, the pixels are assumed to be independent of one another. In this development, a parametric model has been used to characterize the observed data. Again, this parametric model is in the form of a pdf, $p(Y|\Phi)$, where Y is a vector of the data and Φ is a vector of the model parameters. Furthermore, given the nonstationarity of images, we would expect the model parameters to be spatially varying. For these reasons we take the approach of estimating these parameters using a sliding processing window of data surrounding the center pixel. Figure 4.1 illustrates the sliding window technique used in our algorithm.

In this chapter, we take a Bayesian approach to the restoration of noisy images. We model the image intensities as composed of a number of classes. Each such class is characterized by a corresponding probability density function, in particular, by a Gaussian of given mean and variance. The observed data are assumed to be samples from these class

densities. Thus, the data as a whole can be characterized by a **Gaussian mixture density** as was defined in Chapter 2 by equation (2.1). For each pixel we find the *maximum a posteriori* (MAP) estimate of corresponding pixel class and replace the value at that pixel with the mean of the corresponding class. In practice, we do not really know the parameters of the class densities, or even the number of such classes. For these reasons, we find the Maximum Likelihood (ML) estimates of the pixel class density parameters based on the data in the window. Since the class correspondence of each pixel in the window of data is not known *a-priori,* ML estimation for the class parameters is a challenging nonlinear problem. We resort to the expectation maximization (EM) algorithm for its solution. In addition, since the number of classes at each location is also unknown and, in general, spatially varying, we combine the ML approach with the Akaike Information Criterion (AIC) [28] to determine the number of classes directly from the data. Finally, based on these estimated class parameters we process the center pixel as described above.

Figure 4.1: Schematic of Novel Nonlinear Filter



4.1.2 Outlier Detection

Since our method reestimates the class parameters for each pixel, it adapts to the spatially-varying structure of the image. Since the EM algorithm effectively performs classification of every pixel in the data window, we also obtain information about the

local distribution of the pixel intensities. We use this information to identify local outlier classes, which are then excluded from further processing. This makes our procedure robust to such effects as impulsive noise. Impulsive noise can be considered as a noise which makes a pixel an outlier. We shall use the following definition for an outlier class:

Definition 1.1 Let **Y** be a vector of data points within a window. Then, let **W** be the subset of the data points in **Y** which are members of certain class, $\mathbf{k_w}$, that is characterized by a Gaussian pdf with a given mean, μ_w, and variance, σ_w^2. Class $\mathbf{k_w}$ is an outlier class if and only if the prior probability of $\mathbf{k_w}$, $p(\mathbf{k_w})$ is estimated to be less than α as shown by equation (4.2). Thus, an outlier class has a low prior probability.

$$p(k_w) < \alpha \tag{4.2}$$

4.1.3 MAP Estimation Applied to Image Restoration

First we seek the MAP estimate, $\hat{k}$, of the class of pixel i as:

$$\hat{k} = \arg\max_k p\langle k|y_i\rangle \tag{4.3}$$

$$= \arg\max_k p\langle y_i|y_i \in k\rangle p\langle k\rangle \tag{4.4}$$

Where we assume that there are up to K classes, and $p\langle y_i|y_i \in k\rangle$ is the probability of pixel i given it is from class k given by:

$$p\langle y_i|y_i \in k\rangle = N(m_k,\sigma_k^2) \tag{4.5}$$

and $p(k_i)$ is the prior probability of class k_i. Assuming that $p(k_i)$ and the class parameters, $\{m_k,\sigma_k^2\}$, for each class k are known, we can easily compute this MAP estimate of the class at each pixel. Given this MAP estimate of the class at pixel i, we then replace the value of the pixel by the mean of the corresponding class.

4.1.4 ML Parameter Estimation of Spatially Varying Image Models

Unfortunately, in practice, $p(k_i)$ and the class parameters $\{m_k,\sigma_k^2\}$ are not known *a priori*. Further, given the nonstationarity of images, we would expect these parameters to

be spatially varying. The approach we thus take is to estimate these quantities for each pixel using a window of data surrounding the given pixel i as depicted in Figure 4.1. In particular, we find the Maximum Likelihood estimates of the pixel class density parameters. Let $\Phi = [m_1, \sigma_1^2, m_2, \sigma_2^2, .., m_k, \sigma_k^2]^T$ be a vector of all the class parameters (assuming the number of classes K is known) and Y a vector of all the pixel intensities in the given window. Then, assuming that the pixels are independent, we estimate the elements of Φ as:

$$\widehat{\Phi} = \arg max_{\Phi} p\langle Y|\Phi\rangle \tag{4.6}$$

where $p(Y|\Phi)$ is defined as:

$$p\langle Y|\Phi\rangle = \prod_{i=1}^{N} \sum_{k=1}^{K} p(k)\, p\langle y_i|\Phi, y_i \in k\rangle \tag{4.7}$$

Since the densities themselves depend on Φ, ML estimation for the class parameters is a challenging nonlinear problem. We resort to the expectation maximization (EM) algorithm for its solution [20]. The EM algorithm also finds the ML estimate of the prior probabilities, $p(k_i)$. Now, in reality we also do not know the number of classes K at each pixel, which itself will in general be spatially varying. Again, our approach is to estimate this value from the data in the window about pixel y_i by combining the ML estimation approach above with the Akaike Information Criterion (AIC) [28].

4.1.5 Estimation of Optimal Number of Classes

As previously mentioned, the "windowing" approach used in this thesis presents a new problem in estimating the model parameters. In particular, the number of classes of data within a window are not known *a priori.* In homogenous regions there will typically be only one class. However, near edges, there may be two or more classes. In the context of image restoration, the AIC poses the problem as finding the best number of classes, K,

to fit the data model within the window. We seek to choose K that will minimize the following cost function:

$$AIC(K) = -2\log\{p_k(Y|\Phi^k_{ML})\} + 2K' \tag{4.8}$$

where Φ^K_{ML} is the Maximum Likelihood(ML) estimate of the model parameters for that given K. K' is the number of independently adjustable parameters of the K-class model. K' in general is equal to K. Y is a vector of the pixel intensities within that window. At each pixel, y_i, in the image, we estimate the optimal number of classes for the data surrounding y_i.

4.1.6 Edge-preserving Property of Bayesian EM (BEM) Filter

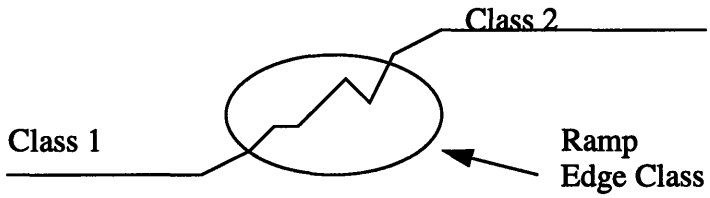
The edge-preserving property of the BEM filter is due to allowing for multiple classes within a window. The "mean" or averaging filter, is often used to reduce additive Gaussian noise, blurs edges by replacing the center pixel with the average of the pixel intensities located on both sides of the edge. However, our filter will only use the mean of one class of pixels. Furthermore, those classes whose *a prior* probabilities, $p(k_i)$, are estimated to be relatively small in comparison with other classes can be considered outliers. Identification of outliers will reduce the effects of impulsive noise.

4.1.7 Edge Detection Property of BEM Filter

We have also utilized the estimated variances of inlier classes to gain an additional insight into the homogeneity of an image. The variance of the class to which the center pixel is estimated to belong will be high in regions near edges and other singularities in the image, and low in homogenous regions. An example in one dimension is illustrated in figure 4.2. Using the estimated inlier class variance, we propose that the BEM filter can perform effective edge-detection on noisy images while simultaneously restoring images. Our restored image utilizes the mean of the inlier class as the output of the BEM

filter. To generate edge maps, we use the variance of the inlier class as the replacement value of the center pixel.

Figure 4.2: High Variance edges. .



Pixel populations located on edges will in general, have higher variances. Extensions of this example to two dimensions is straightforward

4.2 Experimental Results

Our algorithm was applied on the test image "Lena" as shown in figure (4.3). The two classes of noise used to corrupt the image are the additive Gaussian noise and the impulsive (salt&pepper) noise processes. The peak signal-to-noise ratio (PSNR) was measured with the original uncorrupted image serving as the reference. We judge the quality of the restoration by measuring the improvement in PSNR and the subjective quality of the restored image. PSNR is defined by the following equation

$$PSNR = \left(10\log\left(\frac{\sum_{i=1}^{n} 255^2}{\sum_{i=1}^{n} (y(i) - v(i))^2} \right) \right) \tag{4.9}$$

Here, n represents the total number of pixels in an image. Here, $y(i)$ is the original (reference) image, and $v(i)$ is the restored image. PSNR is among the most popular metrics used to judge the performance of restoration methods [3].

Figures (4.4), (4.5), and (4.6) include the results of our simulations for the median filter, mean filter, and the BEM filter. Our algorithm is demonstrated to be superior to the

mean and median filters both in terms of PSNR and subjective image quality in presence of Gaussian noise. Its performance is comparable to the median filter in the presence of impulsive noise in terms of PSNR, but has a better subjective quality by producing sharper images. In the presence of both impulsive and Gaussian noise processes our filter outperforms both median and mean filters in terms of PSNR and subjective image quality as illustrated in figures (4.4) and (4.5). The images are displayed using a grayscale map, where each image is normalized so that its peak intensity is 255 and its lowest intensity is 0.

Moreover, the BEM filter can be used to detect the edges in presence of noise as illustrated in figure (4.7). We compare the edge map of the "Lena" image obtained via a Marr-Hildreth edge detector with the edge map obtained using the BEM filter. From figure (4.7b), we see that the Marr-Hildreth edge detector is sensitive to impulsive noise. However, the BEM filter produces a superior edge map even in the presence of noise.

Figure 4.3: Test Image "Lena," and Noisy Image with Gaussian and Impulsive noise.

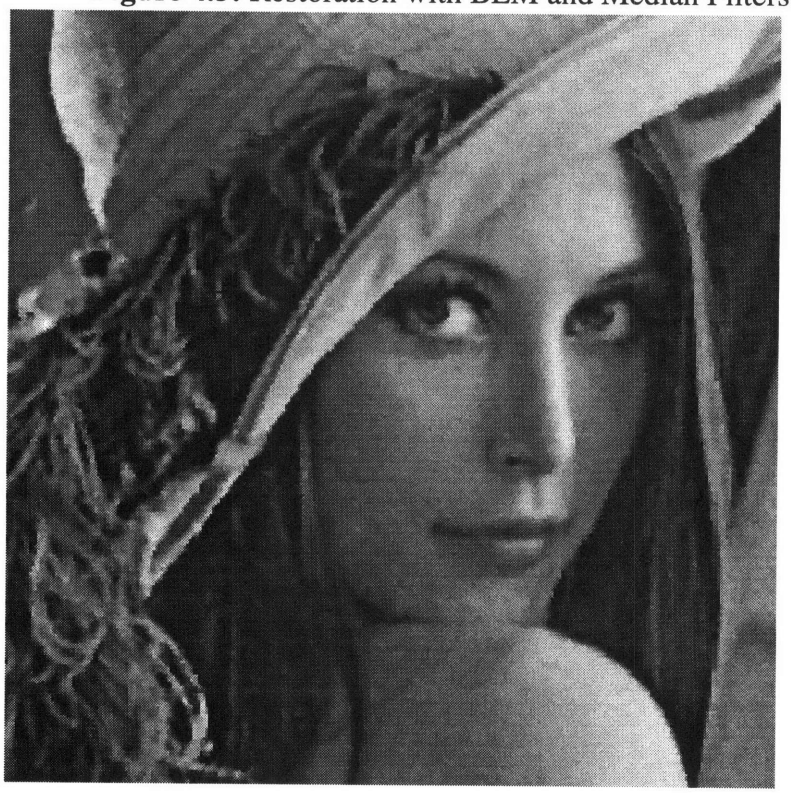
(A) - Original Lena Image

(B)- Corrupted with additive Gaussian and Impulsive Noise

Figure 4.4: Restoration with Mean Filter, PSNR = 23.85 dB

(A) - Restored with Mean Filter

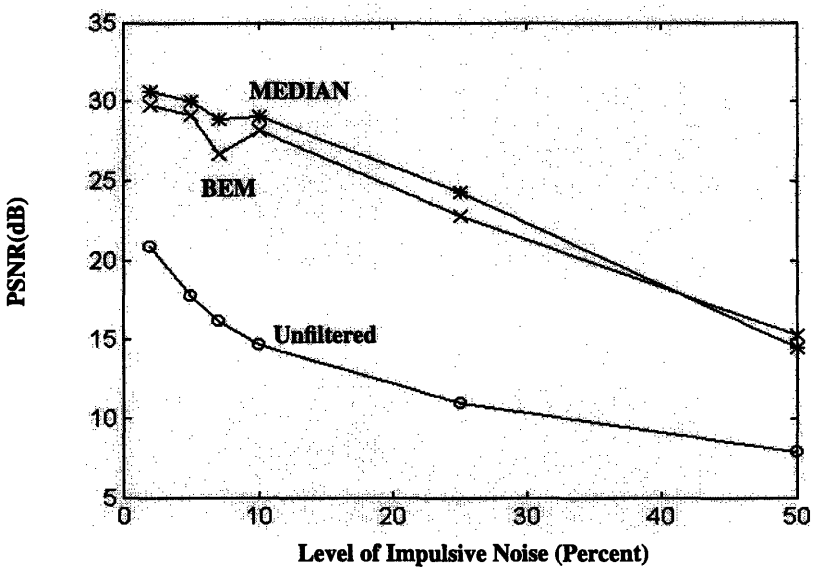
Figure 4.5: Restoration with BEM and Median Filters



(A) - Restored with BEM filter, PSNR = 26.8 dB.

(B) - Restored With Median Filter, PSNR = 24.75 dB.

Figure 4.6: BEM, Median and Mean Filter Restoration of Noisy Images



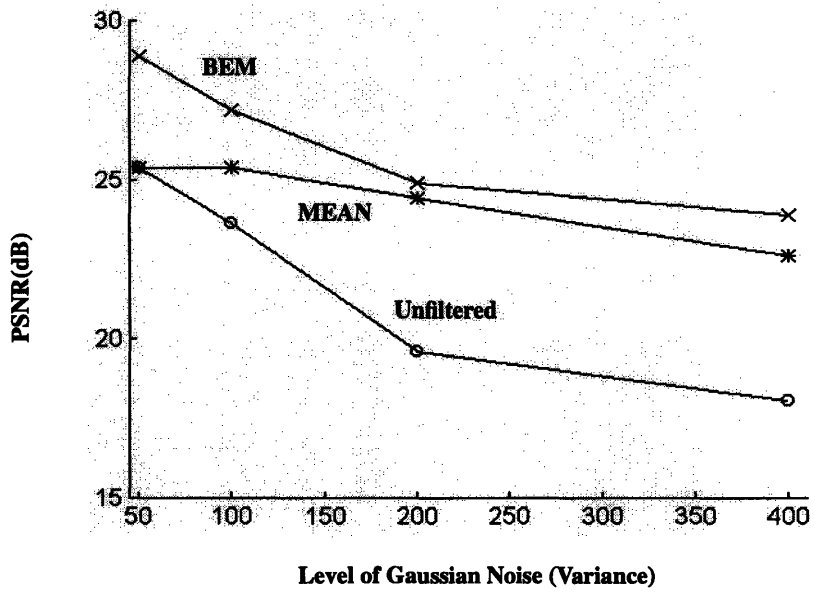
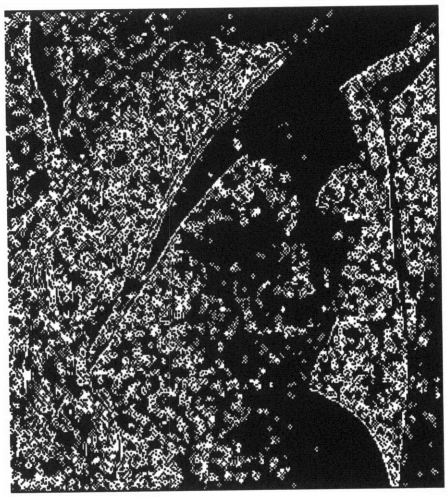
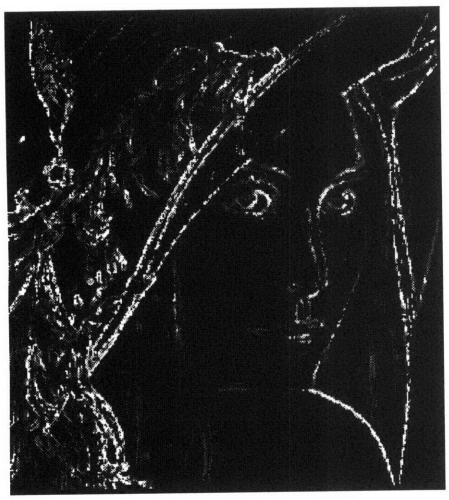


Figure 4.7: Edge Detection Of BEM Filter

(A) - Edge Map of Marr-Hildreth Edge Detector with Original "Lena'



(B) - Edge Map of Marr-Hildreth Edge Detector with Noisy "Lena"



(C) - Localized Variance of BEM Filter with Noisy "Lena"

4.3 Conclusions

In this chapter, we have presented a novel nonlinear filter which utilizes ML parameter estimation via the EM algorithm using a sliding processing window. We model the local

distribution of pixels in a window as a Gaussian Mixture. Using only the data within the window, we estimate the number of classes and the parameters of the classes. We identify outlier classes which makes our filter robust to impulsive noise as well preserves details and edge integrity. Furthermore, this new BEM filter was shown to attenuate Gaussian noise in addition to impulsive noise. In past studies, it has proven difficult to develop a filter which is capable of suppressing both Gaussian and impulsive noise at high levels. Indeed, a challenging task is to suppress these noise processes while preserving edges and fine details.

Moreover, the localized variances computed by our BEM filter provide important information about the level and location of smooth regions in the image and location of edges which could be used in many applications such as edge-detection and segmentation based image coding. We demonstrated that the BEM filter can produce edge maps even in the presence of noise.

Another salient issue is in regards to the window size used for processing. In this study, a three by three window was used because of simplicity of implementation and also because most other nonlinear filters use the three by three window [3]. However, in homogenous regions, a larger window would invariably reduce the error variance. Conversely, near edges, a larger window would cause blurring. Thus, a possible extension for our Gaussian mixture model is to utilize the statistics to make the window size adaptive to fit the data. Similar to how we allowed for changes in the number of classes, K, entropy-based cost functions not unlike the AIC can be used for this purpose. We point out that adaptive windows have been implemented with models other than the GMM [17].

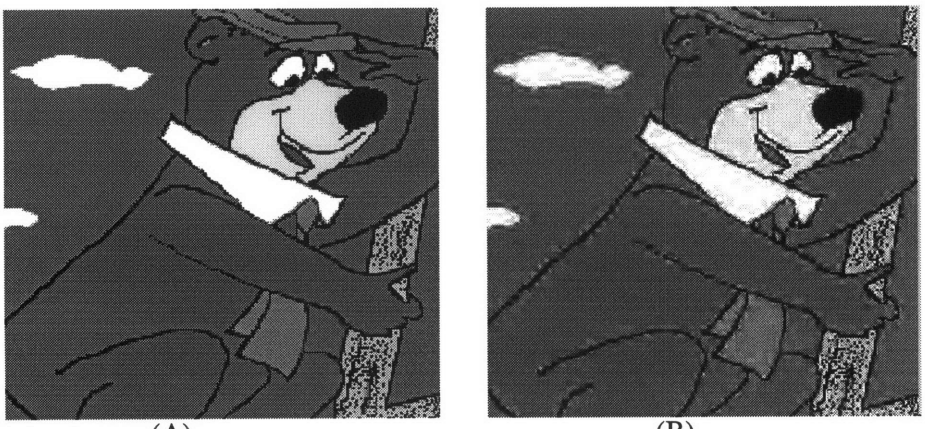
Chapter 5

Reduction of Compression Artifacts with Bayesian Post-processing

5.1 Introduction

In modern multimedia communication systems, data compression is needed to increase the rate at which information is exchanged over band-limited communication channels. In these systems, image and video compression at very low-bit rates are needed for applications such as videophones, teleconferencing, and catalog browsing [26]. Often, very low-bit rate compression techniques are lossy [29]. Lossy image compression at lower bit rates often results in loss of information in the decompressed data, and hence degradation of subjective image quality [26]. The most popular image coding algorithms are transform-based. In this class of algorithms, after applying a transform such as the Discrete-Cosine Transform (DCT) or Discrete Wavelet Transform (DWT), the coefficients of the transform are quantized and entropy encoded to form the compressed bit stream. The quantization step can be considered as the source of noise that causes the loss of information in encoded data [26]. At lower bit rates the loss appears as ringing artifacts in DWT based compression techniques [26], as illustrated in figure 5.1.

Figure 5.1: Ringing Compression Artifacts



(A) (B)

Fig. 5.1 - Ringing compression artifacts at lower bit rates. (a) Original Image (b) Wavelet based encoded image at 0.125 bits per pixel.

Various algorithms have been proposed to reduce these artifacts by postfiltering of the decoded bit streams. Among the most prominent algorithms to date are low pass filtering (LPF), projection onto convex sets (POCS), and maximum a posteriori (MAP) techniques [30]. These methods are mainly used to reduce the artifacts in the spatial domain after decoding, whereas the noise is introduced in the transform domain. The recent results in [29] suggest that the reduction of quantization noise in the transform domain or joint transform and spatial domains, would be more efficient than the conventional spatial domain approaches. In this chapter, we introduce a new transform domain postprocessing algorithm that uses our Bayesian filter, described in the precious chapter, to filter the transform coefficients prior to decoding. In the previous chapter, we assumed that the noise statistics were not known a priori. In this chapter, we assume that the quantization noise statistics are known and that the noise is uncorrelated with the image data. However, we still have no knowledge of the image statistics. Hence, we must estimate the image model parameters from the data. Knowledge of the noise statistics, and the assumption that the noise is uncorrelated with the data, represents the key difference with the BEM

filter used in the preceding chapter. Our experimental results have shown that the BEM postprocessing algorithm can effectively reduce quantization noise in wavelet-encoded images. Section 5.2 introduces the BEM postfiltering algorithm. Finally, the experimental results and concluding remarks are presented in section 5.3.

5.2 The Bayesian Postfiltering Algorithm

Given an $N \times N$ digital image, $\boldsymbol{x}$, we will define the quantized transform image of $\boldsymbol{x}$ as

$$\boldsymbol{Y} = Q[T\boldsymbol{x}] \tag{5.1}$$

where T is the transform matrix (e.g. DCT or DWT), and Q[•] represents the quantization function. Assuming a uniform quantizer is used, we can model $\boldsymbol{Y}$ as

$$\boldsymbol{Y} = \boldsymbol{X} + w \tag{5.2}$$

Where $\boldsymbol{X}$ is the transform of $\boldsymbol{x}$, and $\boldsymbol{w}$ is an additive zero-mean uniform noise process with variance of σ^2, that represents the quantization noise. In this chapter, we shall make the common assumption that $\boldsymbol{w}$ is uncorrelated with the signal $\boldsymbol{X}$ [29].

Now, we shall formulate the problem of estimating $\boldsymbol{X}$ from $\boldsymbol{Y}$ by using the BEM filter. We use a sliding processing window of size $L \times L$ in $\boldsymbol{Y}$, and try to estimate the true value of the center pixel i, of the processing window from its neighbors. If the mean m_{x_i}, and variance $\sigma_{x_i}^2$, of i were known a priori, the weighted least-squares (WLS) algorithm could have been used to estimate $\boldsymbol{X}$ [29]. Let $\boldsymbol{x}$ and $\boldsymbol{y}$ be the lexicographical ordering of the pixels in the processing window, then to find the WLS estimate, we should search for $\hat{x}_i$ that minimize the functional

$$J = \arg min_{\hat{\boldsymbol{x}}}\left[(\hat{\boldsymbol{x}} - \boldsymbol{m}_x)^t M (\hat{\boldsymbol{x}} - \boldsymbol{m}_x) + (\boldsymbol{y} - \hat{\boldsymbol{x}})^t R (\boldsymbol{y} - \hat{\boldsymbol{x}}) \right] \tag{5.3}$$

where M and R are weighting coefficient matrices. In this case $\hat{x}$ can be found explicitly as

$$\hat{x} = m_x + W(y - m_x) \tag{5.4}$$

Where $W = (M + R)^{-1}R$. We can derive the value of W which minimizes (5.3) and this is given by

$$W = \sigma_{x_i}^2 / \left(\sigma_{x_i}^2 + \sigma_{w_i}^2 \right) \tag{5.5}$$

unfortunately, in practice the mean and variance of each pixel intensity x_i is not known *a priori.* In practice, we cannot estimate the statistics of both the image and the noise from the data alone. Therefore, our model shall make the following simplifying assumptions. Since the noise is zero-mean and uncorrelated with the signal, we can assume

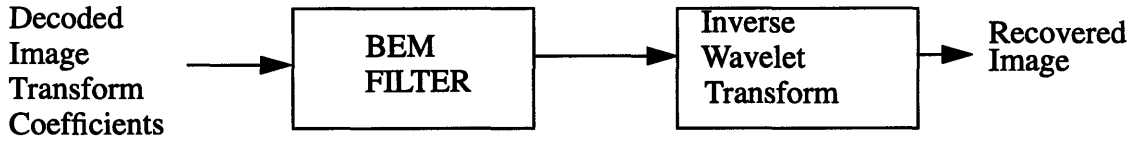
$$m_{x_i} = m_{y_i} \tag{5.6}$$

Here m_{y_i} is the mean of the inlier group for the processing window. The inlier group was defined in chapter 4, and is found using the same mixture model based upon the EM algorithm. Furthermore, we can assume the variance of x_i, is given by

$$\sigma_{x_i}^2 = \sigma_{y_i}^2 - \sigma_{w_i}^2 \tag{5.7}$$

Variance of y is found as the variance of the inlier group in the processing window, where the inlier class is associated with the class of the center pixel. The variance of the quantizer noise is also known *a priori.* The above formulation is similar to the formulation of the BEM filter in chapter 4, and hence we can estimate the essential information to compute the least squares estimation of x, and restore digital images that have severe quantization noise. The BEM postprocessing algorithm is summarized in Fig. 5.2.

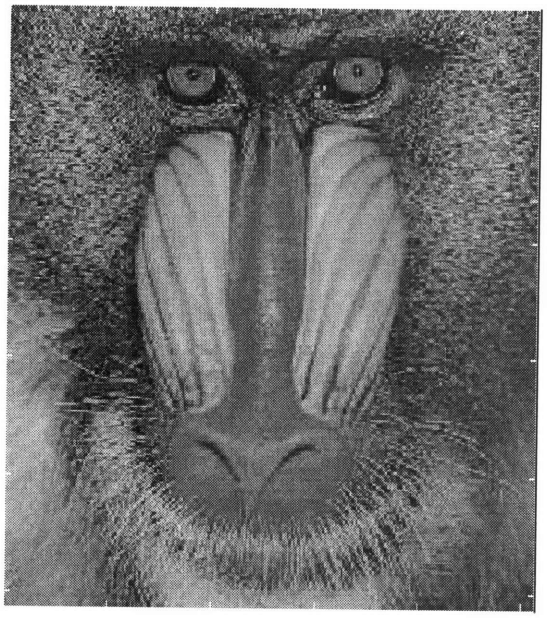
Figure 5.2: The BEM postprocessing algorithm



5.3 Experimental Results and Concluding Remarks

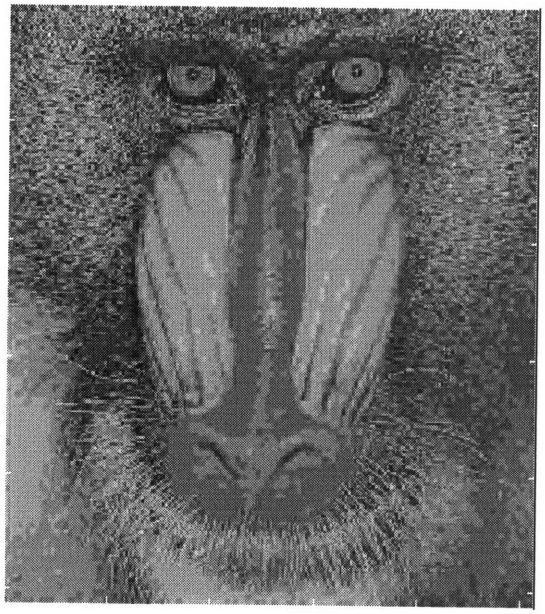
Our BEM post-processing algorithm was applied to real images so that we can illustrate the performance of our algorithm. The Baboon test image is shown in figure 5.3 (a). The DWT of the Baboon was obtained using the DB4 wavelet. The 8 bits per pixel (bpp) DWT coefficients were quantized to 3 bpp each. The inverse transform without postprocessing shows a severely corrupted image as shown in figure 5.3 (b). The BEM algorithm was then applied to the quantized coefficients prior to inverse transformation. Significant improvements in subjective image quality can be observed in figure 5.3 (c). Furthermore the BEM postprocessing resulted in 0.59 dB improvement of the peak signal to noise ratio (PSNR).

Figure 5.3: BEM Processing of Coded Images



(A) Original Image

(B) Compressed Image



(C) BEM Post-Processed Image

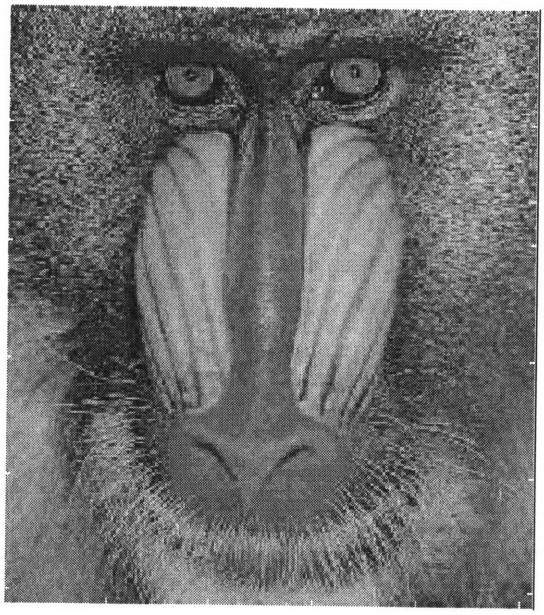


Fig. 5.3 - Reduction of quantization artifacts with BEM filter. (a) Original Baboon (8 bpp), (b) Decoded Baboon without filtering PSNR 20.05 dB (c) Baboon with BEM postprocessing PSNR 20.64

Our experimental results confirmed that the BEM postprocessor can effectively reduce the quantization artifacts in DWT transform based compression techniques. Our algorithm is computationally efficient to implement and produces images that have a better subjective quality and higher PSNR than the current iterative algorithms.

We must emphasize that the simulation used in this thesis did not attempt to use a multilevel wavelet decomposition of images. Multilevel DWT is the most common method used in wavelet-based compression algorithms. Here, we presented the results of a singe level decomposition and subsequent quantization at 3 bpp. To achieve lower bit rates, a multilevel decomposition is antecessor so that we can assign each subband a specified number of bits per pixels. In future studies, we will implement the BEM postfiltering algorithm on multilevel wavelet decomposition of images. This will demonstrate its applications in real-time image and video compression systems.

Chapter 6

Conclusions and suggestions for future research

6.1 Thesis contributions

The focus of this thesis has been on extensions of the Gaussian mixture model to the challenging tasks of image segmentation and restoration. A salient commonality between these two tasks is that uncertainty plays a fundamental role. We have developed probabilistic models to capture this uncertainty. Natural images are typically nonstationary. Furthermore, the value of an image pixel of interest is strongly correlated with its local neighbors. These two key assumptions have allowed us to develop robust methods for segmenting and restoring images. The details of our algorithms were presented in chapters 3,4, and 5. Here we present a summary of our thesis contributions.

6.2 Multiresolution-based Mixture Models: Applications to image segmentation.

In typical applications of the Generalized Gaussian Mixture model to image segmentation, the standard GMM assumes pixels are i.i.d, which yields segmentations that are sensitive to noise and fails to reflect the contiguous nature of homogenous regions. Based on this i.i.d assumption, ML estimation of the model parameters via the EM algorithm and subsequent MAP classification are used to generate segmentation maps of the image field. In practice, neighboring pixels of natural images are highly correlated. Thus, there is interest in incorporating spatial correlation into a more robust segmentation algorithm. Typical attempts to incorporate such correlation models into the segmentation process (e.g. MRF Models) result in computationally taxing optimization problems. Ambroise et al. modified the EM algorithm to include a neighborhood structure into the likelihood function at a tractable computational cost. In our research, we extended this work by developing a

novel multiresolution neighborhood penalization term based upon a multiscale quadtree structure. Using the multiresolution neighborhood, we modified the EM algorithm E-step such that spatial correlation is incorporated into the estimation of the model parameters, in particular, the tissue class probabilities. We have shown in simulated images as well as MR images of the brain that our algorithm allows for accurate segmentations of the image field. Moreover, in comparative studies, our multiresolution approach was more robust than the NEM algorithm of Ambroise. By utilizing information at many resolutions, our algorithm advantageously utilized the information in the coarser resolutions to segment the prominent features of an image. The information in the coarser resolutions was utilized to make our algorithm insensitive to Gaussian noise in images. Conversely, the NEM algorithm was shown to be sensitive to Gaussian noise. Furthermore, we utilized the information in the finest resolutions to segment fine details in images.

For the examples presented in this thesis, the weightings used for the scale-space neighborhood were heuristic. The weights were chosen based upon the knowledge of what the correct segmentation of the image field should be. While the same weights were used for all the images in this thesis, we believe that a systematic approach to determining the weights will yield a more robust segmentation algorithm.

Another important extension of this research deals with choosing the best wavelet for multiresolution analysis. Our research demonstrated that different wavelets yield different segmentation maps. A possible study can involve the implementation of our MEM algorithm using wavelets of different smoothness. Images tend to differ in the sharpness of their edges. Many images have sharp "step" edges whereas others have gradual ramp edges. Accurate segmentation along ramp edges is very challenging. We propose that segmentation along ramp edges can be more accurate by choosing the wavelet with the appropriate smoothness.

6.3 BEM Filter

In this study, we developed a windowed model of the image field using the Generalized Gaussian Mixture model. Using a sliding processing window, we estimated the model parameters using ML estimation via the EM algorithm. In contrast to existing methods, our algorithm exploited the local statistics that we estimated in order to detect outliers and preserving edges and fine details in images.

In future studies, the size of the processing window should be adaptive to the structure of the image. For example, near edges and areas of fine detail, a smaller window is needed to preserve these features. However, in homogenous regions, a larger window can be used to reduce the error variance of the estimate and suppress additive Gaussian and impulsive noise more effectively.

In compression schemes where uncorrelated quantization noise is responsible for significant reduction in subjective image quality, we utilized the *a prior* statistics of the noise to restore wavelet-based compressed images. This novel application of the BEM filter was made possible as a result of the ML parameter estimates that the BEM filter outputs. Artifact reduction in compressed images can perhaps dramatically reduce the bit rates that images are coded at currently. If an effective post-processing algorithm is used at the receiver of a communications channel, very-low bit rate wavelet-based image/video coding can be an attractive alternative to the current JPEG algorithm used in image coding.

References

[1] K. Popat, and R.W. Picard. "Cluster-Based Probability Model and Its Application to Image and Texture Processing." IEEE Transactions on Image Processing. Vol. 6, No. 2, February 1997. pp. 268-284.

[2] M.G. Bello, "A Combined Markov Random Field and Wave Packet Transform-Based Approach for Image Segmentation." IEEE Transactions on Image Processing. Vol.3, No. 6, November 1994, pp. 834-846.

[3] H. R. Rabiee and R.L. Kashyap, "GMLOS: A New Robust Nonlinear Filter for Image Processing Applications," Twenty Seventh Asilomar Conf. on Signals, Systems & Computers, Pacific Grove, CA, Nov. 1-3, 1993.

[4] I. Pitas and A. N. Venetsanopoulos, "Order statistics in digital image processing," Proceedings of the IEEE, vol. 80, no. 12, Dec. 1992.

[5] W.M. Wells, E.L. Grimson, R. Kikinis, and F.A. Jolesz, "Adaptive Segmentation of MRI Data." IEEE Transactions on Medical Imaging, Vol. 15, No 4, August 1996. pp 429-441.

[6] E.A. Ashton, M.J. Berg, K.J. Parker, J. Weisberg, C.W. Chen, and L. Ketonen, "Segmentation and Feature Extraction Techniques, with Applications to MRI Head Studies." MRM 33:670-677. 1995

[7] A.J. Worth, S. Lehar, and D.N. Kennedy, "A recurrent cooperative/competitive field for segmentation of magnetic resonance brain images." IEEE Transaction on Knowledge and Data Engineering, Vol. 4, No. 2, 1992, pp. 156-161.

[8] C. Ambroise, and G. Govaert, "Spatial Clustering and the EM Algorithm," Conference Proceedings of International Workshop on Mixtures, Aussois, France, Sept. 17-21, 1995.

[9] S. Mallat, and S. Zhong, "Characterization of Signals from Multiscale Edges," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol 14, No. 7, July 1992, pp. 710-732.

[10] G. Bongiovanni, L. Cinque, S. Levialdi, and A. Rosenfeld, "Image Segmentation By A Multiresolution Approach," Pattern Recognition, Vol. 26, No. 12, pp. 1845-1854, 1993

[11] D.J. Park, K.M. Nam, and R. Park, "Multiresolution Edge Detection Techniques," Pattern Recognition, Vol. 28, No. 2, pp. 211-229, 1995.

[12] K.L. Vincken, A.S. Koster, and M.A. Viergever, "Probabilistic Multiscale Image Segmentation," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 19, No. 2, February 1997, pp. 109-120.

[13] C.H. Fosgate, Multiscale Segmentation and Anomaly Enhancement of SAR Imagery, S.M. Thesis, Massachusetts Institute of Technology, June 1996.

[14] J. W. Tukey, Exploratory Data Analysis, Addison-Wesley, Reading, MA, 1970.

[15] R. J. Crinon, "The Wilcoxon filter: a robust filtering scheme," in Proc. IEEE ICASSP' 85, pp. 668-971, 1985.

[16] A. C. Bovik, T. S. Huang, and D. C. Munson, "A generalization of median filtering

using linear combination of order-statistics," IEEE Trans. on ASSP, Vol. 31, No. 6, Dec. 1983. pp. 1342-1350

[17] R. Bernstein, "Adaptive nonlinear filters for simultaneous removal of different kinds

of noise in images," IEEE Trans. on CAS, vol. 34, no. 11, Nov. 1987.

[18] I. Pitas and A.N. Venetsanopoulos, "Order Statistics in Digital Image Processing," Proceedings of the IEEE, vol. 80, no. 12, Dec. 1992.

[19] J. Park, and L. Kurz, "Image Enhancement Using the Modified ICM Method," IEEE Transactions on Image Processing, Vol. 5, No. 5, May 1996, pp. 765-771.

[20] J. Zhang, J.W. Modestino, and D.A. Langan, "Maximum-Likelihood Parameter Estimation for Unsupervised Stochastic Model-Based Image Segmentation," IEEE Transactions on Image Processing, Vol. 3, No. 4, July 1994, pp. 404-419.

[21] A. Papoulis, "Probability, Random Variables, and Stochastic Processes," McGraw-Hill, third ed., 1991.

[22] A.P. Dempster, N.M. Laird, and D.B. Rubin, "Maximum Likelihood from incomplete date via the EM algorithm, "Journal of Royal Society of Statistics, Series B., no. 1, pp 1-38, 1977

[23] J.C. Rajapakse, J.N. giedd, and J.L. Rapoport, "Statistical Approach to Segmentation of Single-Channel Cerebral MR Images," IEEE Transactions on Medical Imaging. Vol. 16, No. 2, April 1997, pp. 176-186.