DEVELOPMENT OF IMAGE RESTORATION METHOD USING HIERARCHICAL MRF MODEL

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

Master of Technology
in
Electronic Systems and Communication

By

R. MAITHRI
(Reg.No:211EE1322)

Under the Guidance of
Prof. DIPTI PATRA

Department of Electrical Engineering
National Institute of Technology, Rourkela
Rourkela-769008

(2011-2013)

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

CERTIFICATE

This is to certify that the thesis entitled, "**DEVELOPMENT OF IMAGE RESTORATION METHOD USING HIERARCHICAL MRF MODEL"** submitted by **Ms. R MAITHRI** in partial fulfillment of the requirements for the award of Master of Technology Degree in **Electrical Engineering** with specialization in **"ELECTRONIC SYSTEMS AND COMMUNICATION"** at the National Institute of Technology, Rourkela is an authentic work carried out by him under my supervision and guidance.

To the best of my knowledge, the matter embodied in the thesis has not been submitted to any other University / Institute for the award of any Degree or Diploma.

Date:

Place:

Prof. Dipti Patra
Department of Electrical Engineering
National Institute of Technology
Rourkela-769008

ACKNOWLEDGEMENT

Firstly, I thank my guide Prof. Dipti Patra for her continual support and confidence in me due to which I have been able to complete my work successfully. I am grateful to her for the teaching and the guidance, she has been a source of inspiration.

I am grateful to professors of my specialization, Professors P. K. Sahu, S. Das, K.R. Subhashini, and S. Gupta for their reviews and care. I am also grateful to the Head of the Department of Electrical Engineering, NIT Rourkela for providing the facilities for the work. Thanks to other faculty members in the department as well.

Thanks to Ms. Smitha, Mrs. Prajna, Mr. Yogananda for their reviews and constructive criticism on my work.

I thank all my friends for being there for me. I would like to specially thank my friend Vipin Kamble for helping me with my work. I also thank my friends Medhavi Mahobe, S.Monalisa, Sushant Sawant for their continual support.

R.Maithri

Abstract

In this thesis, hierarchical Markov random field model-based method for image denoising and restoration is implemented. This method employs a Markov random field (MRF) model with three layers. The first layer represents the underlying texture regions and the second layer represents the noise free image. The third layer represents the observed noisy image. Iterated conditional modes (ICM) algorithm is used to find the maximum *a posterior* (MAP) estimation of the noise free image as well as the texture region field. This method can effectively suppress additive noise and restore image details.

A noise-free gray-scale image is considered. Then Gaussian noise is applied to the image so that the image becomes noisy. The aim is to remove this noise from the image. Image is considered as the combination of disjoint texture regions, and a three-layered hierarchical MRF is used to model the image. The first layer represents the region labels. The second layer represents the noise free image color, and the third layer represents the noisy color.

The algorithm starts with choosing the number of the regions l, iteration time T and a MRF neighborhood system. Initially, the local variance of all the pixels is calculated considering a (3*3) window sliding through the image. K-means clustering is applied to the local variance feature image. The clustering result is set as s_i. The MRF parameters are estimated and then the clustered images and the noise-free image are updated using the ICM algorithm and the process is repeated till the MRF parameters become constant. The output obtained is the noise-free image.

The method used employs a three-layered MRF model which can express both smooth and texture signals. The advantage of hierarchical MRF model is that the texture information of the image is considered while the process of denoising, so that the edge information and other interesting structures of the image are not lost and the image is restored efficiently. Thus, the proposed image denoising method performs better than when the simple MRF model is considered.

Table of Contents

List of Figures

List of Tables

Chapter 1 Introduction

1.1 Introduction

Image denoising is a basic problem in image processing. Generally, denoising algorithms such as mean and median filters use some smooth image models, which assume that the image is smooth everywhere. Some of the methods uses piecewise smooth models to represent edges. These methods consider image as the combination of disjoint uniform regions. MRF is presented in smooth image models due to the simple expression of image segmentation.

Most of the MRF models used in image denoising applications are smooth MRF models where the neighboring pixels have the same colors. Because image contains textures besides smooth regions, the commonly used smooth MRF models lead to piecewise smooth results and lose some of the image details. In order to preserve textures of the image, a model that is consistent with texture is required. As image textures are often self-similar, based on this property some new methods restore images. These methods could suppress noise in both smooth regions and texture regions because these two types of regions are both self-similar.

We consider that image is made up of disjoint texture regions. So, we have to segment the image while denoising for which we go for hierarchical MRF image model. Hierarchical MRF model is widely used for texture image segmentation. The most popular model consists of two layers, one layer represents the underlying region labels which is characterized by a logical MRF (LMRF) and the other layer represents image texture. The latter one may model the image color directly or model some texture features such as the output of Gabor filtering of the image. Because modelling the image color directly such as Gaussian MRF (GMRF) and Simultaneous Autoregressive (SAR) is more suitable for the image restoration problem, we choose GMRF to model the noise free image. In order to solve image denoising problem, a third layer is added to the hierarchical MRF model which is the observed noisy image. Finally, the noise free image is obtained by MAP (maximum *a posteriori*) estimation.

1.1 Image Restoration

Image Restoration has been widely studied in image processing. It is used to restore degraded images which is the result of digital picture processing.

Efficient restoration is very useful for many image processing applications. It means the removal or reduction of degradations in the concerned image, and the degradations generally include blurring and noise. In many of the cases, the noise is modelled by a zero-mean white Gaussian process. In some cases, a non-Gaussian process is used as a more accurate characterization of the noise. Because image restoration is a prerequisite to many applications, numerous algorithms have been generated to restore the corrupted image. Bayesian image reconstruction is one of the most powerful approaches in image restoration. Geman and Geman(1984) gave a theory on Bayesian framework for image restoration using Markov random fields. MRF uses some general priors. These priors provide regularization to the restoration problem and smoothing is more powerful and robust. In practical situations, however, this information may not be directly obtained from the image formation process.

Images are to record or display useful information. Due to imperfections in the imaging and capturing processes, the recorded image represents a degraded version of the original image. The undoing of these imperfections is crucial for the performance of many image processing tasks. There exists a wide range of degradations that need to be taken into account, for example, noise, geometrical degradations, illumination and color imperfections (under/over-exposure, saturation),and blur. The field of image restoration (sometimes referred to as image deblurring or image deconvolution) is nothing but the reconstruction or estimation of the uncorrupted image from a blurred and a noisy one.

Image restoration algorithms are different from image enhancement methods in that they are based on models for the degrading process and for the ideal image. The assumption in most of the existing image restoration methods is that the degradation process can be explained using a mathematical model.

1.2 Image Denoising

Image denoising is to recover a digital image that has been contaminated by noise. Digital images play an important role in fields such as satellite television, magnetic resonance imaging, computer tomography, geographical information systems and astronomy. Data collected by image sensors are usually affected by noise. Data can be degraded due to imperfect instruments, problems with the data acquisition process, and interfering natural phenomena. Also, noise can be introduced due to transmission errors and compression. Thus, denoising is an essential step to be taken before the images data is analysed. It is important to apply an efficient denoising technique to reduce such data corruption. Image Denoising is a basic problem in image processing. Presence of noise in images cannot be avoided. It may have been introduced by the image formation process, image recording, image transmission etc.

For image denoising, the most common problem is that some interesting structures in the image will be eliminated from original image during noise suppression Such interesting structures in an image correspond to discontinuities or edges of the image. One pixel value is not independent but it has spatial dependencies on the values of its neighbors. This contextual prior knowledge must be used in our model. MRFs give a description of the interactions between neighboring pixels. Because there are many discontinuities in images, especially in the areas near the edges, we need to control the interaction between neighboring pixels to avoid over-smoothed solutions of images

In all the real applications, measurements are perturbed by noise. In the process of acquiring, transmitting, or processing a digital image, for example, the noise-induced degradation may be dependent or independent of data. The noise is generally described by its probabilistic model, e.g., Gaussian noise is characterized by two moments. A degradation yields an image observation model and is application dependent. The most commonly used model is the additive one $u_0 = u + \eta$ where the observed image u_0 includes the original image u and the independent and identically distributed noise process η.

Image denoising means to recover an image contaminated by noise as shown in Fig 1.1. The challenges of the problem are to faithfully recover the underlying image u from u_0, and to further the estimation by making use of prior knowledge or assumptions about the noise process η.

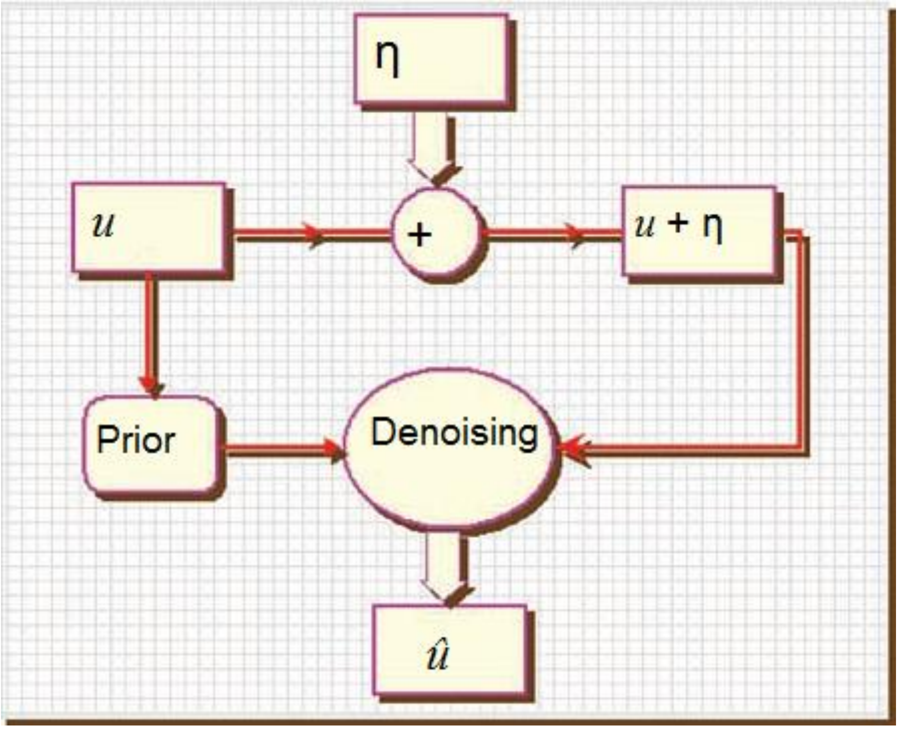


Fig 1.1. Block diagram of image denoising process.

1.3 Literature review:

Image denoising is a basic problem in image processing. Traditional denoising algorithms such as mean filter and median filters use some smooth image models. MRF modelling is used for image segmentation and denoising for its efficiency. Smooth MRF model in which neighboring pixels have similar colors was proposed by Maroquin *et al*. [10]. Tonazzini and Bedini *et.al* presented a coupled MRF model, consisting of a color field and an edge field [13].

Zhang *et.al.* added some geometrical constraints to the edge field with the similar model [15]. It also proposed a Markov random field (MRF) model-based EM(expectation-maximization) procedure for simultaneously estimating the degradation model and restoring the image is described. The MRF is a coupled one. Results on synthetic and real-world images show that this approach provides good blur estimates and restored images are more visually pleasing.

A pixon-based multiresolution method for image denoising was proposed by Qing Lu *et.al.* [9]. In this method, a pixon map is embedded into a MRF model under a Bayesian framework. A simulated annealing algorithm is implemented to find the MAP solution. The techniques of the complex wavelet transform and Markov random fields (MRF) model was combined to restore natural images in white Gaussian noise by Fu Jin *et.al.* [5]. A constrained optimization type of numerical algorithm for removing noise from the images was proposed [16]. The numerical algorithm is simple and relatively fast.

For a degraded binary scene, the image with the MAP estimate can be evaluated exactly using the efficient variants of the Ford-Fulkerson algorithm for finding the maximum flow in a certain capacitated network [4]. A new variational model for image denoising and decomposition was given which combines the total variation minimization model of Rudin, Osher and Fatemi [16] from image restoration, with spaces of oscillatory functions [17].

Since smooth image models don't give efficient results i.e. loss of information (interesting structures) of the image, we consider that an image is composed of disjoint texture regions. So a hierarchical MRF image model was proposed. Hierarchical MRF is a widely used model in texture image segmentation applications. It has been used in Noda *et al*., and Kim *et al*. [11,7].

Hierarchical MRF model considering the texture information for image restoration was proposed by [14]. This shows better results than the previous image denoising methods because of hybridization of MAP estimation using ICM algorithm and the hierarchical MRF model increased the efficiency.

1.4 Motivation

Existing methods for image restoration focusing on image denoising don't consider the texture information of the image which is very essential for preserving the edge and other discontinuities in the image. So we employ a hierarchical MRF model which considers texture information for a better restored image. An optimization algorithm, in our case, Iterated Conditional Modes(ICM) algorithm is combined with the hierarchical MRF model for improved results of a restored image from the noisy image.

1.5 Objectives

The objectives of the thesis are:

- Obtaining an image affected by white Gaussian noise.
- Employing a hierarchical MRF model on the image to estimate the MRF parameters.
- Using these parameters, MAP estimation and ICM algorithm, we restore the nosie-free image from the noisy image with good efficiency.

1.6 Summary of the thesis

Hierarchical Markov random field model-based method for image denoising and restoration is studied in this thesis. This method employs a Markov random field (MRF) model with three layers. Iterated conditional modes (ICM) is used to find the maximum a posterior(MAP) estimation of the noise free image and also the texture region field. Experimental results show that this method can effectively suppress additive noise and restore image details.

We use an image denoising algorithm in which we are given a noisy gray level image degraded by an additive white Gaussian noise of known variance, we have to estimate the noise-free image. Initially, we calculate the local variance of all the pixels considering a 3*3 window moving through the image. Cluster the local variance using k-means. Set S_i as the clustering result. Estimate the MRF parameters using the given equations and update Si and x_i (pixels of the noise-free image) using the ICM algorithm and repeat the process till MRF parameters become constant. The output obtained is the noise-free image.

1.7 Thesis Organization:

The thesis is organized as follows:

- In chapter 2, the concepts of Markov Random Field(MRF), Iterated Conditional Modes(ICM) is discussed.
- Chapter 3 explains the procedure of implementation of denoising using hierarchical MRF.
- Chapter 4 gives the results and its discussion.
- Chapter 5 concludes the thesis.

Chapter 2

2.1 Probabilistic models

For success in our approach, we need to find a class of stochastic models ie random fields that have the following characteristics:

1. The probabilistic dependencies between the elements of the field must be local. This condition is necessary because generally field that is used to model surfaces is only piecewise smooth; and if it is satisfied, the reconstruction algorithms are likely to be efficiently implementable in parallel hardware.
2. The class of models should be rich enough for a wide variety of qualitatively different behaviors to be modelled.
3. The relationship between the parameters of the models and the characteristics of the corresponding sample fields should be comparatively transparent so that the models are easy to specify.
4. It should be possible to represent the prior probability distribution clearly.
5. It should be possible to specify efficient Monte Carlo procedures.

A class of random fields that satisfies these requirements is the class of Markov Random Fields(MRF) on finite lattices.

2.2 Markov Random Field

A Markov random field is a probabilistic model defined by the local conditional probabilities. Consider the discrete 2-D random fields defined over a finite $M*N$ rectangular lattice of pixels. Consider a set of sites, $S = \{(i,j)|1 \leq i \leq M, 1 \leq j \leq N\}$

Firstly, a definition of a neighborhood system on lattice S and the associated cliques is given. A neighbourhood system is used to relate the pixels in S. A neighborhood system over S is defined as $N = \{N_t | for\ all\ t \in S\}$ where N_t is the neighboring sites of pixel t, i.e., a neighbourhood of pixel t. The neighborhood system has the following properties:

- a pixel is not neighboring to itself;
- the neighboring relationship is mutual.

In image modeling, a hierarchically ordered sequence of neighboring systems is most commonly used. In the first-order neighborhood system, every pixel has four neighbors. In the second order neighborhood system, there are eight neighbors for each pixel, as shown in Fig. 2.1b, where the 1's and 2's are its neighbors. A clique *c* of the pair (S,N) is defined as a subset of sites in S such that

- *c* consists of a single site, or
- for $t \neq \Gamma$, $t \in c$ and $\Gamma \in c$ implies that $\Gamma \in N_t$ and $t \in N_t$

The collections of single-, double- and triple-site are denoted by C_1, C_2 and C_3 respectively, where

$$C_1 = \{t | t \in S\}$$

$$C_2 = \{(t, \Gamma) | \Gamma \in N_t, t \in S$$

$$C_3 = \{(t, \Gamma, \omega) | t, \Gamma, \omega \in S \text{ are neighbors to one another}\}$$

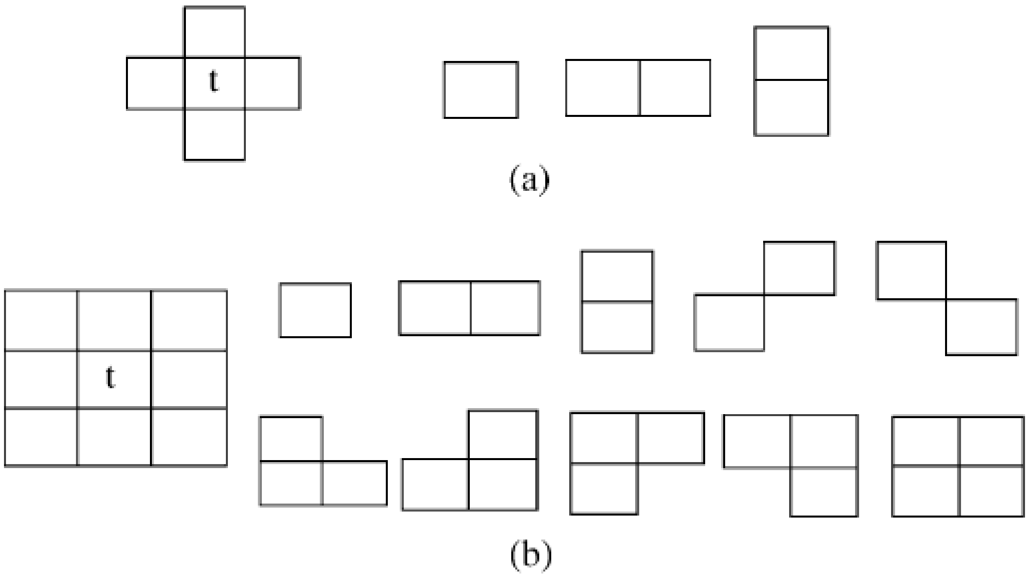


Fig 2.1 a) The first order neighbourhood system and its associated cliques b) The second order neighbourhood system and its associated cliques.

Let $X = \{X_{ij} | (i, j) \in S\}$ denote a family of random variables over the lattice S, in which each random variable X_{ij} takes a value x_{ij} in *G*. *G* denotes the set of all the possible gray levels for a pixel. Let Ω be the state space of the random field *X*, which is said to be a Markov random field on S with respect to a neighbourhood system *N* if and only if

$$P(X = x) > 0 \text{ , } for\ all\ \ x \in \Omega$$

$$P(X_{ij} = x_{ij} | X_{kl} = x_{kl}, (k, l) \in S - \{i, j\}) = P(X_{ij} = x_{ij} | X_{kl} = x_{kl}, (k, l) \in N_{ij}) \tag{2.1}$$

(2.2)

The Hammersley Clifford theorem states that a random field X is an MRF on lattice with respect to neighborhood system N if and only if X is a Gibbs random Field on S with respect to N. Therefore, the joint probability distribution of an MRF can be written as a Gibbs distribution

$$P(X = x) = \frac{1}{Z}\exp\big(-U(x)\big),\ \forall x \in \Omega$$

(2.3)

where

$$Z = \sum_{x\in\Omega} \exp\,(-U(x))$$

(2.4)

is a normalization constant, which is also called the partition function. *U(x)* is the energy function given as

$$U(x) = \sum_{c\in C} V_c(\,x_{ij}, (i,j) \in c)$$

(2.5)

where $C = C_1 \cup C_2 \cup C_3$ .... is the set of cliques and V_c is the clique potential associated with the clique c.

Thus, the joint probability $P(X = x)$ is determined by specifying the clique potential functions $V_c\,(x)$.

MRF was introduced into the image processing field in the mid-1980s and was widely used in low-level computer vision problems. Markov random field (MRF) image models are popular in application to image reconstruction problems like deconvolution, denoising, interpolation, segmentation, etc. MRF models are flexible in finding the expected solution constraints derived from available *a priori* information. This information which is expressed in the form of Gibbs priors, can be used in a Bayesian framework to derive *a posteriori* probability which accounts for both data consistency and the *a priori* constraints. The solution is usually calculated as the maximizer of this posterior probability (maximum *a posteriori* (MAP) estimate) or, as the minimizer of the associated posterior energy.

There is an analogy between images and statistical mechanical systems ie pixel gray levels and the presence and orientation of edges are seen as states of atoms or molecules in a lattice-like physical system. The assignment of energy function in a physical system determines its Gibbs distribution. Because of the Gibbs distribution-MRF equivalence, this assignment also gives us the MRF image model. The energy function is a more convenient and natural mechanism for representing the image attributes than the local characteristics of MRF.

The concept of MRF is one way of extending Markovian dependence from 1-D to a general setting. A natural way of incorporating spatial correlations into segmentation process is to use Markov random fields, as *a priori* models. The MRF is a stochastic process that specifies the local characteristics of an image and is combined with the given data to reconstruct the true image. The MRF of prior contextual information is a powerful method for modeling spatial continuity and other features. Even simple MRF modeling can give useful information for the process of segmentation. The MRF itself is a conditional probability model, where the probability of a voxel depends on its neighborhood. It is equivalent to a Gibbs joint probability distribution determined by an energy function. This energy function is a more convenient and natural mechanism for modeling contextual information than by using the local conditional probabilities of the MRF. The MRF on the other hand is the appropriate method to sample the probability distribution.

MRF theory gives a convenient and consistent way of modeling context-dependent entities such as image pixels and correlated features. This is achieved through characterizing mutual influences among such entities using conditional Markov random field distributions. The practical use of MRF models is largely ascribed to a theorem stating the equivalence between MRF's and Gibbs distributions that was established by Hammersley and Clifford and further developed by Besag. This is because the joint distribution is required in most applications but deriving the joint distribution from conditional distributions turns out to be very difficult for MRF's. The MRF-Gibbs equivalence theorem points out that the joint distribution of an MRF is a Gibbs distribution, the latter taking a simple form. From the computational perspective, the local property of MRF's leads to algorithms that can be implemented in a local and massively parallel manner.

MRF theory tells us how to model the a priori probability of context-dependent patterns, such as textures and object features. Maximum *a posteriori*(MAP) probability is one of the most popular statistical criteria for optimality and in fact has been the most popular choice in MRF vision modeling. MRF's and MAP criterion together give rise to the MAP-MRF framework.

2.3 Gibbs Distribution:

Gibbs models were first introduced into image modeling by Hassner and Sklansky. A Gibbs distribution relative to {S, G} is a probability measure π on Ω given as:

$$\pi(\omega) = \frac{1}{Z} e^{-\frac{U(\omega)}{T}}$$

(2.6)

where Z and T are constants and U is the energy function which is of the form

$$U(\omega) = \sum_{c \in \rho} V_c(\omega)$$

(2.7)

ρ denotes the set of cliques for G. Each V_c is a function on Ω with the behaviour that $V_c(\omega)$ depends only on those coordinates x_s of ω for which s ∈ C. Such a family { V_c, c ∈ ρ} is called a potential. Z is the normalizing constant:

$$Z = \sum_{\omega} e^{-\frac{U(\omega)}{T}}$$

(2.8)

and is called the partition function. Finally, T stands for temperature; for our purposes, T controls the degree of peaking in the density π. The V_c functions represent contributions to the total energy from external fields (singleton cliques), pair interactions(doubletons), and so on. Typically, several free parameters are involved in the specification of U, and Z is then a function of those parameters uncomputable.

The most general form of U is

$$U(\omega) = \sum V_{\{i,j\}}(x_{i,j}) + \sum V_{\{(i,j),(i+1,j)\}}(x_{i,j}, x_{i+1,j}) + \sum V_{\{(i,j),(i,j+1)\}}(x_{i,j}, x_{i,j+1})$$

(2.9)

where the sums extend over all (i, j) ∈ Z_m for which the indicated cliques make sense.

The MRF's are practical represented in terms of Gibbs distributions. The MAP estimate minimizes the posterior energy function**,** which is an energy function of a Gibbs distribution or maximizes the conditional probability distribution function.

The equilibrium states of large-scale discrete physical systems are described by the Gibbs distributions. There is an analogy to statistical physics. Because, for many physical systems the equilibrium states at very low temperatures have desirable properties, what is the state of the matter at these temperatures is the fundamental question. The physical systems are close to ground states (the lowest energy states) at those temperatures. A way to explore such states is trough lowering the temperature until a lowest energy state is reached. But just lowering the temperature is not enough; the system has to be kept in equilibrium too. The cooling process is thus very delicate. The chemical annealing is a method for obtaining low energy states of a material which consists of two steps: first the substance is melted at high temperature (so the equilibrium is reached fast), then the temperature is lowered gradually; enough time is spent at low temperatures for the system to reach equilibrium states.

2.4 Maximum a posterior estimation

In Bayesian statistics, maximum *a posteriori* probability(MAP) estimation is a mode of the posterior distribution. MAP can be used to obtain a point estimate of an unobserved quantity on the basis of empirical data.

MAP estimates are computed in several ways:

1. Conjugate priors are used when the mode(s) of the posterior distribution can be given in closed form.
2. Via a numerical optimization method such as the conjugate gradient method or newton's method. This usually requires first or second derivatives, which can be evaluated analytically or numerically.
3. Via the modification of an expectation-maximization algorithm. It does not require any derivatives of the posterior density.
4. Via monte-carlo method using simulated annealing.

With x_i denoting the category and value of pixel I in the image $x = (x_1, \dots., x_n)$, a Bayesian formulation specifies an a priori distribution p(x) over all available images. Usually, p(x) is taken to be a locally dependent markov random field(MRF), a convenient model for quantifying the belief that the unknown true image x* consists of, for example, large homogeneous patches, or smoothly varying grey levels which occasionally change levels discontinuously. With $y = (y_1, \dots, y_n)$ denoting the observed records of x*, the likelihood *l(y|x)* of any image x is combined with p(x), in accordance with Bayes's theorem to form an a posteriori distribution $p(x|y) \propto l(y|x)p(x)$. The map estimate of x* is that image $\hat{x}$ which maximizes p(x|y). direct calculation of $\hat{x}$ is generally computationally prohibitive and therefore, as an approximation, Geman and Geman have proposed the use of simulated annealing algorithm.

2.5 Optimization based approach

The main reason for the extensive use of optimization is the existence of uncertainties in every vision process. Noise and other degradation factors such as those caused by disturbances and quantization in sensing and signal processing, are sources of uncertainties.

2.5.1 Iterated Conditional Modes(ICM) Algorithm:

In statistics, iterated conditional modes is a deterministic algorithm for obtaining the configuration that maximizes the joint probability of a Markov Random field. It does this by iteratively maximizing the probability of each variable conditioned on the rest.

The Iterated Conditional Modes algorithm was proposed by Besag as a computationally feasible alternative in computing the maximum *a posteriori* probability (MAP) for the actual image given the observations. Indeed, it is known that MAP algorithms make enormous computational demands due to the inherent difficulty in computing the MAP estimate. Further, close related to Markov random fields (MRF), the ICM algorithm is not only computationally undemanding but also ignores the large-scale deficiencies of the *a priori* probability for the true image. It is an iterative procedure and it is easily shown that for each iteration, the MAP estimate never decreases and eventual convergence is assured.

The method is based on the equation for *a posteriori* probability of the value of the pixel *i*, given the observations g and current values of all pixels in the neighborhood of pixel *i*. In the above equation *S* \ *i* represents the set of all neighbours of the pixel *i* and *∂i* a small set of neighbors of the same pixel, defined by a neighborhood system. The usual in image analysis defines the first-order neighbors of a pixel as the four pixels sharing a side with the given pixel. Second-order neighbors are the four pixels sharing a corner.

Higher order neighbors are defined in an analogous manner. In this sense, $f_{S \setminus i}$ is the vector of all current values of the image excluding the pixel *i* and $f_{\partial i}$ is a vector of some neighbors of *fi*, following a neighborhood system. Although it is proposed inside a Bayesian framework, the ICM is a deterministic algorithm and it is given by

1. Choose a MRF model for the true values of *fi*;
2. Initialize f by choosing *fi* as the intensity, *fi* that maximizes $p(gi \mid fi)$ for each *i*;
3. For *i* from 0 to $M2 - 1$, update *fi* by the value of *fi* that maximizes $p(gi \mid fi) \cdot p(fi \mid \hat{}f\partial i)$
4. Repeat item (3) τ_{iter} times.

From the ICM algorithm, the MAP estimations of s_i and x_i can also be known.

Chapter 3

Image denoising based on hierarchical Markov random field model

3.1 Hierarchical MRF image model

An image is regarded as the combination of disjoint texture regions, and a three-layered hierarchical MRF is used to model the image. The first layer represents the region labels. The second layer represents the noise free image color. And the third layer represents the noisy color. Figure 3.1 shows our model, where i is a pixel site; j and k are two neighbors of i; s represents the texture region label; x is the noise free image; and y is the observed noisy image.

For a simple MRF, the parameters can be estimated by the least squares method. In our model, there are l regions, each of which has its own parameters. If the noise free image x and the region segmentation s are known, with the notation Xc denoting the set of pixels in region c and n_c denoting the number of pixels, parameters can be estimated by the same methods using the pixels in region c.

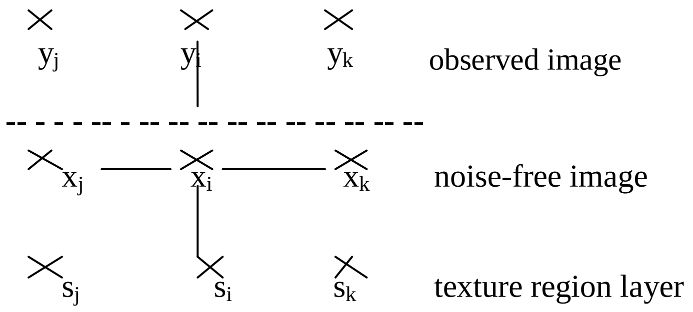


Fig 3.1: Hierarchical MRF Image Model

In the figure, y_i represents the region label, x_i represents the corresponding noise-free pixel and s_i represents the corresponding texture region.

The hierarchical MRF model is very often used in texture analysis using the MRFmodel. Until now, most hierarchical MRF model consists of one MRF model representing the image intensity and the other MRF model representing the region. The hierarchical MRF model can be briefly represented as following equation:

$$p\big(y_{(i,j)}, z_{(i,j)}\big|\theta\big) = p\big(y_{(i,j)}\big|z_{(i,j)}, \theta\big)p(z_{(i,j)}|\theta)$$

(3.1)

where $y_{(i,j)}$denotes the intensity value of the pixel (i, j) and MRF of which parameter is θ, and $z_{(i,j)}$ denotes the label representing the region of the pixel (i, j) which is also an MRF.

3.2 Procedure for image denoising using hierarchical Markov random field and an optimization algorithm(ICM):

Image Denoising is a basic process of image processing. It is one of the main processes of image restoration. Our main aim is to restore an image affected by noise. In our method, image restoration is done efficiently because we consider the texture information of the image while processing it. This helps in restoring the edges and other discontinuities of the image without any loss of information.

A noise-free synthetic gray-scale image is considered. Then Gaussian noise is applied to the image so that the image becomes noisy. The aim is to remove this noise from the image.

We regard an image as the combination of disjoint texture regions, and use a three-layered hierarchical MRF to model the image. The first layer represents the region labels. The second layer represents the noise free image color. And the third layer represents the noisy color. Fig 3.1 shows the model, where i is a pixel site; j and k are two neighbors of i's represents the texture region label, x is the noise free image, and y is the observed noisy image. Suppose the image has l texture regions, i.e., si ∈ L = {1, 2, . . . , l}. According to different types of conditional probability distribution, MRFs are classified into many subclasses such as LMRF and GMRF.

LMRF is often used to model image region labels. With the notation $\delta\big(x_i, x_j\big) = 0\ if\ \big(x_i = x_j\big); 1\ if\ \big(x_i \neq x_j\big)$; the conditional distribution of LMRF is

$$p\big(s_i\big|s_{N_i}, \psi\big) = \frac{1}{Z_i} \exp\big(-\sum_{j \in N} \beta_j \delta\big(s_i, s_{i+j}\big)\big)$$

(3.2)

where ψ denotes all MRF parameters, Z_i is the partition function and the parameter $\beta_j = \big||j|\big|$.

GMRF is widely used in texture classification and texture segmentation. It is supposed that the variables in the field have Gaussian conditional distributions. We suppose the noise free image is a GMRF, and the noise is additive white Gaussian noise with variance of σ^2 i.e.,

$$p\left(x_i \middle| x_{N_i}, s_i, \psi\right) = (2\pi\sigma_{s_i}^2)^{-1/2} \exp\left(-\frac{1}{2\sigma_{s_i}^2}\left(x_i - \mu_{s_i} - \theta_{s_i}^T q_i\right)^2\right) \tag{3.3}$$

$$p(y_i|x, s, \psi) = p(y_i|x_i) = (2\pi\sigma^2)^{-1/2} \exp\left(-\frac{1}{2\sigma^2}(y_i - x_i)^2\right) \tag{3.4}$$

Local variance is evaluated for every pixel in the image. The local variance is calculated by considering a (3*3) window and sliding it over the entire image. The local variance feature matrix is used to detect texture regions of the image so that this information can be used for image denoising. Then the local variance feature image is used for image clustering. K-means clustering is used for clustering the image. The image is divided into a given number of regions using k-means clustering. The clustering result is set as s_i. Since the local variance image matrix was used for k-means clustering, the texture information of the image is considered for splitting it into regions. Then the iterative process is started. The MRF parameters are estimated i.e. the mean, theta and variance of every region are estimated using the following formulae:

$$\widehat{\mu_C} = \frac{1}{n_c} \sum_{i \in \Omega_c} x_i \tag{3.5}$$

$$\hat{\theta_c} = \left(\sum_{i \in \Omega_c} q_i q_i^T\right)^{-1} \left(\sum_{i \in \Omega_c} q_i x_i\right) \tag{3.6}$$

$$\hat{\sigma_c}^2 = \frac{1}{{n_c}^2} \sum_{i \in \Omega_c} (x_i - \hat{\mu_c} - \hat{\theta_c}^T q_i)^2$$

(3.7)

Where Ω_c denoting the set of pixels in region c and nc = $|\Omega_c|$ denoting the number of pixels, noise free image x and q_i being the vector formed by the sum of symmetric neighboring variables, for instance, and $q_i = (x_{i+(1,0)} + x_{i-(1,0)} - 2\mu, x_{i+(0,1)} + x_{i-(0,1)} - 2\mu)^T$ in the first order neighbour system.

Now S_i(clustered image) and x_i(noise-free image) are updated using the ICM algorithm which optimizes the result. In most of the applications, the aim is to maximize the probability or the likelihood of a specialized MRF. Due to the complex dependency of the neighboring variables, these problems are very difficult to optimize. Various methods were proposed to solve the problem, but the global optimum is not tractable except in some simple cases. If an approximate solution is acceptable, there are many works focusing on more general MRFs. ICM uses a greedy scheme, updates sites one by one to achieve a local maximum.

If ψ is known, x and s can be calculated by maximizing a posteriori probability:

$$(x, s) = \arg max_{x,s}\, p(x, s|y, \psi)$$

(3.8)

After solving the above equation, the final result we get is

$$p(x, s|y, \psi) = p(y_i|x_i)p\left(x_i\middle|s_i, x_{N_i}, \psi\right)p(s_i|s_{N_i})\frac{p(x_{\Omega\setminus i}, s_{\Omega\setminus i}|y, \psi)}{p(y_i|x_{\Omega\setminus i}, s_{\Omega\setminus i}, y_{\Omega\setminus i}, \psi)}$$

(3.9)

Substituting LMRF and GMRF equations in the above equation, we get the MAP estimations of s_i and x_i.

ICM updates one pixel at a time assuming all other information is known, and estimates MRF parameters after every iteration according to the current data. ICM needs to initialize x, s and solve $(x_i, s_i) = argmax_{x_i,s_i} p(x, s|y, \psi)$ iteratively. Set s_i means to segment the image according to the texture. S_i is calculates as follows:

$$\hat{s_i} = max_{s_i \in L}(-\frac{1}{2\sigma_{s_i}^2}\left(x_i - \mu_{s_i} - \theta_{s_i}^T q_i\right)^2 - \log(\sigma_{s_i}) - \sum_{j \in N} \beta_j \delta(s_i, s_{i+j}) + C_1$$

(3.10)

After the iterative process, say around 5 iterations, we get the noise-free output image x.

$$\hat{x_i} = \frac{\sigma_{s_i}^2 y_i + \sigma^2(\mu_{s_i} + \theta_{s_i} q_i)}{\sigma_{s_i}^2 + \sigma^2}$$

(3.11)

3.3 Image Denoising Algorithm:

The algorithm for image denoising is as follows:

Given a noisy gray level image y degraded by an additive white Gaussian noise of variance σ^2, the noise free image is estimated as follows:

- choose the number of the regions l, iteration time T and an MRF neighbourhood system,
- x = y, calculate the local variance v_i for all x_i,
- cluster v_i with k-means, set s_i as the clustering result,
- for t = 1:T do,

  estimate MRF parameters.

  for $i \in \Omega$ do,

  update s_i,

  update x_i,

  end for,
- end for, the output is x.

3.4 Image Quality Indexes

PSNR:

The peak-signal-to-noise ratio (PSNR) is a most common measure of picture quality. Another more popularly used measure is Mean-Squared Error (MSE). The mean-squared error (MSE) between two images *g(n,m)* and *g^(n,m)* is as follows:

$$E_{MSE} = \frac{1}{MN} \sum_{n=1}^{N} \sum_{m=1}^{M} [\hat{g}(n,m) - g(n,m)]^2$$

(3.12)

The drawback with mean-squared error is that it depends strongly on the image intensity scaling. PSNR avoids this problem by scaling the MSE according to the image range given as:

$$PSNR = -10\log_{10}\frac{E_{MSE}}{S^2}$$

(3.13)

where S is maximum pixel value. PSNR is generally measured in decibels (dB). The advantages are that they are simple to calculate, have clear physical meanings, and are mathematically convenient in the context of optimization. But they are not very well matched for perceiving visual quality.

SSIM:

Structural SIMilarity (SSIM) index is a method for measuring the similarity between two images. The SSIM index can be defined as a quality measure of one of the images being compared provided the other image is considered as of perfect quality. The difference with respect to other techniques such as MSE or PSNR, is that these approaches estimate perceived errors on the other hand SSIM considers image degradation as perceived change in structural information. Structural information is the idea that the pixels have strong inter-dependencies especially when they are spatially close. These dependencies carry important information about the structure of the objects in the image. At a high level, SSIM attempts to measure the change in luminance, contrast, and structure in an image. Luminance is modeled as average pixel intensity, contrast by the variance between the reference and distorted image, and structure by the cross-correlation between the 2 images.

$$SSIM(x,y) = \frac{\left(2\mu_x\mu_y + C_1\right)(2\sigma_{xy} + C_2)}{\left(\mu_x^2 + \mu_y^2 + C_1\right)(\sigma_x^2 + \sigma_y^2 + C_2)}$$

(3.14)

where x, y are image patches, μ_x and μ_y are mean averages of x and y, $\sigma_x^2$and σ_y^2 are variances of x and y; σ_{xy} is covariance of x and y.

Universal Image Quality Index:

Universal objective image quality index is easy to calculate and applied to various image processing applications. Instead of considering the traditional error summation methods, this index is designed by modelling any image distortion as a combination of three factors: loss of correlation, luminance distortion, and contrast distortion. Though the new index is mathematically defined and no human visual system model is externally employed, it performs significantly better than the widely used distortion metric mean squared error. Let x and y be the original and test image signals. The value of Quality index is generally 0.3 in images affected by additive white Gaussian noise.

Let x = $\{x_i | i = 1,2, \dots \dots N\}$ and y = $\{y_i | i = 1,2, \dots \dots N\}$ be the original and the test image signals respectively. Then the universal image quality index is defined as

$$Q = \frac{4\sigma_{xy}\bar{x}\bar{y}}{(\sigma_x{}^2 + \sigma_y{}^2)[(\bar{x})^2 + (\bar{y})^2]}$$

Where

$\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i$, $\bar{y} = \frac{1}{N}\sum_{i=1}^{N} y_i$

$\sigma_x{}^2 = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})^2$, $\sigma_y{}^2 = \frac{1}{N-1}\sum_{i=1}^{N}(y_i - \bar{y})^2$

$\sigma_{xy} = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})(y_i - \bar{y})$.

Chapter 4 Simulation Results and Analysis

4.1 Simulation Results:

In simulation, synthetic textured images and real time grayscale images are considered to validate the proposed algorithm. Grayscale images consisting of 3 and 5 classes are considered for the simulations. One of the real-time images used for experimentation is taken from "http://www.flickr.com." The code is written in Matlab language and the execution of code is done using the Matlab software R2012a.

Firstly, a 3-class synthetic image is considered. Fig 4.1a shows the original noise-free grayscale image and then when Gaussian noise is added to it, noisy image is obtained as shown in fig 4.1b. K-means clustering is performed on the variance matrix of the noisy-image to obtain the clustered regions shown in fig 4.1c-e. Fig 4.1f shows the noise-free image after all iterations obtained after the application of the combination of hierarchical MRF image model and MAP estimation using ICM algorithm. The number of iterations to be performed to obtain an efficient noise-free image is till the values of the mean and variance are obtained constant for two consecutive iterations. Generally, the algorithm converges after five iterations. This noise free image is compared with the noise-free image obtained after the application of the ICM algorithm to MRF image model (shown in fig 4.1g). Fig 4.1f is more noise-free compared to fig. 4.1g because texture information is considered while calculating fig 4.1f and it is not considered while calculating fig. 4.1g. Thus, the hierarchical MRF model with ICM algorithm is more efficient than the MRF model because the texture information is considered in the previous case.

Fig 4.2 and fig 4.3 represents the same set of figures of a 3-class synthetic image but with difference noise variance values. And in all the cases it is proved that the MRF model with texture(hierarchical MRF) with ICM algorithm works more efficiently than an MRF model with ICM algorithm.

Fig 4.4 shows a 5-class synthetic image. It shows the original noise-free grayscale image and then when Gaussian noise is added to it, noisy image is obtained as shown in fig 4.4b. K-means clustering is performed on the variance matrix of the noisy-image to obtain the clustered regions shown in fig 4.4c-g. Fig 4.4h shows the noise-free image after every iteration obtained after the application of the combination of hierarchical MRF image model and MAP estimation using ICM algorithm. The number of iterations to be performed to obtain an efficient noise-free image is till the values of the mean and variance are obtained constant for two consecutive iterations. This noise free image is compared with the noise-free image obtained after the application of ICM algorithm to MRF image model (shown in fig 4.4i). Fig 4.4h is more noise-free compared to fig.4.4i because texture information is considered while calculating fig 4.4h and it is not considered while calculating fig. 4.4i. This shows that hierarchical MRF model with ICM algorithm is more efficient than the MRF model because the texture information is considered in the previous case. Fig 4.5 represents the same set of figures as fig 4.4 but with different noise level (noise variance = 0.001).

The same process is performed on real-time images as well. Fig 4.6-4.7 represent a 5 class real-time image and where the denoising results are obtained with different levels of noise. Fig 4.8 represents a 3 class real-time image on which the process of denoising is performed.

For a 3-class synthetic image:

i) Noisy-gray scale image is of variance 0.0005

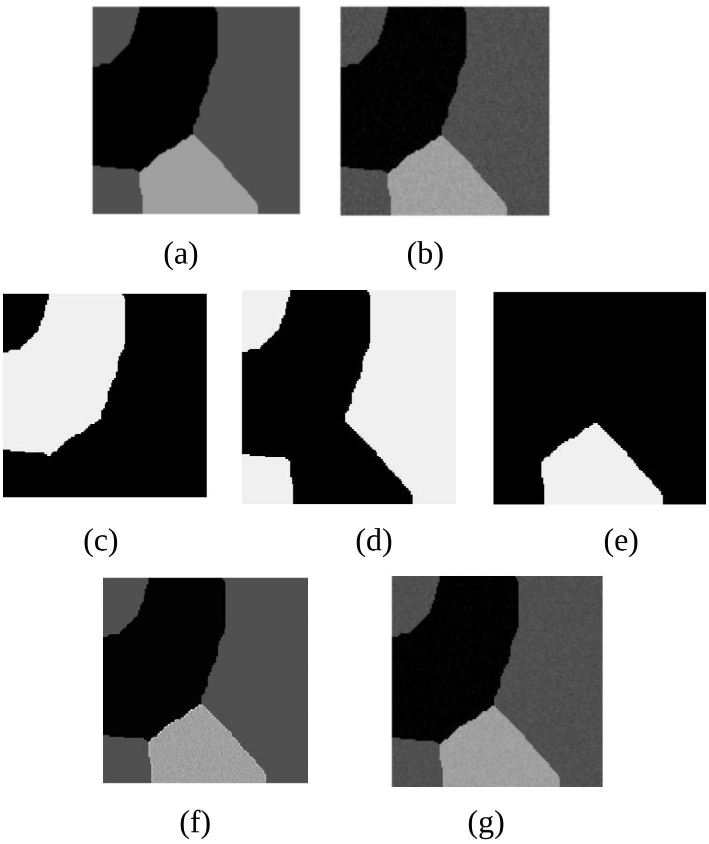
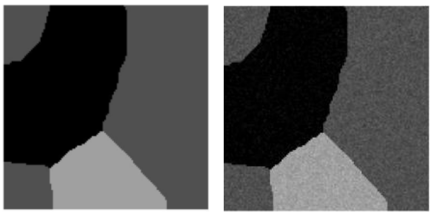


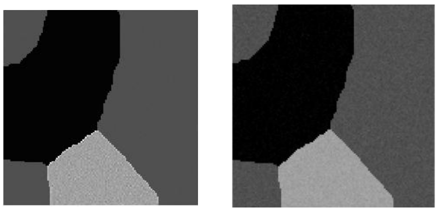
Fig 4.1 (a) original noise-free gray-scale image (b) noisy image with a noise variance of 0.0005 (c)-(e) outputs of k-means clustering when the number of regions is 3
(f) noise-free image (g) noisefree image when texture is not considered

ii)Noisy gray scale image is of variance 0.001



(a) (b)

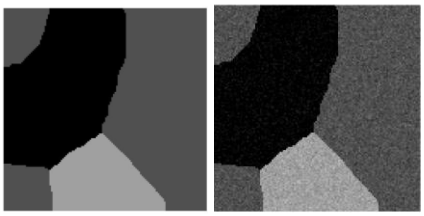
(c) (d) (e)



(f) (g)

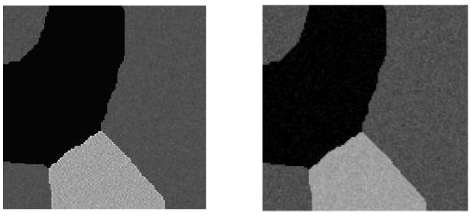
Fig 4.2 (a) original noise-free gray-scale image (b) noisy image with a noise variance of 0.001 (c)-(e) outputs of k-means clustering when the number of regions is 3
(f) noise-free image (g) noisefree image when texture is not considered

iii) Noisy gray scale image is of variance 0.002



(a) (b)

(c) (d) (e)

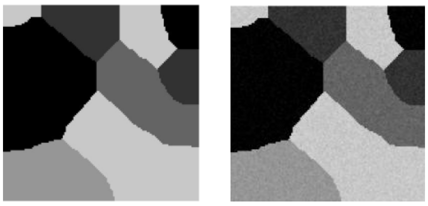


(f) (g)

Fig 4.3 (a) original noise-free gray-scale image (b) noisy image with a noise variance of 0.002 (c)-(e) outputs of k-means clustering when the number of regions is 3
(f) noise-free image (g) noisefree image when texture is not considered

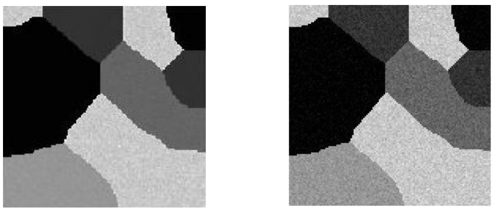
For a synthetic 5-class image:

i) Noisy image is of variance 0.0005



(a) (b)

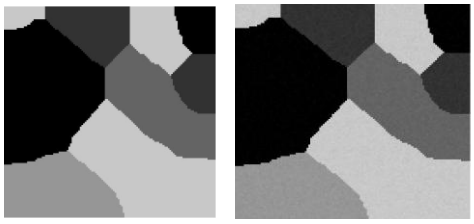
(c) (d) (e) (f) (g)



(h) (i)

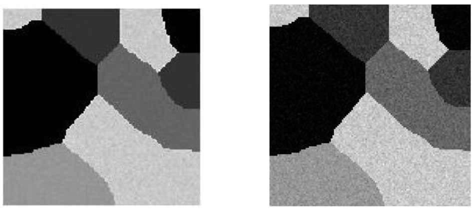
Fig 4.4 (a) original noise-free gray-scale image (b) noisy image with a noise variance of 0.0005 (c)-(g) outputs of k-means clustering when the number of regions is 5 (h) noise-free image (i) noisefree image when texture is not considered

ii) Noisy image is of variance 0.001



(a) (b)

(c) (d) (e) (f) (g)

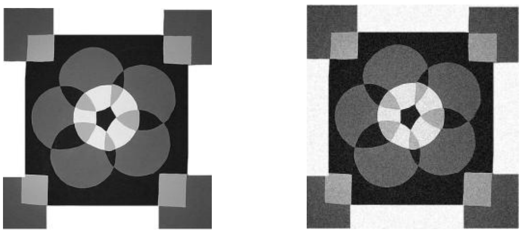


(h) (i)

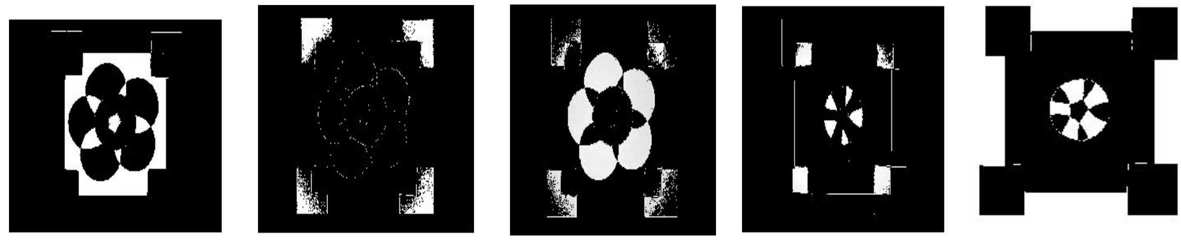
Fig 4.5 (a) original noise-free gray-scale image (b) noisy image with a noise variance of 0.001 (c)-(g) outputs of k-means clustering when the number of regions is 5 (h) noise-free image (i) noisefree image when texture is not considered

For a real-time 5-class image :

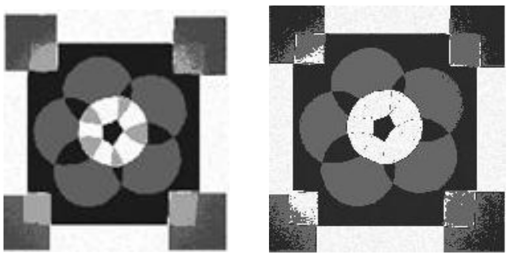
i) Noisy image is of variance 0.004



(a) (b)



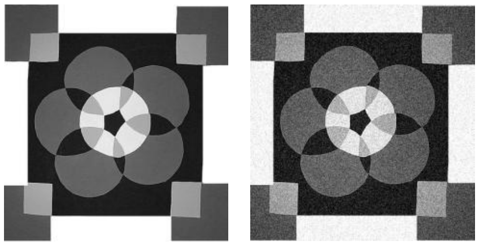
(c) (d) (e) (f) (g)



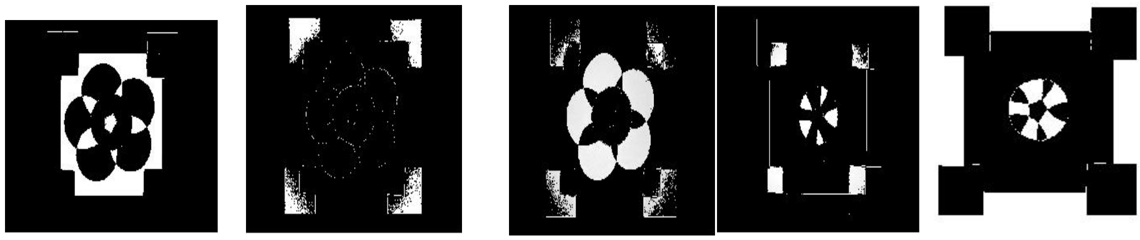
(h) (i)

Fig 4.6 (a) original noise-free gray-scale image b) noisy image with a noise variance of 0.004 (c)-(g) outputs of k-means clustering when the number of regions is 5 (h) noise-free image (i) noisefree image when texture is not considered

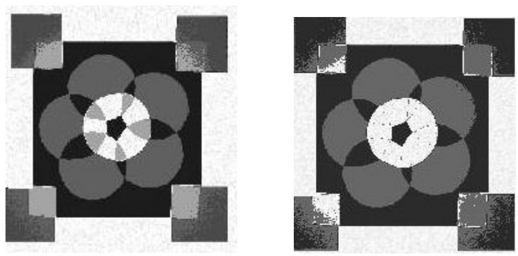
ii) Noisy-image is of variance 0.008



(a) (b)



(c) (d) (e) (f) (g)



(h) (i)

Fig 4.7 (a) original noise-free gray-scale image (b) noisy image with a noise variance of 0.008 (c)-(g) outputs of k-means clustering when the number of regions is 5 (h) noise-free image (i) noisefree image when texture is not considered

Real-time 3-class image:

i) Noisy image is of variance 0.002

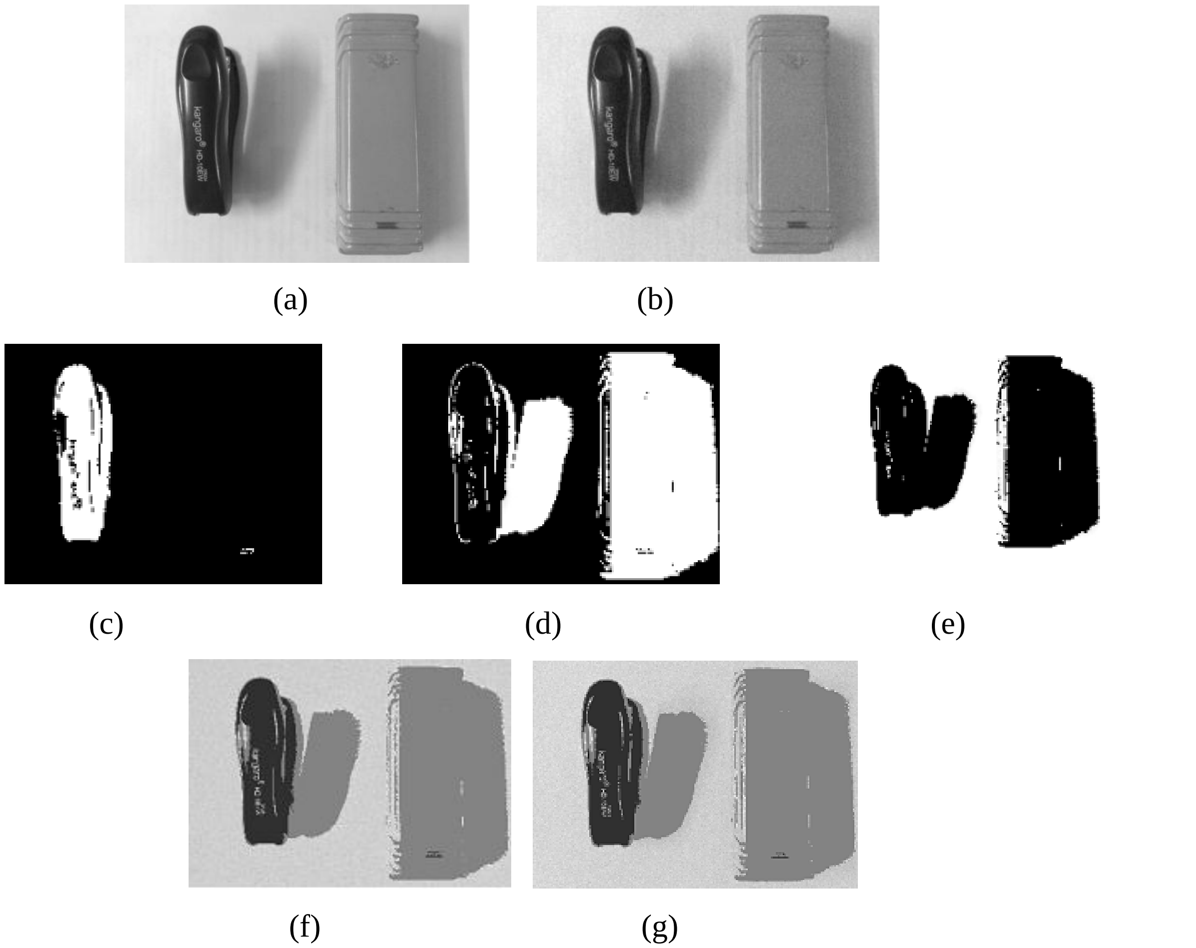


Fig 4.8 (a) original noise-free gray-scale image (b) noisy image with a noise variance of 0.002 (c)-(e) outputs of k-means clustering when the number of regions is 3 (f) noise-free image (g) noise-free image when texture is not considered

Noise variance	PSNR of noisy image in dB	PSNR of restored image in dB	SSIM	Universal image quality index
0.0005	15.58	15.86	0.7864	0.1536
0.001	15.43	15.58	0.7102	0.1521
0.002	15.06	15.40	0.5647	0.1488
0.002(without texture)	14.62	14.96	0.47	0.124

Table-4.1: PSNR Comparisons, SSIM and Universal image quality index calculations for 3-class synthetic image with different noise variance.

Noise variance	PSNR of noisy image in dB	PSNR of restored image in dB	SSIM	Universal image quality index
0.0005	15.417	15.63	0.7519	0.2609
0.001	15.18	15.35	0.6723	0.2580
0.002	14.86	15.39	0.5863	0.2533
0.002(without texture)	14.42	14.88	0.49	0.236

Table-4.2: PSNR Comparisons, SSIM and Universal image quality index calculations for a 5-class synthetic image with different noise variance.

Noise variance	PSNR of noisy image in dB	PSNR of restored image in dB	SSIM	Universal image quality index
0.0005	13.48	13.90	0.9038	0.1136
0.001	13.47	13.87	0.9033	0.1138
0.002	13.45	13.85	0.9026	0.1134
0.002(without texture)	13.15	13.39	0.87	0.1028

Table-4.3: PSNR Comparisons, SSIM and Universal image quality index calculations for a 5-class realtime image with different noise variances.

4.2 Analysis of PSNR and SSIM results:

For the 3-class synthetic image, PSNR of the image decreases as noise increases. From table 4-1, it is seen that the PSNR of the restored image is more than that of the noisy image. PSNR of the restored image is in the range of 20-30, implies that the signal has been restored successfully from the noisy image. SSIM is in the range of 0-1. More the SSIM, more is the restored image structurally similar to the reference image. From table 4-1, it is seen that as the noise increases, SSIM decreases. For noise variance of 0.0005, SSIM is 0.7864 which implied that the image is 78.6% restored structurally.

For the 5-class synthetic image, the PSNR is less compared to the 3-class synthetic image. This is because, as the number of regions increases, the denoising efficiency slightly decreases. From table 4-2, it is seen that PSNR of restored image is in the range 36-30 for noise variance 0.0005-0.004. This implies that the signal strength of the restored image is 36 times higher than the strength of the reference image. From table 4-2, it is seen that as the noise increases, SSIM decreases. For noise variance of 0.0005, SSIM is 0.7519 which implied that the image is 75.2% restored structurally.

4.3 Analysis of Universal Image Quality Index results:

From table 4-1 we can see that as the noise increases, the value of image quality index decreases. It is seen that for a 3-class synthetic image and real-time image(from table 4-1 and table 4-3), it is of the range 0.1 and for a 5-class synthetic image, it is of the range 0.2(from table 4-2). So, from experimental results, we can say that the quality of the restored image is much better than that of the noisy image.

Chapter 5

Conclusion and Future Work

A new hierarchical MRF model based image denoising method is implemented. The method we used employs a three-layered MRF model along with an optimization algorithm(ICM), which can express both smooth and texture signals. The advantage of the hierarchical MRF model is that the texture information of the image is considered while the process of denoising, so that the edge information and other interesting structures of the image are preserved and the image is restored efficiently.

An image is composed of disjoint texture regions. This requires to segment the image while denoising. In order to take care of texture regions in restoration, hierarchical MRF image model is employed. Hierarchical MRF has been widely used in the application of texture image segmentation. The most popular model consists of two layers. One represents the underlying region labels characterized by a logical MRF (LMRF). The other one represents the image texture. In order to solve image denoising problem, a third layer is added to model the observed noisy image. Finally, the noise free image is calculated by a MAP estimation.

The hierarchical MRF model is compared with the one-layered MRF model for analysis and the PSNR, SSIM, Universal Image Quality Index parameters are calculated. Though the computational complexity increases in the case of hierarchical MRF model, the results are more accurately obtained.

Along with the general information of the noisy image, if the texture information is also considered, the image can be restored in an efficient manner which is being done using hierarchical MRF models. And the optimization algorithm used is ICM algorithm which is a greedy algorithm used to calculate maxima/minima of a function.

The Markov random field model can be extended to conditional random field model and an optimization algorithm which performs better than the ICM algorithm for MAP estimation can be used for obtaining the denoised image in future. The work may be extended by considering the problem in unsupervised framework.

References:

[1] Geman, S., Geman, D., “Stochastic relaxation, Gibbs distribution, and the Bayesian restoration of images,” *IEEE Trans. Pattern Anal. Machine Intell.* Vol.6 (6), 721-741, 1984.

[2] Buades, A., Coll, B., Morel, J, “ A non-local algorithm for image denoising,” *IEEE CVPR*, vol. II, pp. 60–65, 2005.

[3] Deng, H., Clausi, D.A, “Gaussian MRF rotation-invariant features for image Classification,” *IEEE Trans. PAMI*, vol. 26 (7), pp. 951–955, 2004.

[4] Greig, D., Porteous, B., Seheult, A,“ Exact maximum a posteriori estimation for binary images,” *Journal of the Roy. Statist. Soc.*, vol. 51 (2), pp. 271–279, 1989.

[5] Jin, F., Fieguth, P., Winger, L, “ Image denoising using complex wavelets and Markov prior models,” *Image Anal. Recognition*, pp.73–80, 2005.

[6] Kashyap, R., Chellappa, R., “Estimation and choice of neighbors in spatial interaction models of images,” *IEEE Trans. Information. Theory*, vol. 29 (1), pp.60–72, 1983.

[7] Kim, D.H., Yun, I.D., Lee, S.U., “New MRF parameter estimation technique for texture image segmentation using hierarchical GMRF model based on random spatial interaction and mean field theory,” *ICPR*, vol. 2, pp. 365–368, 2006.

[8] Li, X., Zhu, S.,“ Image denoising based on wavelet domain spatial context Modelling,” *WCICA*, vol. 2, pp. 9504–9508, 2006.

[9] Lu, Q., Jiang, T., “Pixel-based image denoising with Markov random field,” *Pattern Recognition* vol. 34 (10), pp. 2029–2039, 2001.

[10] Maroquin, J., Milter, S., Pogio, T., “Probabilistic solution of illposed problems in computer vision,” *J. Ann. Statist. Assoc.* vol. 82, pp. 76-89, 1987.

[11] Noda, H., Shirazi, M.N., Zhang, B., Kawaguchi, E., “Mean field decomposition of a posteriori probability for MRF-based unsupervised textured image segmentation,” *ICASSP*, vol. 6, pp. 3477–3480, 1999.

[12] Szeliski, R., Zabih, R., Scharstein, D., Veksler, O., Kolmogorov, V., Agarwala, A., Tappen, M., Rother, C, “A comparative study of energy minimization methods for Markov random fields with smoothness-based priors,” *IEEE Trans.Pattern Anal. Machine Intell.* Vol. 30 (6), pp.1068–1080, 2008.

[13] Tonazzini, A., Bedini, L, “Monte carlo Markov china techniques for unsupervised MRF-based image denoising,” *Pattern Recognition Lett.* Vol. 24 (1–3), pp. 55–64, 2003.

[14] Yang, C, Yupin, L, Shiyuan, Y, “Image denoising based on hierarchical Markov random field,” *Pattern Recognition Lett.* Vol. 32 (1–3), pp. 368-374, 2011

[15] Zhang, J., “The mean fields theory in EM procedures for Markov random fields.” *IEEE Trans. Signal Process* vol. 40 (10), 1992.

[16] Leonid, I., Stanley, O., Fatemi, E., “Nonlinear total variation based noise removal algorithms,” *Physica D* vol.60, pp. 259-268, 1992.

[17] Luminita, A., Stanley, J., “Image denoising and decomposition with Total Variation Minimization and Oscillatory functions,” *Journal of Mathematical Imaging and Vision* vol. 20, pp. 7–18, 2004.