COLOR IMAGE SEGMENTATION USING MARKOV RANDOM FIELD MODELS

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

Master of Technology

In

Electronics System and Communication

By

RAHUL DEY

Department of Electrical Engineering

National Institute of Technology

Rourkela

2007

COLOR IMAGE SEGMENTATION USING MARKOV RANDOM FIELD MODELS

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

Master of Technology

In

Electronics System and Communication

By

RAHUL DEY

Under the Guidance of

Prof. P. K. NANDA

Department of Electrical Engineering

National Institute of Technology

Rourkela

2007

National Institute of Technology

Rourkela

CERTIFICATE

This is to certify that the thesis entitled, "**Color Image Segmentation Using Markov Random Field Models**" submitted by Sri **Rahul Dey** in partial fulfillment of the requirements for the award of Master of Technology Degree in **Electrical Engineering** with specialization in "**Electronics System and Communication**" at the National Institute of Technology, Rourkela (Deemed University) 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:

Dr. Pradipta Kumar Nanda
(Professor & Head)
Dept. of Electrical Engg.
National Institute of Technology
Rourkela - 769008

ACKNOWLEGMENT

I wish to express my deep and sincere gratitude towards my respectable guide and mentor ***Prof. P.K Nanda*** *for his consistent and enthusiastic support throughout my thesis work, without his suggestions and ideas, this thesis would not be an asset for me.*

I would like to thanks all the faculties and staff of the Department of Electrical Engineering for their valuable suggestions and help during my M.Tech carrier.

I am indebted to the research scholars Mr. Priyadarshi Kanungo and Mrs. Sucheta Panda for their help towards the understanding of my thesis problem and work.

I would like to thank M. Tech scholars Mr. Badri Narayan Subudhi and Mr. Saibal Dutta for their appreciation and valuable questions which increases my understanding towards the problem. I am also grateful to all my batch mates and friends for their support and help.

I would like to thanks my IPCV lab which has provided me a great environment not only to build but to explore and grow.

Last but by no means least, as always I would like to thanks all those who are directly or indirectly supported and advised me during my thesis work.

Rahul Dey

Roll No.20502001

Contents

List of Figures

List of Tables

Abstract

In this thesis, the problem of color image segmentation is address in stochastic framework. The problem is formulated as pixel labelling problem. The pixel labels are estimated using maximum a Posteriori(MAP) criterion. The observed image is viewed as the degraded version of the true labels. The degradation process is assumed to be Gaussian process. The image labels are modeled as Markov Random Field(MRF) model and the Ohta (I_1, I_2, I_3) model is used as the color model. A hybrid algorithm is proposed to obtain the MAP estimates. The performance of the algorithm is compared with that of Simulated Annealing (SA) algorithm and the proposed algorithm is found to outperform that of SA. A new constrained MRF model is used as the Image model and the segmented results obtained are found to be with less percentage of missclassification error than that of using MRF model. A new double MRF model is proposed to take care of the segmentation while using RGB color model. The double MRF model(DMRF) is proposed to take care of the interactions in different planes in the RGB color model. The DMRF model and constrained MRF outperformed the MRF model. Simulation results are presented in each of the above cases to substantiate our proposed model and algorithm.

Chapter 1

INTRODUCTION

Segmentation is an important process in automated image analysis. It is during segmentation that regions of interest are extracted from an image for subsequent processing such as surface description and object recognition. It is the low level operation concerned with partitioning images by determining disjoint and homogeneous regions, or, equivalently, by finding edges or boundaries. The homogeneous regions, or the edges are supposed to correspond to actual objects or parts of them within the images. Thus, in a large number of applications in image processing and computer vision, segmentation plays a fundamental role as the first step before applying to images for higher level operations such as recognition, semnatic interpretation and representation. Segmentation can be defined as follows:

Let I denote an image and H define a certain homogenity predicate, then the segmentation of I is a partition P of I into a set of N regions $R_n,\ n = 1, 2, ...N$ such that:

1. $\bigcup_{n=1}^{N}\ R_n = I$ with $R_n \cap R_m \neq 0, n \neq m$

2. $H(R_n) = TRUE \quad \forall n$

3. $H(R_n \cup R_m) = FALSE \quad \forall R_n$ and R_m adjacent[1].

Condition 1) states that partition has to cover the whole image; condition 2) states that each region has to be homogeneous with respect to predicate H; condition 3) states that no two adjacent region cannot be merged into a single region that satisfies the predicate H. Regions of image segmentation should be uniform and homogeneous with respect to some characteristics such as gray tone, texture or color. Region interiors should be simple and without many small holes. Adjacent regions of segmentation should have significantly different values with respect to the charactristic on which they are uniform. Boundaries of each segment should be simple, not ragged and must be spatially accurate[1].

Markov random fields posses several characteristics which make them useful in color image segmentation. Properties such as smoothness and continuity of color regions over an entire image can be enforced using only dependencies among local neighbours, discontinuties which separate regions of constant color may be computed while smooth regions are being found. In addition, the inclusion of both the prior and posterior distributions(through Baye's rule) establishes a relationship between noisy observed imagery and the color segmentation results. Details about the MRF model is provided in the susbsequent chapters.

For color image segmentation color models are having vital role, the purpose of a color model is to facilitate the specification of colors in some standard , generally accepted way. In essence, a color model is a specification of a 3-D coordinate system and a subspace within that systems where each color is represented by a single point.There are so many color models used for segmentation i.e RGB, HSI, CIE Lab, CIE Luv, Ohta etc.Each model is having

its own specifications which will be dealt in detail in coming chapter.In color image segmentation, the choice of a particular color model plays a vital role as segmentation quality very much depends on the color model being used. The transformation from one color model to another is equivalent to feature extraction in pattern recognition [2, 8].In this dissertation, Ohta and RGB model is used as the color model

Until very recently, attention has been focused on segmentation of gray level images since these have been the only kind of visual information that acquisition devices were able to take and computer resources to handle. Nowadays, color imagery has definitely supplanted monochromatic information and computational power is no longer a limitation in processing large volume of data. The attention has accordingly been focused in recent years on algorithms for segmentation of color images and various techniques, ofted borrowed from the background of gray level image. In this thesis, we are concerned with color image segmentation in statistical frame work which is becoming increasingly important in many applications such as content based retrival of information in images and videos, image compression, surface description, object recognition and many more.

The existing segmentation technique can be grouped into two categories: statistical and structural. The technique considered for color image segmentation falls into statistical category. In this technique, the images are modeled as realization of random fields and for segmentation, statistically optimal estimation techniques, such as Minimum Mean Square Error (MMSE), Maximum Likilihood (ML) and Maximum a Posteriori (MAP) estimation are used. Markov Random Field (MRF) and Gibbs Random Field (GRF) image models have been used to obtain MAP and other related optimal segmentations. For obtaining MAP and other related segmentations different

techniques have been used, namely: dynamic programming, stochastic relaxation/simulated annealing, split and merge and deterministic relaxation. The statistically segmentation technique can be divided into two categories.

1. Supervised segmentation

2. Unsupervised segmentation

In supervised segmentation technique, the model parameters are assumed to be known. But in case of unsupervised segmentation, the model parameters are unknown and need to be estimated. Hence, unsupervised segmentation is one of the classical problem in image processing [6][7].

As we know color provides more information about the Image than that of gray image, so recent research focuses on color image segmentation due to its growing need. Zolan Kato et al [9, 12] proposed a MRF image segmentation model which aims at combining color and texture features. Zoltan Kato et al [10] also proposed a unsupervised segmentation algorithm using MRF. They have proposed a new method to estimate initial mean vectors even if the histrogram does not have clearly distinguishable peaks. In[16] an unsupervised segmentation algorithm was proposed , which uses MRF models for color textures. These models characterize a texture in terms of spatial interaction between different color planes. The models are used by a segmentation algorithm based on agglomerative hierarchical clustering. Recently, Gaussian Mixture Model (GMM) have been proposed to model color textures on various feature spaces for image segmentation[14, 13]. In order to take into account the dependencies of the color planes as well as the interactions across the scales, [15] have proposed a wavelet domain hidden markov model (WDHMM) which has yielded a promising results for color texture. In this dissertation the problem of color image segmentation has been investi-

gatde . The problem is cast as a pixel labelling problem. The image model as well as color model plays a crucial role in image segmentation. We have used Ohta(I_1, I_2, I_3) color model and formulated the problem in stochastic framework. The observed image is considered as the degraded version of the true image labels. We have assumed the degradation to be gaussian process. The true labeled are modeled as Markov Random Field(MRF) model. The chapter 3 deals with the image label estimation using MAP estimation criterion. The MAP estimate of the image labels and hence segmentation is obtained using the proposed Hybrid algorithm. The performance of the Hybrid algorithm is compared with that of the SA algorithm. We have used to clique potential function namely weak memberane model and reward punishment model

The performance of both the models is compared with synthetic as well as real images A wide variety of images have been consider to validate the approach.

The accuracy of segmentation has been improved by using constrained MRF model. This model is found to have better grouping attribute than that of MRF model. This is covered in chapter 4. The problem is again formulated as a MAP estimation problem. Hybrid algorithm is uaed to obtain the MAP estimates. Hybrid algorithm with CMRF model is found to yield segmented images with less percentage of misclassification error that that of using MRF model. In the above case, Ohta(I_1, I_2, I_3) color model is used as the color model. It is observed that RGB color model does not yield better segmentation results . In this regard, a double MRF model is proposed to take care of the interaction among the individual R, G and B planes . A MRF model is used to model the labels of each plane and second

MRF model is used to model the inter plane interactions. This is called as double MRF model for intra and inter plane labels. Hybrid algorithm is used to obtain the MAP estimation. The DMRF model with RGB color model is yielded better segmentation than that of using a single MRF model. Thus scheme could be succesfully tested for real images.

Chapter 2

COLOR MODELS AND MARKOV RANDOM FIELD MODEL

The objects in a visual scene differ in brightness and also in the spectral characteristics of the light that emanates from them. This is true whether they generate light themselves as in astronomy and fluorescene microscopy, or if, they merely reflect a portion of the incident illumination. The human eye responds to their overall brightness and also the spectral content of the emanation. This perceptual phenomenon is referred to as "*color vision*"[3].

2.1 MONOCHROME IMAGING

When digitizing an image, one measures and quantizes the brightness at each pixel. The value obtained at each pixel depends upon the following factors:

1. how brightly the object is lighted
2. how reflective it is and
3. the sensitivity of the sensor that makes this measurement

The sensor has a spectral sensitivity characteristic that can be plotted as a spectral sensitivity curve $S(\lambda)$. In addition, the illumination has a particular *spectral power density* (SPD) function $I(\lambda)$, and the reflecting surface also has its own spectral characteistics. Thus, the gray level at the position x, y in the digitized image is given by

$$f(x,y) = \kappa \int I(\lambda) R(\lambda, x, y) S(\lambda) d\lambda \tag{2.1}$$

where λ is wavelength, $I(\lambda)$ is the SPD of the illumination source, $R(x, y, \lambda)$ is the reflectance spectrum of the object at the location that maps to (x, y) in the image, $S(\lambda)$ is the sensitivity spectrum of the sensor, amd κ is a callibration constant. The integration is taken over the speactral range in which both $I(\lambda)$ and $I(\lambda)$ are nonzero.

Hence, one could obtain very different images of the same scene by substituing a different light source or sensor. The salient characteristics of monochrome digital imaging is that only a single gray-level value is recorded at each pixel location. The spectral characteristics of the illumination and the sensor dramatically affect the way a given scene will be rendered in the digital image, but no record of these characteristics is kept in the digitized image. It is impossible to reconstruct the spectral characteristics of the object given only the digitized image $f(x, y)$, even if the illumination and sensor spectra are known. Many different $R(x, y\lambda)$ functions would yield the same gray level $f(x, y)$. Thus, monochrome imaging inherently loses the spectral information in a scene[3].

2.2 MULTISPECTRAL IMAGING

One can extend the process of Eq. (2.1) by measuring more than a single gray level at each pixel location. Then Eq. (2.1) becomes

$$f_i(x,y) = \kappa \int I(\lambda)R(\lambda,\ x,\ y)S_i(\lambda)d\lambda \tag{2.2}$$

where $1 \leq\ i\ \leq\ N$ is an index, N is the number of spectral channels used, and $S_i(\lambda)$ is the spectral sensitivity of the ith channel. Then, we obtain an N-element vector of gray levels at each pixel. This is equivalent to recording a series of monochrome images, each taken with a different sensor spectral characteristics. The different $S_i(\lambda)$ are commonly chosen to be bandpass filters spaced side by side throughout the range of intrest of the electromagnetic spectrum. The resulting digital image can be viewed as three dimensional, two spatial dimension and one spectral dimension. Such a digitizing scheme preserves, to some extent, the spectral characteristic of the scene[3].

2.2.1 Color Imaging

The human visual system employs the technique of Eq.(2.2), where $N = 3$, to implement the perception of the spectral characteristics of a scene. The light sensing cells on the retina of the human eye come in two ways, rods and cones. The rods are more sensitive and give us night vision, but they respond only to brightness and do not see color. The less sensitive cone cells come in three types that differ primarily in three spectral sensitivity. The cells that have peak absorption around 420nm sense what we know as blue light. Those that peak at about 530nm respond to green light, and those that peak at about 560nm see red. This is the red/green/blue ("RGB") color system. Red, green and blue are called the *primary colors* of human vision[3].

Tricolor Systems

Tricolor RGB imaging equipment is manufactured in huge quantities at relatively low cost to serve the imaging needs of humans. Thus, the availability of equipments alone dictates the one at least consider tricolor RGB for any image processing application that can not be performed with monochrome equipment. Three spectral channels are relatively convenient to accommodate in an imaging problem, although spectral information is thereby quantized rather crudely[3].

2.2.2 Color Image Processing and Analysis

There are two distinct purposes for which one might use digital color imaging. One is to produce processed color images for human consumption, the other is to analyze color images toward some quantitative goal. These are direct generalization of monochrome digital image processing and analysis, respectively[3].

2.2.3 Definitions

The field of color science has developed a set of strict definitions of the terms used to describe color, particularly in documents published by the organizations mentioned.

In a *color matching* experiment, a human observer adjusts the intensity of three primary colors (red, green and blue) until it is perceived that their mixture matches given color. The intensities of the three primaries, relative to those required to match a reference white, from the set of three `tristimulus` values that specify the color. A color matching experiment requires prior selection of three primary colors and a reference white that corresponds to the

chosen illumination source.

Brightness refers to how strongly a lighted or light emitting object stimulates the visual system. Scaled between bright and dim, it is the perception of how much light is being radiated. Related terms are intensity, luminance, luma, lightness, and value.

Intensity is the total emitted radiant power per unit solid angle. It can be obtained by integrating SPD of the source across the wavelength of the spectrum.

Luminance, as defined by the CIE, is the emitted radiant power, weighted by the spectral sensitivity function of human vision. This latter curve is called the luminous efficiency of the standard observer.

The human eye has a nonlinear response to light intensity. One must reduce the intensity of light source to about 18% for it to appear half as bright. *Lightness* models this behavior and measures brightness relative to a reference white. Usually denoted L^*, it is proportional to the cube root of the luminance ratio of the test source over the reference white.

Luma, usually denoted Y', is computed from RGB tristimulus values in a manner similar to luminance, except that non-linear values of R, G, and B are used. It is similar to lightness, except that the nonlinear transformation is done before averaging the RGB components.

Value is a scale of ten perceptually equal steps of brightness developed

by Munsell which will be discussed later in this chapter.

The *hue* of a color refers to the spectral color to which it is perceptually closest. Hue is an attribute associated with the dominant wavelength in a mixture of light waves. Hue represents dominant color as perceived by an observer. Thus, when we call an object red, orange, or yellow, we are specifying its hue.

Saturation refers to the relative purity or the amount of white light mixed with a hue. The pure spectrum colors are purely saturated. Colors such as pink and lavender are less saturated, with the degree of saturation being inversely proportional to the amount of white light added.

Hue and saturation together are called *chromaticity*, and, therefor, a color may be characterized by its brightness and chromaticity[3].

2.3 COLOR MODELS

Several color models or coordinate systems or color spaces have been developed for specifying color. A few of the more important color systems are decribed below.

2.3.1 RGB Color Spaces

The RGB system of color specification was based on the trichromatic theory of color vision and the color matching experiments carried out by Guild and Wright. It has been suggested by Maxwell in 1860 that if the specifications of the red, green and blue lamps used in such experiments were known, then the amounts of each light required to match a sample color would provide

a numerical specification of that color. These amounts may be found either by passing the light from the sample through filters matching the RGB primaries, or by calculation from the reflectance valus of the sample at a series of wavelengths through the visible spectrum[4].

$$R = \int_{\lambda} E(\lambda) S_R(\lambda) d\lambda \tag{2.3}$$

$$G = \int_{\lambda} E(\lambda) S_G(\lambda) d\lambda \tag{2.4}$$

$$B = \int_{\lambda} E(\lambda) S_B(\lambda) d\lambda \tag{2.5}$$

where S_R, S_G, S_B are the color filters on the incoming light , $E(\lambda)$ is the light radiance and λ is the wavelength.

The RGB color space can be geometrically represented in a 3-dimensional cube. The coordinates of each point inside the cube represent the vlaues of red, green and blue components, respectively. RGB color space is the most prevalent choice for computer graphics because color displays use red, green and blue to create the desired color. The choice of the RGB color space simplifies the architecture and design of the system[5].

But the RGB is not good for color image segmentation and analysis because of the high correlation among the R,G, and B components. By high correlation, we mean that if the intensity changes , all the three components will change accordingly. Also, the measurement of a color in RGB space does not represent color differences in a uniform scale, hence , it is impossible to evaluate the similarity of two colors from their distance in RGB space[2].

The $R_N G_N B_N$ and $R_S G_S B_S$ Color Spaces

The NTSC developed the $R_N G_N B_N$ colorimetry system using as primaries three phosphors that are typical of those used in commercial color television in 1953. The NTSC tristimulus values for a particular color can be obtained from the CIE values by the linear transformation

$$\begin{bmatrix} R_N \\ G_N \\ B_N \end{bmatrix} = \begin{bmatrix} 0.842 & 0.156 & 0.091 \\ -0.129 & 1.320 & -0.203 \\ 0.008 & -0.069 & 0.897 \end{bmatrix} \begin{bmatrix} R_c \\ G_c \\ B_c \end{bmatrix} \tag{2.6}$$

Here, R_c, G_c, B_c is the CIE developed color coordinate system based on the color matching experiment data obtained by Guild and Wright. The CIE 1931 system forms the basis for much of modern colorimetry.

Later the SMPTE developed the R_s, G_s, B_s system that better matches CRT phosphors in use today. These two RGB system are linearly related by[3][17]

$$\begin{bmatrix} R_s \\ G_s \\ B_s \end{bmatrix} = \begin{bmatrix} 1.609 & -0.447 & -0.104 \\ -0.058 & 0.977 & 0.051 \\ -0.025 & -0.037 & 1.162 \end{bmatrix} \begin{bmatrix} R_N \\ G_N \\ B_N \end{bmatrix} \tag{2.7}$$

2.3.2 Color Difference Systems

For color transmission to remain compatible with existing black-and- white receivers, it was necessary to establish color standards that

1. continue to transmit the luminance signal as before and
2. encode the color information so that it would not affect existing monochrome rceivers.

The RGB signals are combined to form the luminance signal, and two new signals are designed to carry the color information. These were color difference signal ($B-Y$ and $R-Y$) from which R, G and B could be recovered by suitable processing[3].

YUV and YIQ Coordinate Systems

The PAL and SECAM color television broadcast systems used in many countries transmit three color component signals called Y, U and V. Y is the luminance signal again formed by summing the R_N,G_N and B_N values. The chrominance (color difference) signals U and V that contain the color information are ignored by black-and-white television receivers. The YUV coordinates of a color can be obtained from its NTSC components by the linear relationship

$$\begin{bmatrix} Y \\ U \\ V \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.147 & -0.289 & -0.437 \\ 0.615 & -0.515 & -0.100 \end{bmatrix} \begin{bmatrix} R_N \\ G_N \\ B_N \end{bmatrix} \tag{2.8}$$

The NTSC color broadcasting system used in the United States employs three color component signals called Y, I and Q. Y is again the luminance signal, but I and Q are chrominance signals obtained by rotating the U and V signals by 33°. They are given by

$$I = -U\ sin(33^\circ)\ +\ V\ cos(33^\circ) \tag{2.9}$$

and

$$Q = U\ cos(33^\circ)\ +\ V\ sin(33^\circ) \tag{2.10}$$

The I and Q signals can be subjected to more bandwidth reduction than the U and V signals without visibly degrading the transmitted image. The

YIQ signals can be obtained directly from the NTSC components by the linear relationship[3]

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & -0.274 & -0.322 \\ 0.211 & -0.253 & 0.312 \end{bmatrix} \begin{bmatrix} R_N \\ G_N \\ B_N \end{bmatrix} \tag{2.11}$$

In YIQ or YUV color model, the Luminance (Y) is decoupled from the color information. Hence, it can partly get rid of the correlation of the red, green and blue components in an image. The linear transformation needs less computation time than nonlinear one, which makes the YIQ or YUV space more preferable to nonlinear transformation[2].

The XYZ Coordinate System

In the CIE spectral primary system, the tristimulus valus required to achive a color match are sometimes negative. The CIE has developed a standard artifical primary coordinate system in which all tristimulus valus required to match colors are positive[17]. It is designed so that the Y component is the luminance. The XYZ tristimulus values can be obtained from the NTSC tristimulus values by the linear relationship.

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.607 & 0.174 & 0.200 \\ 0.299 & 0.587 & 0.114 \\ 0.000 & 0.066 & 1.116 \end{bmatrix} \begin{bmatrix} R_N \\ G_N \\ B_N \end{bmatrix} \tag{2.12}$$

The XYZ valus can be normalized to sum to one to form the chromaticity coordinates

$$x = \frac{X}{X+Y+Z} \tag{2.13}$$

$$y = \frac{Y}{X+Y+Z} \; and \tag{2.14}$$

$$z = \frac{Z}{X + Y + Z} \tag{2.15}$$

Then, the chrominance of the color, independent of intensity, can be shown in an x,y chromaticity diagram. x, y, and z are known as *Trichromatic Coefficients*. The trichromatic coefficients reflect pure color and do not depend upon brightness but the percentage of red, green and blue components.

The distribution of color as a function x and y is known as *CIE Chromaticity* diagram. This diagram gives the amount of x and y to form a color $z = 1 - (x + y)$[5].

The CIE system for colorimetry has limitations. It was designed for color specification not for color apperance. The system was perceptually nonuniform. The perceptual color difference between two color with a fixed distance in XYZ space can vary by an order of magnitude[5].

2.3.3 Cube Root System

It is desirable that perceptually equal changes in color (hue, saturation, and intensity) should produce equal changes in color coordinates. Several color spaces that have been developed for this, use the cube root function, which approximates the sensitvity of the human eye[3].

The UVW and $U^*V^*W^*$ Coordinate Systems

In 1960, the CIE adopted a coordinate system, called *Uniform Chromaticity Scale*, in which, to a good approxiamtion, equal changes in the chromaticity coordinates result in equal, just noticeable changes in the perceived hue and saturation of a color. The V coordinate of the UCS system represents luminance. The u, v chromaticity coordinates are related to the x, y chromaticity

coordinates by the following relations[17].

$$u = \frac{4x}{-2x\ +\ 12y\ +3} \tag{2.16}$$

$$v = \frac{6y}{-2x\ +\ 12y\ +3} \tag{2.17}$$

$$x = \frac{3u}{2u\ -\ 8v\ -4} \tag{2.18}$$

$$x = \frac{2v}{2u\ -\ 8v\ -4} \tag{2.19}$$

The tristimulus values of the uniform chromaticity scale coordinate system UVW are related to the tristimulus values of the spectral coordinate primary system by the following relation[17].

$$\begin{bmatrix} U \\ V \\ W \end{bmatrix} = \begin{bmatrix} 0.327 & 0.206 & 0.133 \\ 0.177 & 0.812 & 0.011 \\ 0.020 & 1.068 & 0.411 \end{bmatrix} \begin{bmatrix} R_C \\ G_C \\ B_C \end{bmatrix} \tag{2.20}$$

The $U^*V^*W^*$ color coordinate system, adopted by the CIE in 1964, is an extension of the UVW coordinate system in an attempt to obtain a color solid for which unit shift in luminance and chrominance are uniformly perceptible. The $U^*V^*W^*$ coordinates are defined as follows[17].

$$U^* = 13W^*(u - u_0) \tag{2.21}$$

$$V^* = 13W^*(v - v_0) \tag{2.22}$$

$$W^* = 25(100Y)^{\frac{1}{3}}\ -\ 17 \tag{2.23}$$

where the luminance Y is measured over a scale of 0.0 to 1.0 and u_0 and v_0 are the chromaticity coordinates of the reference illuminant[17].

The UVW and $U^*V^*W^*$ coordinate system were rendered obsolete in 1976 by the introduction by the CIE of the more accurate $L^*a^*b^*$ and $L^*u^*v^*$ color coordinate systems. Although depricated by the CIE, much valuable data has been collected in the UVW and $U^*V^*W^*$ color systems[3][17].

$L^*a^*b^*$ Color Coordinate System

The $L^*a^*b^*$ cube root color coordinate system was developed to provide a computationally simple measure of color in agreement with Munsell color system[17]. The color coordinates are

$$L^* = \begin{cases} 903.3 * \left(\frac{Y}{Y_0}\right) & \text{if } 0 \leq Y < 0.008856 \\ 116 * \left(\frac{Y}{Y_0}\right)^{\frac{1}{3}} - 16 & \text{otherwise} \end{cases} \tag{2.24}$$

$$a^* = 500 * \left[\left(\frac{X}{X_0}\right)^{\frac{1}{3}} - \left(\frac{Y}{Y_0}\right)^{\frac{1}{3}}\right] \tag{2.25}$$

$$b^* = 500 * \left[\left(\frac{X}{X_0}\right)^{\frac{1}{3}} - \left(\frac{Z}{Z_0}\right)^{\frac{1}{3}}\right] \tag{2.26}$$

The term X_0, Y_0 Z_0 are the tristimulus values for the reference white. Basically, L^* is correlated with brightness, a^* with redness-greenness, and b^* with yellowness-blueness.

$L^*u^*v^*$ Color Coordinate System

The $L^*u^*v^*$ coordinate system, which has evolved from the $L^*a^*b^*$ and $U^*V^*W^*$ coordinate system, became a CIE standard in 1976[17]. It is defined as

$$L^* = \begin{cases} 903.3 * \left(\frac{Y}{Y_0}\right) & \text{if } 0 \leq Y < 0.008856 \\ 116 * \left(\frac{Y}{Y_0}\right)^{\frac{1}{3}} - 16 & \text{otherwise} \end{cases} \tag{2.27}$$

$$u^* = 13L(u^{'} - u_0) \tag{2.28}$$

$$v^* = 13L(v^{'} - v_0) \tag{2.29}$$

where,

$$u' = \frac{4X}{X + 15Y + 3Z} \tag{2.30}$$

$$v' = \frac{6Y}{X + 15Y + 3Z} \tag{2.31}$$

Here, Y_0, u_0, v_0 are the values for standard white light.

Each point on these two color spaces can be regarded as point in the $(L^*a^*b^*)$ or $(L^*u^*v^*)$ three dimensional color space, so that the difference of two color can be calculated as the Euclidean distance between two color points. The Euclidean distance between two colors L_1^*, a_1^*, b_1^* and L_2^*, a_2^*, b_2^* is given by

$$\triangle E_{ab}^* = [(L_1^* - L_2^*)^2 + (a_1^* - a_2^*)^2 + (b_1^* - b_2^*)^2]^{\frac{1}{2}} \tag{2.32}$$

The color difference is directly proprotional to the color difference as percieved by humans. The ability to express color difference of human perception by Euclidean distance is very important to color segmentation. Direct color comparison can be pereformed based on geometric separation within the color space. Hence these models can be used for measurement of small color difference. But other color space does not have this kind of property[4][2].

2.3.4 Cylindrical Color Spaces

Separating Luminance and Chrominance

The three types of color spaces mentioned before do not truly separate luminance from chrominance. In color image processing, it is often necessary, for example, to segment the image into regions that correspond to different

objects, such as forest and cropland in an aerial photograph. These regions differ in luminance, hue and saturation due to the different reflective properties of the vegetation. They will also differ in luminance due to variations in illumination, such as shadows from clouds. Thus luminance, like R, G, and B, is affected by nonuniform illumination, but hue and saturation, to a first approximation, are not. Thus, it is often useful to use a color space in color image processing wherin hue and saturation are truly independent of luminance[3].

Color Circle

Because the hue scale is periodic, artists have used the concept of a "color circle" for centuries. The fully saturated colors of the rainbow are spaced around the perimeter of the circle in the sequence red, yellow, green, cyan, blue, purple, and black to red. Note that the rainbow extends only from red through blue. Between blue and red fall the nonspectral colors that the eye percieves when red-and-blue sensitive cones , but not the green-sensitive cones are stimulated simultaneously. This is not realizable with monochromatic light, but it does occur in nature. The color circle is described in Fig 2.1.

The colors on the color circle become less saturated toward the centre, but still retain their hue. The color circle, which has long proved useful in illustrating similarities and differences among colors, forms the basis for cylindrical color spaces[3].

The Ostwald System

About 100 years ago, Wilhelm Ostwald developed a three dimensional generalization of the color circle. He placed a perpendicular gray axis through the centre of the gray circle to form a double cone solid. White fell at the

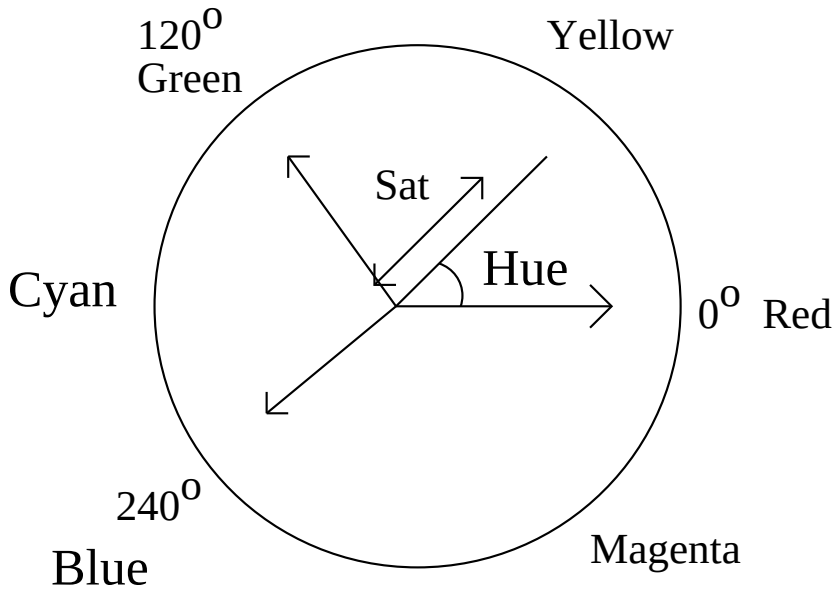


Figure 2.1: The Color Circle

upper apex, and black at the lower apex. The hues, numbered 1 through 24, are arranged around through the periphery. The fully saturated colors fall on the perimeter of the circle where the two cones intersect. Adding white to a paint color causes it to move toward the upper apex, and adding black moves it toward the lower apex[3].

The Munsell System

The color system developed by H.A. Munsell at about same time as Ostwald's is a cylindrical coordinate system, wherein ten hues are spaced around the color circle and nine values (intensities) cover from black to white along the vertical axis. Munsell established scales of chroma and for each hue, illustrated different chroma/value combinations with painted chips. The hue, value, and chroma scales were desinged to have perceptually equal steps[3].

Cylindrical Coordinates

Intensity. For quantitative purposes, the three color coordinates, hue, saturation, and intensity, define a cylidrical color space as shown in Fig 2.2. Intensity specifies the overall brightness of a color, without regard to its spectral wavelength, and it is the vertical axis of cylindrical color space. Different intensities fall along the vertical axis from black at the bottom to white at the top.

Hue. The two chrominance parameters, hue and saturation, are illus-

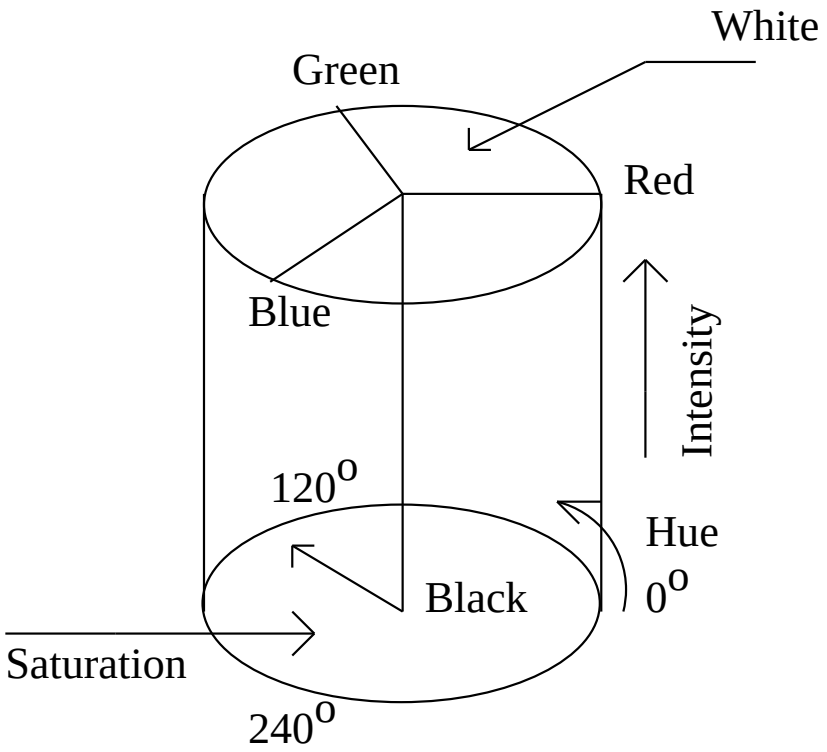


Figure 2.2: Cylindrical Color Space

trated by the color circle in Fig 2.1. Hue is an angle, and saturation is a vector length in the hue-saturation plane. Arbitrarily, a hue of 0^o is red, 120^o is green and 240^o is blue. Hue traverses the colors of the visible spectrum as it goes from zero to 240 degrees. The nonspectral colors fall between 240^o

and 360^o[3].

Saturation. The saturation parameter is the radius of the point from the origin of the color circle. The vivid (saturated) colors fall around the periphery of the cylinder and their saturation valus are unity. At the centre lie neutral shades, those with zero saturation. The fully bright, fully saturated colors fall on the perimeter of the circulat top surface[3].

The HSI Coordinate System

HSI (Hue-Saturation-Intensity) resembles the way in which human view and interpret color. The HSI model is an ideal tool for developing image processing algorithms , based on color desciptors. It separates the color information of an image from its intensity information. The hue, saturation and intensity can be obtained from RGB as follows[2].

$$H = cos^{-1}\frac{\frac{1}{2}[(R-G)+(R-B)]}{[(R-G)^2+(R-B)(G-B)]^{\frac{1}{2}}} \tag{2.33}$$

The saturation component is given by

$$S = 1 - \frac{min(R,G,B)}{I} \tag{2.34}$$

Finally, the intensity component is given by

$$I = \frac{R+G+B}{3} \tag{2.35}$$

It is assumed that the RGB values have been normalized to the range [0,1] and that angle H is measured with respect to the red axis of the HSI space. It is assumed that the RGB values have been normalized to the range [0,1] and that angle H is measured with respect to the red axis of the HSI space.

The major advantage in HSI space is that color information is decoupled from intensity information. One can use the intensity component of HSI to

apply the gray level algorithm. In order to segment objects with different colors, one may apply the segmentation algorithm to the Hue component only. The Hue space is efficient when the images have nonuniform illumination such as shades, since hue is independent on intensity vlaues.Hue information is useful in the case where the illumination level varies from point to point or image to image.

The main disadvantage of HSI model is due to its non-linearity. Hue component has a nonremovable singularity near the axis of the cylinder, where a slight change of input R,G and B values can cause large jump in the transformed values. The singularities may create discontinuities in the representation of color. In order to overcome this problem, pixels having low saturation are left unassigned to any region in many segmentation algorithm. Again, if the intensity of the color lies very close to white or black, then hue and saturation play little role in distinguishing colors[2].

The variants of HSI model are

1. HLS(Hue-Saturation-Lightness)
2. HSV(Hue-Saturation-Value)

The HLS Coordinate System

It is a variant of HSI color space. The three components of the color space are:

1. Hue
2. Lightness
3. Saturation

Here also Hue defines the color, saturation indicates the amount that Hue differs from a neutral gray and lightness indicates the level of illumination.

The equation in terms of RGB system are given below.

$$H = \begin{cases} \frac{60(G-B)}{max(R,G,B)-min(R,G,B)} & \text{if } R = max(R,G,B) \\ 120 + \frac{60(B-R)}{max(R,G,B)-min(R,G,B)} & \text{if } G = max(R,G,B) \\ 240 + \frac{60(R-G)}{max(R,G,B)-min(R,G,B)} & \text{otherwise} \end{cases} \tag{2.36}$$

$$L = \frac{max(R,G,B) + min(R,G,B)}{2} \tag{2.37}$$

$$S = \begin{cases} \frac{max(R,G,B)-min(R,G,B)}{max(R,G,B)+min(R,G,B)} & \text{if } 0 < L < 0.5 \\ \frac{max(R,G,B)-min(R,G,B)}{2*255-max(R,G,B)-min(R,G,B)} & \text{otherwise} \end{cases} \tag{2.38}$$

HSV Color Space

This color space is similar to HLS model. It produces color components that more closely match perception compared to HSI model. The three components of the model are Hue, Saturation, and Value. Brightness corresponds to the value. But the Hue component is not very reliable when saturation is

low. The equation in terms of RGB system are given below[3].

$$H = \begin{cases} \frac{60(G-B)}{max(R,G,B)-min(R,G,B)} & \text{if } R = max(R,G,B) \\ 120 + \frac{60(B-R)}{max(R,G,B)-min(R,G,B)} & \text{if } G = max(R,G,B) \\ 240 + \frac{60(R-G)}{max(R,G,B)-min(R,G,B)} & \text{otherwise} \end{cases} \tag{2.39}$$

$$S = \begin{cases} \frac{max(R,G,B)-min(R,G,B)}{max(R,G,B)+min(R,G,B)} & \text{if } max(R,G,B) \neq 0 \\ 0 & \text{if max(R,G,B)=0} \end{cases} \tag{2.40}$$

$$V = max(R,G,B) \tag{2.41}$$

2.3.5 Ohta Color Space

This color space also known as I_1, I_2, I_3 has been derived by performing systematic experiments of region segmentation to derive a set of effective color fatures. At each step of the recursive region splitting, new color features are calculated by Karhunen-Loeve transformation of R, G, and B. It applied eight kinds of color pictures, analyzed over 100 color feature, and found a set of effective color features as follows.

$$\begin{bmatrix} I_1 \\ I_2 \\ I_3 \end{bmatrix} = \begin{bmatrix} 0.333 & 0.333 & 0.333 \\ 0.500 & 0.000 & -0.500 \\ -0.250 & 0.500 & -0.250 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{2.42}$$

Comparing Ohta's color space with other spaces, it has been found that Ohta's color space is more effective in terms of the quality of segmentation and the computational complexity of the transformation[2].

2.4 MARKOV RANDOM FIELD

Random fluctuation in intensity, color, texture, object boundary, or shape can be seen in most real world images. The causes for these fluctuations are diverse and complex, and they are often due to factors such as non-uniform lighting, random fluctuations in object surface orientation and texture, complex scene geometry, and noise. Consequently, the processing of such images become a problem of statistical inference, which requires the definition of a statistical model corresponding to the image pixels.

Although simple image models can be obtained from image statistics such as the mean, variance, histogram and correlation function, a more general approach is to use random fields. Indeed, as a two dimensional extension of the one-dimensional random process, a random field provides a complete statistical characterization for given class of images. Combined with various frameworks for statistical inference, such as Maximum Likelihood (ML) and Bayesian estimation, random field models in recent years led to significant advances in many statistical image processing applications. A landmark paper by Geman and Geman in 1984 addressed Markov Random Field models and has attracted great attention and invigorated research in image modeling. Indeed the MRF, coupled with the Bayesian framework, has been the focus of many studies[7].

MRF theory provides a convenient and consistent way for modeling context dependent entities such as image pixels and correlated features. This is achived through characterizing mutual influences among such entities using conditional MRF distributions. The MRF theory tells us how to model the a priori probability of contextual dependent patterns, such as textures

and object features. A particular MRF model favors the class of patterns encoded by itself by associating them with larger probabilities than other pattern classes. MRF theory is often used in conjuction with statistical decision and estimation theories, so as to formulate objective functions in terms of established optimality principles. 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. MRFs and the MAP criterion together give rise to the MAP-MRF framework. This frame work, advocated by Geman and Geman and others, enables us to develop algorithms for a variety of vison problems systematically using rational principles rather than relying on ad hoc heuristics.

An objective function is completely specified by its form, i.e. the parametric family, and the involved parameters. In the MAP-MRF framework, the objective is the joint posterior probability of the MRF labels. Its form and parameters are determined according to the Bayes formula, by those of the joint prior distribution of the labels and the conditional probabilty of the observed data[6].

2.5 VISUAL LABELING

Many vision problems can be posed as labeling problems in which the solution to a problem is a set of labels assigned to image pixels or features.

2.5.1 Sites and Labels

A labeling is specified in terms of a set of sites and a set of labels. Let S index a discrete set of m sites.

$$S = \{1, 2, 3, ..., m\}$$

in which $1, 2, .., m$ are indices. A site often represents a point or a region in the Euclidean space such as an image pixel or an image feature such as a corner point, a line segment or a surface patch.A set of sites may be categorized in terms of their regularity. Sites on a lattice are considered as spatially regular. A rectangular lattice for a 2-D image of size $n \times n$ can be denoted by

$$S = \{(i, j)|1 \leq (i, j) \leq n\} \tag{2.43}$$

We normally treat the sites in MRF models as unordered.

A label is an event that may happen to a site. Let L be a set of labels.A label set may be categorised into two categories[6].

1. Continuous case

2. Discrete case

In the continuous case, a label set may correspond to the real line $\mathbf{R}$ or a compact interval of it

$$L_c = [X_l, X_h] \subset \mathbf{R} \tag{2.44}$$

In the discrete case, a label assumes a discrete value in a set of M labels.

$$L_d = \{l_1, l_2, l_3, ..., l_M\} \tag{2.45}$$

For example in case of edge detection, the label set is

$$L = \{edge, non - edge\} \tag{2.46}$$

2.6 THE LABELING PROBLEM

The labeling problem is to assign a label from the label set L to each of the sites in S. Edge-detection in an image, for example, is to assign a label z_i

from the set $L = \{edge, non-edge\}$ to site $i \in S$ where elements in S index the image pixels. The set

$$z = \{z_1, z_2, z_3, ..., z_m\} \tag{2.47}$$

is called labeling of the sites in S in terms of the labels in L. Whem each site is assigned a unique label $z_i = z(i)$ can be regarded as a function with domain S and image L. Because the support of the function is the whole domain S, it is a mapping from $S\, to\, L$ i.e.

$$z : S \longrightarrow L \tag{2.48}$$

In the terminology of random fields, a labeling is called a configuration.In vision, a configuration or labelling can correspond to an image, an edge map, an interpretation of image features in terms of object features, and so on[6].

2.7 LABELING WITH CONTEXTUAL CONSTRAINTS

Contextual constraint may be expressed locally in terms of conditional probabilities $p\{z_i|\{z_{i'}\}\}$, where $\{z_{i'}\}$ denotes the set of labels at the other sites $i^{'} \neq i$ or globally as the joint probability $p(z)$. Because local information is more directly observed, it is normal that a global inference is made based on local properties.

In situations where labels are independent of one another (no context), the joint probability is the product of the local ones.

$$p(z) = \prod_{i \in S} p(z_i) \tag{2.49}$$

The above equation implies conditional independence. Hence,

$$p(z_i/\{z_{i'}\}) = p(z_i) \qquad i' \neq i \tag{2.50}$$

Therefore, a global labelling z can be computed by considering each label z_i locally. This is advantageous for problem solving.

In the presence of context, labels are mutually dependent. The simple relationship expressed in (2.49) and (2.50) do not hold any more. Now the problem is to make a global inference using local infortmation. MRF theory provides a mathematical foundation for solving this problem[6].

2.8 MARKOV RANDOM FIELD AND GIBBS DISTRIBUTION

MRF theory is a branch of probability theory for analysing the spatial or contextual dependencies of physical phenomena. It is used in visual labeling to establish probabilistic distributions of interacting labels.

2.8.1 Neighborhood System and Cliques

The site in S are related to one another via a neighborhood system. A neighborhood system for S is defined as

$$N = \{N_i \mid \forall i \in S\} \tag{2.51}$$

where N_i is the set of sites neighboring i. The neighboring relationship has the following properties:

1. a site is not neighboring to itself: $i \notin N_i$

2. the neighboring relationship is mutual: $i \in N_{i'} \Leftrightarrow i' \in N_i$

For a regular lattice S, the set of neighbors of i is defined as the set of sites within a radius of $\sqrt{r}$ from i.

$$N_i = \{i' \in S \mid [dist((x_{i'}, y_{i'}), (x_i, y_i))]^2 \leq r, \ i' \neq i\} \tag{2.52}$$

where $dist(A, B)$ denotes the Euclidean distance between A and B and r takes an integer value. The Fig 2.3 shows the first order (η^1) and second order (η^2) neighborhood system.

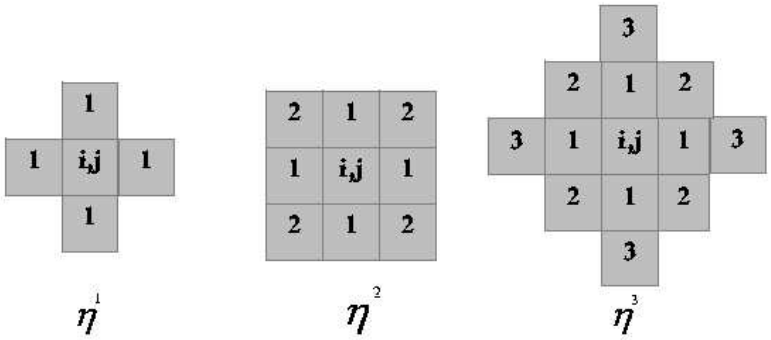


Figure 2.3: Figure showing first order (η^1), second order (η^2) and third order (η^3) neighborhood structure

The pair $(S, N) = G$ constitutes a graph in the usual sense; S contains the nodes and N determines the links between the nodes according to the neighboring relationship. A *clique* c for (S, N) is defined as a subset of sites $c = \{i, i'\}$, or a triple of neighboring sites $c = \{i, i', i"\}$, and so on. The collections of singl-site, pair-site and triple-site cliques will be denoted by C_1, C_2, C_3, respectively, where

$$C_1 = \{i \mid i \in S\} \tag{2.53}$$

$$C_2 = \{\{i, i'\} \mid i' \in N_i, \ i \in S\} \tag{2.54}$$

$$C_3 = \{\{i, i', i"\} \mid i, i', i'' \in S \ \textit{are neighbors to one another}\} \tag{2.55}$$

The sites in a clique are ordered, and $\{i, i'\}$ is not the same clique as $\{i', i\}$, and so on. The collection of all cliques for (S, N) is

$$C = C_1 \cup C_2 \cup \ C_3 \cup ..... \tag{2.56}$$

The type of a clique for (S, N) of a regular lattice is determined by its size, shape and orientation. Fig 2.4 shows the clique types for the first order and second order neighborhood systems for a lattice[6][7].

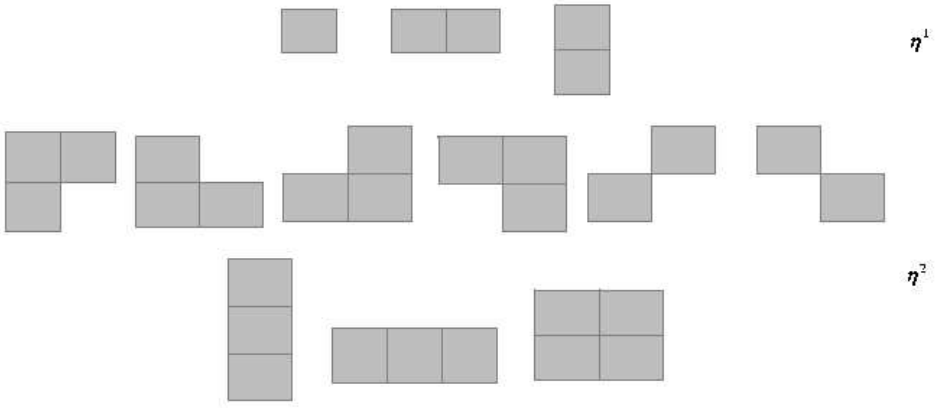


Figure 2.4: Cliques on a lattice of regular sites

2.8.2 Markov Random Field(MRF)

Let $Z = \{Z_1, Z_2, ..., Z_m\}$ be a family of random variables defined on the set S, in which each random variable Z_i takes a value z_i in L. The family Z is called a *random field*. We use the notion $Z_i = z_i$ to denote the event that Z_i takes the value z_i and the notion $(Z_1 = z_1, Z_2 = z_2, ...Z_m = z_m)$ to

denote the joint event. For simplicity a joint event is abbreviated as $Z = z$ where $z = \{z_1, z_2, ...\}$ is a configuration of z, corresponding to realization of a field. For a discrete label set L, the probability that random variable Z_i takes the value z_i is denoted $P(Z_i = z_i)$,abbreviated $P(z_i)$, and the joint probability is denoted as $P(Z = z) = P(Z_1 = z_1, Z_2 = z_2, ...Z_m = z_m)$ and abbreviated $P(z)$.

F is said to be a Markov random field on S with respect to a neighborhood system N if and only if the following two conditions are satisfied:

$$P(Z = z) > 0, \quad \forall z \in \mathbf{Z} \qquad (Positivity) \tag{2.57}$$

$$P(z_i|z_{S-i}) = P(z_i|z_{N_i}) \qquad (Markovianity) \tag{2.58}$$

where $S - i$ is the set difference, z_{S-i} denotes the set of labels at the sites in $S - i$ and

$$z_{N_i} = \{z_{i'}|i' \in N_i\} \tag{2.59}$$

stands for the set of labels at the sites neighboring i.

The positivity is assumed for some technical reasons and can usually be satisfied in practice. The Markovianity depicts the local characteristis of Z. In MRF, only neighboring labels have direct interactions with each other[6][7]. The concept of MRF is a generalization of that of Markov processes(MPs) which are widely used in sequence analysis. An MP is defined on a domain of time rather than space. It is a sequence of random variables $.....Z_1, ..., Z_m$ defined in the time indices $....1, ..., m, ...$ It is generalised into MRFs when the time indices are considered as spatial indices.

There are two approaches for specifying an MRF:

1. Conditional probability

2. Joint probability

According to Besag, the conditional approach has the following disadvantages:

1. No obvious method is available for deducing the joint probability from the associated conditional probability.

2. The conditional probability themselves are subject to some non-obvious and highly restrictive consistency conditions.

3. The natural specification of an equlibrium of statistical process is in terms of the joint probability rather than the conditional distribution of the variables.

A theoritical result about the equivalence between MRF and Gibbs distribution (Hammersley and Clifford Theorem) provides a mathemtical tractable means of specifying the joint probability of an MRF[6].

2.8.3 MRF models

In the 20's, mostly inspired by the Ising model, a new type of stochastic process appeared in the thory of probability called Markov Random Field. MRF's become rapidly a broadly ued tool in a variety of problems not only in stastical mechanics. Its use in image processing became popular with the famous paper of S.Geman and D.Geman in 1984 but its first use in the domain dates in the early 70's. Here we briefly give introduction to the theory of some MRF models.

Weak memberane model

The weak memberane model have been introduced in image restoration by A.Blake and A.Zisserman[19] . The problem is to reconstruct surfaces which

are continuous almost everywhere or, in other words , continuous in patches. To reach a satisfactory formalization of this principle, they have developed a memberane model: Imagine an elastic membrane which we are trying to fit to a surface, the edge will appear as tears in the memberane. Depending on how elastic is the memberane there may be more or less edges. The memberane is described by an energy function (the elastic energy of the memberane) which has to be minimized in order to find an equilibrium state. The energy has three components:

D: A measure of faithfulness to the data:

$$D = \int (u - d)^2 dA$$

where $u(x, y)$ represents the memberane and $d(x, y)$ represents the data.

S: A measure of how the function $u(x, y)$ is deformed

$$S = \lambda^2 \int (\nabla u)^2 dA$$

P: The sum of penalties α levied for each break in the memberane

$$p = \alpha Z$$

where Z is a measure of the set of contours along which $u(x, y)$ is discontinuous The elastic energy of the memberane is then given by

$$E = D + S + P = \int (u - d)^2 dA + \lambda^2 \int (\nabla u)^2 dA + \alpha Z$$

There is astrong relation between the *weak memberane* model and MRF models. An elastic system can also be considered from a probalistic view point. The link between the elastic energy and probability P is

$$P \propto exp(\frac{-E}{T})$$

that is the Gibbs distribution. however the *weak membrane model* operates with mechanical analogies, representing *a priori* knowledge from a mechanical point of view, while MRF modelization is purely probabilistic.

Reward Punishment(RP) model

The auto logistic model can be generalized to multi level logistic(MLL) model, also called Strauss process and generalized Ising model. There are $M(>2)$ discrete label set, $L = 1, 2, ..., M$. In this type of models, a clique potential depends on the type c (related to size, shape and possibly orientation) of the clique and local configuration $f_c \cong f_i | i \epsilon c$. For cliques containing more thanone site $(c > 1)$, the MLL clique potentials are defined by

$$V_c(f) = \begin{cases} +\alpha_c & \text{if all sites on c have the same label} \\ -\alpha_c & \text{otherwise} \end{cases} \tag{2.60}$$

where α_c is the potential for type-c cliques.

We have chosen a simple case of Ising model, In our case we have studied the behaviour of reward and punishment given by the model, depending on the homogeniety of the class. If the adjacent pixel is same as that of the center pixel then a reward is assingned to the energy function, otherwise punishment and the amount of reward and punishment is dependent on the homogeniety of the given image. So the clique potential of the model is given by:

$$V_c(z) = \begin{cases} +\delta_c & \text{if } |z_i - z_j| = 0 \\ -\delta_c & \text{if } |z_i - z_j| \neq 0 \end{cases} \tag{2.61}$$

where δ_c is selected on adhoc manner in our case.

2.8.4 Gibbs Random Field

A set of random variables Z is said to be a *Gibbs random field (GRF)* on S with rspect to N if and only if its configuration obey a *Gibbs distribution.* A

Gibbs distribution takes the following form.

$$P(Z = z) = \frac{1}{Z'} \times e^{-\frac{U(z)}{T}} \tag{2.62}$$

where

$$Z' = \sum_{z \in \mathbf{Z}} e^{-\frac{U(z)}{T}} \tag{2.63}$$

Z is a normalizing constant called the partition function, T is a constant called the temperature which shall be assumed to be 1 unless otherwise stated, and $U(Z)$ is the energy function. The energy

$$U(Z) = \sum_{c \in C} V_c(z) \tag{2.64}$$

is a sum of clique potentials $V_c(z)$ over all possible cliques C. The value of $V_c(z)$ depends on the local configuration on the clique C[7][6].
A GRF is said to be homogeneous if $V_c(z)$ is independent of the relative position of the clique c in S. It is said to be isotropic if V_c is independent of the orientation of c. It is considerably simpler to specify a GRF distribution if it is homogeneousor isotropic than one without such properties. The homogeneity is assumed in most MRF vision modes for mathematical and computational convenience. The isotropy is a property of direction-independent blob-like regions[6].

To calculate a Gibbs distribution, it is necessary to evaluate the partition function Z' which is the sum over all possible configurations in $\mathbf{Z}$. $P(Z = z)$ measures the probability of the occurance of a particular configuration, or pattern, z. The more probable configuration are those with lower energies. The temperature T controls the sharpness of the distribution. When the temperature is high, all configurations tend to be equally distributed. Near the zero temperature, the distribution concentrates around the global energy

minima.

For discrete labeling problem, a clique potential $V_c(z)$ can be specified by a number of parameters. For example, letting $z_c = (z_i, z_{i'}, z_{i''})$ be the local configuration on a triple clique $c = \{i, i', i''\}$, z_c takes finite number of states and therefore $V_c(z)$ takes a finite number of values. Sometimes, it may be convenient to express the energy of a Gibb's distribution as the sum of several terms, each ascribed to cliques of a certain size, that is,

$$U(z) = \sum_{\{i\} \in C_1} V_1(z_i) + \sum_{\{i,i'\} \in C_2} V_2(z_i, z_{i'}) + \sum_{\{i,i',i''\} \in C_3} V_3(z_i, z_{i'}, z_{i''}) \qquad (2.65)$$

The above implies a homogeneous Gibbs distribution because V_1, V_2, V_3 are independent of the locations of i, i', i". For nonhomogeneous Gibbs distributios, the clique functions should be written as $V_1(i, z_i)$, $V_2(i, i'', z_i)$, and so on[6].

2.8.5 Markov-Gibbs Equivalence

An MRF is characterised by its local property whereas a GRF is characterised by its global property. The Hammersley-Clifford theorem establishes the equivalence of these two types of properties. The theorem states that *Z is an MRF on S with respect to N if and only if Z is a GRF on S with respect to N.*

The practical value of the theorem is that it provides a simple way of specifying the joint probability. One can specify the joint probability $P(Z = z)$ by specifying the clique potential functions $V_c(z)$ and choosing appropriate potential functions for desired system behavior. How to choose the forms and parameters of the potential functions for proper encoding of the constraints is

a major issue in MRF modeling. The forms of the potential functions determine the forms of the Gibbs distribution. When all the parameters involved in the potential functions are specified, the Gibbs distribution is completely defined.

To calculate the joint probability of an MRF, which is a Gibbs distribution, it is necessary to evaluate the partition function (3.19). Because it is the sum over a combinatorial number of configurations, the computation is usually intractable. The explicit evaluation can be avoided in maximum probability based MRF vision models when $U(z)$ contains no unknown parameters. But this is not true when the parameter estimation is also a part of the problem. In the latter case, the enrgy function $U(z) = U(z/\theta)$ is also a function of parameters θ and so is the partition function $Z' = Z'(\theta)$. The evaluation of $Z'(\theta)$ is required. To circumvent the formidable difficulty therein, the joint probability is often approximated in practice[6][7].

2.9 LINE PROCESS

Smoothness is a generic assumption in MRF models which characterizes the spatial coherence and homogeneity of image lattice. However improper imposition of it can lead to undesirable, over-smoothed solutions. It is necessary to take care of discontinuities when using smoothness prior. To avoid the problem of over-smoothing Geman and Geman proposed the underlying MRF (label process) with an additional line process. The line process is neither a data nor the target of estimation. Rather, it is an auxiliary process which is coupled to the label process in such a manner that the joint probability distribution of intensity function is locally smooth with line process for discontinuities. The prior on the line process is often selected to emphasize

continuous line and to reject spurious edge elements. Such a model has the desirable property of promoting structure within the image without causing over-smoothing. A couple of MRFs are defined on the image lattice, one is for intensity or label field, other is the dual lattice for the edge field or "line field" . A line process comprises a lattice $S^{'}$ of random variable $f \in F$, whose sites $i^{'} \in S^{'}$ corresponded with vertical and horizontal boundaries between adjacent pixels of the image lattice. It takes the values from 0, 1 which signifies the absence or occurrence of edges. $z_{i'} = 1$ of the line process variable indicates that a discontinuity is detected between the neighbouring pixels j and i, i.e. $V_{i,j}(z_i, z_j)$ is taken same before.

Another neighborhood N is defined over the dual lattice $S^{'}$ for line sites. Each pixel has four line site neighbors. Image lattice can be represented as $S \cup S^{'}$. The (3.18) can be represented with the incorporation of the line fiels as

$$P(Z = z,\ F = f)\ =\ \frac{1}{Z'} e^{-\frac{U(z,f)}{T}} \tag{2.66}$$

The resulting MAP estimation can therefore defined using a Gibbs posterior distribution whose prior energy function is

$$U(z, f)\ =\ U(z|f)\ +\ U(f) \tag{2.67}$$

Assignment of line field is preferred as it results in smaller energy and better estimation[7][6].

2.10 BAYE'S LABELING OF MRFs

Bayes statistics is a theory of fundamental importance is estimation and decision making. According to this theory, when both the prior distribution

and the likelihood function of a pattern are known, the best that can be estimated from these sources of knowledge is the Bayes labeling. The maximum a posterior (MAP) solution, as a special case in Bayes framework, is sought in many vision works[6].

2.10.1 Bayes Estimation

In Bayes estimation, a risk is minimized to obtain the optimal estimate. The Bayes risk of estimate z^* is defined as

$$R(z^*) = \int_{z \in Z} C(z^*, z) P(z, x) dz \tag{2.68}$$

where x is the observation,$C(z^*, z)$ is a cost function and $P(Z = z|X = x)$ is the posterior distribution. Hence, we need to compute the posterior distribution from the prior and the likelihood. According to the Bayes rule, the posterior probability can be computed by using the following formulation

$$P(Z = z|X = x) = \frac{P(X = x|Z = z)P(Z = z)}{P(X = x)} \tag{2.69}$$

Where $P(Z = z)$ is the prior probability of labeling z, $P(X = x|Z = z)$ is the conditional probability density function of the observation x, also called the likelihood function of z for x fixed, and $P(X = x)$ is the density of x which is a constant when x is given.

The cost function $C(z^*, z)$ determines the cost of estimate z when the truth is z^*. It is defined according to the preference. Two popular choice are the quadratic cost function

$$C(z^*, z) = \| z^* - z \|^2 \tag{2.70}$$

where $\| a - b \|$ is a distance between a and b, and the $\delta(0-1)$ cost function

$$C(z^*, z) = \begin{cases} 0 & \text{if } \| z^* - z \| \leq \delta \\ 1 & \text{otherwise} \end{cases} \tag{2.71}$$

where $\delta > 0$, is any small constant.

The Bayes risk under the quadratic cost function measures the variance of

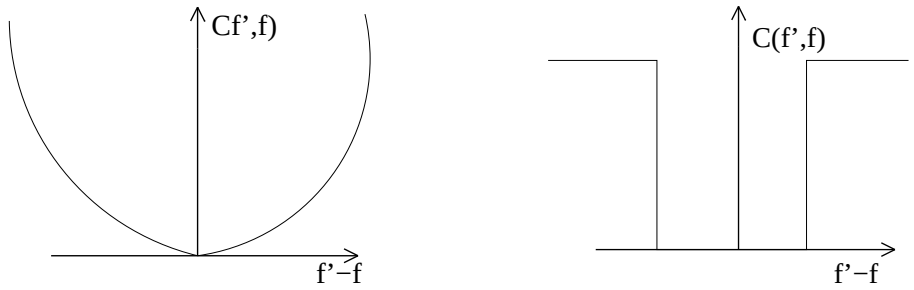


Figure 2.5: Two choices of cost function

the estimate

$$R(z^*) = \int_{z \in Z} \| z^* - z \|^2 \, P(Z = z | X = x) dz \tag{2.72}$$

Letting $\partial R(z^*)/\partial z^* = 0$, we obtain the minimal variance estimate

$$z^* = \int_{z \in Z} z \, P(Z = z | X = x) dz \tag{2.73}$$

The above is the mean of the posterior probability.

For the δ cost function, the Bayes risk is

$$R(Z^*) = \int_{z:\|z^*-z\|>\delta} P(Z = z | X = x) dz = 1 - \int_{z:\|z^*-z\|\leq\delta} P(Z = z | X = x) dz \tag{2.74}$$

When $\delta \to 0$, the above is approximated by

$$R(z^*) = 1 - \kappa P(Z = z | X = x) \tag{2.75}$$

where κ is the volume of the space containing all points z for which $\| z^* - z \| \leq \delta$. Minimizing the above is equivalent to maximizing the posterior probability. Therefore, the minimal risk estimate is

$$z^* = arg\, max_{z \in Z} P(Z = z | X = x) \tag{2.76}$$

which is known as the Maximum a posterior(MAP) estimate[6][18].

Chapter 3

SEGMENTATION USING MAP FRAMEWORK

Color image segmentation is formulated as a pixel labeling problem and the true labels are modeled as the MRF model. The observed color image is assumed to be the degraded version of the true labels. The label estimates are obtained by using Bayesian framework and MAP criterion. The MAP estimate is obtained by Simulated Annealing (SA) and as well as by the proposed hybrid algorithm and compared to SA.

Here we define all the images to be defined on a discrete rectangular lattice $M = N \times N$. Let Z denote the label process associated with the given noise free image X and z be a realization of Z. The label process Z of the image is modeled as MRF. The observed image X can be considered as a transformed and degraded version of an MRF realization z. The transformation may include geometric transformation and blurring and the degradation is due to random factors such as noise. These determine the conditional distribution $p(X = x|Z = z)$, or the likelihood of z[7][6].

Let H denote the blurring matrix corresponding to a shift invariant point-

spread function. The formation of Z give rise to a blurred image $H(Z)$ which is recorded by a sensor. The later often involves a non-linear transformation of $H(Z)$ denoted here by φ, in addition to random sensor noise $W = \{w_{i,j}\}$, discrete or continuous. The degraded image is then a function of $\varphi(H(Z))$ and W, say $\psi(\varphi(H(Z)), W)$, for example , addition or multiplication. For notational case, one can write

$$X = \varphi(H(Z)) \odot W \tag{3.1}$$

In practice, a simple observation model of no blurring, linear transformation and independent additive Gaussian noise is often assumed. Each observed pixel is assumed to be the sum of the true pixel value and independent Gaussian noise. Hence the image model considered is

$$X_{i,j} = Z_{i,j} + W_{i,j} \quad \forall (i,j) \in (N \times N) \tag{3.2}$$

Which with a lexicographical ordering (4.2) will be $X = Z + W$. The following assumptions are made:

1. $W_{i,j}$ is a white Gaussian sequence with zero mean and variance σ^2.

2. $W_{i,j}$ is statistically independent of $Z_{k,l}$, for all (i,j) and (k,l) belonging to $N \times N$.

3. $z_{i,j}$ takes any value from the label set $M = (1, 2, ..., M_m)$, (typically $M_m = 256$). In general the unknown parameter vector is $\Theta = [\phi^T, \sigma^2]$.

The noise free label process Z is assumed to be a Markov random field with respect to a neighborhood system η, and is described in terms of its local characteristics.

$$\begin{aligned} P(Z_{i,j} = z_{i,j} | Z_{k,l} = z_{k,l}, \quad k,l \in N \times N, (k,l) \neq (i,j)) \\ = P(Z_{i,j} = z_{i,j} | Z_{k,l} = z_{k,l}, (k,l) \in \eta) \end{aligned} \tag{3.3}$$

Here Z is an MRF or equivalently Gibb's distributed which is considered as *a priori* distribution. This is expressed as $P(Z = z|\phi) = \frac{1}{Z'} e^{-U(z,\phi)}$ where $Z' = \sum_z e^{-U(z,\phi)}$ is the partition function, ϕ represents the clique parameter vector, and the exponent term $U(Z, \phi)$ is called the energy function and is of the form $U(z, \phi) = \sum_{C:(i,j)\in C} V_c(z, \phi)$, where $V_c(z, \phi)$ being referred to as the potential.

3.1 IMAGE LABEL ESTIMATION

Since in supervised framework the model parameters are assumed to be known, it is required to estimate the pixel labels using the model parameters. Let X be the observed color image which is considered as a realization of a random field. The problem is to find the true label of X, hidden in it. The segmentation problem can be formulated as finding the true label of the given observed image X.

Let us suppose that the observed image $X = \{\mathbf{x_s}|s \in S\}$ consists of three spectral component values at each pixel $s \in S$ denoted by the vector $\mathbf{x_s}$. Let Z be the segmented image which represents the true label process, which is assumed to be an MRF. Let $z \in Z$ is a realization of the label process and $x \in X$ is a realization of the observed process. The labeling $\hat{z}$ is obtained by maximizing the posteriori probability $P(Z = z|X = x, \theta)$[7][6]. Hence, the optimality criterion is

$$\hat{z} = \underset{z}{arg\,max}\ P(Z = z|X = x, \theta) \tag{3.4}$$

Since z is unknown the above can not be computed. So, by using Baye's

theorem, (3.4) can be expressed as

$$\hat{z} = \underset{z}{arg\,max}\ \frac{P(X=x|Z=z,\theta)P(Z=z)}{P(X=x|\theta)} \tag{3.5}$$

Since X corresponds to the given image $P(X=x|\theta)$ is a constant quantity and $P(Z=z)$ is the a priori probability of the labels. According to Hammersley-Clifford theorem

$$P(Z=z) = \frac{1}{Z'}e^{-\frac{U(z)}{T}} \tag{3.6}$$

where, Z' = normalizing constant or partition function and T = Temperature which controls the sharpness of the distribution. Usually T is taken as 1. The only unknown remaining in (3.5) is the $P(X=x|Z=z,\theta)$. Here we assume that the observed image can be considered as a transformed and degraded version of an MRF realization. Using the degradation image modle of (3.2), $P(X=x|Z=z,\theta)$ can be written as

$$\begin{aligned} P(X=x|Z=z,\theta) &= P(X=z+w|Z,\theta) \\ &= P(W=x-z|Z,\theta) \end{aligned} \tag{3.7}$$

Since W is a Gaussian process and using the assumption stated earlier we obtain

$$P(W=x-z|Z,\theta) = \frac{1}{\sqrt{(2\pi)^n det[\bar{\mathbf{K}}]}}e^{-\frac{1}{2}\mathbf{(x-z)^T\bar{K}^{-1}(x-z)}} \tag{3.8}$$

where $\bar{\mathbf{K}}$ is the covariance matrix. The covariance matrix is defined as follows.

$$\bar{K} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \ldots & \sigma_{1n} \\ \sigma_{21} & \sigma_{22} & \ldots & \sigma_{2n} \\ . & . & . & . \\ . & . & . & . \\ \sigma_{n1} & \sigma_{n2} & \ldots & \sigma_{nn} \end{bmatrix} \tag{3.9}$$

The covarinace matrix describes the second order relationship between the components of the random vectors X. The components are said to be uncorrelated when

$$\sigma_{ij} = 0,\ i \neq j \tag{3.10}$$

As there are three spectral components present in a color image and assuming that the components are uncorrelated as in Ohta color space, the covariance matrix reduces to a diagonal matrix as mentioned below.

$$\bar{K} = \begin{bmatrix} \sigma_1^2 & 0 & 0 \\ 0 & \sigma_2^2 & 0 \\ 0 & 0 & \sigma_3^2 \end{bmatrix} \tag{3.11}$$

If we assume that, the variance is same, then the above equation reduces to

$$\bar{K} = \begin{bmatrix} \sigma^2 & 0 & 0 \\ 0 & \sigma^2 & 0 \\ 0 & 0 & \sigma^2 \end{bmatrix} \tag{3.12}$$

Hence (4.8) reduces to

$$P(W = x - z|Z, \theta) = \frac{1}{\sqrt{(2\pi)^3}\sigma^3} e^{-\frac{1}{2\sigma^2}(\mathbf{x}-\mathbf{z})^2} \tag{3.13}$$

Substituting (3.13) and (3.6) in (3.5) we have

$$\hat{z} = \underset{z}{arg\,max}\ \frac{1}{Z'\sqrt{(2\pi)^3}\sigma^3} e^{-[\frac{(x-z)^2}{2\sigma^2} + U(z)]} \tag{3.14}$$

The maximization of the above function is equal to the minimization of the following.

$$\hat{z} = \underset{z}{arg\,min}\ [\frac{(\mathbf{x}-\mathbf{z})^2}{2\sigma^2} + U(z)] \tag{3.15}$$

As color image has three spectral components, it can be represented as

$$\hat{z} = \underset{z}{argmin}\ [\frac{(x^{(1)} - z^{(1)})^2 + (x^{(2)} - z^{(2)})^2 + (x^{(3)} - z^{(3)})^2}{2\sigma^2} + \sum_{c \in C} V_c(z^{(1)}, z^{(2)}, z^{(3)})] \qquad (3.16)$$

where V_c is the clique potential function for all the three spectral components, C is the set of all the cliques and $1, 2$ and 3 represents the three spectral components.

3.2 SIMULATED ANNEALING

Simulated annealing is inspired by an analogy between the physical annealing of solids (crystals) and combinatorial optimization problems. In the physical annealing process a solid if first melted and then cooled very slowly, spending a long time at low temperatures, to obtain a perfect lattice structure corresponding to a minimum energy state. SA transfers this process to local search algorithms for combinatorial optimization problem. It does so by associating the set of solutions of the problem attacked with the states of the physical system, the objective function with the physical energy of the solid, and the optimal solution with the minimum energy states[11].

Metropolis in the earliest days of scientific computing, introduced a simple algorithm that can be used to provide an efficient simulation of a collection of atoms in equilibrium at a given temperature. In each step of this algorithm, an atom is given a small random displacement and the resulting change ΔU, in the energy of the system is computed. If $\Delta U \leq 0$, the displacement is accepted, and the configuration with the displaced atom is used as the start-

ing point of the next step. The case $\Delta U > 0$ is treated probabilistically: the probability that the configuration is accepted is $P(\Delta U) = exp(-\frac{\Delta U}{k_B T})$. Random numbers uniformly distributed in the interval (0, 1) are a convenient means of implementing the random part of the algorithm. One such number is selected and compared with $P(\Delta U)$. If it is less than $P(\Delta U)$, the new configuration is retained , if not , the original cinfiguration is used to start the next step. By repeating the basic step many times, one simulates the thermal motion of atoms in thermal contact with a heat bath at temperature T. The choice of $P(\Delta U)$ has the consequence that the system evolves into a Boltzman distribution[11].

Using the cost function inplace of the energy and defining configuration by a set of parameters, it is straightforward with the Metropolis procedure to generate a population of configurations of a given optimization problem at some effective temperature. This temperature is simply a control parameter in the same units as the cost function. The simulated annealing process consists of first *melting* the system being optimized at a high effective temperature, then lowering the temperature at slow stages until the system freezes and no further changes occur. At each temperature, the simulation must proceed long enough for the system to reach a stedy state. The sequence of temperature and the number of rearrangements of the parameters attempted to reach equilibrium at each temperature is known as an *annealing schedule* [11][6].

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

1. Initialize the temperature T_{in}.

2. Compute the energy U of the configuration.

3. Perturb the system slightly with suitable Gaussian disturbance.

4. Compute the new energy $U^{'}$ of the perturbed system and evaluate the change in energy $\Delta U = U^{'} - U$.

5. If $(\Delta U < 0)$, accept the perturbed system as the new configuration. Else accept the perturbed system as the new configuration with a probability $exp(-\Delta U)/t$, where t is the cooling schedule.

6. Decrease the temperature according to the cooling schedule.

7. Repeat steps 2-7 till the stopping criterion is met. The stopping criterion used here is the energy.

3.3 ITERATED CONDITIONAL MODE ALGORITHM

Since it is difficult to maximize the joint probability of an MRF, Besag proposed a deterministic algorithm called Iterated Conditional Modes (ICM) which maximizes local conditional probabilities sequentially. The ICM algorithm uses the greedy strategy in the iterative local maximization. Given the data x and the other labels $z_{S-i}^{(k)}$, the algorithm sequentially updates each $z_i^{(k)}$ into $z_i^{(k+1)}$ by maximizing $P(z_i \mid x, z_{S-i})$, the conditional probability, with respect to z_i. Two assumptions are made in calculating $P(z_i \mid x, z_{S-I})$:

1. The observation components $x_1,\ x_2,\ x_3 ... \ x_m$ are conditionally independent given z and each x_i has the same knowm conditional density function $p(x_i \mid z_i$ dependent only on z_i. Thus

$$p(x \mid z) = \prod_i p(x_i \mid z_i) \tag{3.17}$$

2. The second assumption is that z depends on the labels in the local neighborhood, which is the Markovianity.

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

$$P(z_i \mid x, z_{S-i}) \propto p(x_i \mid z_i) P(z_i \mid z_{N_i}) \tag{3.18}$$

Obviously, $P(z_i \mid x_i,\ z_{N_i}^k)$ is much easier to maximize than $P(z \mid x)$, which is the point of ICM. Maximizing (4.18) is equivalent to minimizing the corresponding posterior energy using the following rule.

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

The result obtained by ICM depends very much on the initial estimator $z^{(0)}$ and the ICM is locally convergent[6].

3.4 HYBRID ALGORITHM

A new hybrid algorithm has been proposed to obtain the MAP estimate. In the hybrid algorithm SA algorithm is first run for some prespecified amounts of epochs and then ICM algorithm is run until the stopping criterion has been satisfied. The combination of global and local concept leads to a faster converging algorithm. The proposed hybrid algorithm is described below.

1. Initialize the temperature T_{in}.

2. Compute the energy U of the configuration.

3. Perturb the system slightly with suitable Gaussian disturbance.

4. Compute the new energy $U^{'}$ of the perturbed system and evaluate the change in energy $\Delta U = U^{'} - U$.

5. If ($\Delta U < 0$), accept the perturbed system as the new configuration Else accept the perturbed system as the new configuration with a probability $exp(-\Delta U/t)$, where t is the cooling schedule.

6. Decrease the temperature according to the cooling schedule.

7. Repeat steps 2-7 till some prespecified number of epochs of SA.

8. Compute the energy U of the configuration.

9. Perturb the system slightly with suitable Gaussian disturbance.

10. Compute the new energy $U^{'}$ of the perturbed system and evaluate the change in energy $\Delta U = U^{'} - U$.

11. If ($\Delta U < 0$), accept the perturbed system as the new configuration, otherwise retain the original configuration.

12. Repeat steps 8-12, till the stopping criterion is met. The stopping criterion is the energy ($U < threshold$).

3.5 RESULT DISCUSSION

Simulation is carried out for different class of images. The algorithms were tested on indoor as well as out door images. We have used I_1 I_2 I_3 (ohta) color model to model the images . The a priori and the a posteriori energy function considered for the above images are given as follows.

$$\begin{aligned} U(z) = \alpha[&(z_{i,j}^{(1)} - z_{i-1,j}^{(1)})^2(1 - v_{i,j}^{(1)}) + (z_{i,j}^{(1)} - z_{i,j-1}^{(1)})^2(1 - h_{i,j}^{(1)}) + \\ &(z_{i,j}^{(1)} - z_{i+1,j}^{(1)})^2(1 - v_{i,j}^{(1)}) + (z_{i,j}^{(1)} - z_{i,j+1}^{(1)})^2(1 - v_{i,j}^{(1)}) + \\ &(z_{i,j}^{(2)} - z_{i-1,j}^{(2)})^2(1 - v_{i,j}^{(2)}) + (z_{i,j}^{(2)} - z_{i,j-1}^{(2)})^2(1 - h_{i,j}^{(2)}) + \\ &(z_{i,j}^{(2)} - z_{i+1,j}^{(2)})^2(1 - v_{i,j}^{(2)}) + (z_{i,j}^{(2)} - z_{i,j+1}^{(2)})^2(1 - v_{i,j}^{(2)}) + \end{aligned}$$

$$(z_{i,j}^{(3)} - z_{i-1,j}^{(3)})^2(1 - v_{i,j}^{(3)}) + (z_{i,j}^{(3)} - z_{i,j-1}^{(3)})^2(1 - h_{i,j}^{(3)}) +$$
$$(z_{i,j}^{(3)} - z_{i+1,j}^{(3)})^2(1 - v_{i,j}^{(3)}) + (z_{i,j}^{(3)} - z_{i,j+1}^{(3}{}^2(1 - v_{i,j}^{(3)})] +$$
$$\beta_1[h_{i,j}^1 + v_{i,j}^1] + \beta_2[h_{i,j}^2 + v_{i,j}^2] + \beta_3[h_{i,j}^3 + v_{i,j}^3] \quad (3.20)$$

and

$$U_p(z) = \frac{[(x^{(1)} - z^{(1)})^2 + (x^{(2)} - z^{(2)})^2 + (x^{(3)} - z^{(3)})]^2}{2\sigma^2} + U(z) \quad (3.21)$$

Where (1), (2) and (3) in $z_{i,j}^{(1)}$, $z_{i,j}^{(2)}$, $z_{i,j}^{(2)}$ represents the three components of the I_1, I_2, I_3 color model considered and $h_{i,j}^1$, $h_{i,j}^2$and $h_{i,j}^3$ represents the horizontal line process in the respective three planes whereas $v_{i,j}^1$, $v_{i,j}^2$and $v_{i,j}^3$ represent the vertical line process. Here α and β are constants. If β_1, β_2 and β_3 are equal, then the second part of (3.20) becomes

$$U(f) = \beta[(h_{i,j}^1 + v_{i,j}^1) + (h_{i,j}^2 + v_{i,j}^2) + (h_{i,j}^3 + v_{i,j}^3)] \quad (3.22)$$

The line fields are used separately for three different planes. $h_{i,j}^c$ and $v_{i,j}^c$ for a given plane are given by

$$If \ |z_{i,j}^c - z_{i-1,j}^c| > Threshold$$
$$h_{i,j} = 1$$
$$Else \, h_{i,j} = 0 \quad (3.23)$$

and

$$If \ |z_{i,j}^c - z_{i,j-1}^c| > Threshold$$
$$v_{i,j} = 1$$
$$Else \, v_{i,j} = 0 \quad (3.24)$$

Where c indicates the spectral component of the color image.

In simulation different images are considered to validate the two models(1)Reward punishment model (2)Weak memberane model. The corresponding clique

potential functions for the two models are given below for reward and punishment model it is

$$V_c(x) = \begin{cases} +\delta_c & \text{if } |x_r - x_s| = 0 \\ -\delta_c & \text{if } |x_r - x_s| \neq 0 \end{cases} \tag{3.25}$$

where $x(r)$ and $x(s)$ denote the pixel values of the sites r and s respectively. The clique potential function for weak memberane model can be expressed as follows.

$$V_c(x) = \alpha |x_r - x_s|^2 \tag{3.26}$$

Initially the model given by (3.25) is validated for four different images as shown in Fig 3.1, 3.2, 3.3, and 3.4. The corresponding δ parameters are tabulated in Table 3.1. Checkerboard image is shown in Fig 3.1(a) and the corresponding low noise case is shown in Fig 3.1(b). Both isotropic and anisotropic fields are considered in simulation. Fig 3.1(c) is the image corresponding to the ground truth drawn manually. The MAP estimates of the pixel labels are obtained by the proposed hybrid algorithm and also are compared with that of using SA algorithm, Ohta model is used as color model for all the simulations. Fig 3.1(d), (e), (f) and (g) correspond to four different methods, namely (i)Isotropic with SA, (ii)Isotropic with Hybrid, (iii)Anisotropic with Hybrid algorithm (iv)Isotropic 2nd order with Hybrid. it is observed from Fig 3.1(a)that for isotropic case SA converges to the solution after 1000 iterations, where as the isotropic 1st order case with Hybrid algorithm converges to the solution around 350 iterations .The 2nd order, anisotropic case, the Hybrid algorithm converges at 350 iterations to energy level of 1000 which much lower as compared to isotropic case. SA algorithm converges after 2000 iterations as shown in Fig 3.1(i). The misclassification error with reference to the ground truth image is tabulated in Table 3.4. The misclassification error is 0.04 and 0.09% and in isotropic

2nd order is zero.Hence perfect segmentation is obtained by this.

Results obtained from Jellybeans image, Searock image and Redhouse image are presented in Fig 3.2,3.3, and 3.4 respectively. It is observed from Fig 3.2 that the segmentation results obtained by isotropic, anisotropic first order and 2nd order are with misclassification error of about 19% and is minimum in the case of isotropic first order i.e 19.43%. Visual observation also reveals that the image shown by fig 3.2(c) is more appealing As, observed the same portion of the background has a different class. However the beans have been segmented appreciably. In this case also, the Hybrid algorithm converges at around 200 iterations that is much lower than that of SA which converges at around 3000 iterations. As observed previously the energy level settles at 100 in each case of isotropic 2nd order case. This implies that 2nd order markov random yields better results and hence is more than that of first order case.

Fig 3.3 shows the searock image and the corresponding segmented results The ground truth image is shown in Fig 3.3(b). There are some missclassified pixels in case of Fig 3.2(d), (e), and (f). The percentage of misclassification error as tabulated in Table 3.4 is around 4% and is minimum in case of isotropic first order case i.e. 3.46% , As regards convergence, the observation, thet is Hybrid algorithm converges much faster as same as those of previous examples. This phenomenon is dipicted in fig 3.3(g) and 3.3(h).

The house image and the corresponding ground truth image is shown in Fig 3.4. As observed from Fig 3.4(c), (d), (e) and (f), proper classification is obtained in different parts of the house except some misclassified pixels in the grass portion associated with the image Because of this phenomenon the percentage of misclassification error is more around 25% for this image. However, visually the segmented image is akin to the ground truth image. As

expected the hybrid algorithm converges faster than that of SA and converges at around 200 iterations. This is shown in Fig 3.4(a) and Fig 3.4(b).

The four image considered for the reward-punishment model is also considered for validating weak memberane model as given in (eq 3.20, 3.26). The segmented results as obtained using this model is shown in Fig 3.5, 3.6, 3.7, and 3.8. In all the cases , the proposed hybrid algorithm converges faster than that of using SA. As seen from Fig 3.5, there are some missing edges and hence misclassification error,as presented in Table 3.5. is more as compared to that of reward punishment model. This is around 6.53% in case of 2nd order isotropic field. As seen from Table 3.5, the misclassification error for searock and jellybeans images are almost close to that of reward punishment model. This is also visually observed from Fig 3.6, Fig 7. In case of the house image the %age of missclassification error is less than that of using reward punishment model. This implies that, the weak memberane model is an appropriate choice for real image, while the reward punishment model performs well in case of synthetic images.

Images	α	δ	β	σ_{RP}	σ_{WM}	T_{in}	$Thresh$
CheckerBoard	0.012	0.012	2.2	3.22	0.6	0.38	2.5
Searock	0.019	0.019	2.4	1.3	1.3	0.38	3.5
Jellybeans	0.012	0.012	2.2	0.6	0.6	0.38	2.0
Redhouse	0.009	0.009	2.4	1.2	1.2	0.38	3.0

Table 3.1: Parameters for Isotropic 1st order and 2nd order weak memberane and reward punishment model

Images	δ_1	δ_2	δ_3	δ_4
CheckerBoard	0.010	0.011	0.009	0.012
searock	0.019	0.020	0.022	0.021
Jellybeans	0.012	0.011	0.013	0.010
Redhouse	0.009	0.010	0.012	0.014

Table 3.2: Parameters for Anisotropic 1st order Reward punishment model

Images	α_1	α_2	α_3	α_4
CheckerBoard	0.010	0.011	0.009	0.012
searock	0.021	0.023	0.025	0.027
Jellybeans	0.02	0.009	0.010	0.015
Redhouse	0.009	0.010	0.011	0.012

Table 3.3: Bonding Parameters for Anisotropic 1st order weak memberane model

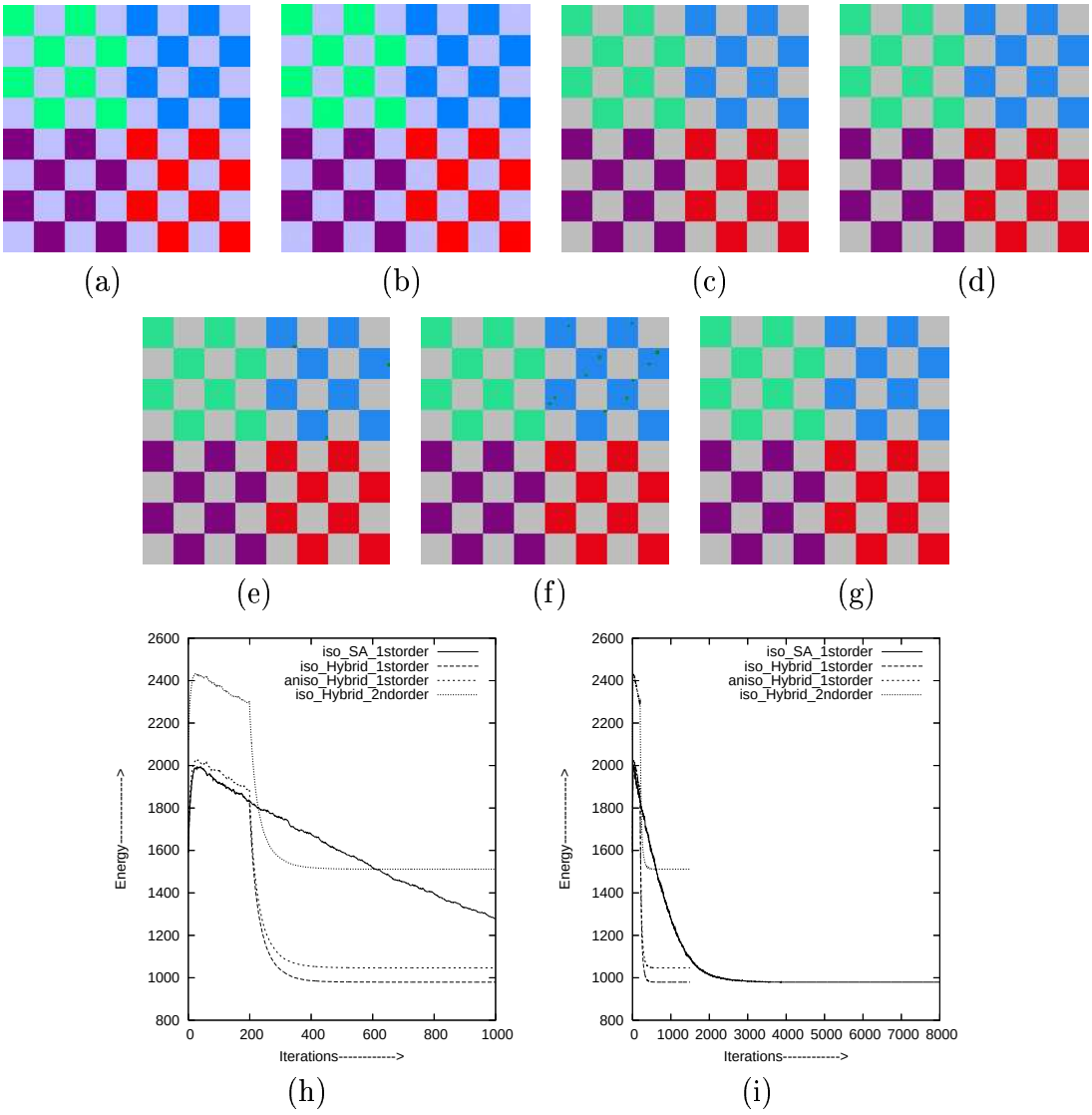


Figure 3.1: Segmentation of checkerboard image of size 200x200 using reward and punishment model: (a) Original image (b) Noisy Image with SNR=35dB (c) Ground Truth (d) Isotropic 1st order using SA (e) Isotropic 1st order using Hybrid (f) Anisotropic 1st order using Hybrid (g) Isotropic 2nd order using Hybrid (h) Comparison of Energy convergence of all the above model for first 1000 iterations (i) Comparison of Energy convergence of all the above model for whole process

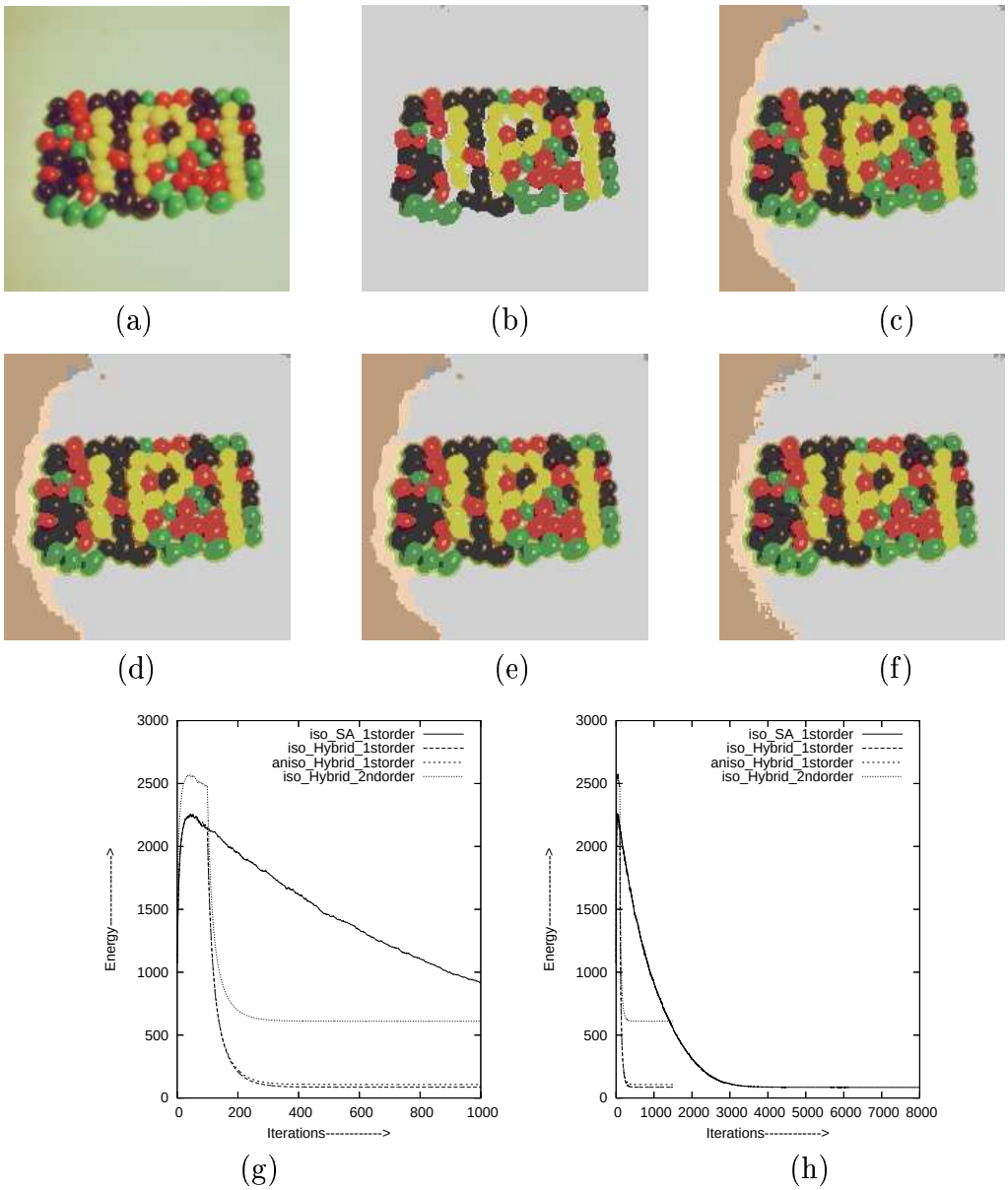


Figure 3.2: Segmentation of Jellybeans image of size 200x200 using reward and punishment model: (a) Original image (b) Ground Truth (c) Isotropic 1st order using SA (d) Isotropic 1st order using Hybrid (e) Anisotropic 1st order using Hybrid (f) Isotropic 2nd order using Hybrid (g) Comparison of Energy convergence of all the above model for first 1000 iterations (h) Comparison of Energy convergence of all the above model for whole process

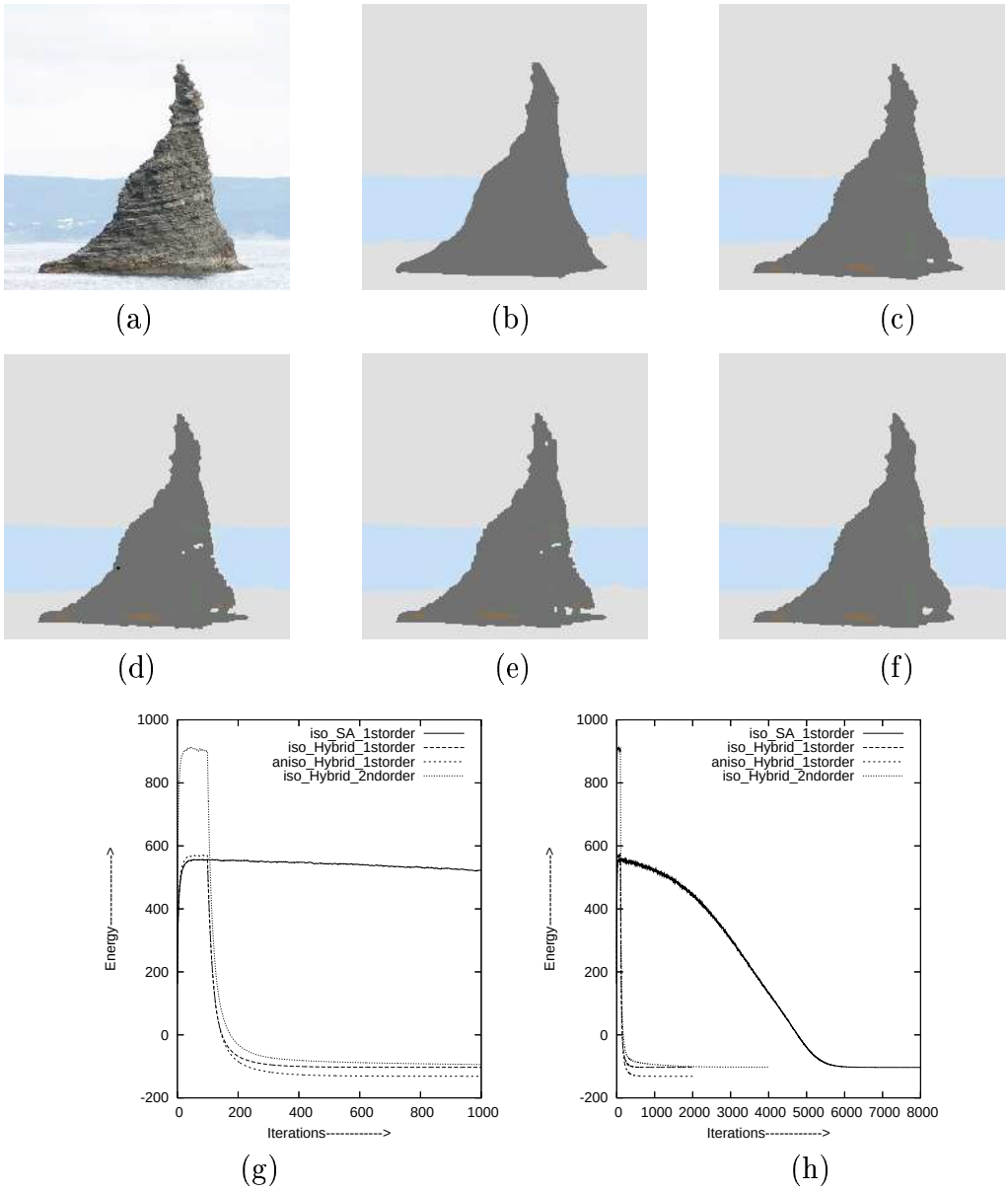


Figure 3.3: Segmentation of Searock image of size 220x220 using reward and punishment model: (a) Original image (b) Ground Truth (c) Isotropic 1st order using SA (d) Isotropic 1st order using Hybrid (e) Anisotropic 1st order using Hybrid (f) Isotropic 2nd order using Hybrid (g) Comparison of Energy convergence of all the above model for first 1000 iterations (h) Comparison of Energy convergence of all the above model for whole process

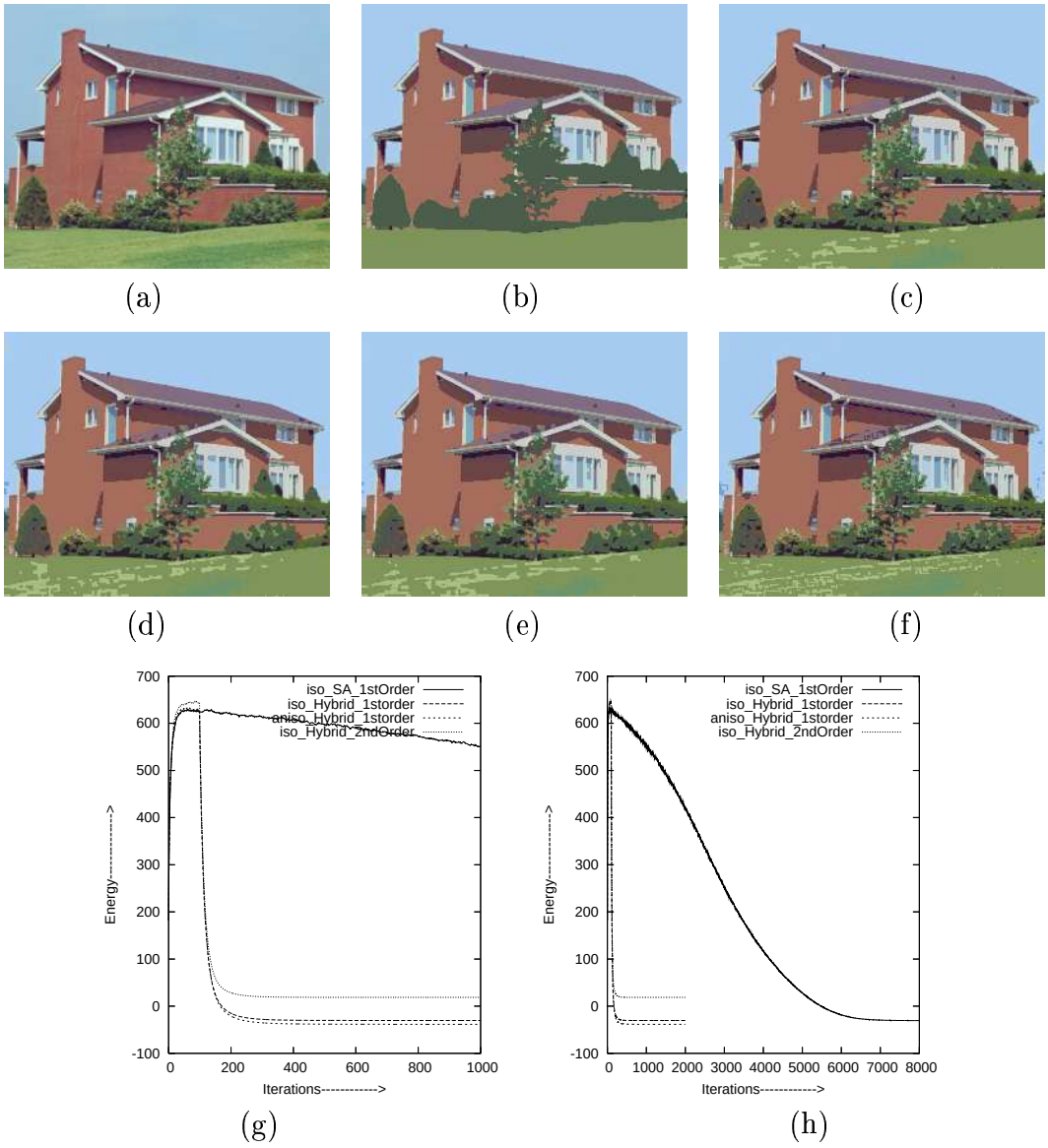


Figure 3.4: Segmentation of Redhouse image of size 210x256 using reward and punishment model: (a) Original image (b) Ground Truth (c) Isotropic 1st order using SA (d) Isotropic 1st order using Hybrid (e) Anisotropic 1st order using Hybrid (f) Isotropic 2nd order using Hybrid (g) Comparison of Energy convergence of all the above model for first 1000 iterations (h) Comparison of Energy convergence of all the above model for whole process

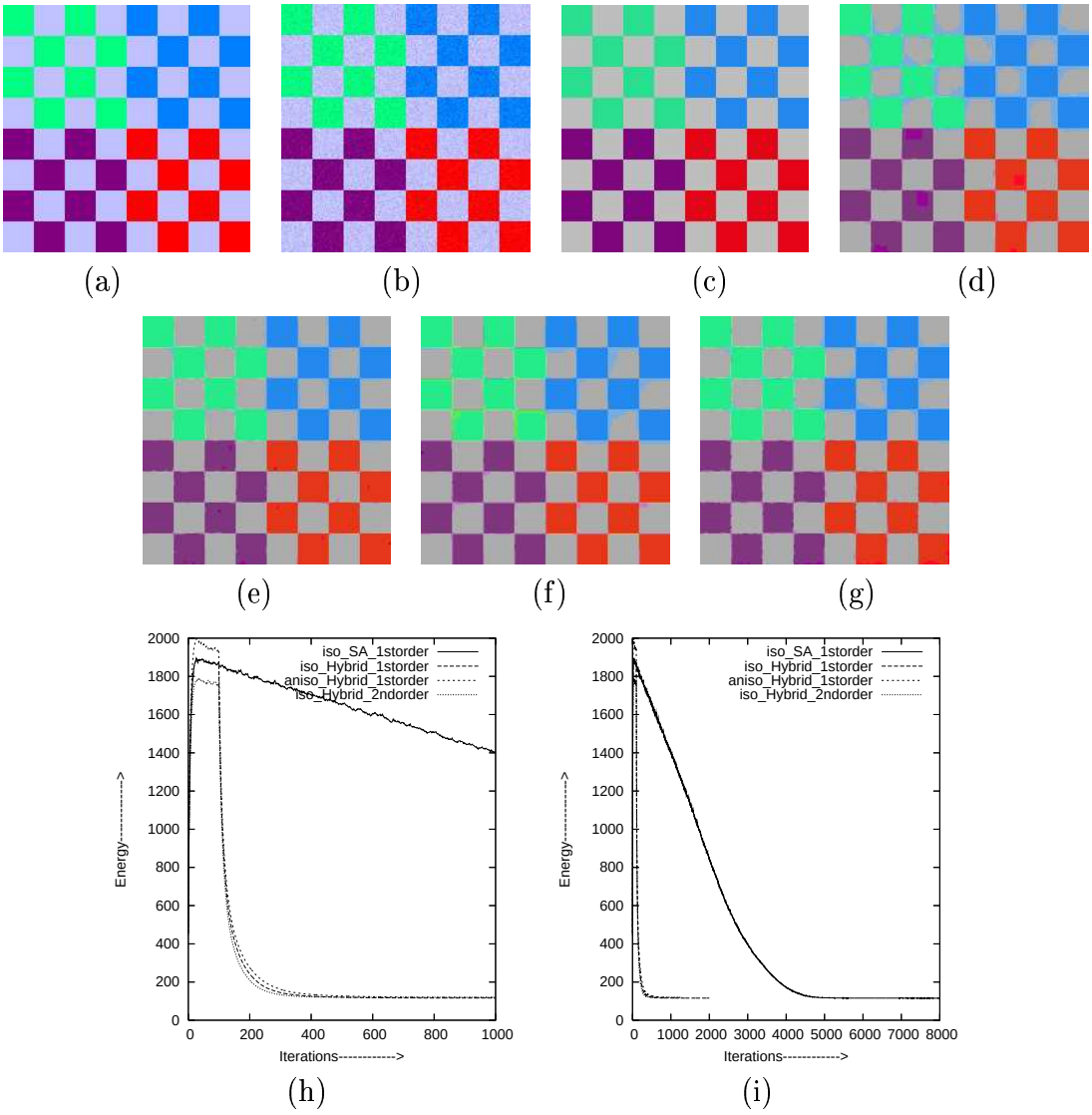


Figure 3.5: Segmentation of checkerboard image of size 200x200 using weak memberane model: (a) Original image (b) Noisy Image with SNR=20dB (c) Ground Truth (d) Isotropic 1st order using SA (e) Isotropic 1st order using Hybrid (f) Anisotropic 1st order using Hybrid (g) Isotropic 2nd order using Hybrid (h) Comparison of Energy convergence of all the above model for first 1000 iterations (i) Comparison of Energy convergence of all the above model for whole process

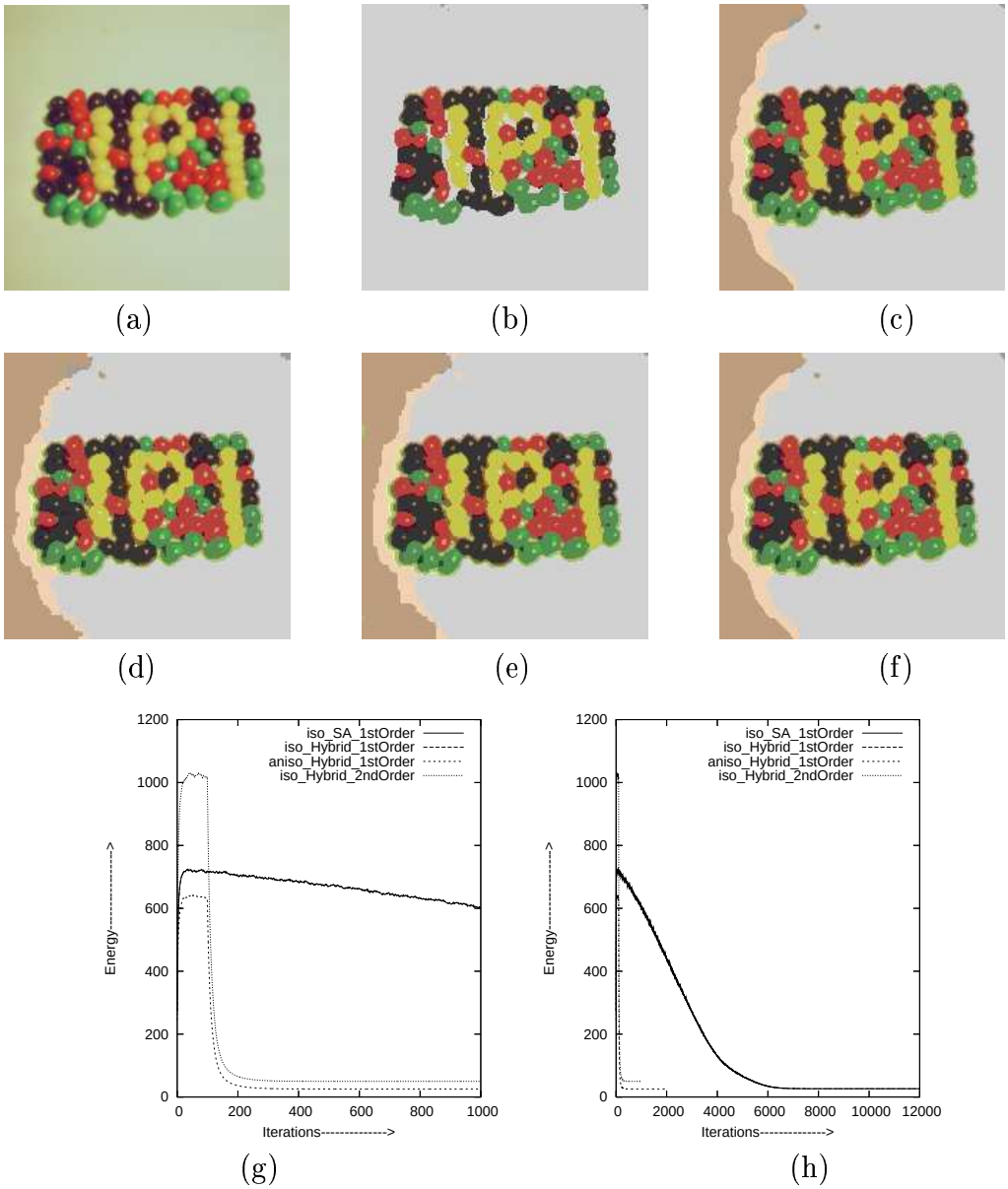


Figure 3.6: Segmentation of Jellybeans image of size 200x200 using weak memberane model: (a) Original image (b) Ground Truth (c) Isotropic 1st order using SA (d) Isotropic 1st order using Hybrid (e) Anisotropic 1st order using Hybrid (f) Isotropic 2nd order using Hybrid (g) Comparison of Energy convergence of all the above model for first 1000 iterations (h) Comparison of Energy convergence of all the above model for whole process

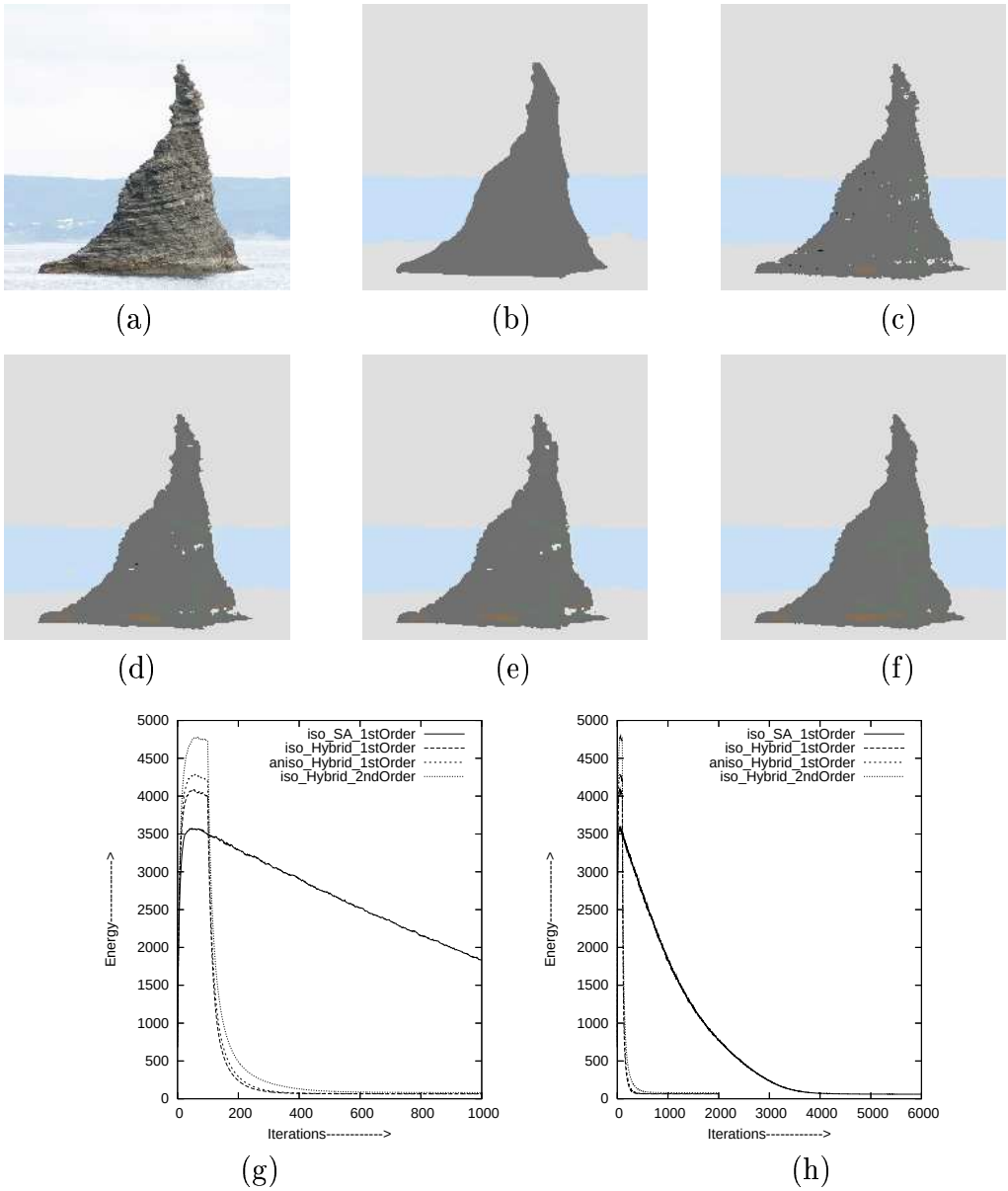


Figure 3.7: Segmentation of Searock image of size 220x220 using weak memberane model: (a) Original image (b) Ground Truth (c) Isotropic 1st order using SA (d) Isotropic 1st order using Hybrid (e) Anisotropic 1st order using Hybrid (f) Isotropic 2nd order using Hybrid (g) Comparison of Energy convergence of all the above model for first 1000 iterations (h) Comparison of Energy convergence of all the above model for whole process

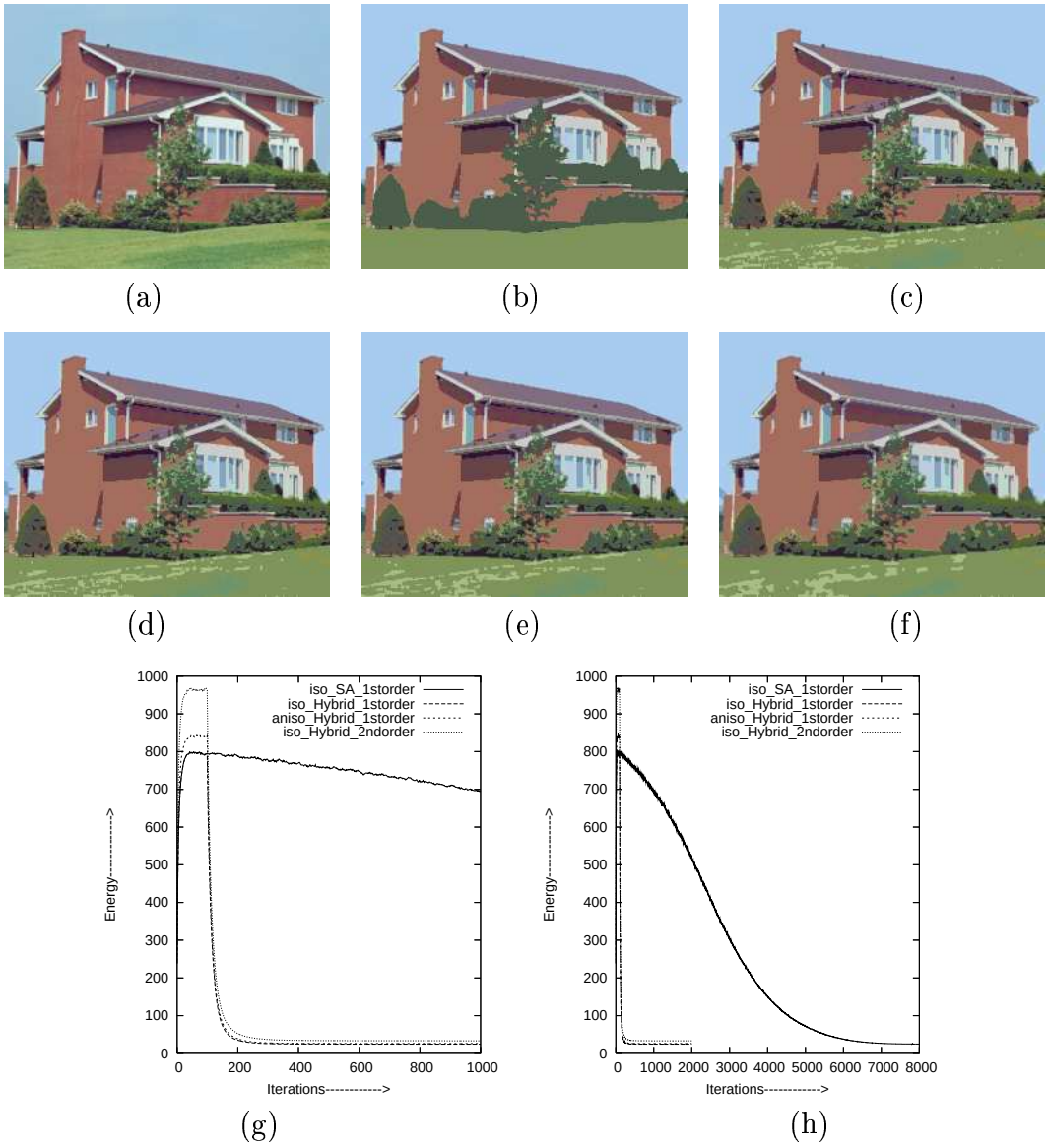


Figure 3.8: Segmentation of Redhouse image of size 210x256 using weak memberane model: (a) Original image (b) Ground Truth (c) Isotropic 1st order using SA (d) Isotropic 1st order using Hybrid (e) Anisotropic 1st order using Hybrid (f) Isotropic 2nd order using Hybrid (g) Comparison of Energy convergence of all the above model for first 1000 iterations (h) Comparison of Energy convergence of all the above model for whole process

Images	%*age* of misclassification Error			
	Iso_SA 1stOrder	Iso_Hybrid 1stOrder	Aniso_Hybrid 1stOrder	Iso_Hybrid 2ndOrder
Checkerboard	0.0	0.04	0.09	0.0
Searock	3.46	4.23	3.99	3.94
Jellybeans	19.43	19.84	19.97	20.34
Redhouse	25.22	25.63	26.32	24.87

Table 3.4: Comparison of %*age* of Misclassification Error of segmentation results obtained using Reward and Punishment(RP) model

Images	%*age* of misclassification Error			
	Iso_SA 1stOrder	Iso_Hybrid 1stOrder	Aniso_Hybrid 1stOrder	Iso_Hybrid 2ndOrder
Checkerboard	7.07	3.6	6.89	6.53
Searock	4.73	5.33	5.51	6.37
Jellybeans	20.4	20.45	21.3	20.13
Redhouse	18.83	18.2	19.31	19.46

Table 3.5: Comparison of %*age* of Misclassification Error of segmentation results obtained using Weak Memberane(WM) model

Images	Time taken(in seconds)			
	Iso_SA 1stOrder	Iso_Hybrid 1stOrder	Aniso_Hybrid 1stOrder	Iso_Hybrid 2ndOrder
Redhouse_{WM}	1419	78	78	99
searock_{WM}	1379	152	152	190
Redhouse_{RP}	1296	73	73	77
searock_{RP}	1277	135	135	148

Table 3.6: Comparison of Timing for each model

Chapter 4

THE CONSTRAINED MRF MODEL AND DOUBLE MRF MODEL

In probability theory, a martingale is a stochastic process (i.e., a sequence of random variables) such that the conditional expected value of an observation at some time t, given all the observations up to some earlier time s, is equal to the observation at that earlier time s. Precise definitions are given below.

Originally, martingale referred to a class of betting strategies popular in 18th century France. The simplest of these strategies was designed for a game in which the gambler wins his stake if a coin comes up heads and loses it if the coin comes up tails. The strategy had the gambler double his bet after every loss, so that the first win would recover all previous losses plus win a profit equal to the original stake. Since as a gambler's wealth and available time jointly approach infinity his probability of eventually flipping heads approaches 1, the martingale betting strategy was seen as a sure thing by those who practiced it. Of course in reality the exponential growth of the bets would eventually bankrupt those foolish enough to use the martingale for a long time. The concept of martingale in probability theory was introduced by Paul Pierre Lvy, and much of the original development of the theory was

done by Joseph Leo Doob. Part of the motivation for that work was to show the impossibility of successful betting strategies.

A discrete-time martingale is a discrete-time stochastic process (i.e., a sequence of random variables) $X_1, X_2, X_3...$ that satisfies for all n

$$E(|X_n|) < \infty$$

$$E(X_{n+1}|X_1, X_2, X_3..., X_n) = X_n$$

i.e., the conditional expected value of the next observation, given all of the past observations, is equal to the last observation.

Examples

1. Suppose X_n is a gambler's fortune after n tosses of a "fair" coin, where the gambler wins Rs 1 if the coin comes up heads and loses Rs 1 if the coin comes up tails. The gambler's conditional expected fortune after the next trial, given the history, is equal to his present fortune, so this sequence is a martingale. This is also known as D'Alembert system.

2. Let $Y_n = X_n^2 - n$ where $X_n$is the gambler's fortune from the preceding example. Then the sequence $Y_n : n = 1, 2, 3, ...$ is a martingale. This can be used to show that the gambler's total gain or loss varies roughly between plus or minus the square root of the number of steps.

4.1 CONSTRAINED MARKOV RANDOM FIELD

Capturing the salient spatial properties of an image lead to the development of image models. Though the MRF model takes into account the local spatial interactions, it has its limitations in modeling natural scenes of distinct

regions. In order to incorporate a stronger local dependence, we constrain this model based on the notion of martingale. The motivation behind the new model is as follows:

Let $Z(i), i = 1, 2, \cdots n$ be a martingale sequence, namely for all $i = 1, 2, \cdots, n$ $E[|Z(n)|] < \infty$ and $E[Z(n+1)/Z(1), \cdots Z(n)] = Z(n)$. Now, let $Z_1, Z_2, \cdots, Z_n$ be the random variables associated with the image of size $n = N^2$. Therefore, $E[Z_{i,j}/Z_{k,l}, k, l \neq i, j] = Z_{i-1,j}$ for any $k, l \in \eta_i, j$, where $\eta_{i,j}$ is the neighborhood of i, j. Consider,

$$E[Z_{i,j} \mid Z_{k,l}, k, l \neq i, j] = \sum_{z_{i,j} \in G} z_{i,j} P[Z_{i,j} = z_{i,j} \mid Z_{k,l} = z_{k,l}, k, l \neq i, j]$$

Assuming further that Z is a Markov process, we have

$$= \sum_{z_{i,j} \in G} z_{i,j} P[Z_{i,j} = z_{i,j} \mid Z_{k,l} = z_{k,l}, k, l \in \eta_{i,j}]$$

$$= \sum_{z_{i,j} \in G} z_{i,j} \frac{P(Z = z)}{\sum_{z_{i,j}} P(Z = z)}$$

Since Z is a MRF,

$$= \sum_{z_{i,j} \in G} z_{i,j} \frac{e^{-U(Z)}}{\sum_{z_{i,j}} e^{-U(Z)}}$$

$$z_{k,l} \forall k, l \in \eta_{i,j} = \sum_{z_{i,j} \in G} z_{i,j} \frac{e^{-U(Z)}}{\sum_{z_{i,j} \in G} e^{-U(Z)}}$$

For a specific case,

$$z_{i-1,j} = \sum_{z_{i,j} \in G} z_{i,j} \frac{e^{-U(Z)}}{\sum_{z_{i,j} \in G} e^{-U(Z)}}$$

Instead of taking the pixel $z_{i-1,j}$, we take the average of the neighborhood pixels. Now, $\max_z P(Z = z) = e^{-U(Z)}$ is equivalent to $\min_z U(Z)$ such that for every i, j

$$\left[z_{i-1,j} - \sum_{z_{i,j} \in G} z_{i,j} \frac{e^{-U(Z)}}{\sum_{z_{i,j} \in G} e^{-U(Z)}} \right]^2$$

is minimized. Or, equivalently,

$$\min_{z} \sum_{i,j} U(z_{i,j}) + \lambda \left\{ z_{i,j} - \sum_{z_{i,j} \in G} z_{i,j} \frac{e^{-U(Z)}}{\sum_{z_{i,j} \in G} e^{-U(Z)}} \right\}^2 \tag{4.1}$$

Where,

$$\sum_{i,j} U(z_{i,j}) = \sum_{i,j} \alpha\{(z_{i,j} - z_{i-1,j})^2(1 - v_{i,j}) + (z_{i,j} - z_{i,j-1})^2(1 - h_{i,j})\} + \beta(h_{i,j} + v_{i,j}) \tag{4.2}$$

Because this is an additional constraint in the clique potential, we name this model as the constrained MRF model(CMRF).

4.2 DOUBLE MARKOV RANDOM FIELD

Here we consider the RGB color model and in each plane we consider the observed image is assumed to consists of a label process and the degraded process.The degraded process in each plane (say for example R plane) is a Gaussian process. Let Z denote the label process and Z is assumed to MRF in case of each plane.

The pixel in one plane (say for e.g R-plane) is assumed to have interaction with G and B planes.This interaction model I is assumed to be MRF. Thus, the inter plane interaction (R-G-B) is assumed to be MRF. Thus, the label process in RGB plane consists of two MRFs, one Z corresponding to intraplane and other I, corresponding to interplane. Thus, the double MRF model is denoted by (Z,I) and the image X is represented by (Z,I) i.e X is modeled as, $X = (Z, I)$ and the Gibbs equivalent of both interplane and intraplane interaction is given by,

$$Z = \frac{1}{Z'} exp^{\frac{-U(z)}{T}} \tag{4.3}$$

$$I = \frac{1}{I'} exp^{\frac{-U(I)}{T}} \tag{4.4}$$

This notion is dipicted in Fig 4.1, where the intra plane and inter plane interactions are shown. The model parameters for both Z and I processes

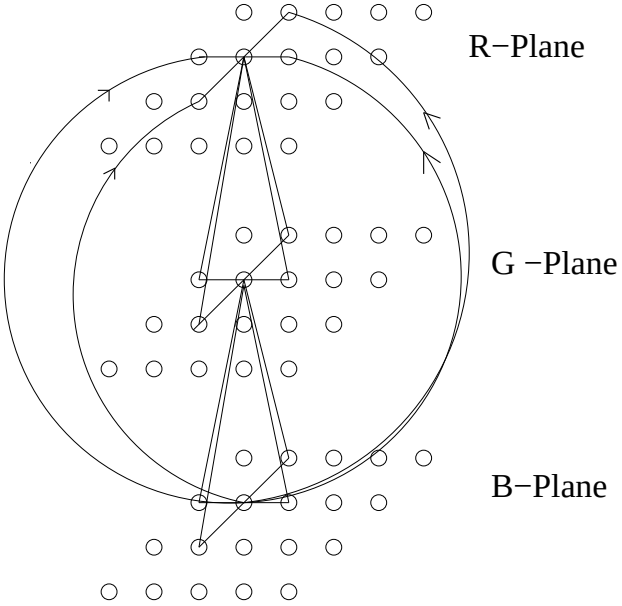


Figure 4.1: RGB Plane Interaction

are selected on an ad hoc manner. The weak memberane model is considered for intra as well as inter plane clique potential functions.

4.3 RESULTS AND DISCUSSION

4.3.1 CMRF model

Scene as well as color textured images of different classes have been considered in simulation.The *a priori* energy function for the CMRF model and the weak membrane model as given by (4.1) and (3.21), are considered . The *a priori* image model parameters of CMRF model α, β, σ and λ are selected on trial and error basis.We have taken the initial temperature $T_{in} = 0.38$ for all the simulations. The MRF and CMRF model parameters considered for different images are given in Table 4.1. The second, third and fourth column

of the table correspond to the MRF model parameters. For the CMRF model,with the same MRF model parameters, the 'λ' parameter is selected on an ad hoc manner for texture as well as scene images. For example, for images having two textures as shown in Fig 4.2(a), the parameters are α=0.16, β=2.0, σ=1.20 and λ=1.0. The segmented images are shown in Fig 4.1(b), (c) and (d). As observed from Fig 4.1(b), the MRF model using SA could not yield proper segmentation and there are many misclassified pixels in different classes. Use of the proposed CMRF model, which has the grouping attribute, could produce segmented images with appreciably less misclassified pixel.This is evident from Fig 4.1(c). This is corroborated from the percentage of misclassification presented in Table 4.2.

For CMRF model the percentage of misclassification is 3.2% while using MRF it is 9.2% which is almost three times that of the case of CMRF. The percentage of misclassification is computed based on the ground truth image obtained manually. Based on the texture and the texture boundary, we have drawn the ground truth image for e.g for the image shown in Fig 4.1(a) there are two distinct texture and hence two labels. The two labels are drawn with the distinct boundary and thus a ground truth image is created as shown in Fig 4.1(b). Hybrid algorithm with CMRF model accelerate convergence as shown in Fig 4.2(g) and Fig 4.2(h). Fig 4.2(g) shows that convergence of SA algorithm around 4000 iterations while Hybrid algorithm converges around 200 iterations, thus a substantial computational saving is obtained. This is reflected in Fig 4.2(g) which shows a part of the figure of Fig 4.2(h). The segmented image obtained using hybrid and CMRF model is shown in Fig 4.2(c) and the percentage of misclassification is 3.5 % which is close to that obtain using CMRF with SA. The quality of the solution almost remains same while the speed is approximately 20 times. Use of Hybrid algorithm for other

images as shown in Fig 4.4 and Fig 4.7 has exhibited faster convergence than those of using SA. A color texture having five different textures is shown in Fig 4.4(a) and the segmented images obtained are shown in Fig 4.4(c),4.4(d) and 4.4(e). The segmented image with CMRF model has misclassification error of 2.67% where as with MRF it is found to be 26.8%. This is quite evident visually from the segmented images shown in Fig 4.4(c),4.4(d) and 4.4(e). The CMRF model with different values of λ could model efficiently the scene image, which is shown in Fig 4.7. In case of the texture images, CMRF model could produce visually better segmented result than that of using MRF model.This is also corroborated from the percentage of misclassification given in Table 4.2. Similar results are obtained in Fig 4.3, 4.5 and 4.6, In Fig 4.3 the mat region of the image which is having good amount of texture in it is also classified appreciably by CMRF model and in Fig 4.5 the grass part of the image is segmented properly by the new method and results are better in terms of percentage of mis classification as shown in Table 4.2.In case of carhouse image the roof part and the grass part which MRF could not able to labeled properly are segmented nicely by CMRF models. Thus the proposed CMRF model posses the unifying modelling attributes to model scene as well as textures. The Ohta model together with CMRF model yield better segmented result as opposed to MRF model.

Images	α	β	σ	λ
Tex_two	0.16	2.0	1.2	1.0
Tex_four	0.05	2.2	1.72	0.06
Tex_five	0.038	2.5	1.58	0.06
Scene	0.03	2.5	1.137	0.07
Redhouse	0.04	2.5	1.0	0.048
Carhouse	0.04	2.5	1.0	0.014

Table 4.1: Parameters for CMRF model

Images	$\%age$ of misclassification Error		
	MRF using SA	CMRF using Hybrid	CMRF using SA
Tex_two	9.167	3.21	3.5
Tex_four	14.33	1.43	0.79
Tex_five	26.8	2.67	2.91
Scene	22.28	14.75	14.3
Redhouse	20.48	16.24	12.8
Carhouse	23.56	15.81	10.33

Table 4.2: Comparison of $\%age$ of Misclassification Error for segmentation results of three texture and three scene Image

4.3.2 DMRF model

The result obtained using double MRF model is shown in Fig 4.9(d). This is also corroborated from the percentage of misclassification error as tabulated in Table 4.4. In case of MRF model without interaction, the percentage of error is 33.471%, while the error is 2.554% for double MRF model. Thus, the proposed double MRF model yielded better segmentation results. The model parameters for all the images are presented in table 4.3. The parameters remain same for the interaction among different planes. The segmented

Images	α	β	σ	$Thresh$
HillRiver	0.01241	2.3	1.11	2.4
Plane	0.01	2.93	2.51	2.4
Scene	1.095	2.93	1.9	2.2
Searock	0.6	2.93	1.4	2.4
Redhouse	0.09	2.93	0.1	2.2

Table 4.3: Parameters for segmentation results obtained using MRF model and DMRF model with RGB as color model

Images	$\%age$ of misclassification	
	MRF model	DMRF model
Hillriver	19.36	4.037
Plane	33.47	2.554
Scene	9.305	8.34
Searock	21.08	2.76
Redhouse	9.4	7.9

Table 4.4: Comparison of $\%age$ of Misclassification Error of segmentation results obtained using MRF model and DMRF model with RGB as color model

results obtained for scene and house image is shown in Fig 4.10 and 4.11. The results obtained in both the cases are almost close to each other. The percentage of misclassification error is presented in Table 4.4, are very close to each other

The double MRF model in case of searock image yielded promising results. This is evident from Fig 4.12(c) and (d), where there are more missclassified pixelsas compared to that of using DMRF model. Fig 4.10(d) shows the results obtained for the DMRF. The percentage of misclassification is 21.08% in case of single MRF and 2.76% in case of DMRF. Thus for RGB model yielded proper segmentation results in case of use of proposed DMRF model that takes care of interaction among the planes.

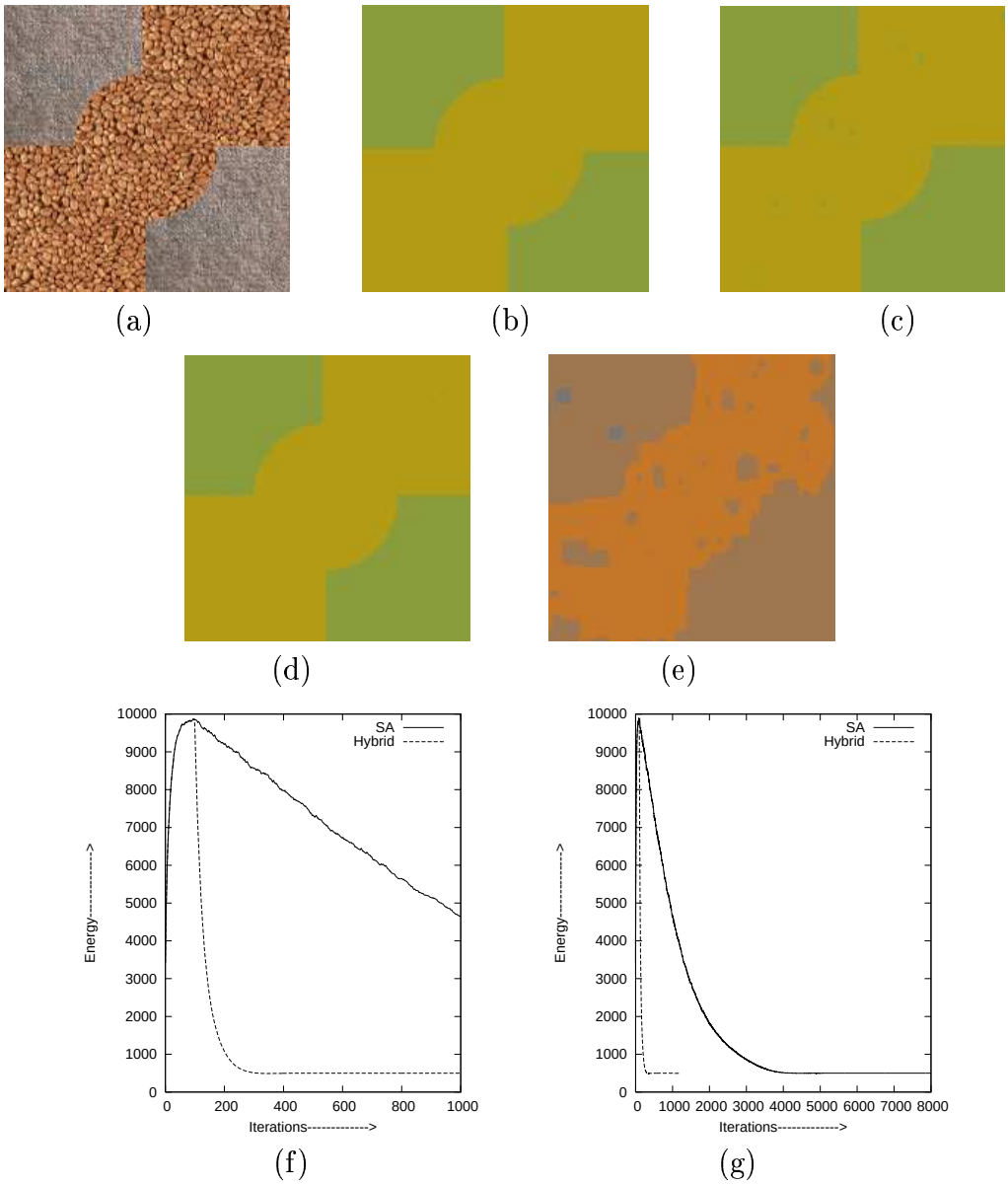


Figure 4.2: Segmentation of two class texture image(Tex_two) of size 128x128 using weak memberane model: (a) Original image (b) Ground Truth (c) CMRF optimized using SA (d) CMRF optimized using Hybrid (e) MRF optimized using SA (f) Comparison of Energy convergence curve for CMRF using SA and Hybrid for first 1000 Iterations (g) Comparison of Energy convergence curve for CMRF using SA and Hybrid for whole process

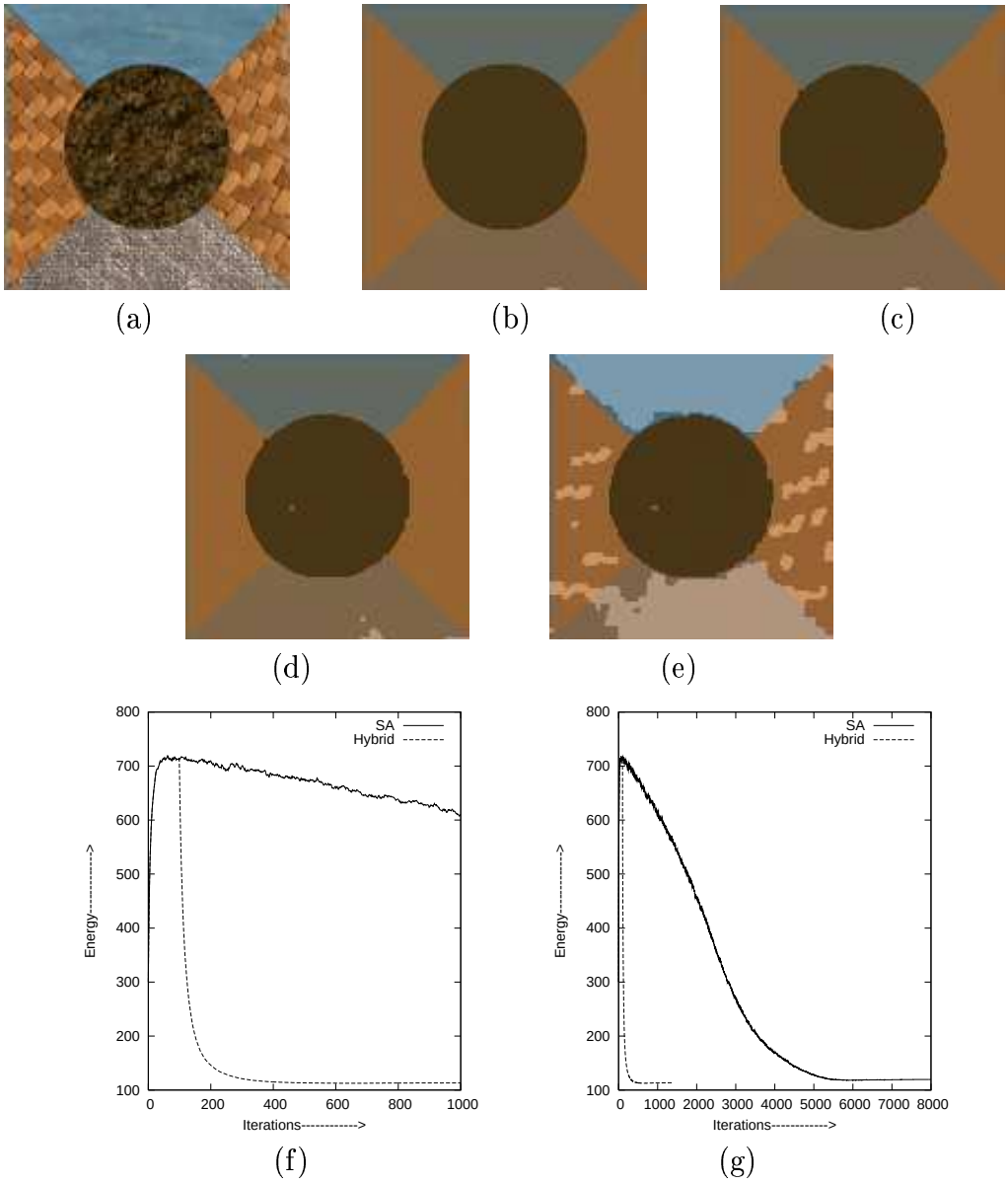


Figure 4.3: Segmentation of four class texture image(Tex_four) of size 128x128 using weak memberane model: (a) Original image (b) Ground Truth (c) CMRF optimized using SA (d) CMRF optimized using Hybrid (e) MRF optimized using SA (f) Comparison of Energy convergence curve for CMRF using SA and Hybrid for first 1000 Iterations (g) Comparison of Energy convergence curve for CMRF using SA and Hybrid for whole process

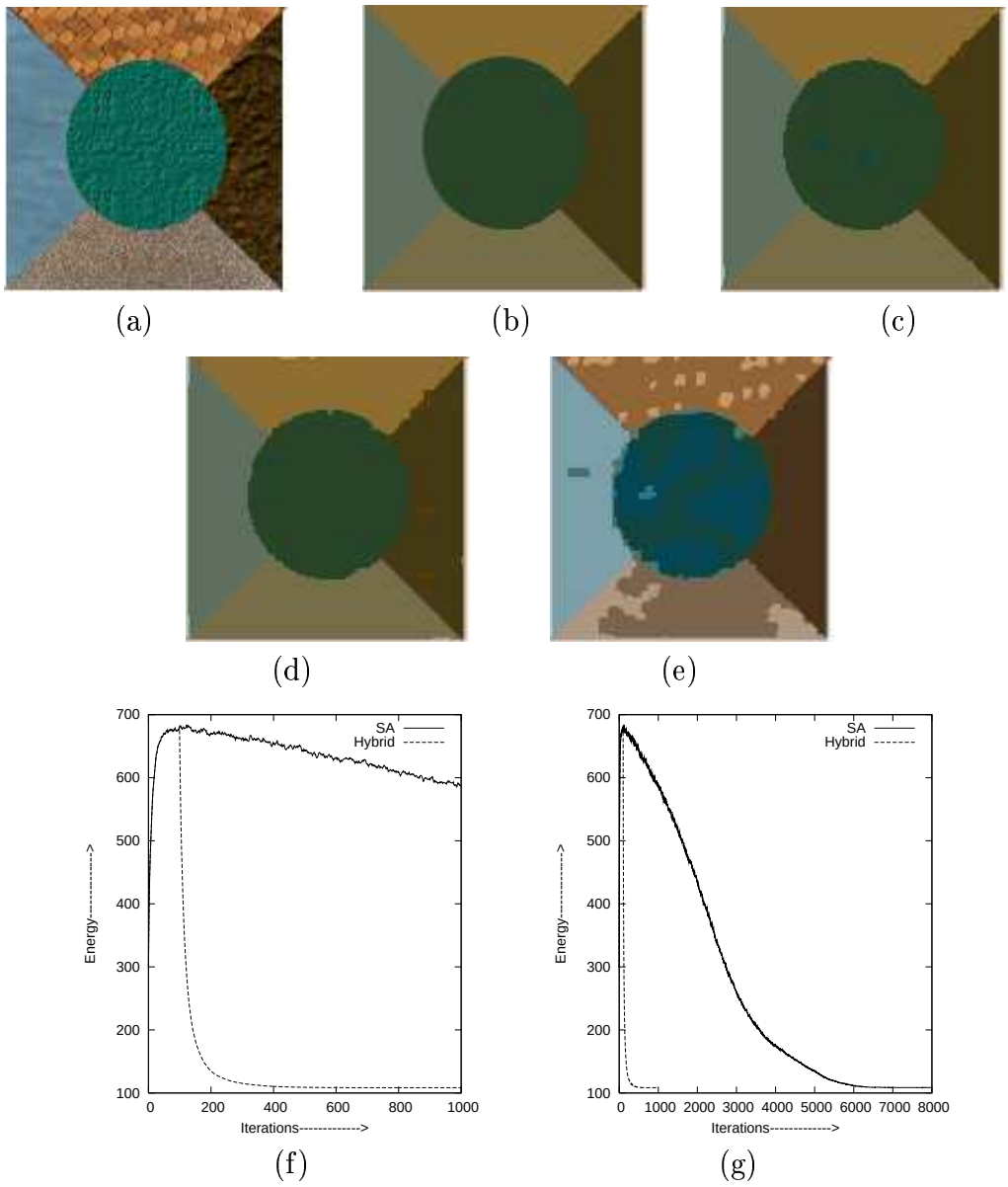


Figure 4.4: Segmentation of five class texture image(Tex_five) of size 128x128 using weak memberane model: (a) Original image (b) Ground Truth (c) CMRF optimized using SA (d) CMRF optimized using Hybrid (e) MRF optimized using SA (f) Comparison of Energy convergence curve for CMRF using SA and Hybrid for first 1000 Iterations (g) Comparison of Energy convergence curve for CMRF using SA and Hybrid for whole process

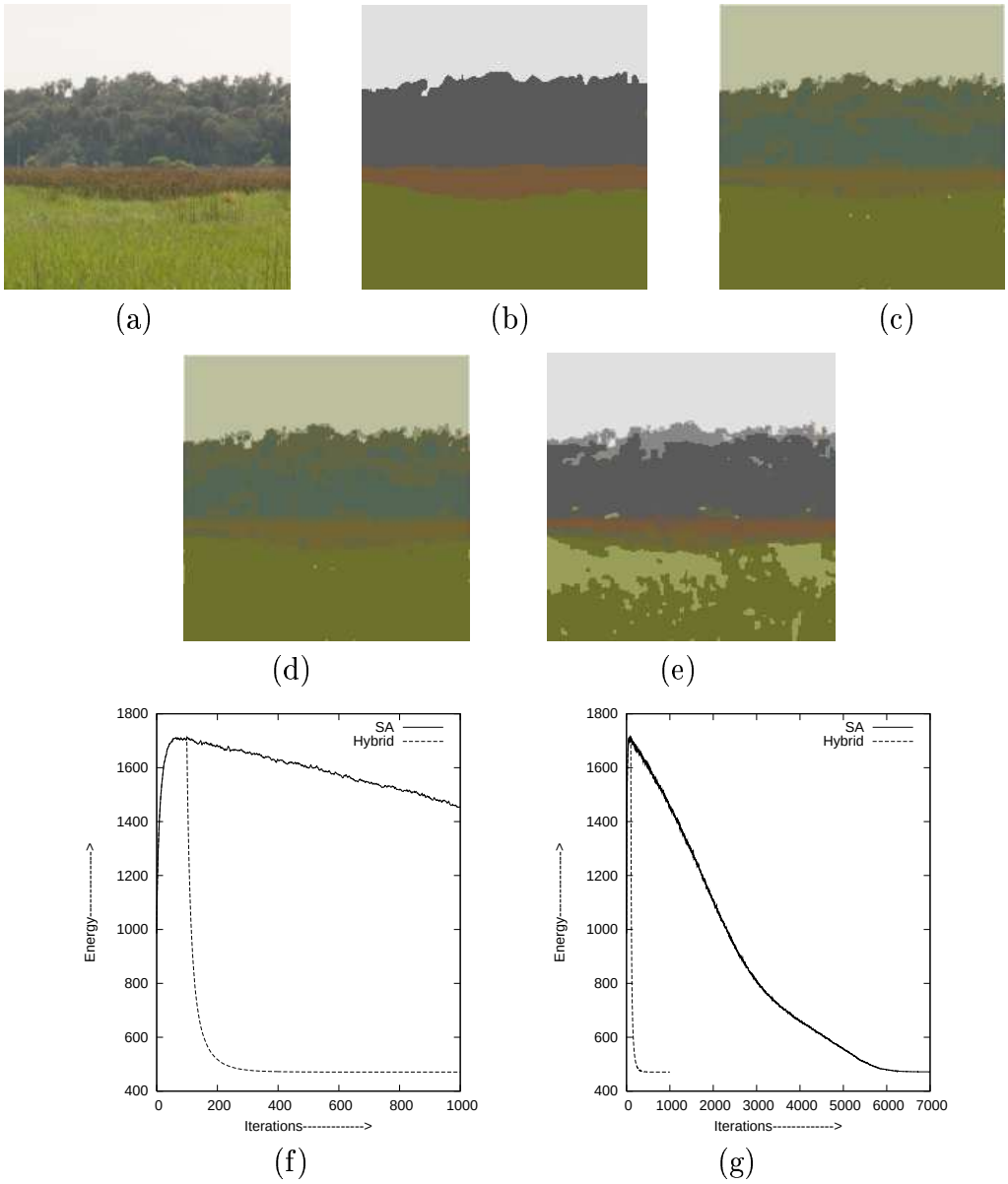


Figure 4.5: Segmentation of Scene image of size 200x200 using weak membrane model: (a) Original image (b) Ground Truth (c) CMRF optimized using SA (d) CMRF optimized using Hybrid (e) MRF optimized using SA (f) Comparison of Energy convergence curve for CMRF using SA and Hybrid for first 1000 Iterations (g) Comparison of Energy convergence curve for CMRF using SA and Hybrid for whole process

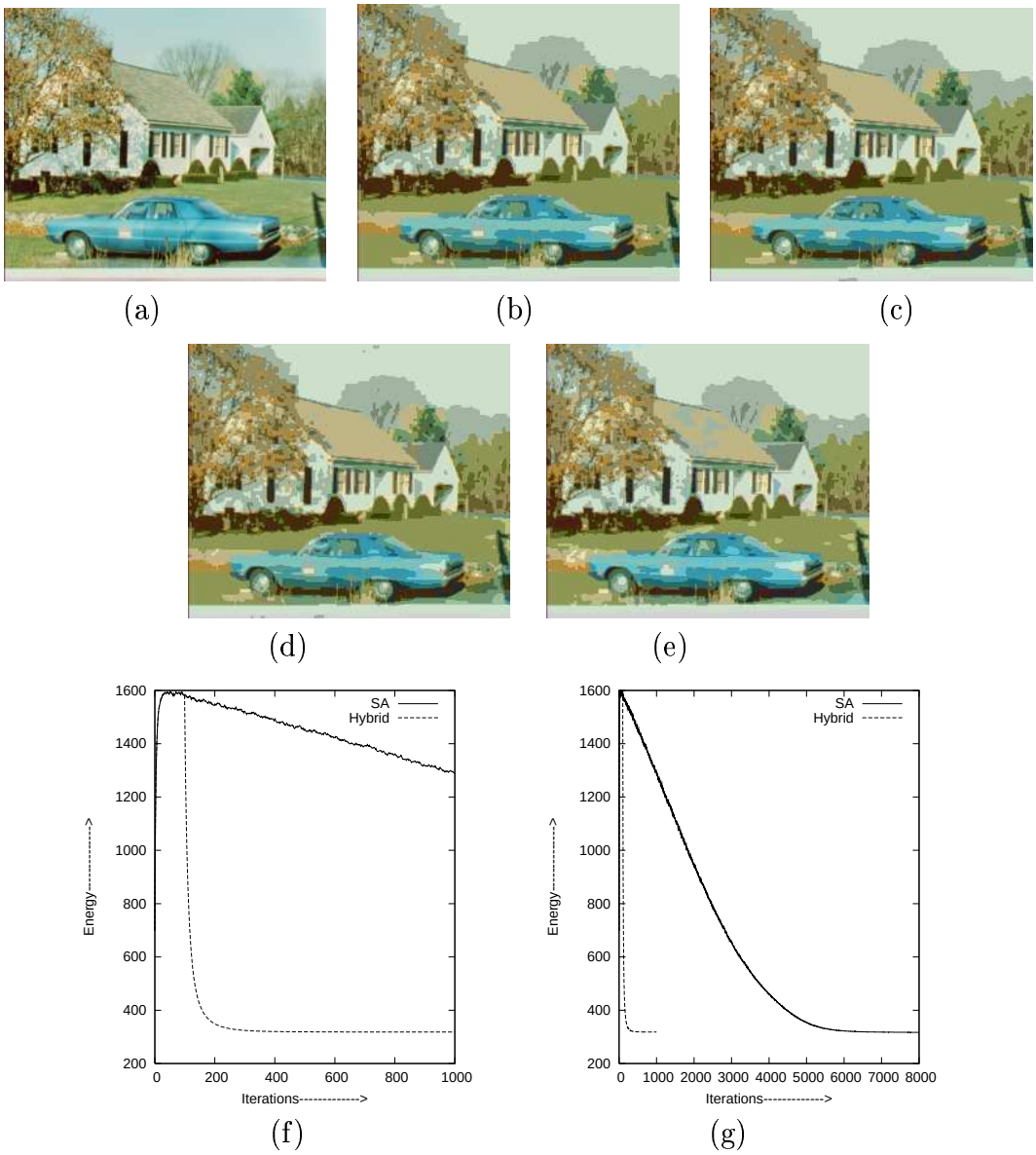


Figure 4.6: Segmentation of carhouse image of size 256x220 using weak memberane model: (a) Original image (b) Ground Truth (c) CMRF optimized using SA (d) CMRF optimized using Hybrid (e) MRF optimized using SA (f) Comparison of Energy convergence curve for CMRF using SA and Hybrid for first 1000 Iterations (g) Comparison of Energy convergence curve for CMRF using SA and Hybrid for whole process

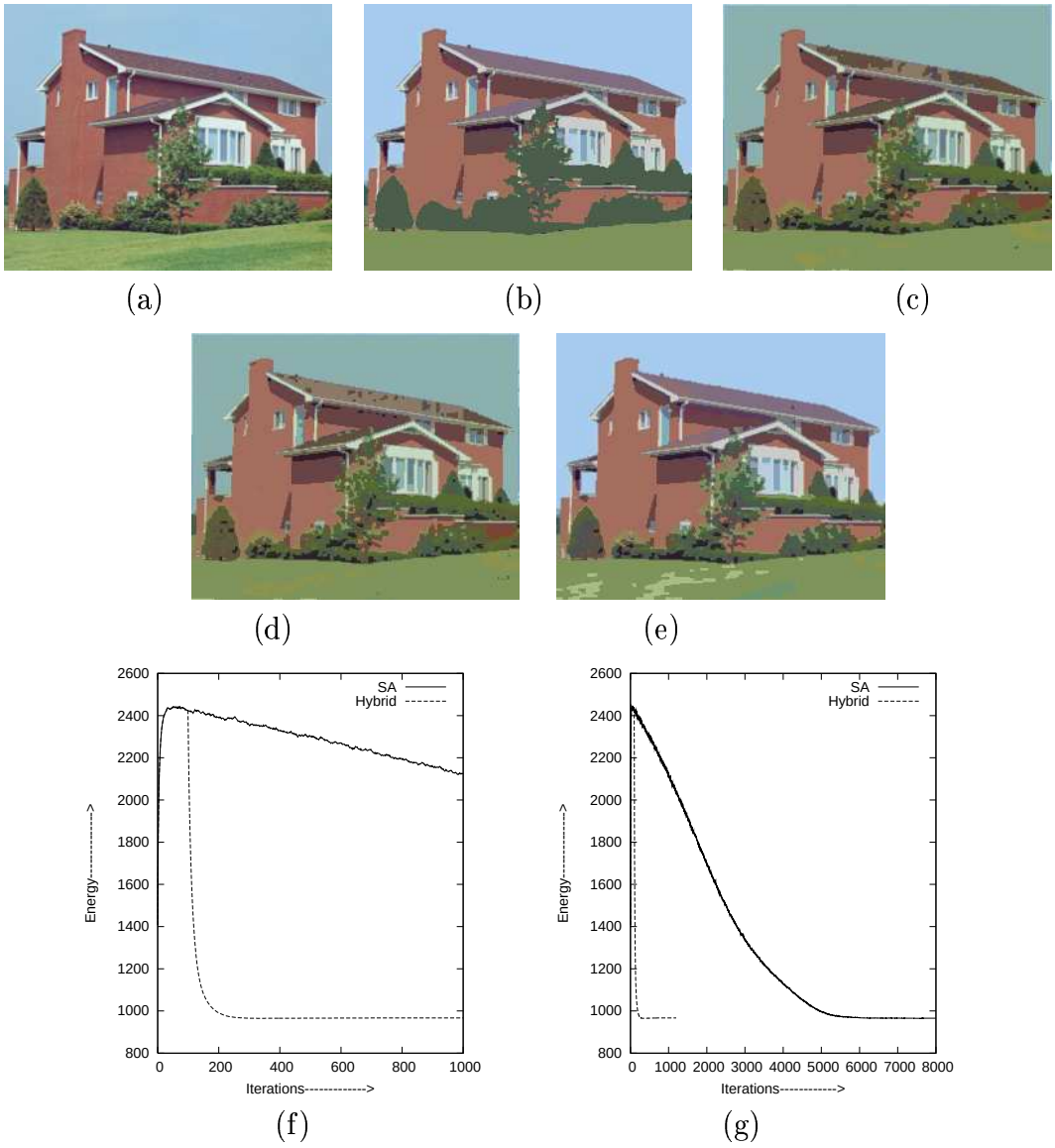
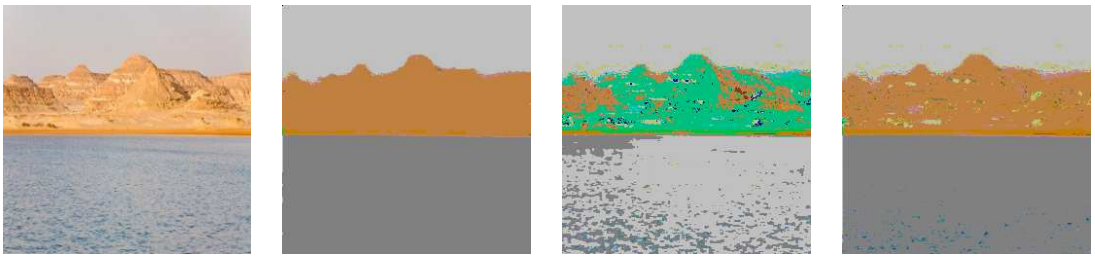
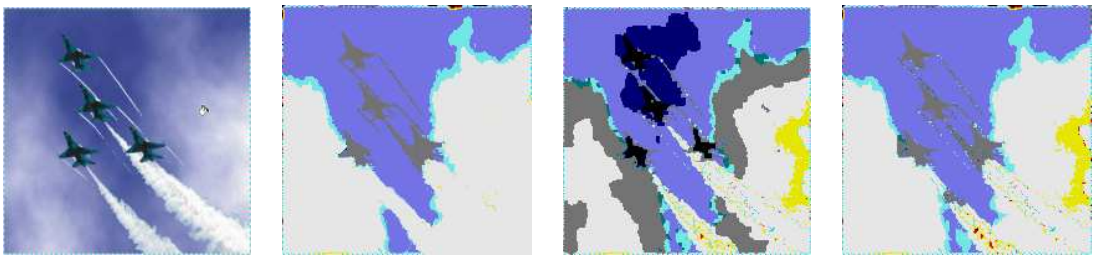


Figure 4.7: Segmentation of redhouse image of size 210x256 using weak membrane model: (a) Original image (b) Ground Truth (c) CMRF optimized using SA (d) CMRF optimized using Hybrid (e) MRF optimized using SA (f) Comparison of Energy convergence curve for CMRF using SA and Hybrid for first 1000 Iterations (g) Comparison of Energy convergence curve for CMRF using SA and Hybrid for whole process



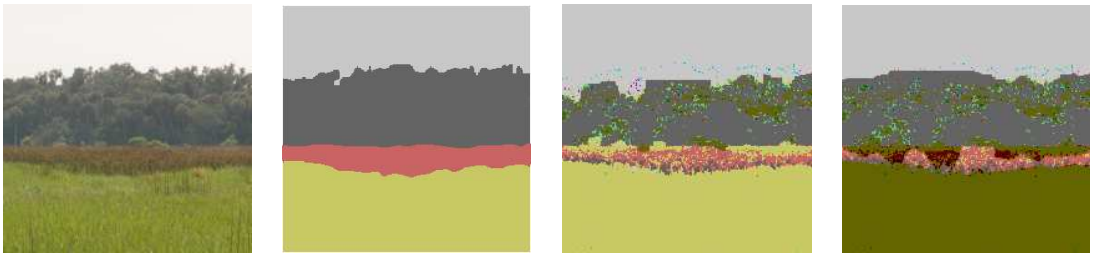
(a) (b) (c) (d)

Figure 4.8: Segmentation of Hillriver image of size 256x256 using weak memberane model: (a) Original image (b) Ground Truth (c) MRF with Hybrid algorithm (d) DMRF with Hybrid algorithm



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

Figure 4.9: Segmentation of Jet fighters plane image of size 200x200 using weak memberane model: (a) Original image (b) Ground Truth (c) MRF with Hybrid algorithm (d) DMRF with Hybrid algorithm

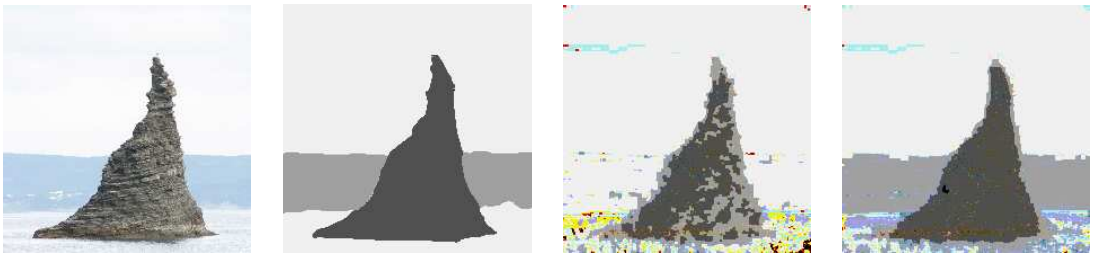


(a) (b) (c) (d)

Figure 4.10: Segmentation of Scene image of size 200x200 using weak memberane model: (a) Original image (b) Ground Truth (c) MRF with Hybrid algorithm (d) DMRF with Hybrid algorithm

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

Figure 4.11: Segmentation of Scene image of size 256x256 using weak membrane model: (a) Original image (b) Ground Truth (c) MRF with Hybrid algorithm (d) DMRF with Hybrid algorithm



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

Figure 4.12: Segmentation of Searock image of size 200x200 using weak membrane model: (a) Original image (b) Ground Truth (c) MRF with Hybrid algorithm (d) DMRF with Hybrid algorithm

Chapter 5

CONCLUSIONS

The problem of color image segmentation is addressed in stochastic frame work. The problem is cast as pixel labeling problem. The pixel labels are estimated using MAP estimation criteria. By and large SA algorithm has been used to obtain the MAP estimates. The bottle neck of using SA is the heavy computational burden thus a hinderence for implementation. A hybrid algorithm is proposed to reduce the computational burden. The rate of convergence is found to be around 10 times faster than that of SA algorithm. The two models, namely the weak membrane model and reward punishment model has been used for this. The segmentation accuracy is computed with referance to the ground truth image constructed manualy in each case. The image model parameter is selected on adhoc manner.

In order to improve the segmentation accuracy, a constrained MRF model is used as image model. The CMRF model is found to preserve the improve grouping attribute and hence better segmentation could be found than that of using MRF model. In this case also hybrid algorithm is used to obtain MAP esimates. The model parameters in this case also are selected on an adhoc manner. The parameter estmation problem for constrained MRF model and MRF model is worth pursuing.

It has been found that RGB model yields poor segmentation results as compared to ohta model. A new double MRF model is proposed to improve the segmentation accuracy for RGB model. The inter and intra model outperformed the MRF model. This model could be successsfully tested for a wide variety of images. Parameter estimation problem for the double MRF model is also worth pursuing.

Bibliography

[1] L. Lucchese and S.K. Mitra , "Color image segmentation: A state of the art survey," Proc. of the Indian National Science Academy, volume 67, No.2, (feb 2006): pp. 207-221.

[2] Cheng H.D., Jiang X.H., Sun Y., and Wang J., "Color image segmentation: advances and prospects," Pattern Recognition, volume 34, No.12, Dec 2001: pp. 2259-2281.

[3] Encyclopedia of imaging science and technology, 2002, volume 1, "color vision", pp. 100-122.

[4] Encyclopedia of spectroscopy and spectrometry, 2000, volume 1, "color", pp. 337-342

[5] Gonzalez R.C. and Woods R.E. Digital image processing. Singapore:Pearson Education, 2001.

[6] Stan Z. Li , "Markov field modeling in image analysis," Japan:Springer, 2001.

[7] Geman. S and Geman D, "Stochastic relaxation, Gibbs distributions, and the bayesian restoration of images," IEEE Transactions on Pattern Analysis and Machine Intelligence, volume 6, No. 6, November 1984: 721-741.

[8] Ohta. Y.I., Kanade. T. and Sakai. T.: "Color information for region segmentation". Academic Press, Computer Graphics and Image Processing. volume 13, No.6, 1980: pp. 222-241.

[9] Kato. Z., Pong. T. and Lee J.C., "Color image segmentation and parameter estimation in a Markovian framework," Pattern Recognition Letters, volume 22, No. 3-4, March 2001: pp. 309-321.

[10] Kato. Z. and Pong. T. , "A Markov random field image segmentation model using combined color and texture features," Proceedings of International Conference on Computer Analysis of Images and Patterns. (September 2001): pp. 547-554.

[11] Kirkpatrick S., Gelatt C. D., and Vecchi M.P., "Optimization by Simulated Annealing," Science, volume 220, No. 4598, 13 May 1983: pp. 671-679.

[12] Z.Kato., Pong T.C., and Qiang S. G., "Multicue MRF Image segmentation:Combining Texture and Color Features". IEEE computer Society, International Conference on Pattern Recognition , volume 21, 11-15 August 2002, 660-663.

[13] Kato Z., and Pong T.C., "A Markov random field image segmentation model for color textured images". Image and vision computing, volume 24, No. 10, Oct 2006: pp. 1103–1114.

[14] Permuter H., Francos J., Jermyn I., "A study of Gaussian mixture models of color and texture features for image classification and segmentation." Pattern Recognition, volume 39, No.4, April 2006: pp 695-706.

[15] Xu Q., Yang J., Ding S., "Color texture analysis using the wavelet-based hidden Markov model." Pattern Recognition Letter, volume 26, No.11, April 2005: pp 1710-1719.

[16] panjawani D. K. and healey G., "Markov random field models for unsupervised segmentation of textured color images." IEEE transactions on pattern analysis and machine intelligence, volume 17, No. 10, (oct 1995): pp. 939-954.

[17] Pratt. William K. "Digital Image Processing." John Wiley and Sons, 2004.

[18] K. S. Shanmugam and A. M. Breipohl,"Random signals detection, estimation and data Analysis," John Wiley and Sons, 1988.,pp.50-55,

[19] Blake. A., and Zisserman A., "Visual Reconstruction". London: MIT press, 1987.