On the applications of Gibbs random field in image processing: from segmentation to enhancement

Jiebo Luo, Chang Wen Chen, Kevin J. Parker

Department of Electrical Engineering
University of Rochester
Rochester, NY 14627

Telephone: (716)275-2112 *Fax:* (716)275-2073 *e-mail:* luo@ee.rochester.edu

ABSTRACT

The Gibbs random field (GRF) has been proved to be a simple and practical way of parameterizing the Markov random field (MRF) which has been widely used to model an image or image related process in many image processing applications. In particular, Gibbs random field can be employed to construct an efficient Bayesian estimation that often yields optimal results. In this paper, we describe how the Gibbs random field can be efficiently incorporated into optimization processes in several representative applications, ranging from image segmentation to image enhancement. One example is the segmentation of CT volumetric image sequence in which the GRF has been incorporated into K-means clustering to enforce the neighborhood constraints. Another example is the artifact removal in DCT based low bit rate image compression in which GRF has been used to design an enhancement algorithm that smooths the artificial block boundary as well as ringing pattern while still preserve the image details. The third example is an elegant integration of GRF into a wavelet subband coding of video signals in which the high-frequency bands are segmented with spatial constraints specified by a GRF while the subsequent enhancement of the decompressed images is accomplished with the smoothing function specified by another corresponding GRF. With these diverse examples, we are able to demonstrate that various features of images can all be properly characterized by a Gibbs random field. The specific form of the Gibbs random field can be selected according to the characteristics of an individual application. We believe that Gibbs random field is a powerful tool to exploit the spatial dependency in various images, and is applicable to many other image processing tasks.

Keywords: Gibbs random field, MAP estimation, image processing, image segmentation, image enhancement.

1. INTRODUCTION

Markov random field has been used to model various images in many image processing applications. As proved by the Hammersley-Clifford theorem[1], the Gibbs distribution provides us with a simple and practical way of specifying MRFs through certain potential functions. An appropriate form of the potential function will enable the Gibbs random field to be efficiently used to construct an optimal Bayesian estimation in a variety of image processing tasks. In this paper, we will present several applications in which GRF has been incorporated into specific problems to yield optimal results. With these diverse examples, we are able to demonstrate that various features of images can all be properly characterized by a Gibbs random field.

The objective of this paper is to show how versatile the Gibbs random field can be with its simple and practical way of parameterization. In general, a GRF can be described by a potential function such that the characteristics of the image are appropriately modeled. Often the choice of a specific form of the potential function depends not only on the type of image data being used, but also on the individual application of image processing techniques. The applications presented in this paper serve as examples as how a specific form of GRF can be selected according to natures of each estimation problem. The successful application of GRF to these individual estimation problems also reveals that many ill-posed inverse problems become solvable with the incorporation of GRF into their regularization processes.

In the segmentation of CT volumetric images, GRF is employed to specify spatial constraints in order to avoid mis-clustering caused by the impulse noise that is often introduced in the process of image acquisition. Without GRF based spatial constraints, the K-means clustering is not able to correctly label a pixel whose gray level is significantly different from the cluster mean due to the noise in the image. In addition, different clusters at different locations may have similar intensity appearance while same cluster may have different intensity appearance at different locations due to the inhomogeneities introduced in the imaging process. The GRF based 3D spatial constraints employed in the

0-8194-1638-X/94/$6.00

segmentation of the volumetric CT images enable us to develop an adaptive clustering algorithm capable of segmenting noisy images with spatial varying inhomogeneities. The extracted left ventricle chamber is consistent with both given image data and the left ventricle anatomy.

In the removal of artifacts in DCT coded images at low bit rate, GRF is used as an image model to distinguish artificial block boundary from image details. At low bit rate, the block-based DCT compression schemes generate artifacts known as the blocking effect in general and the ringing effect for high contrast images, such as text. In addition to a novel scheme which attempts to recover the DC components of each block from the coded ones, a special form of potential function, Huber minimax function, is used in the subsequent smoothing algorithm. A simple parameterization of the Huber minimax function allows the appropriate differentiation of artifacts from image details. The characteristics of this function enables the smoothing of the artifacts while preserving the image details. With such GRF modeling, we are able to obtain good reconstruction based on the decompressed image in terms of both visual observation and PSNR improvement.

In the 3D subband video coding, GRF is used in both encoding and postprocessing of decoded images. At the encoding end, a video sequence is decomposed into temporal subbands and each temporal subband is further decomposed into spatial subbands respectively. According to the characteristics of the high frequency subbands, a segmentation based adaptive quantization is designed to reduce the activity of these subbands while still preserve the visually significant components. By a proper parameter selection for the GRF incorporated in the segmentation, quantization of the high frequency subbands yields large homogenous regions by eliminating the isolated singular value pixels and thus achieves higher compression ratio. At the receiving end, the reconstructed images using the segmented high frequency subbands can be postprocessed with the same GRF based approach to remove the resultant impulsive noise in the reconstructed images while still preserving the image details. The elegant integration of the segmentation and the enhancement, both based on GRF, makes it possible for the transmission of a high quality video signal with high compression ratio.

Following notations will be used throughout the subsequent sections: Uppercase letters are used for random variables and lowercase letters for the corresponding realizations; A random field X will be defined on a set of sites S, i.e. a set of N×N points; A pixel at site s is denoted by $X_s \in \mathcal{R}$; Bold uppercase letter is used for matrices or transformations.

2. BAYESIAN ESTIMATION BASED ON GIBBS RANDOM FIELD

In image processing applications, many problems require the estimation of an image or other 2D field X from corrupted data Y. These inverse problems are generally ill-posed. Prior information is often very useful in forming a regularized process so that optimal results can be obtained by solving the regularized problem.

A widely used approach to these problems is the Bayesian estimation that incorporates prior information through an *a priori* distribution for the random field X. The estimation usually can be significantly improved by exploiting prior information about the characteristics of X. The prior information is often represented by a GRF which can be incorporated into the Bayesian estimation schemes. An appropriate choice of the potential function for the GRF would generally enable the estimation to be implemented efficiently.

2.1. General form of Gibbs random field

Gibbs distribution is an explicit expression of the MRF in the form of

$$g(x) = \frac{1}{Z} \exp\left\{ -\sum_{c \in C} V_C(x) \right\} \tag{1}$$

where Z is a normalization constant and V_C is certain clique potential for clique c. In image processing applications, clique c is a local group of pixels while C is the set of all such local groups. The neighbors of s is denoted by $\partial s \subset S$ where S is the set of all sites. If c is defined such that $\forall s, r \in c$, s and r are neighbors, then the clique has the property that

$$\forall s, r \in S \ , \ s \notin \partial s \ , r \in \partial s \Leftrightarrow s \in \partial r \ . \tag{2}$$

Moreover,

$$\forall s \in S,\ p(x_s \mid x_r,\ r \neq s) = p(x_s \mid x_r,\ r \in \partial s) \tag{3}$$

is the essential property of an MRF upon which we build the image model.

Recently, because of the equivalence between MRF and GRF, many researchers have focused on MRF's in the form of the general Gibbs distribution[2]

$$\log g(x) = -\sum_{\{s,r\} \in C} b_{sr}\rho(|x_s - x_r|) + constant \tag{4}$$

where λ is a scaling factor and $\rho(.)$ is preferably monotonous. In the following, we will discuss two major types of potential functions: non-convex and convex potential functions.

2.1.1. Non-convex potential functions

Many non-convex potential functions used in GRF based image modeling are of simple forms. One such function used to represent image continuity is written as

$$V_c(x) = \begin{cases} \beta, & \text{if } x_s = x_t \text{ and } s,\ t \in c \\ -\beta, & \text{if } x_s \neq x_t \text{ and } s,\ t \in c\ . \end{cases} \tag{5}$$

Note that here x_t and x_s represent the clusters to which pixels t and s belong, c represents a designated neighborhood system. This function is very effective, especially when employed in the segmentation algorithms[3]. By switching between the parameters β and $-\beta$, this potential function can be used to enforce desired spatial constraints if appropriate neighborhood system c and parameter β are chosen.

Another nonconvex function[4] is given as

$$\rho(\Delta) = \min\{|\Delta|,\ T\} \tag{6}$$

where Δ represents the difference of gray level values between two pixels and T is a parameter such that the equal penalty region beyond T allows sharp edge to be preserved. However, this function only belongs to $\mathbf{C}^0$ and the transition is not smooth. The theoretical and practical disadvantages may lead to some unnatural results, e.g. edges of magnitude larger than a threshold are sharp yet those of lower magnitude are smooth.

2.1.2. Convex potential functions

Convex functions are often chosen for practical as well as theoretical reasons. A convex constrained optimization problem is usually desired because there exists a unique, stable solution to such problem and it can be optimized efficiently. In addition, convex functions with smooth transition result in desired continuity. Some examples are shown in Figure 1.

•Generalized Gaussian Markov random field

It has been proved[2] that a proper model of scale-invariant property is of the form

$$\rho(\Delta) = |\Delta|^p \tag{7}$$

where Δ also represents the difference of gray level values between two pixels. This is called Generalized Gaussian Markov random field (GGMRF) with parameter p controlling the character of the GRF. Large values of p tend to smooth the discontinuities while smaller ones tend to preserves them. The derivative of $\rho(.)$ represents the attraction between two pixels with gray levels separated by Δ and is called the influence function. Influence function is also an indication of image smoothness. Generally, $\rho(.)$ belongs to $\mathbf{C}^2$ except when p is 1. Note that p=2 reduces the model to Gaussian Markov random field (GMRF). Because of the analytical advantages, GMRF has been widely used. However, the linear lowpass filtering nature of GMRF tends to blur the image edges and other details excessively and indiscriminately. This is due to the quadratic term which grows too quickly with respect to the difference of pixel values and therefore imposes excessive penalty to edges.

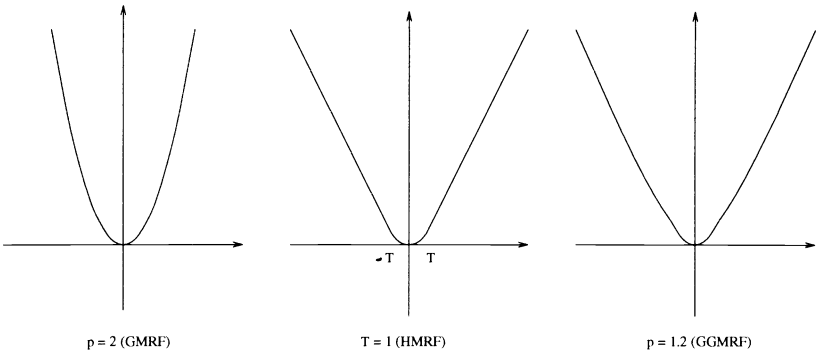

Figure 1: Some potential functions

•Huber Markov Random Field

Non–Gaussian MRFs are of great interest because they can potentially model different contents and features in an image. One such convex function known as Huber minimax function has been investigated.

$$\rho(\Delta) = \begin{cases} \Delta^2, & |\Delta| \leq T \\ T^2 + 2T||\Delta| - T| & |\Delta| > T \end{cases} \tag{8}$$

For Δ's greater than T, the linear segment of the function imposes lighter penalty and thus allows sharp transitions such as edges. Note that Huber function is in $\mathbf{C}^1$, so the influence function is continuous. The MRF characterized by Huber minimax function is referred to as Huber Markov random field (HMRF)[5] (see Figure 1). The parameter T controls not only the threshold to switch from quadratic to linear, but also the slope of the linear weighting. This is due to the $\mathbf{C}^1$ continuity imposed by such convex function.

2.2. Maximum *a posteriori* estimation

Maximum *a posteriori* (MAP) has been a powerful estimation tool based on the observed image data Y and a reasonable *a priori* distribution of the 2D random field Z. An MAP estimation can be written as:

$$\hat{z} = \arg\max_{\mathbf{Z}} p(z \mid y) \,. \tag{9}$$

Using Bayes' rule, the *a posteriori* probability can be expressed as

$$p(z|y) \;\propto\; p(y|z)\, p(z) \;. \tag{10}$$

The *a priori* distribution $p(z)$ is often modeled by a Gibbs random field which can be characterized by a neighborhood system and a potential function. The optimization can also be conveniently expressed using the log-likelihood function

$$\hat{z} = \arg\max_{\mathbf{Z}} \{\log p(y \mid z) + \log p(z)\} \;. \tag{11}$$

Note that, without the *a priori* distribution $p(z)$, the estimation scheme becomes a maximum likelihood estimation (MLE). MLE is often used as the initial estimate in the iterative MAP estimation. Various distributions of $p(y|z)$ and $p(y)$ have been proposed for different types of applications. However, all MAP estimation schemes share the same principles of optimization.

2.3. Optimization technique: iterative conditional mode (ICM)

The *local* characteristics of an image is combined with the given image data to construct an MAP estimation of the original image. Certain large-scale characteristics of the model is often induced in choosing a non-degenerate field to describe the local properties of the scene[6]. In general, Markov random field exhibits positive correlations over arbitrarily large distances when adjacent pixels are very likely to be of same color or intensity. To avoid such large-scale characteristics is one of the major concerns in choosing the optimization techniques.

Some of the optimization techniques may lead to undesired computation and convergence difficulties[1]. Simulated annealing, while it can guarantee the convergence to the global optimum, is computationally demanding and may be impracticable in many applications. Gradient descent is less computationally demanding but it can not guarantee the convergence to the global optimum. The computation burden using both these techniques is enormous while the reconstruction may suffer from some undesired large-scale properties of the random field because of the *simultaneous* optimization of the objective function.

Fortunately, the difficulties in the GRF based optimization processes can be overcome by selecting an optimization technique known as *Iterative Conditional Mode* (ICM)[6]. This method is computationally inexpensive and invulnerable to the large–scale characteristics. It was first proposed as an approximation to the Bayesian estimation that has overall maximum probability and later established as a distinct optimization method in its own right yielding an estimation that has maximum probability at each individual pixel. A single cycle of ICM only requires the *successive* minimization of the objective function at each pixel. Note that each pixel has only a few neighbors and is highly restricted by the consistency constraint identified by the H-C theorem. As Besag pointed out, the dependence of the estimation on only the local characteristics is ensured by the rapid convergence[6] of the ICM implementation and the undesired large-scale characteristics is minimized.

3. SEGMENTATION OF CT VOLUMETRIC CARDIAC IMAGES

The 3D image data used in this research are a sequence of CT volumetric data of 16 volumes each contains 95 90×90 slices. Each volumetric element, or voxel, represents a 0.9 $(mm)^3$ cube of tissue. To bring out the left ventricle chamber as a bright object, a Roentgen contrast agent is injected into the right atrium several seconds prior to the scanning of the heart. The left ventricle chamber appears in the CT volumetric images as a large, bright, smooth, solid region, varying in size and shape over time, approximately attached to the left atrial chamber and aorta through the valves, and separated from the myocardium by a strong, but blurred and noisy, interface.

3.1. Adaptive K-means clustering with GRF based spatial constraints

Traditional statistical image segmentation algorithms, from thresholding to K-means and even Fuzzy K-means clustering, all classify the pixels into clusters based only on their intensity values. Each cluster is usually characterized by a constant intensity and no spatial constraint is imposed. In practice, images are usually noise contaminated version of the reflected density function, and the image intensity of the same class may change over space due to some physical constraints of the imaging system. In many biomedical applications, even though the relative intensity is evident for different clusters within a small neighborhood, different clusters at different locations may have similar intensity appearance due to the inhomogeneous of the imaging media. Traditional K-means fails here because of the low signal-noise-ratio in these CT images. The ability of being adaptive to the local intensity distribution is generally required for a robust image clustering algorithm to obtain the correct clustering results. In addition, certain spatial constraints are needed to avoid mis-clustering caused by the noise introduced in the imaging process, since a pixel generally tends to belong to same cluster as most of its neighbors unless it is on the edge of a sharp region transition.

The proposed adaptive K-means algorithm is based on the segmentation algorithm proposed recently by Pappas[3]. His algorithm includes the 2D spatial constraints characterized by Gibbs random fields, and the adaptive estimation of local means of each region. We have extended Pappas' algorithm in two important aspects. We have developed 3D spatial constraints to suit the volumetric nature of the image data. We have also enhanced the adaptive clustering algorithm to account for the varying characteristics of both the local means and local variances.

3.1.1. Incorporation of Gibbs random field

In image segmentation, the goal is to produce a robust labeling of each pixel. Therefore, a non-convex potential function defined by Equation (5) is suitable to enforce the spatial continuity of the labeled regions. For a 2D image defined on the Cartesian grid, a simple neighborhood system of a pixel is represented by its 4 nearest pixels. For a 3D image, an extension of 2D neighborhood system indicates that the neighborhood of a voxel can be represented by its 6 nearest neighbors. If we model the conditional density as a Gaussian process with spatially varying mean μ_s and

variance σ_s at a pixel location s, and denote a given image by y and a segmentation of this image by x, then the overall *a posteriori* probability function will be:

$$p(x \mid y) = \exp\left\{ -\sum_s \frac{1}{2\sigma_s^2}(y_s - \mu_s)^2 - \sum_c V_c(x) \right\} \tag{12}$$

The first term corresponds to the adaptive capability that forces the segmentation to be consistent with local image distribution with locally estimated mean μ_s and variance σ_s. The second term corresponds to the spatial continuity constraint characterized by the clique potentials within a given 3D lattice.

In biomedical image segmentation tasks, the *a priori* knowledge of the structure–of–interest is usually available since we often study certain biomedical structures with known anatomical information. The anatomical information is then used in the design of K-means clustering to set the value K and to parameterize the GRF. In the case of CT volumetric image data, K is set to 4 according to the available knowledge of the cardiac structure with the brightest cluster corresponding to potential left ventricle chamber. We assign the same β to the clique potentials both within a cross section and between cross sections, since the sampling lattice of the CT volumetric data is uniformly structured. β is also set to be inverse proportional to the variance of each cluster because it makes sense that a cluster with large variance needs a large β to bind its samples together.

3.1.2. The segmentation and beyond

The proposed adaptive clustering algorithm applied to the CT volumetric data is implemented using ICM. First, an initial segmentation x is acquired through the simple K-means algorithm. Then, overall probability function is maximized on a point-by-point basis, with the μ_s and the σ_s being updated after each iteration. Therefore, the optimization is accomplished through alternating between MAP estimation of the clustered regions and iterative update of the cluster means and variances. Such alternating process is repeated until no pixel changes classes.

There exist a number of differences between this algorithm and that of Pappas. One important difference is the introduction of iterative estimation of cluster variances in the process of optimization. The assumption of the changing variance and the implementation of estimation scheme allow us to account for the noise levels to change from one local area to another, and from one cluster to another. This additional feature of the proposed scheme enhances the flexibility of the adaptive K-means clustering algorithm, since, in practice, the variances of different clusters are generally different and the variance of a specific cluster also changes with location. Second one is the choice of the parameter of the Gibbs random field. It is evident that the parameter is related to the image contents as well as imaging conditions. According to biomedical structure and known imaging condition, we are able to choose the parameter β such that the spatial constraint is strong enough to smooth out the noise while still preserving the structural details. Upon the completion of the adaptive K-means clustering, subsequent processing may be necessary if the given biomedical images contain certain structures that are anatomically separate but statistically indistinguishable. In the case of cardiac images, we have also designed a knowledge-based morphological operations in order to distinguish the left ventricle from left atrium and aorta. The discussion of such operations is beyond the scope of this paper[7].

3.2. Results and discussion

We have successfully applied the proposed segmentation algorithm to the volumetric cardiac images. The volumes of left ventricle extracted using this approach compare favorably with those obtained using traditional K-means method (Figure 2). It can be seen that the segmented regions are not be as homogeneous if the GRF based spatial constraints are not enforced. In addition, the 3D spatial constraints also help to propagate structural information from a slice to its neighboring slices, allowing the algorithm to handle the spatial varying gray level distributions.

This automatic extraction is fast, reproducible, without operator bias and suitable for further processing and analysis. In addition, the temporal changes of the shape of the biomedical object will undoubtedly provide another constraint which can be used to better resolve the image ambiguity.

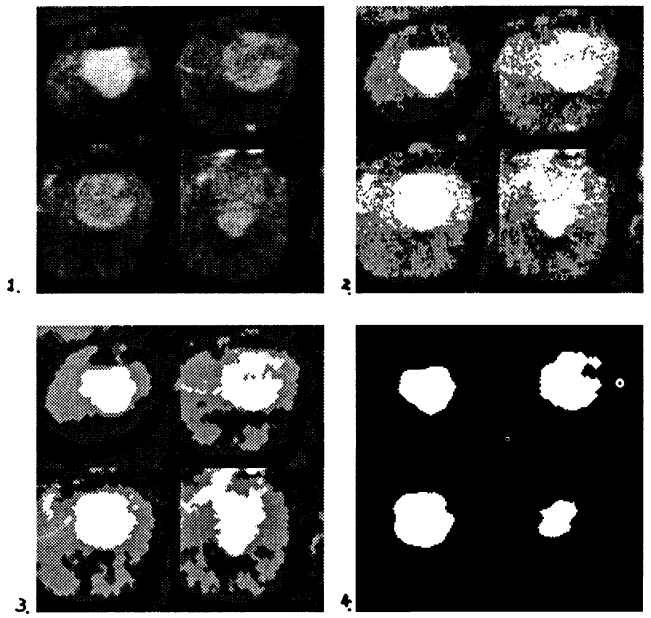

Figure 2: Comparison of the segmentation results: 1. original CT images;
2. K-means segmentations; 3. Adaptive K-means segmentations; 4. Final segmentations (left ventricle).

4. ARTIFACTS REMOVAL IN LOW BIT RATE DCT-BASED IMAGE CODING

Block-based Discrete Cosine Transform has been the most popular transform in a variety of image and video compression applications. In low bit rate applications, high compression ratio is desired and usually achieved by coarse quantization and truncation of the high frequency coefficients that are considered visually less significant[8, 9]. Consequently, two major artifacts known as blocking effect and ringing effect are generated and they severely degrade the image quality.

One major artifact is the blocking effect which appears as perceptible rectangular block structures in the reconstructed image. Several techniques have been proposed to remove blocking effect[10]. Many filtering techniques are essentially applying smoothing algorithms to the blocky image which also smooth the original image details. To avoid oversmoothing of the edges, a scheme was proposed by first estimating the edge segments in the compressed image before smoothing[11]. However, the estimation is very challenging, especially in the case of low bit rate coded images when it is very hard to differentiate true edges from the artifacts. Other techniques attempt to formulate the removal of blocking effect as a constrained image restoration problem[12, 13, 14]. Another type of artifact is the ringing effect which appears as ringing pattern around sharp edges in the image. In order not to blur these images, an edge-preserving nonlinear filtering is desired. However, there has been little investigation on this aspect until recently[15].

Our artifacts removal technique is based on the convex constrained restoration with GRF model. The Huber minimax function is chosen to distinguish artificial block boundary from image details. With GRF, in particular the HMRF, we are able to devise an enhancement technique suitable for the artifact removal in low bit rate coded images.

4.1. Convex constrained restoration with HMRF model

In general, the DCT based image coding can be modeled as

$$Y = \mathbf{Q}[\mathbf{H}(Z)] \tag{13}$$

where Z is the original image, $\mathbf{H}$ is the unitary DCT transform, $\mathbf{Q}[.]$ is a scalar quantization operation. Y now consists of quantized coefficients which generally need fewer bits to represent. Note that $\mathbf{Q}[.]$ is a many-to-one mapping as well

as the source of distortion in the compressed image. The conditional probability of the compressed Y given the original image Z, $P(Y|Z)$, is a uniform density and can be written as

$$P(Y \mid Z) = \begin{cases} 1, & Y = \mathbf{Q}[\mathbf{H}Z] \\ 0, & Y \neq \mathbf{Q}[\mathbf{H}Z] \end{cases} \tag{14}$$

The Huber-Markov random field model has been utilized to formulate a convex constrained restoration problem. Such technique has been shown able to smooth the artificial discontinuity across the block boundary[13] while preserving the remaining details of the original image. Since the conditional density is of the form given in Equation (14), the optimization is simplified to

$$\begin{aligned} \widehat{Z} &= \arg\min_{Z \in \mathcal{Z}} \sum_{c \in C} V_c(Z) \\ &= \arg\min_{Z \in \mathcal{Z}} \sum_{1 \le m,n \le N} \sum_{k,l \in c_{mn}} \rho(z_{mn} - z_{kl}) \end{aligned} \tag{15}$$

where $\rho(.)$ is a proper Huber minimax function as given in Equation (8), c_{mn} is the 8-neighborhood of the current pixel at (m,n) and N is the dimension of the image, $\mathcal{Z}$ is the constraint space

$$\mathcal{Z} = \{Z : \ y = \mathbf{Q}[\mathbf{H}Z]\} \ . \tag{16}$$

The parameter T in the potential function $\rho(.)$ controls how much discontinuity is allowed. Below the threshold, the quadratic term represents a least square fit smoothing of the areas with relatively similar intensities. If the difference is above the threshold, a linear cost function is used to preserve the discontinuity of the original image.

4.2. Implementation of HMRF based restoration

A major advantage of the HMRF in the restoration over other type of GRF is its ability to switch the weighting according to the difference of gray level between current pixel and its neighbors. However, this switch property is still inadequate when we need to distinguish image details from the artifacts. A single value of T can not accurately describe all the discontinuities, including both artifacts and true image edges. In DCT-based coding, the mechanism that generates the artifacts and the locations of these artifacts are known. We can use these information to develop a variation of the HMRF model. The discontinuity inside each image block is produced in a different way from those along the block boundary regions. Therefore, two kinds of discontinuities should be treated separately. In this research, larger parameter $T1$ is chosen for the local HMRF model for those pixels in the boundary regions in order to smooth the artifacts, while a moderate $T2$ is applied to the inner block regions[16].

Because of the convexity of Huber function, a convex constrained restoration problem is formed based on MAP estimation and the received data. The MAP estimation produces a smoothed update of the initial image obtained using standard decompression. Then, the estimated result is projected back to the constraint space by forcing the coefficients fall into the original quantization interval. The projected image is then obtained by taking inverse DCT of the projected coefficients. Improvement of image quality is obtained through the iterative ICM reconstruction.

4.3. Results and discussions

We adopt a specific DCT coding scheme[9] in all the experiments leading to Table 1. The coding of DCT coefficients is done by first applying zonal sampling and then quantization using uniform quantizers. We use this quantization scheme because it is able to produce more severe blocking effect and ringing effect than other tables[8, 12] while still retaining relatively high PSNR at the same bit rate.

We have applied our approach to two groups of test images[16]. Group 1 consists typical gray scale images and is used to verify the capability of this algorithm in reducing the blocking effect. Group 2 consists of high contrast images to verify the ability of the algorithm in reducing the ringing effect. By using ICM to implement the smoothing algorithm, the localized spatial constraint can be efficiently enforced. With $T1$=5 and $T2$=10, this algorithm usually reaches convergence within 10–20 ICM iterations. Table 1 shows the bit-rate and the corresponding PSNR for each test image. The PSNR improvement is significant in both two groups, and the preservation of image details is evident using this HMRF based restoration. Note that this algorithm succeeded in the very challenging image "resolution chart" which contains details in a wide range of scales. Without GRF based filtering, linear lowpass filtering will blur the image details and even cause the degradation of PSNR.

image	bitrate (bpp)	DCT (dB)	enhanced (dB)	image	bitrate (bpp)	DCT (dB)	enhanced (dB)
lena	0.30	27.61	28.35	boats	0.35	27.39	27.77
peppers	0.30	27.69	28.75	text	0.43	16.59	19.48
USC girl	0.20	30.03	31.25	chart	0.80	23.21	27.76

Table 1 : The PSNR evaluation of the results

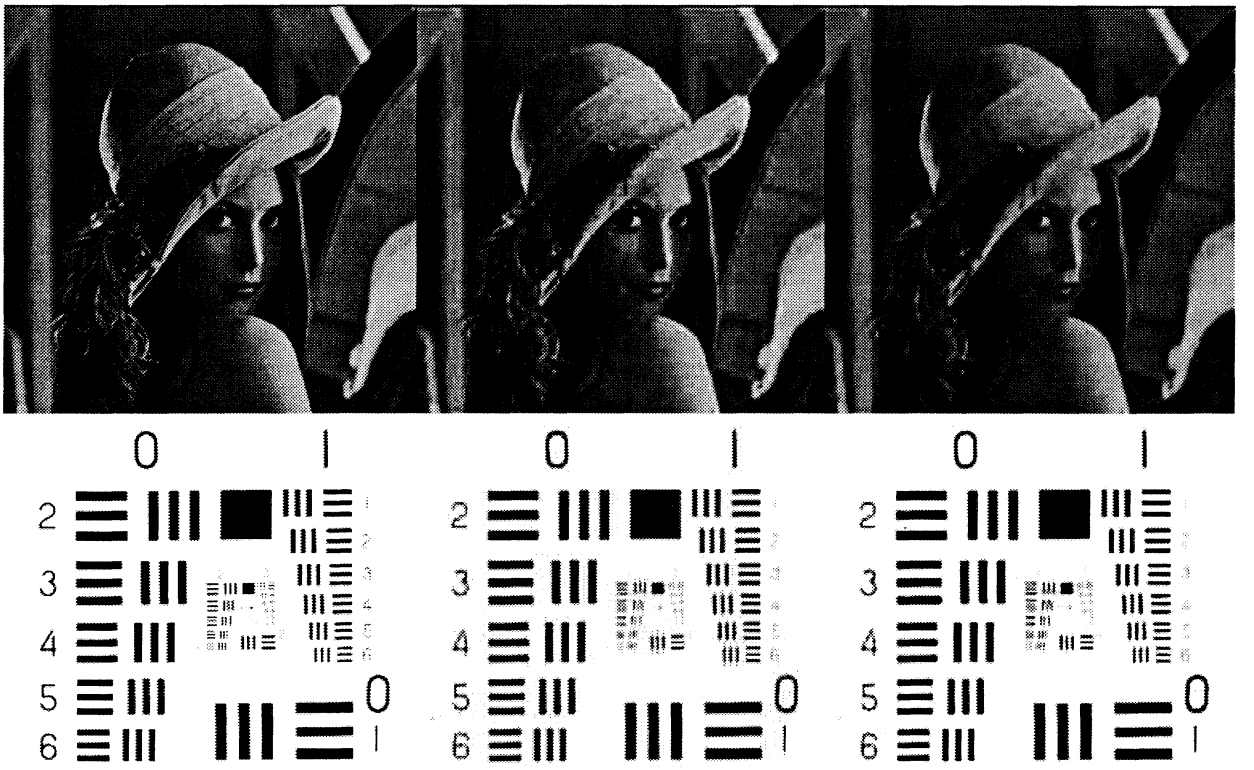

Figure 3: Comparison of the artifacts reduction: 1. original "lena" image; 2. standard decompression; 3. enhanced decompression; 4. original "resolution chart" image; 5 standard decompression; 6. enhanced decompression

5. 3D SUBBAND VIDEO CODING WITH SEGMENTATION IN HIGH FREQUENCY SUBBANDS

Video coding and communication has received great attention recently. We developed a novel scheme based on a 3D subband decomposition of video signal targeting the ISDN applications such as videoconferencing at 384 kbps bit rate. The advantage of 3D subband coding is that it avoids motion estimation and motion compensation, and does not generate visually annoying "blocking effect" even at low bit rate.

In 3D subband coding, a video sequence is decomposed into temporal subbands and each decomposed temporal subband is further decomposed into spatial subbands respectively[17]. We adopted an 11-band tree-structured decomposition scheme for video signals[18]. To minimize the computational burden of the temporal filtering needed to process the video signal, temporal decomposition is based upon the 2-tap Haar filterbank. The temporal filtering results in two subbands: high pass temporal (HPT) band and low pass temporal (LPT) band. Spatial decomposition, both horizontal and vertical, is based on multi-tap perfect reconstruction wavelet filterbanks[19]. To achieve potential high compression rate, the lowest frequency band is further decomposed in a tree structure. In our scheme, HPT band is decomposed into 4 spatial subbands, and LPT band is decomposed into tree-structured 7 spatial subbands.

5.1. GRF based segmentation as adaptive quantization

Upon subband decomposition, each subband would exhibit certain distinct features due to the characteristics of frequency responses of a particular class of bandpass filters. To best exploit these features, we developed corresponding coding strategies: baseband is the low resolution representation of the original image and has similar characteristics in histogram, but its bandwidth has been significantly reduced. It can be efficiently coded using DPCM or DCT. The coding of the high frequency subbands is critical to the overall coding performance if both high compression ratio and high quality are expected.

High frequency subbands are sparse and highly structured. Many existing schemes using conventional quantization methods have tried to exploit the characteristics of high frequency subbands[19, 20, 21], but found it difficult to code the impulse like pixels that often appear in higher frequency subbands.

We proposed an adaptive quantization based on K-means clustering with GRF as spatial constraints. The segmentation of the high frequency subbands and the representation of each pixel by the cluster mean is equivalent to a data-adaptive quantization process. However, because of the spatial constraints, we are able to generate larger homogeneous regions by forcing those impulse like pixels to be of the same cluster as their neighbors. Such spatial constraints can be adjusted to each resolution level and preferential orientation in order to preserve the visually significant structures in these subbands. The isolated impulse like pixels, which would otherwise require considerable bits to code, are eliminated in the process of adaptive segmentation through the incorporation of Gibbs random field spatial constraint. The compression ratio of these segmented high frequency subbands can be greatly increased because of the reduced entropy due to the smoother spatial distribution of each cluster contained in these subbands.

5.2. Coding and synthesis with postprocessing

Different scanning scheme can be performed for individual subband to increase the runlength coding efficiency since these segmented high frequency subbands are composed of well defined "edges" whose directions usually correspond to the directions of the highpass filters. Another scheme of increasing the runlength is to partition the subbands into non-overlapping blocks. Through such partition, local area of zero values can be better exploited to improve the coding efficiency[20]. Hilbert-Peano scan is also proper for this purpose. The performance of different coding schemes with respect to the qualities of the reconstructed image and the bit rate is currently under investigation.

The reconstructed images from these segmented high frequency subbands would generally contain impulse like noise because of the previous removal of isolated impulse like pixels due to spatial constraints. Fortunately, a corresponding GRF can be incorporated in the postprocessing to remove the reconstruction noise while preserving the image details. The principle of such enhancement is similar to the artifacts removal scheme described in Section 4.

Preliminary results are shown in Figure 4 and 5. The average back-to-back reconstruction PSNR with only 5-level segmentation is about 33 dB. Both subjective observation and objective measurement show the promise of this method in high quality video communication[22]. This is also an excellent example demonstrating that GRF can be integrated into both subband analysis and subsequent synthesis of video signals, and therefore is versatile for many different applications.

6. CONCLUSION

We have addressed the application of Gibbs random field in several image processing problems. Through these diverse applications of Gibbs random field, the intrinsic connection between seemingly different image processing problems is revealed. Although implementation issues may differ from one case to another, Gibbs random field provides a general form to characterize various spatial dependency and localized features in images. The applications presented in this paper also serve as examples as how a specific form of GRF can be selected according to the nature of an estimation problem. Many ill-posed inverse problems become solvable with the incorporation of Gibbs random field into their regularization processes.

Gibbs random field can be versatile with its simple and practical way of parameterization. We believe that Gibbs random field is a powerful tool to exploit the spatial dependency in various images, and is applicable to many other image processing tasks.

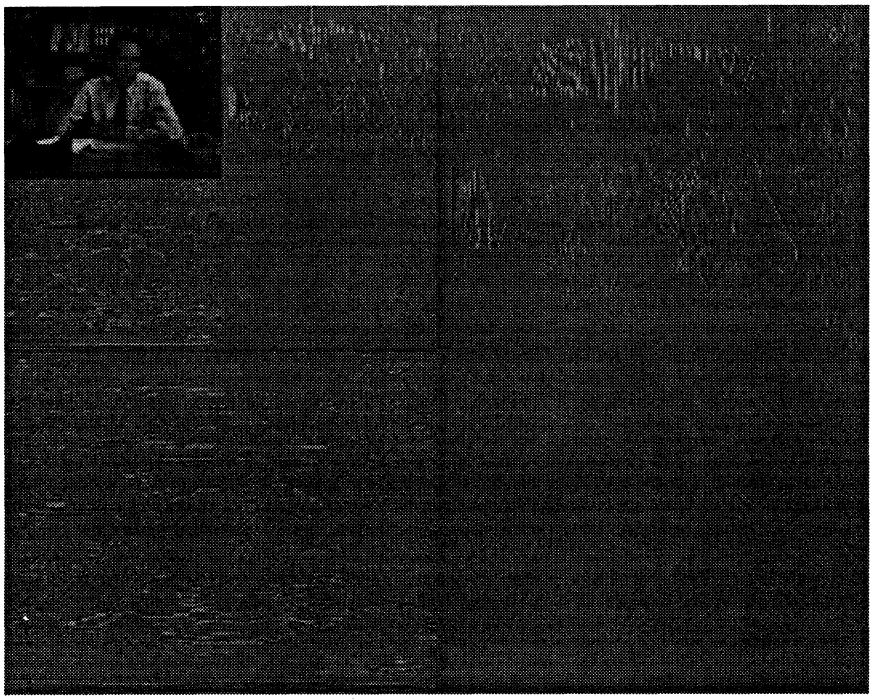

Figure 4: 7-band decomposition of LPT band of a typical "salesman" frame (high frequency subbands segmented)

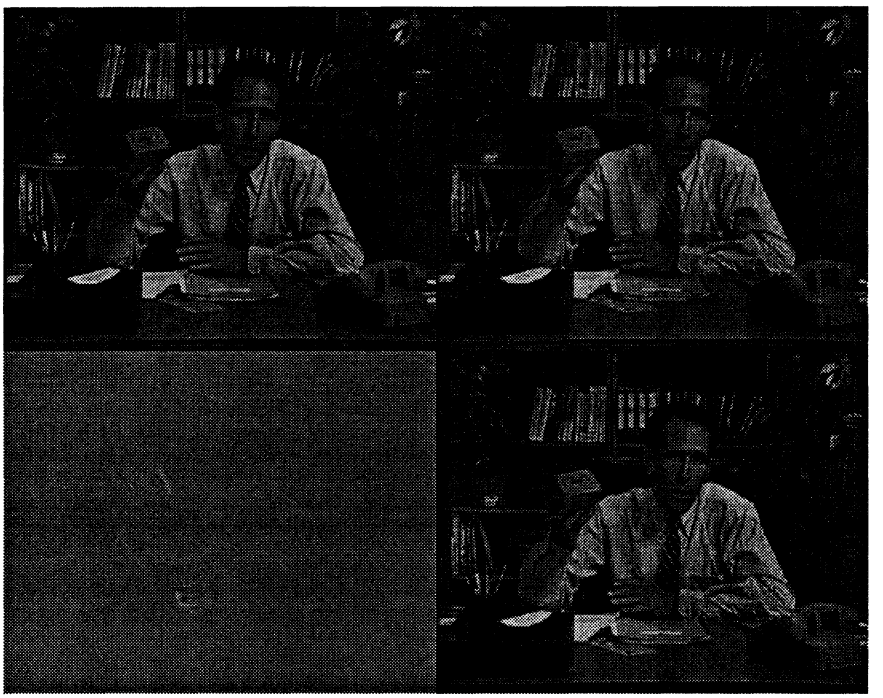

Figure 5: Upper-left: original frame; Upper-right: LPT synthesis after segmentation;
Lower-left: HPT synthesis after segmentation; Lower-right: overall reconstruction. (all half-sized)

References

[1] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distribution, and the Bayesian restoration of images," *IEEE Trans. Patt. Anal. Machine Intell.*, vol. 6, pp. 721–741, 1984.

[2] C. Bouman and K. Sauer, "A generalized gaussian image model for edge-preserving map estimation," *IEEE Trans. Image Processing*, vol. 2, pp. 296–310, July 1993.

[3] T. Pappas, "An adaptive clustering algorithm for image segmentation," *IEEE Trans. Signal Proc.*, vol. 40, pp. 901–914, 1992.

[4] R. Kinderman and J. L. Snell, *Markov Random Feilds and Their Applicantions*. Providence, RI: Amer. Math. Soc., 1980.

[5] R. L. Stevenson, "A nonlinear filtering structure for smoothing discontinuous signals corrupted with gaussian noise," in *Proc. SPIE Nonlinear Image Processing III*, vol. 1658.

[6] J. Besag, "On the statistical analysis of dirty pictures," *Journal of the Royal Statistical Society*, vol. 48, pp. 259–302, 1986.

[7] C. W. Chen, J. Luo, and K. J. Parker, "3D segmentation via adaptive k-mean clustering and knowledge-based morphological operations," submitted to *IEEE Trans. Image Processing*, February 1994.

[8] W. B. Pennebaker and J. L. Mitchell, *JPEG: still image data compression standard*. Bellingham, WA: SPIE Optical Engineering Press, 1993.

[9] M. Rabbani and P. W. Jones, *Digital Image Compression Techniques*. Bellingham, WA: SPIE Optical Engineering Press, 1991.

[10] H. C. Reeve and J. S. Lim, "Reduction of blocking effect in image coding," in *Proc. International Conf. Acoustics, Speech and Signal Processing*, (Boston, MA), pp. 1212–1215, 1983.

[11] K. Sauer, "Enhancement of low bit-rate coded images using edge detection and estimation," *CVGIP: Graphical Models and Image Processing*, vol. 53, pp. 52–62, 1991.

[12] A. Zakhor, "Iterative procedures for reduction of blocking effects in tranform image coding," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 2, pp. 91–94, March 1992.

[13] R. L. Stevenson, "Reduction of coding artifacts in transform image coding," in *Proc. International Conf. Acoustics, Speech and Signal Processing*, (Minneapolis, MN), pp. V401–404, 1993.

[14] Y. Yang, N. P. Galatsanos, and A. K. Katsaggelos, "Regularized reconstruction to reduce blocking artifacts of block discrete cosine transform compressed images," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 3, pp. 421–432, December 1993.

[15] Z. Fan and R. Eschbach, "JPEG decompression with reduced artifacts," in *Proc. of IS&T/SPIE Symposium on Electronic Imaging: Image and Video Compression*, (San Jose, CA), February 1994.

[16] J. Luo, C. W. Chen, and K. J. Parker, "A new method for artifacts removal in low bit rate DCT based image compression," submitted to *IEEE Trans. Image Processing*, May 1994.

[17] J. W. Woods and S. D. O'Nell, "Sub-band coding of images," *IEEE Trans. ASSP*, vol. ASSP-34, pp. 1278–1288, 1986.

[18] C. Podilchunk and A. Jacquin, "Subband video coding with a dynamic bit allocation and geometric vector quantization," in *Human Vision, Visual Processing, and Digital Display III*, vol. SPIE 1666, (Boston, MA), pp. 241–252, 1992.

[19] M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies, "Image coding using wavelet transform," *IEEE Trans. Image Processing*, vol. 1, pp. 205–220, 1992.

[20] H. Gharavi, "Subband coding of video signals," in *Subband Image Coding* (J. W. Woods, ed.), pp. 229–272, Boston, MA: Kluwer Academic Publishers, 1991.

[21] G. Karlsson and M. Vetterli, "Three dimensional sub-band coding of video," in *Proc. International Conf. Acoustics, Speech and Signal Processing*, pp. 1110–1113, 1988.

[22] J. Luo, C. W. Chen, and K. J. Parker, "Three dimensional subband video coding with segmentation in high frequency subbbands," in *International Conference on Image Processing*, (Austin, TX), 1994.