PRIOR MODEL FOR THE MRF MODELING OF MULTI-CHANNEL IMAGES

Hyung Il Koo and Nam Ik Cho

School of Electrical Engineering
Seoul National University, Seoul 151-744 Korea
email: hikoo@ispl.snu.ac.kr, nicho@snu.ac.kr

ABSTRACT

In multi-channel images (e.g. color images with R, G, B channel, and multi-spectral images), there exist higher-order correlations among the channels. We develop a new MRF-MAP (Markov random field - maximum a posteriori) framework that can be used for various multi-channel image processing. Main features of the proposed framework is that the higher-order correlation between the channels is considered, whereas it is not well addressed in the conventional works[4, 8]. Given a channel image, the prior probability of another channel is computed based on the MRF modeling that the channel correlation is described as piecewise linear relationship. An optimization algorithm for the MAP estimation is also developed. The effectiveness of the proposed priors is demonstrated with a simple application, i.e., image denoising.

Index Terms**—** Markov Random Field, Prior Model, Color Image Denoising

1. INTRODUCTION

The spatial prior model of an image is needed in many image processing tasks, for example, image denoising, image interpolation, super resolution, etc. Hence many prior models of natural images have been developed to encode the prior knowledge. In general, the prior model of natural images is designed to encode the smoothness constraints without the excessive smoothing of edge regions. However, conventional priors are designed in ad hoc manner and do not take higher order statistics into consideration. Only some of more recent works, the learned prior model (Field of Experts)[7, 8] consider this rich higher order statistics. However, these approaches consider only 2×2 cliques, and can not be easily generalized to cover larger cliques due to its computational complexity. Moreover the strong inter-channel correlation has not much been considered. Note that there is high correlation between the channels, especially in the intensity transition zones and texture regions. From this correlation, much information useful for image processing tasks can be obtained. The channel correlations have been justified and formulated in many color image formation models and statistical analyses, and they have already been used in image processing applications (demosaicing and chrominance interpolation[1, 2, 6]). However, their models are oversimplified and noise-sensitive, and thus they cannot be easily generalized to other applications. For the use of channel correlation in wider applications, we develop a more general conditional probability $P(I^i|I^j)$ based on the MRF modeling, where I^i and I^j are images of different channels in a multi-channel image. Also, in order to solve the problems related to image processing, we develop a MAP estimation algorithm for the proposed model.

For this purpose, we begin from the color image formation model in [1]. We first extend this simplified model, because this model is too simple to explain higher order statistics of natural images. Then we derive the conditional probability $P(I^i|I^j)$ in terms of clique potential. Because there are many possibilities in I^i (although I^j is given), many free parameters are introduced in $P(I^i|I^j)$, which severely complicates the optimization process. Hence, we develop an alternating optimization algorithm that is similar to EM (Expectation Maximization) algorithm.

The rest of paper is organized as follows. Section 2 describes a conventional MRF-MAP framework and our MAP model. We explain a new model of color image formation, conditional probability $P(I^i|I^j)$ and its optimization method in Section 3. Finally, We demonstrate the effectiveness of proposed priors with a simple application, image denoising in Section 4, and conclude the paper in Section 5.

2. MAP ESTIMATION

To reconstruct an image I from its noisy or low-resolution observation J (as in image interpolation or denoising), the MRF-MAP framework has been extensively used. This framework can be expressed as

$$\hat{I} = \arg\max_I P(I|J) = \arg\max_I P(J|I)P(I) \qquad (1)$$

where the likelihood term $P(J|I)$ is the result of observation model and the prior term $P(I)$ is the result of prior knowledge (e.g., MRF modeling). In the MRF framework, the prior probability of an image I can be expressed as a product of clique potential $\Phi(c)$ for all maximal cliques $c \in \mathcal{C}$,

$$P(I) = \frac{1}{Z} \prod_{c \in \mathcal{C}} e^{-\frac{1}{T}\Phi(I_c)} \tag{2}$$

where Z is normalization constant, T is the temperature and I_c is the region that corresponds to the clique c. In the context of image enhancement, the MRF model is designed to preserve discontinuity. For examples, four second-order cliques are considered and their potentials are designed in [4, 5]. In [7], 2×2 cliques are considered and their clique potentials are modeled using Fields-of-Experts framework [8], where the parameters are learned from training data. However, it seems impractical to derive a higher order model form these approaches due to increased computation complexity.

In this paper, we develop a new MRF model for multi-channel images $I = (I^1, I^2, I^3, \cdots, I^N)$. Given a reference image I^j and a noisy observation J^i of I^i $(i \neq j)$, our MAP framework is to find

$$\hat{I}^i = \arg\max_{I^i} P(I^i | J^i, I^j). \tag{3}$$

or equivalently,

$$\hat{I}^i = \arg\max_{I^i} P(J^i | I^i, I^j) P(I^i | I^j) \tag{4}$$

by Bayesian rule with the denominator $P(J^i|I^j)$ neglected. Since J^i is the observation of I^i, Markov property can be applied and this further reduces to

$$\hat{I}^i = \arg\max_{I^i} P(J^i | I^i) P(I^i | I^j). \tag{5}$$

Hence, in order to obtain the estimate $\hat{I}^i$ from its observation J^i and another channel image I^j, we need to compute $P(J^i|I^i)$ and $P(I^i|I^j)$. The former is simply a likelihood term that depends on the observation model. Hence, what we need to do is to formulate "the guided prior term" $P(I^i|I^j)$, by exploiting the relationship between two channels.

3. THE INTER-CHANNEL PRIOR MODEL

3.1. Color Formation Model

The relation between different channels (RGB vectors or the results of affine transformation of the RGB vectors) has been used in many applications, e.g., demosaicing or interpolation [1, 2, 6]. For the case of RGB channels[1], the model for color images is derived in the context of the "Mondriaan world", i.e. the world consisted of Lambertian nonflat surface patches with fixed light direction **l**. This means that the RGB values at 3-D point **z** in the space is described as

$$\begin{bmatrix} I^1(\mathbf{z}) \\ I^2(\mathbf{z}) \\ I^3(\mathbf{z}) \end{bmatrix} = (\mathbf{N}(\mathbf{z}) \cdot \mathbf{l}) \begin{bmatrix} \rho_1(\mathbf{z}) \\ \rho_2(\mathbf{z}) \\ \rho_3(\mathbf{z}) \end{bmatrix} \tag{6}$$

where $\mathbf{N}(\mathbf{z})$ is the normal vector at **z** and ρ_i $(i = 1,2,3)$ is albedo of the material that changes according to spectral channels. Under the assumption that albedo ρ_i is constant in an object, the ratio of channel values at a pixel position **x** is

$$\frac{I^i(\mathbf{x})}{I^j(\mathbf{x})} = \frac{\rho_i(\mathbf{x})}{\rho_j(\mathbf{x})} = \frac{\rho_i}{\rho_j} = constant \tag{7}$$

within an object or in the local neighborhood [1]. Image demosaicing can also be performed using a similar approach as shown in [2]. In [6], the relation between luminance and chrominance channels is exploited statistically. Since the Luminance-Chrominance representation (e.g., YCbCr, YUV) is simply an affine transformation of RGB vectors, it can be seen that these two approaches share the same assumption in eq-(7). In order to correctly estimate pixels from this relation, the values of constant in eq-(7) from region to region are needed. These values can be estimated from the gradient of neighborhood pixels, although not explicitly mentioned in the above mentioned works. However, this approach is suitable only for the limited applications, because the parameter estimation from the neighboring pixel is very sensitive to noise. Hence eq-(7) can only be used in the restricted applications that estimate missing pixel values using noise-free neighbor pixels.

3.2. Extended Color Formation Model

From the channel correlation, we develop a new prior model that can be used in noisy environment. On the contrary to the generic prior model $P(I^i)$ developed in many literatures, the proposed prior model $P(I^i|I^j)$ exploits the information of another channel. Hence, the proposed approach naturally reflects another channel information. Especially when a higher resolution channel or higher SNR channel is available (Luminance in 4:2:0 format or panchromatic in multi-spectral image), the image enhancement can be performed more successfully.

Let an N-channel image $I = (I^1, I^2, \cdots, I^N)$ be the affine transformation of another multi-channel image that satisfies eq-(7) and the assumptions mentioned in the previous subsection. Also, let us use an abstract index k for a rectangular lattice of a 2-D image, and denote the set of lattice points as $\mathcal{L}$. Then, we have more generalized linear relationship of two channels as

$$I^i(\mathbf{x}_k) = a(\mathbf{x}_k) + b(\mathbf{x}_k) I^j(\mathbf{x}_k), \tag{8}$$

where $\Gamma_a = \{a(\mathbf{x}_k) | k \in \mathcal{L}\}$ and $\Gamma_b = \{b(\mathbf{x}_k) | k \in \mathcal{L}\}$ are piecewise constant. But this model is still too oversimplified to encode rich higher order statistics of an image. Moreover, it can not model an abrupt changes including object boundaries, without an accurate segmentation result. So we relax the model by assuming that Γ_a and Γ_b are approximately linear in the local area with additive white Gaussian noise n_G. In other words, we fit the relation parameters to a piecewise linear plane with additive white Gaussian noise. To be precise, in the neighborhood of $\mathbf{x}_k$, the channels are related as

$$I^i(\mathbf{x}) = a_k(\mathbf{x}) + b_k(\mathbf{x})I^j(\mathbf{x}) + n_G \tag{9}$$

for $\mathbf{x} \in \mathcal{B}(\mathbf{x}_k, R) = \{\mathbf{x} \,:\, ||\mathbf{x} - \mathbf{x}_k||_\infty \leq R\}$ where $a_k(\mathbf{x})$ and $b_k(\mathbf{x})$ are some linear functionals of $\mathbf{x}$. Actually, $\Gamma = (\Gamma_a, \Gamma_b)$ field is slowly varying within the same object. Although it may experience abrupt changes at the object boundaries, these discontinuities are not modeled as step functions but as sigmoid functions that have transition zones. The reason for this is that there is also some blurring at the edges in actual image acquisition devices. Thus the linear approximation can be applied to not only within the objects but also on the object boundaries.

3.3. MRF Model

Our main objective is to find a tractable form of $P(I^i|I^j)$ using MRF and then make a multi-channel image processing into an MRF-MAP problem. We define a neighborhood system in an image I^i that connects all the nodes in the $(2R+1) \times (2R+1)$ rectangular region. So a maximal clique in this system can be represented as $c_k = \mathcal{B}(\mathbf{x}_k, R)$. From the aforementioned multi-channel image model and observation, a clique potential is defined as

$$\psi(c_k;\theta_k) = \sum_{\mathbf{x}\in\mathcal{C}_k} (I^i(\mathbf{x}) - a_k(\mathbf{x}) - b_k(\mathbf{x})I^j(\mathbf{x}))^2 \tag{10}$$

where $a_k(\mathbf{x}) = a_k + \mathbf{h}_k^{\mathrm{T}}(\mathbf{x} - \mathbf{x}_k)$, $b_k(\mathbf{x}) = b_k + \mathbf{g}_k^{\mathrm{T}}(\mathbf{x} - \mathbf{x}_k)$ for some $a_k, b_k \in \mathcal{R}$, $\mathbf{g}_k, \mathbf{h}_k \in \mathcal{R}^2$ and θ_k is $[a_k, b_k, \mathbf{g}_k, \mathbf{h}_k]$. From the GRF (Gibbs Random Field) - MRF equivalence, we have

$$P(I^i|I^j) = \frac{1}{Z(\Theta)} \prod_{k\in\mathcal{L}} e^{-\frac{1}{T}\ \psi(c_k;\theta_k)} \tag{11}$$

where $\Theta = \{\theta_k\}_{k\in\mathcal{L}}$, T is the temperature and $Z(\Theta)$ is the partition function. If we have some specific values of the parameter Θ, MAP estimation problem can be solved using some standard techniques, e.g., steepest descent algorithm. But since they are not available, the optimization process becomes intractable; because the computation of a partition function $Z(\Theta)$ needs a high dimensional integration[9, 10].

Hence, we propose an approach based on the greedy algorithm, which is consisted of two alternating stages in the MAP estimation. The first stage is to update I^i using the conjugate gradient algorithm[3], with fixed parameter Θ. Then the second is to fix I^i, and optimize Θ to be the most probable one, i.e.,

$$\hat{\Theta} = \arg\max_{\Theta} \prod_{k\in\mathcal{L}} e^{-\frac{1}{T}\ \psi(c_k;\theta_k)}. \tag{12}$$

This estimation can be computed using a pseudo inverse. Although the computational load in eq-(12) may seem to be large, it can be performed efficiently using the pre-computed pseudo inverse. Experimentally, this optimization scheme converges well in spite of its greediness. If we assume that the partition function has little effect on the optimization process[11], this algorithm can be regarded as a generalized EM (Expectation Maximization) algorithm that decreases the *Helmholtz* free energy with a point estimate on I^i[9].

4. EXPERIMENTAL RESULTS

4.1. Application to Denoising

A denoising problem that reconstructs I^i from J^i and I^j, where J^i is a noisy observation of I^i and I^j is a noisy free another channel image, can be formulated to

$$\hat{I}^i = \arg\max_{I^i} P(J^i|I^i)P(I^i|I^j), \tag{13}$$

as explained in Section. 2. If we assume that the noise is additive white Gaussian noise, the log-likelihood term becomes

$$\log P(J^i|I^i) \propto -\sum_{k\in\mathcal{L}} (J^i(\mathbf{x}_k) - I^i(\mathbf{x}_k))^2. \tag{14}$$

First, we test the proposed algorithm in the ideal situation. The chrominance image with PSNR 24.8dB is used as J^i and the luminance is used as I^j. The results on the pepper image are shown in Fig. 1. In contrast to conventional MRF models that show excessive blurring or many artifacts as in Fig. 1-(c),(d), the proposed algorithm successfully reduces noise with less blurring as shown in Fig. 1-(e),(f). The objective quality (PSNR) of the proposed method is $1.5 \sim 2$ dB higher than that of the conventional MRF model[4] on Pepper and Lena images. The proposed algorithm also shows much better subjective visual quality, mainly because the conventional MRF-model does not use the information of luminance channel (high-SNR or high-resolution).

In practice, we do not have the noise-free I^j. But since the luminance channel has higher energy and also generally higher resolution than the chrominance channels, we can assume that the luminance channel (25.0dB) is the noise-free channel I^j and apply the proposed algorithm. From the experiments, it is shown that the performance degradation from the ideal case is small (about 0.5 dB), and the PSNR is still 1-dB higher than that of the conventional MRF model (Pepper image=34.50dB, Lena image=35.00dB). Finally, effects

Table 1. Chrominance channel denoising results(dB).

Radius	3	4	5	6
Lena (Perfect Lum.)	36.06	36.66	36.88	36.89
Lena (Noisy Lum.)	36.13	36.32	36.12	35.93
Pepper (Perfect Lum.)	36.18	36.61	36.70	36.63
Pepper (Noisy Lum.)	36.63	35.55	35.34	35.10

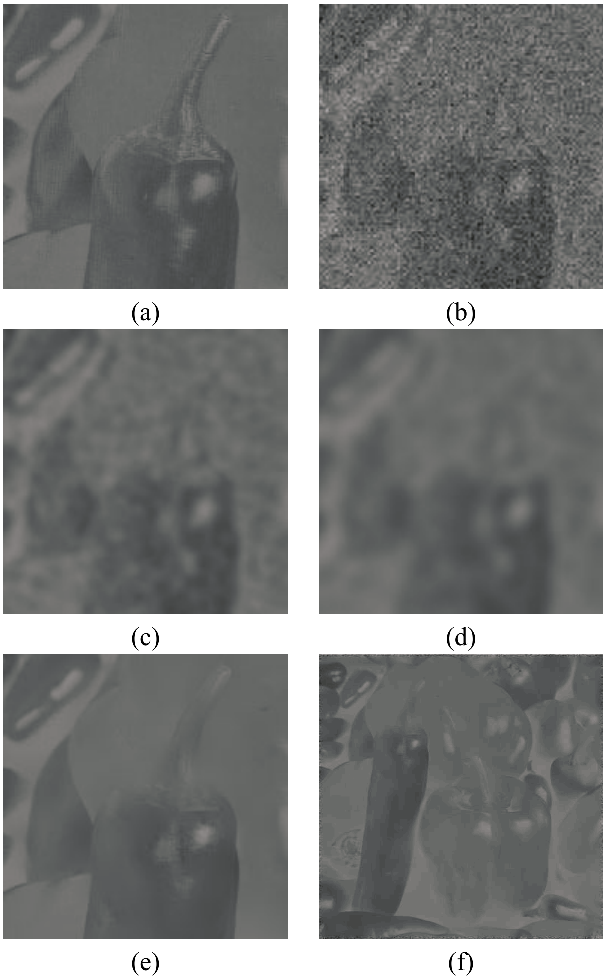

Fig. 1. (a) the original chrominance image, (b) the corrupted chrominance image (24.8 dB), (c),(d) the reconstructed image using traditional MRF model[4, 5] with varying parameters, (e),(f) the reconstructed image using proposed MRF model.

of clique size (radius R) on the performance are shown in Table. 1. In the experiments, we use the noisy observation J^i as the initial estimate of the I^i, for the fast convergence. We get the virtually same results when we use the initial estimate of the I^i as a black image.

5. CONCLUSIONS

In the conventional MRF model based image processing algorithms, rich information between the channels (in multichannel images) has not been taken into consideration. Hence, we develop an approach that exploits the inter-channel correlation in terms of the prior probability. We also developed an alternating optimization scheme, that can deal with not only inter-channel correlation but also much larger cliques than the conventional MRF models.

We have tested the effectiveness of the proposed priors with two applications, image denoising and image fusion (although not shown in this paper due to page limit). In image fusion, the proposed method shows competitive visual quality to that of the state of art image fusion techniques. In chrominance denoising, the proposed method shows better objective and subjective quality.

6. REFERENCES

[1] Ron Kimmel, “Demosaicing: Image Reconstruction from Color CCD samples ,” *IEEE Transactions on Image Processing,* vol. 8, no. 9, September, 1999.

[2] D. Darian Muresan and Thomas W. Parks, “Demosaicing Using Optimal Recovery,” *IEEE Transactions on Image Processing,* vol. 14, no. 2, February, 2005.

[3] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C : The Art of Scientific Computing*, Camibridge University Press, 1992.

[4] R. R. Schultz, R. L. Stevenson, “A Bayesian approach to image expansion for improved definition,” *IEEE Transactions on Image Processing,* May, 1994.

[5] R. R. Schultz, R. L. Stevenson, “DCT Qunatization Noise in Compressed Images,” *IEEE Transactions on Circuits and Systems for Video Technology,* January 2005.

[6] Maciej Bartkowiak, “Improved Interpolation of 4:2:0 Colour Images to 4:4:4 Format Exploiting Inter-Component Correlation,” *European Signal Processing Conference,* 2004.

[7] Xiangyang Lan, Stefan Roth, Daniel Huttenlocher, and Michael J. Black, “Efficient Belief Propagation with Learned Higher-order Markov Random Fields,” *European Conference on Computer Vision*, 2006.

[8] S. Roth, M.J. Black, “Fields of experts : A framework for learning image priors,” *in CVPR*, 2005.

[9] B. Frey and N. Jojic, “Advances in algorithms for inference and learning in complex probability models,” accepted, *IEEE Transactions on Pattern Analysis and Machine Intelligence,*” 2003.

[10] M. I. Jordan, Z. Ghahramani, T. S. Jaakkola, and L. K. Saul, “An introduction to variational methods for graphical models,” *in Learning in Graphical Models,* M. I. Jordan, Ed. Kluwer Academic Publishers, Norwell MA., 1998.

[11] A. Blake, C. Rother, M. Brown, P. Perez, and P. Torr, “Interactive Image Segmentation using an adaptive GMMRF model,” *In Proc. European Conf. Computer Vision,* 2004.