Gaussian Mixture Markov Random Field For Image Denoising And Reconstruction

Ruoqiao Zhang and Charles A. Bouman
School of Electrical and Computer Engineering
Purdue University
West Lafayette, IN 47907-0501 USA

Jean-Baptiste Thibault
GE Healthcare Technologies
Waukesha, WI 53188 USA

Ken D. Sauer
Department of Electrical Engineering
University of Notre Dame
Notre Dame, IN 46556-5637 USA

***Abstract*—Markov random fields (MRFs) have been widely used as prior models in various inverse problems such as tomographic reconstruction. While MRFs provide a simple and often effective way to model the spatial dependencies in images, they suffer from the fact that parameter estimate is difficult. In practice, this means that MRFs typically have very simple structure that cannot completely capture the subtle characteristics of complex images.**

In this paper, we present a novel Gaussian mixture Markov random field model (GM-MRF) that can be used as a very expressive prior model for inverse problems such as denoising and reconstruction. This method forms a global image model by merging together individual Gaussian-mixture models for image patches. Moreover, we present a novel analytical framework for computing MAP estimates with the GM-MRF prior model through the construction of exact surrogate functions that result in a sequence of quadratic optimizations. We demonstrate the value of the approach with some simple applications to denoising of dual-energy CT images.

***Index Terms*—Markov random fields, Gaussian mixture, patch-based methods, image model, prior model**

I. Introduction

Model-based inversion methods have been widely applied in applications such as image denoising and reconstruction [1], [2]. One typical approach to model-based inversion is to compute the maximum a posteriori (MAP) estimate given by

$$\hat{x} \leftarrow \arg\max_{x\in\Omega} \left\{ \log p(y|x) + \log p(x) \right\}, \tag{1}$$

where y is the measured data, and x is the unknown image. In this framework, the conditional density $p(y|x)$ comprises the forward model of the measurement process, and the density $p(x)$ comprises the prior model for x.

For model-based inversion methods, it is crucial to have an accurate prior model that captures the representative features of the image. The Markov random field (MRF) [3] has been one of the most popular choices of prior models in many model-based inversion problems [2], [4]. This model limits the dependencies in the image such that only local pixel interaction needs to be taken into account. However, due to the difficulty of parameter estimation, model-based inversion methods typically use simple prior models based on pair-wise interaction with ad hoc potential functions and a small number of parameters. This severely restricts the expressiveness of the prior model. Recently, MRFs with more sophisticated local models have been proposed [5], [6]. In [5], a high-order MRF was designed for CT reconstruction to allow enhancement of high frequency components. In [6], a more complex MRF model with implicit specification has been proposed. This method models the conditional probability of a pixel given its neighborhood as a Gaussian mixture; however, it does not allow for the explicit specification of the distribution $p(x)$.

This research was supported by GE Healthcare.

Meanwhile, there have been a variety of methods proposed to exploit image patches for prior modeling. These methods include dictionary learning [7], field of experts [8], non-local mean [9], kernel regression [10] and Gaussian scale mixture [11]. Perhaps the most closely related work is Zoran and Weiss's method [12]. This work uses Gaussian mixture model for patches and collects all the patches as a model for the whole image. However, in simply performing a hard classification during optimization, it does not fully exploit the mixture information.

In this paper, we propose a novel framework for image modeling, which we call a Gaussian mixture MRF (GM-MRF). The proposed method forms an MRF model by merging together individual Gaussian-mixture models for image patches. More specifically, the proposed model first fits the distribution of all image patches with a multivariate Gaussian mixture model. Then for each tiling of the patches, which covers the whole image with non-overlapping patches, we model the distribution of each tiling as the product of the distributions of all its patches. Finally, the proposed method constructs the MRF for the whole image by taking the geometric average of the distributions of all possible tilings. We also introduce a novel technical method for computing the MAP estimate with a GM-MRF prior model that is based on majorization-minimization using a surrogate function.

A important potential advantage of the proposed GM-MRF model over existing methods is that it allows for joint modeling of both the pixel intensities and spatial correlation. This is particularly important for problems such as dual energy CT reconstruction, since in this application, the specific pixel intensities typically correspond to specific materials, each with their own distinctive spatial structure. Another important advantage of the GM-MRF model is that it allows the MRF parameters to be easily estimated using standard methods for Gaussian-mixture parameter estimation.

II. Proposed Algorithm

A. Gaussian Mixture Markov Random Field (GM-MRF)

Let x be an image with pixels $s \in S$, where S is the set of all pixels in the image. Let z_s be a patch in the image with the pixel s at the upper left corner. More specifically, we can define $z_s = \{x_r : r \in s + W\}$ where W is a window of p pixels. We assume that each patch can be modeled as having a multivariate Gaussian mixture distribution with K subclasses with the form

$$g(z_s) = \sum_{k=0}^{K-1} \pi_k \frac{1}{(2\pi)^p} |B_k|^{\frac{1}{2}} \exp\left\{-\frac{1}{2}\|z_s - \mu_k\|^2_{B_k}\right\}. \quad (2)$$

where parameters π_k, μ_k, B_k represent the mixture probability, mean, and inverse covariance of subclass k.

Let $S_0, \cdots, S_{\eta-1}$ be a partition of the set of all pixels into η sets, each of which tiles the plane. In other words, $\{z_s\}_{s\in S_m}$ contains all the pixels in x. A simple example of this is when each z_s is a square $M \times M$ patch, and S_0 is set of pixels at each M^{th} row and column. Then the set of patches, $\{z_s\}_{s\in S_m}$ tile the plane.

Using this notation, the distribution of the full image can be modeled as the product of the distribution for each patch, so that

$$p_m(x) = \prod_{s\in S_m} g(z_s) = \exp\left\{-\sum_{s\in S_m} V(z_s)\right\}. \quad (3)$$

where $V(z_s) = -\log\{g(z_s)\}$. In this case, $p_m(x)$ is a proper distribution that has the desired distribution for each patch. However, the discrete tiling of the plane introduces artificial boundaries between patches. In order to remove the boundary artifacts, we use an approach similar to [13] and take the geometric mean of each of the tilings of the plane to obtain the resulting distribution

$$p(x) = \frac{1}{z}\left(\prod_{m=0}^{\eta-1} p_m(x)\right)^{1/\eta} = \frac{1}{z}\exp\left\{-\frac{1}{\eta}\sum_{s\in S} V(z_s)\right\}. \quad (4)$$

where the normalizing constant z is introduced to assure that $p(x)$ is a proper distribution after the geometric mean is computed.

We call (4) a Gaussian mixture MRF (GM-MRF) where

$$V(z_s) = -\log\left\{\sum_{k=0}^{K-1} \pi_k \frac{1}{(2\pi)^p}|B_k|^{\frac{1}{2}} \exp\left\{-\frac{\|z_s - \mu_k\|^2_{B_k}}{2}\right\}\right\}. \quad (5)$$

Then for a typical MAP problem as shown in (1), the estimate can be equivalently computed as

$$\hat{x} \leftarrow \arg\min_{x\in\Omega}\left\{\frac{1}{2}\|y - Ax\|^2_\Lambda + \frac{1}{\eta}\sum_{s\in S} V(z_s)\right\}, \quad (6)$$

assuming that $\log p(y|x)$ is modeled by $\frac{1}{2}\|y - Ax\|^2_\Lambda$ with Λ as a diagonal weighting matrix.

We estimate the parameters π_k, μ_k, and B_k for each subclass k in (2) from training images by using the expectation-maximization (EM) algorithm using the software in [14].

B. Optimization

Notice that the second term in (6) is difficult for direct optimization due to the mixture of logarithmic and exponential functions. Therefore, we propose a functional substitution approach to replace this complicated function with a mixture of quadratic functions that is computationally simpler.

For the function

$$u(x) = \frac{1}{\eta}\sum_{s\in S} V(z_s), \quad (7)$$

where $V(z_s)$ is given by (5), we define a quadratic function,

$$u(x; x') = \frac{1}{2\eta}\sum_{s\in S}\sum_{k=0}^{K-1} \tilde{w}_k \|z_s - \mu_k\|^2_{B_k} + c(x'), \quad (8)$$

where x' represents the current state of the image,

$$\tilde{w}_k = \frac{\pi_k |B_k|^{\frac{1}{2}} \exp\left\{-\frac{1}{2}\|z'_s - \mu_k\|^2_{B_k}\right\}}{\sum_{j=0}^{K-1} \pi_j |B_j|^{\frac{1}{2}} \exp\left\{-\frac{1}{2}\|z'_s - \mu_j\|^2_{B_j}\right\}}, \quad (9)$$

$$c(x') = \frac{1}{\eta}\sum_{s\in S}\left\{V(z'_s) - \frac{1}{2}\sum_{k=0}^{K-1}\tilde{w}_k\|z'_s - \mu_k\|^2_{B_k}\right\} \quad (10)$$

with z'_s representing a patch in x'. Then by using the following lemma proved in [15], it can be easily shown that,

$$u(x'; x') = u(x), \quad (11)$$

$$u(x; x') \geq u(x), \quad (12)$$

which implies that $u(x; x')$ is a surrogate function for $u(x)$ so that majorization minimization methods can be used, and reduction of $u(x; x')$ ensures reduction of $u(x)$ [16].

Lemma: surrogate functions for logs of exponential mixtures
Let $f : \Re^N \to \Re$ be a function of the form,

$$f(x) = \sum_k w_k \exp\{-v_k(x)\} \quad (13)$$

where $w_k \in \Re^+$, $\sum_k w_k > 0$, and $v_k : \Re^N \to \Re$. Furthermore $\forall (x, x') \in \Re^N \times \Re^N$ define the function

$$q(x; x') \triangleq -\log f(x') + \sum_k \tilde{\pi}_k (v_k(x) - v_k(x')) \quad (14)$$

where $\tilde{\pi}_k = \frac{w_k \exp\{-v_k(x')\}}{\sum_j w_j \exp\{-v_j(x')\}}$. Then $q(x; x')$ is a surrogate function for $-\log f(x)$, and $\forall (x, x') \in \Re^N \times \Re^N$,

$$q(x'; x') = -\log f(x') \quad (15)$$

$$q(x; x') \geq -\log f(x) \quad (16)$$

Therefore, any optimization problem of the form

$$\hat{x} \leftarrow \arg\min_x\{f(x) + u(x)\} \quad (17)$$

can be implemented as a sequence of optimizations as

$$\text{repeat}\{ \quad \hat{x} \leftarrow \arg\min_x\{f(x) + u(x; x')\} \quad x' \leftarrow \hat{x} \quad \},$$

where we will be solving an optimization problem with quadratic priors at each iteration.

Particularly, we use the iterative coordinate descent (ICD) algorithm [17] to solve this problem. By using ICD, we sequentially update each pixel by solving a 1-D optimization problem, with the remaining pixels fixed. In this case, the surrogate function for a single pixel s can be written as

$$u(x_s; x'_s) = \frac{1}{2\eta} \sum_{r \in \mathcal{P}_s} \sum_k \tilde{w}_k \| z_r - \mu_k \|^2_{B_k} + c(x') \quad (18)$$

where $\mathcal{P}_s$ represents a set of pixels, each being the upper left corner of a patch that contains pixel s. Then derived from (6), we solve the following 1-D optimization problem for pixel s,

$$\hat{x}_s \leftarrow \arg\min_{x_s \in \Omega} \left\{ \sum_i \frac{\lambda_i}{2} (y_i - A_{i,s} x_s)^2 + u(x_s; x'_s) \right\}. \quad (19)$$

where λ_i is the i^{th} diagonal element of Λ. This is a quadratic minimization problem for pixel s and can be solved exactly by any standard rooting algorithm.

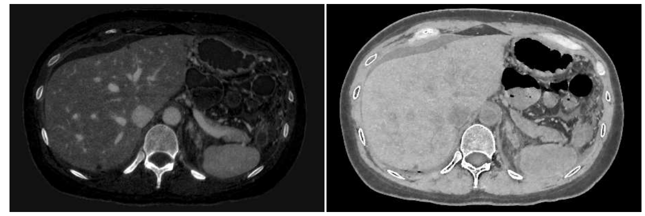
III. Results

We apply the proposed GM-MRF model to denoise the dual-energy CT (DECT) clinical images with additive white noise. The DECT reconstruction produces two values at one single pixel, corresponding to the equivalent densities of water and iodine. The two reconstructed densities typically correspond to specific materials and therefore are highly correlated. The ground truth DECT images were reconstructed from the raw data acquired on a Discovery CT750 HD scanner (GE Healthcare, WI, USA) in the dual-energy acquisition mode, where we used the JDE-MBIR method [18] for reconstruction since it produces images with less noise but higher resolution than traditional filtered back projection. We add white noise to each of the DECT images separately to generate noisy data.

We learn the GM-MRF model from 2×10^5 overlapping patches of size 3×3, which are extracted from one slice of the clean DECT images, as shown in Figure 1. Since each pixel in this case has two density values corresponding to water and iodine respectively, we have an 18-dimensional patch vector. The resulting GMM model is chosen to have 20 subclasses. The denoising is performed jointly on both of the material images by using the GM-MRF model with 3×3 patches.

We compare the proposed model with the state-of-the-art K-SVD method [1]. We experiment with two sets of parameters for K-SVD. In the first set, we use the same patch size as GM-MRF does, which is 3×3; while in the second set, we select the patch size as suggested in [7], which is 8×8. For each of the above cases, we use two distinct dictionaries, one for water and the other for iodine. Each dictionary has 256 atoms, which is trained by using the corresponding material image in Figure 1. The K-SVD denoising is then performed separately on each material density of the testing images, with the regularization tuned to produce the best PSNR.

[1] We use the K-SVD denoising code from http://www.cs.technion.ac.il/ ronrubin/software.



(a) water (b) iodine

Fig. 1. Training data: one slice of dual-energy clinical images from an abdominal scan. Left: iodine-equivalent density; Right: water-equivalent density. For DECT reconstruction, each pixel has two values instead of one.

Figure 2 and 3 show the iodine and water results respectively. The results show that the GM-MRF model with 3×3 patches produces images with much less noise as compared to the K-SVD with same patch size. As compared to the K-SVD with larger patch size, the GM-MRF produces visually sharper images and retains more details with consistently higher PSNR. The K-SVD results with different patch sizes also indicate that increasing the patch size for the GM-MRF model may improve the performance.

IV. Conclusion

In this paper, we presented a novel GM-MRF model for image denoising and reconstruction. The proposed method forms a global image model by merging together individual models for image patches, with each patch following a multivariate Gaussian mixture distribution. We use a surrogate function to replace the complicated potential function so as to make the optimization tractable. We compared this model with the state-of-the-art K-SVD method in a denoising experiment for dual-energy CT images and demonstrated the quality improvement by our model. Future investigation will optimize the patch size and apply the GM-MRF model to more applications such as CT reconstruction.

References

[1] C. R. Vogel and M. E. Oman, "Iterative methods for total variation denoising," *SIAM Journal on Scientific Computing*, vol. 17, no. 1, pp. 227–238, 1996.

[2] J.-B. Thibault, K. D. Sauer, J. Hsieh, and C. A. Bouman, "A three-dimensional statistical approach to improve image quality for multislice helical CT," *Med. Phys.*, vol. 34, no. 11, pp. 4526–4544, Nov. 2007.

[3] J. Besag, "Spatial interaction and the statistical analysis of lattice systems," *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 192–236, 1974.

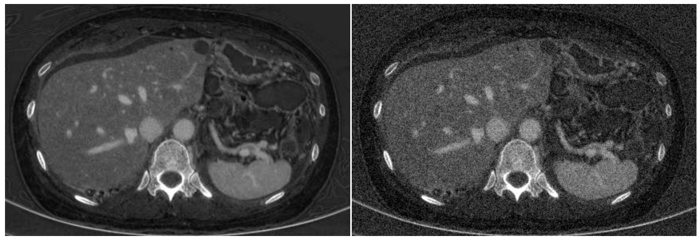
[4] T. Hebert and R. Leahy, "A generalized EM algorithm for 3-d Bayesian reconstruction from Poisson data using Gibbs priors," *IEEE Trans. on Medical Imaging*, vol. 8, no. 2, pp. 194–202, 1989.

[5] J. Wang, K. Sauer, J.-B. Thibault, Z. Yu, and C. Bouman, "Prediction coefficient estimation in Markov random fields for iterative x-ray CT reconstruction," in *SPIE Medical Imaging*, 2012, p. 831444.

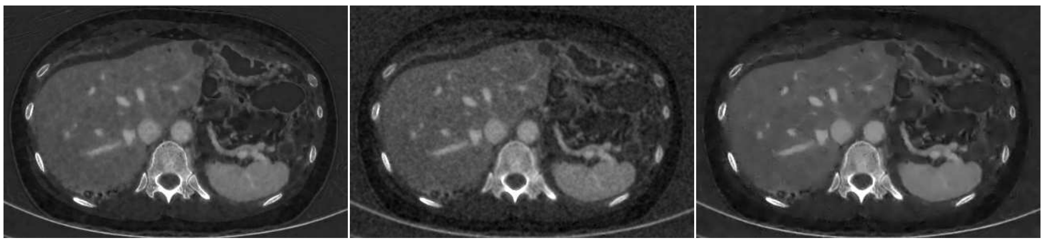
[6] E. Haneda and C. Bouman, "Implicit priors for model-based inversion," in *Proc. of IEEE Int'l Conf. on Acoust., Speech and Sig. Proc.*, 2012, pp. 3917–3920.

[7] M. Elad and M. Aharon, "Image denoising via sparse and redundant representations over learned dictionaries," *IEEE Trans. Image Process.*, vol. 15, no. 12, pp. 3736–3745, 2006.

[8] S. Roth and M. J. Black, "Fields of experts: A framework for learning image priors," in *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, vol. 2, 2005, pp. 860–867.

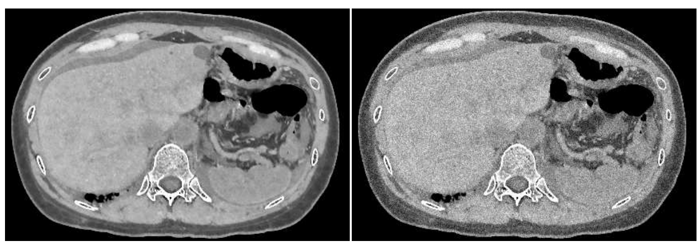


(a) ground truth (b) noisy image, PSNR: 22.33



(c) GM-MRF, 3×3 patch, PSNR: 34.11 (d) KSVD, 3×3 patch, PSNR: 30.67 (e) KSVD, 8×8 patch, PSNR: 33.00

Fig. 2. Denoising results on iodine densities with different priors. With the same patch size, GM-MRF produces much less noisy result than K-SVD, while as compared to K-SVD with larger patch size, GM-MRF retains more details with higher PSNR.



(a) ground truth (b) noisy image, PSNR: 26.59

(c) GM-MRF, 3×3 patch, PSNR: 33.51 (d) KSVD, 3×3 patch, PSNR: 32.46 (e) KSVD, 8×8 patch, PSNR: 33.15

Fig. 3. Denoising results on water densities with different priors. With the same patch size, GM-MRF produces much less noisy result than K-SVD, while as compared to K-SVD with larger patch size, GM-MRF retains more details with higher PSNR.

[9] A. Buades, B. Coll, and J.-M. Morel, "A non-local algorithm for image denoising," in *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, vol. 2, 2005, pp. 60–65.

[10] H. Takeda, S. Farsiu, and P. Milanfar, "Kernel regression for image processing and reconstruction," *IEEE Trans. Image Process.*, vol. 16, no. 2, pp. 349–366, 2007.

[11] J. Portilla, V. Strela, M. J. Wainwright, and E. P. Simoncelli, "Image denoising using scale mixtures of Gaussians in the wavelet domain," *IEEE Trans. Image Process.*, vol. 12, no. 11, pp. 1338–1351, 2003.

[12] D. Zoran and Y. Weiss, "From learning models of natural image patches to whole image restoration," in *Computer Vision (ICCV), 2011 IEEE International Conference on*, 2011, pp. 479–486.

[13] R. Salakhutdinov and G. Hinton, "An efficient learning procedure for deep Boltzmann machines," *Neural Computation*, vol. 24, no. 8, pp. 1967–2006, 2012.

[14] C. A. Bouman, "Cluster: An unsupervised algorithm for modeling Gaussian mixtures," April 1997, available from http://engineering.purdue.edu/~bouman.

[15] E. Haneda and C. A. Bouman, "Implicit priors for model based inversion," Purdue ECE, Tech. Rep., Oct. 2011.

[16] L. Kantorovich and G. Akilov, *Functional Analysis in Normed Spaces*. Moscow: Fitzmatgiz, 1959.

[17] K. D. Sauer and C. A. Bouman, "A local update strategy for iterative reconstruction from projections," *IEEE Trans. on Signal Processing*, vol. 41, no. 2, pp. 534–548, Feb. 1993.

[18] R. Zhang, J.-B. Thibault, C. A. Bouman, K. D. Sauer, and J. Hsieh, "A model-based iterative algorithm for dual-energy X-ray CT reconstruction," in *Proc. 2nd Intl. Mtg. on Image Formation in X-ray CT*, 2012, pp. 439–443.