CONTINUOUS MRF BASED IMAGE DENOISING WITH A CLOSED FORM SOLUTION

Ming Liu[1], *Shifeng Chen*[2], *and Jianzhuang Liu*[1,2]

[1]Department of Information Engineering
The Chinese University of Hong Kong, Hong Kong

[2]Shenzhen Institutes of Advanced Technology
Chinese Academy of Sciences, Shenzhen, China

ABSTRACT

In this paper, we formulate the problem of image denoising as the maximum a posterior (MAP) estimation problem using Markov random fields (MRFs). Such an MAP estimation for MRFs is equivalent to a maximum likelihood estimation constrained on spatial homogeneity and is generally NP-hard in the discrete domain. To make it tractable, we convert it to a continuous label assignment problem based on a Gaussian MRF model and then obtain a closed form globally optimal solution. Since the Gaussian MRFs tend to over-smooth images and blur edges, we incorporate pre-estimated image edge information into the energy function to better preserve image structures. Patch similarity based pairwise interaction is also involved to better preserve image details and make the algorithm more robust to impulse noise. Both quantitative and qualitative comparative experimental results are given to demonstrate the better performance of our algorithm.

Index Terms— Image denoising, Markov random field, label relaxation, closed form solution

1. INTRODUCTION

The goal of image denoising is to remove the noise while maintaining and recovering the details of the image as much as possible. Many denoising approaches have been developed over past decades. They can be grouped into two basic categories: filtering in spatial domain and filtering in transformed domain. The former in essence estimates each pixel's value with its neighboring pixels in some way. The basic idea of the latter is to project an image onto a set of bases, and then to process the coefficients of the transformed representation [1]. In [2], the authors considered the evaluation of a good denoising algorithm. Their analysis indicates that algorithms in both spatial and transformed domains are of great importance with respect to different evaluation aspects. Since our algorithm is developed in spatial domain, we focus on the discussion and comparison of spatial filtering methods in this paper.

Nonlinear models based algorithms, such as *Total Variation (TV) filter* [3], *anisotropic filter* [4], are either incapable of preserving edges well or suffer the loss of image texture information. Two popular algorithms, called *SUSAN filter* [5] and *bilateral filter* [6], take the average value of the pixels close to the reference pixel in both grey-level and spatial location, and they can maintain the structures well. Recently, Buades et al. presented a non-local means (NL-means) algorithm for image denoising [2]. The NL-means approach estimates the "true" value for a pixel as the weighted average grey level of all pixels whose Gaussian neighbors look like the neighbors of the reference pixel with a close neighborhood configuration. This approach is suitable for denoising images with periodic texture patterns, but fails in images with strong noise due to the corruption of image structures.

Markov random field (MRF) models [7] provide a robust and unified framework for many computer vision tasks, including image denoising. However, as a multi-label assignment problem, image denoising modeled as the MRF energy minimization in the discrete domain is generally NP-hard. Many approaches have been proposed to solve the MRF energy optimization, such as simulated annealing [7] and iterated conditional modes [8]. Recently developed algorithms based on graph cuts [9] and belief propagation [10] have demonstrated their good performances to handle MRF optimization problems. Both algorithms work in the discrete domain and usually can only find a local optimum. We also find that these methods cannot well preserve image edges and the large labels in the denoising task cause the optimization procedure to be very slow.

In this paper, we also formulate image denoising as an MRF energy minimization problem with elaborately defined pairwise relationship between neighboring pixels. We solve the label estimation by transforming it to a continuous optimization problem, where the labels are relaxed from discrete values to continuous ones. Compared with related approaches, the contributions of our work are summarized as follows: 1) In the continuous domain, a closed form globally optimal solution can be obtained, which provides a good prerequisite for our final result. 2) Image edges and details can be better preserved in our algorithm since pre-estimated edge information and patch based similarity are incorporated into our energy function. 3) While obtaining promising results, our algorithm is efficient. The experimental results demonstrate the advantages of our approach and show that it outperforms several recent methods both quantitatively and qualitatively.

2. OUR APPROACH

2.1. The Basic MAP-MRF Model

MRFs is popularized in computer vision in [7]. The formulation of the MRF model is justified in terms of the MAP estimation of the MRFs in the Bayesian framework. Let an input image and its labelling be represented by $X = [x_1, x_2, ..., x_n]^T$ and $F = [f_1, f_2, ..., f_n]^T$, respectively, where x_i is the intensity of pixel i, f_i is its corresponding label, $1 \leq i \leq n$, and n is the number of the pixels. F denotes a realization of a MRF. The optimization problem is to maximize the posterior probability $P(F|X)$, and it can be elegantly expressed as an MRF-based energy minimization problem. The energy function is generally defined as:

$$E(F) = \sum_{i=1}^{n} D_i(x_i, f_i) + \sum_{i=1}^{n} \sum_{j \in \mathcal{N}(i)} S_{i,j}(f_i, f_j), \quad (1)$$

where $D_i(x_i, f_i)$ is called data penalty that penalizes the inconsistency between labels and data, $\mathcal{N}(i)$ is the neighborhood of pixel i, and $S_{i,j}(f_i, f_j)$ is the clique potential representing the prior knowledge of the labels, which is also called smoothness penalty imposing the pairwise smoothness. Based on the piecewise smooth assumption on F, a good clique potential should be able to enforce spatial homogeneity in low contrast regions and preserve discontinuity in high contrast regions such as object boundaries.

2.2. Our MRF Energy Function

In this paper, we focus on grey-level images, but the formulation can be directly extended to handling color images without increasing the computational complexity. The data term and the clique potential in our approach are defined as

$$D_i(x_i, f_i) = (f_i - x_i)^2, \;\; S_{i,j}(f_i, f_j) = w_{ij}(f_i - f_j)^2, \quad (2)$$

where w_{ij} denotes the affinity value between pixels i and j and is used to control the smoothness degree for each pairwise interaction. $D_i(x_i, f_i)$ defined here models the additive Gaussian noise and is commonly used in image denoising. The quadratic label difference without truncation in $S_{i,j}(f_i, f_j)$ is not edge preserving. However, with the special design of the affinity value w_{ij}, where pre-estimated edge information and patch based similarity measurement are incorporated into its calculation, we can maintain edge sharpness and preserve image details as much as possible in the denoised image.

Before giving w_{ij} explicitly, we first define the patch based similarity between pixels i and j as $\Delta(i, j) = ||\mathbf{x}_{\mathcal{B}(i)} - \mathbf{x}_{\mathcal{B}(j)}||^2/|\mathcal{B}(i)|$, where $\mathbf{x}_{\mathcal{B}(i)}$ and $\mathbf{x}_{\mathcal{B}(j)}$ represent the grey-level vectors of the pixels in two same-size blocks $\mathcal{B}(i)$ and $\mathcal{B}(j)$ centered at pixels i and j, respectively. $|\mathcal{B}(i)|$ is the cardinality of $\mathcal{B}(i)$.

We design w_{ij} as follows: 1) If the difference between the blocks centered at neighboring pixels i and j is large in the input image, the smoothness penalty $S_{i,j}(f_i, f_j)$ should be small. 2) The farther the distance between pixels i and j, the less effect of them on $S_{i,j}(f_i, f_j)$. 3) If pixels i and j fall into two regions separated by an edge, they have no effect on $S_{i,j}(f_i, f_j)$. Based on these criteria, we define w_{ij} as

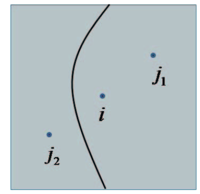

Fig. 1. Explanation of the region indexes. The curve denotes an edge separating the window into two regions. Pixels i and j_1 have the same region index ($C_i = C_{j_1}$), but pixels i and j_2 have different region indexes ($C_i \neq C_{j_2}$).

$$w_{ij} = a \cdot \exp(-\frac{\Delta(i, j)}{b}) \cdot k(i, j) \cdot T(C_i = C_j), \quad (3)$$

where a and b are two positive factors to control the contribution of w_{ij} to the smoothness penalty, $k(i, j) = \exp(-d_{ij}^2/2)$ is a Gaussian kernel function to reach target 2) above, and $T(C_i = C_j)$ is used towards target 3). $T(\cdot)$ is 1 if its argument is true and 0 otherwise. C_i and C_j are region indexes explained with Fig. 1. We use Canny edge detector to find edges in the input image, and then assign indexes to different regions. The design of w_{ij} in (3) makes our algorithm not only be able to denoise but also preserve image details well.

With the data penalty and the smoothness penalty (2), the energy function (1) can be written as

$$E(F) = \sum_{i=1}^{n} (f_i - x_i)^2 + \sum_{i=1}^{n} \sum_{j \in \mathcal{N}(i)} w_{ij}(f_i - f_j)^2. \quad (4)$$

This objective function represents a Gaussian MRF. The approaches in [9], [11], and [12] can solve the energy minimization problem by graph cuts or belief propagation. However, solutions obtained by them are locally optimal in the discrete domain. The algorithm proposed in [13] can exactly optimize the energy function (4) in the discrete domain by converting the problem into a min-cut/max-flow [14] problem on a complicated directed graph. However, it is not suitable for image denoising due to the heavy computational burden caused by the large sets of graph nodes and edges. In the next section, we give our closed form globally optimal solution to this optimization problem, which is based on the label relaxation and can be efficiently obtained.

2.3. A Closed Form Solution

With the smoothness term in (4), the MRF is isotropic and we can construct an undirected weighted graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ is the set of vertices denoting the image pixels and $\mathcal{E}$

is the set of weighted edges. Then the adjacency matrix of $\mathcal{G}$ is $W = [W_{ij}]_{n \times n}$ whose elements are defined as

$$W_{ij} = \begin{cases} w_{ij}, & \text{if } i \neq j, j \in \mathcal{N}(i) \\ 0, & \text{if } i \neq j, j \notin \mathcal{N}(i) \\ c, & \text{if } i = j, \end{cases} \quad (5)$$

where $c > 0$ is some constant. It is clear that c has no effect on the energy function $E(F)$ since the term $c(f_i - f_i)^2 = 0$. Let D be an $n \times n$ diagonal matrix with the (i,i)-th entry $D_{ii} = \sum_{j=1}^{n} W_{ij}$. When $D_{ii} = 0$, we call i an isolated vertex that causes W to be singular. By using the positive constant c, $D_{ii} \neq 0$ and singular W is avoided. Moreover, c builds up numerical stability for our solution. With the design of w_{ij}, D_{ii} may be very small for some vertices that have edges with small weights. Since the following derivation of the closed form solution involves D^{-1} and $D^{-\frac{1}{2}}$, a proper c can make these matrices numerically stable. We also find that it is better for c to be comparable with the value of the parameter a. Thus, we choose $c = a$ in our experiments.

To solve the energy minimization problem, we first relax the labels f_i, $1 \leq i \leq n$, from discrete values to continuous ones, and let $R = [r_1, r_2, ..., r_n]^T$, where $r_i = \sqrt{D_{ii}} f_i$, $1 \leq i \leq n$, form a set of medium variables. Then we have $F = D^{-\frac{1}{2}} R$. With the medium variable R and the corresponding graph $\mathcal{G}$, the energy function $E(F)$ in (4) can be written as

$$\begin{aligned} E(F) &= E(D^{-\frac{1}{2}} R) = E_1(R) \\ &= \sum_{i=1}^{n} \left(\frac{r_i}{\sqrt{D_{ii}}} - x_i\right)^2 + \sum_{i,j=1}^{n} W_{ij} \left(\frac{r_i}{\sqrt{D_{ii}}} - \frac{r_j}{\sqrt{D_{jj}}}\right)^2, \end{aligned} \quad (6)$$

which has a compact matrix-form expression:

$$E_1(R) = ||D^{-\frac{1}{2}} R - X||_{\mathcal{F}}^2 + 2(R^T \overline{L} R), \quad (7)$$

where $\overline{L}$ is the normalized Laplacian matrix of $\mathcal{G}$ defined as $\overline{L} = D^{-\frac{1}{2}}(D - W)D^{-\frac{1}{2}}$, and $|| \cdot ||_{\mathcal{F}}$ is the Frobenius norm of a matrix. To minimize $E_1(R)$, taking its derivative with respect to R and setting it to zero yields

$$(D^{-1} + 2\overline{L})R = D^{-\frac{1}{2}} X. \quad (8)$$

Since W and $\overline{L}$ are positive semi-definite, $(D^{-1} + 2\overline{L})$ is nonsingular. Then a closed form solution with the medium variable R is obtained as $R = (D^{-1} + 2\overline{L})^{-1} D^{-\frac{1}{2}} X$. Thus the globally optimal solution is

$$F = D^{-\frac{1}{2}} R = D^{-\frac{1}{2}} (D^{-1} + 2\overline{L})^{-1} D^{-\frac{1}{2}} X. \quad (9)$$

It is worth noting that the above derivation of the globally optimal F is based on the assumption that F is a continuous vector, which is preferred if the denoised image is used for further higher-level processing. However, if we want to display the denoised image, we quantize the optimal continuous solution to obtain its discrete version by setting its components to their closest discrete values.

From our formulation and the derived solution, it is not difficult to see that our approach is able to handle color image denoising directly by treating the color of each pixel as a three-dimensional vector in the CIE-Lab color space. In this space, we can use the Euclidean distance as the color difference measurement involved in (3), which is consistent with human perception. Therefore, the same closed form solution as in (9) can be achieved for color image denoising.

3. EXPERIMENTATION

To demonstrate the performance of our algorithm, we compare our algorithm with four most related algorithms: swap graph cuts (GC) [9], max-product belief propagation (BP) [10], bilateral filter (BF) [6], and NL-means (NL) [2]. GC [9] and BP [10] are representative MRF-based approaches, BF [6] is a nonlinear local smoothing filter, and NL [2] is a nonlocal method. We test these algorithms on a set of classic grey level images, "Boat", "Pepper", and "Lena" of size 256×256, and all 300 natural images in the Berkeley segmentation benchmark [15]. These images are contaminated by adding Gaussian noise of five levels with standard deviations $\sigma = 10, 20, 30, 50$ and 100. Both visual quality comparisons and quantitative comparisons in terms of the measurement peak signal-to-noise ratio (PSNR) are given.

The energy functions of GC and BP used in our experiments are given in [9] and [10] respectively. The parameters of the five algorithms are tuned best in terms of the largest PSNR at each noise level. Table 1 gives PSNR values, and the visual results corresponding to the best PSNR outputs on the noisy images "Lena" with $\sigma = 20$ and 30 are showed in Fig. 2. From these results, we can see that our algorithm can remove the noise effectively while preserving the image details well and generate comparable or better outputs over the other methods. Another experiment is carried out on the Berkeley segmentation benchmark [15]. Table 2 shows the average PSNR values, and indicates again that our algorithm outperforms the others.

4. CONCLUSION

In this paper, we formulate the image denoising problem as a MRF-based energy minimization problem. With the incorporation of pre-estimated edge information and patch similarity based pairwise interaction into the energy function, our algorithm can effectively reduce noise while maintaining image details well. Furthermore, by relaxing the labels from discrete values to continuous values, a closed form globally optimal solution can be obtained. The experimental results show that our algorithm outperforms the related approaches both qualitatively and quantitatively. Besides, the proposed approach can be extended to handling other computer vision tasks such as image segmentation and stereo vision, and we can construct a spatial-temporal MRF model for video denoising by adding a temporal constraint to the energy function.

Fig. 2. Results of the five algorithms on the "Lena" image with the noise $\sigma = 20$ (the first row), and $\sigma = 30$ (the second row). From left to right: the noisy images, the results by GC, BP, BF, NL, and our algorithm, and the original noise-free image

Table 1. PSNR values obtained by the five algorithms on the three noisy images at five noise levels.

image	σ	PSNR				
		GC	BP	BF	NL	Ours
Boat	10	30.55	30.55	30.42	30.29	**30.79**
	20	26.43	26.60	26.65	**27.27**	27.02
	30	24.24	24.55	24.81	25.71	**25.92**
	50	20.85	22.11	22.34	22.28	**22.54**
	100	18.15	18.70	19.91	14.54	**20.06**
Pepper	10	28.77	27.29	32.79	**33.34**	33.18
	20	27.03	25.04	28.91	30.15	**30.42**
	30	24.64	22.35	26.42	27.79	**28.43**
	50	22.31	21.02	23.45	23.89	**24.92**
	100	15.93	13.58	20.31	14.78	**20.82**
Lena	10	31.65	31.74	32.00	**32.61**	32.25
	20	27.56	27.79	28.35	**29.53**	29.02
	30	25.10	25.80	26.36	27.28	**27.51**
	50	22.89	23.56	24.24	23.88	**25.24**
	100	18.95	19.75	21.14	14.70	**21.37**

Table 2. Average PSNR values on the 300 noisy images in the Berkeley segmentation benchmark.

σ		10	20	30	50	100
PSNR	GC	30.95	27.41	25.65	23.12	18.76
	BP	31.09	27.57	25.95	23.63	19.06
	BF	31.43	27.95	26.36	23.84	20.75
	NL	31.92	28.42	27.02	23.53	15.82
	Ours	**32.22**	**28.93**	**27.49**	**25.17**	**21.26**

5. ACKNOWLEDGEMENT

This work was supported by grants from Key Laboratory of Robotics and Intelligent System, Guangdong Province (2009A060800016), Natural Science Foundation of China (No. 60903117, No. 60975029), and Shenzhen Bureau of Science Technology & Information, China (No. JC20090318 0635A).

6. REFERENCES

[1] D.L. Donoho, "De-noising by soft-thresholding," *IEEE Trans. Information Theory*, vol. 41, no. 3, 1995.

[2] A. Buades, B. Coll, and J.M. Morel, "Nonlocal Image and Movie Denoising," *IJCV*, vol. 76, no. 2, 2008.

[3] L. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D*, vol. 60, 1992.

[4] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Trans. PAMI*, vol. 12, 1990.

[5] S.M. Smith and J.M. Brady, "SUSANŁA new approach to low level image processing," *IJCV*, vol. 23, 1997.

[6] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," in *ICCV*, 1998.

[7] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images," *IEEE Trans. PAMI*, vol. 6, pp. 721–741, 1984.

[8] J. Besag, "On the Statistical Analysis of Dirty Pictures," *Journal of the Royal Statistical Society. Series B*, vol. 48, 1986.

[9] Y. Boykov, O. Veksler, and R. Zabih, "Fast approximate energy minimization via graph cuts," *IEEE Trans. PAMI*, 2001.

[10] R. Szeliski, R. Zabih, D. Scharstein, O. Veksler, V. Kolmogorov, A. Agarwala, M. Tappen, and C. Rother, "A Comparative Study of Energy Minimization Methods for Markov Random Fields with Smoothness-Based Priors," *IEEE Trans. PAMI*, pp. 1068–1080, 2008.

[11] V. Kolmogorov and R. Zabin, "What energy functions can be minimized via graph cuts?," *IEEE Trans. PAMI*, vol. 26, 2004.

[12] P.F. Felzenszwalb and D.P. Huttenlocher, "Efficient Belief Propagation for Early Vision," *IJCV*, vol. 70, 2006.

[13] H. Ishikawa, "Exact optimization for Markov random fields with convex priors," *IEEE Trans. PAMI*, 2003.

[14] Y. Boykov and V. Kolmogorov, "An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision," *IEEE Trans. PAMI*, vol. 26, no. 9, 2004.

[15] D. Martin, C. Fowlkes, D. Tal, and J. Malik, "A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics," in *ICCV*, 2001.