

# Automatic stopping criterion in fields-of-experts image denoising

Zhi Zhang, Peng Zhang, Ping Zhong, and Runsheng Wang

National University of Defense Technology, ATR Laboratory, Changsha, Hunan, 410073 China  
E-mail: zhizh.zhang@gmail.com

**Abstract.** Fields of experts (FoE) image denoising is one of the most promising high-order Markov random field (MRF)-based image denoising methods. However, the original algorithm by Roth and Black did not consider the parameter selection problem in its iteration, so it cannot be directly applied to real image denoising tasks. An automatic stopping criterion in FoE image denoising is introduced, through which the denoised image can be obtained without reference image and noise variance estimation. Experimental results validate its better performance than the classic FoE method, both on synthetic and real noisy images. © 2010 Society of Photo-Optical Instrumentation Engineers.  
[DOI: 10.1117/1.3456368]

Subject terms: image processing; image restoration; image denoising; Markov random field.

Paper 100257LRR received Mar. 30, 2010; revised manuscript received May 23, 2010; accepted for publication May 28, 2010; published online Jun. 23, 2010.

## 1 Introduction

In many scientific applications, images are often corrupted by noise because of either data acquisition or data transmission. Therefore, the problem of recovering an original image from noisy data has received ever-increasing attention in recent years, but it is still very challenging.<sup>1</sup> It has been popular to denoise an image by Markov random field (MRF)-based methods. Recently, a new fields of experts (FoE)-based approach<sup>2</sup> was introduced that considered high-order MRF cliques to grasp the complex structural information in image data and utilized a maximum-*a-posteriori* (MAP) framework to get the clean image. It is one of the most promising methods in generic MRF-based image denoising approaches.<sup>3</sup>

The typical objective in denoising problems is a recovery of an image with minimal mean-squared error (MSE). However, in practice, we do not have the original image to compare to, and thus we cannot know what choice of parameters minimizes the MSE. Therefore, the parameters are often tuned manually by looking at the reconstructed result.

In FoE image denoising algorithms, the tradeoff parameter and stopping criterion must be specified before iteration. However, tuning of the parameters is not an easy task. Assuming that the noise level is known *a priori*, Roth and Black adopted a fixed number of iterations and experientially determined the appropriate tradeoff parameter between the likelihood term and the FoE prior term using a complicate training procedure. But in real image denoising, accurate noise variance estimation is unavailable, even

though there are many types of noise variance estimate algorithms.<sup>4</sup>

In this work, we have developed an automatic blind method in FoE real image denoising, which can take effect without accurate noise variance estimation and automatically terminate at a good result.

## 2 Brief Review of Field of Experts Image Denoising

The basic problem of image denoising is the recovery of a latent clear image  $X$  from an observed noisy image  $Y: Y = HX + N$ , where  $N$  represents additive noise, assumed to be white Gaussian noise (WGN) with zero mean and known standard deviation  $\sigma$ . Using MAP rules to estimate the clear image  $X$ , we can maximize the posterior probability  $\hat{X}_{MAP} = \text{Arg max } P(X|Y) = \text{Arg max } P(Y|X)P(X)$ , in which the conditional possibility is

$$P(Y|X) \propto \prod_k \exp\left[-\frac{1}{2\sigma^2}(y_k - x_k)^2\right], \quad (1)$$

where  $k$  ranges over the pixels in the image. Modeling image  $X$  as a MRF, and according to the Hammersley-Clifford theorem, we can write the probability density of this graphical model as a Gibbs distribution,

$$P(X) = \frac{1}{Z(X)} \exp\left[-\sum_k V_k(X_{(k)})\right],$$

where  $V_k[X_{(k)}]$  is the potential function for clique  $X_{(k)}$ , and  $Z(X)$  is the normalized constant, called the partition function. In the FoE framework, the probability density of image  $P(X)$  can be written as,

$$P_{\text{FoE}}(X) = \frac{1}{Z(\Theta)} \prod_k \prod_{i=1}^N \phi_i(J_i^T X_{(k)}; a_i).$$

From Eq. (1), the gradient of the log-likelihood is written as  $\nabla_x \log p(Y|X) = (Y - X) / \sigma^2$ , and the gradient of the log-prior is

$$\nabla_x \log p(X) = \sum_{i=1}^N J(i)^- * \psi_i[J(i) * X]. \quad (2)$$

$J(i) * X$  denotes the convolution of image  $X$  with filter  $J(i)$ ;  $\psi_i(Y)$  is the log differential of experts, that is,  $\psi_i(Y) = \partial \log \phi_i(Y; a_i) / \partial Y$ ;  $J(i)^-$  denotes the filter obtained by mirroring  $J(i)$  around its center pixel;  $\phi_i$  obeys the function form of the Student t-distribution as in Ref. 2. By introducing an iteration index  $t$ , an update rate  $\eta$ , and an optional weight  $\lambda$ , we can write the gradient descent algorithm as:

$$X^{(t+1)} = X^{(t)} + \eta \left\{ \sum_{i=1}^N J(i)^- * \psi_i[J(i) * X^{(t)}] + \frac{\lambda}{\sigma^2} \cdot (Y - X^{(t)}) \right\}. \quad (3)$$

It could be seen that the denoised image can be obtained by gradient descent optimization in the FoE image denoising method (Roth-FoE). More details can be found in Ref. 2.

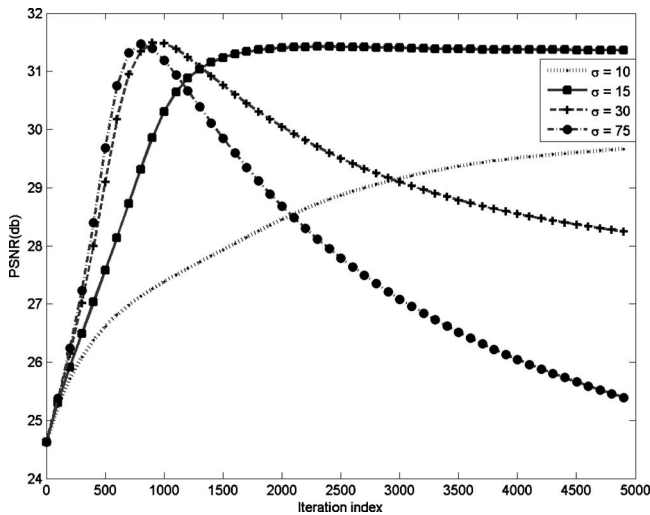


Fig. 1 Denoising performance of noisy image ( $\sigma=15$ ) using the  $5 \times 5$  FoE models with varying noise level ( $\hat{\sigma} \in \{10, 15, 30, 75\}$ ) shown in terms of PSNR.

### 3 Parameter Selection in Roth/Field-of-Experts

In Roth-FoE, two important parameters, the tradeoff parameter  $\lambda$  and stopping criterion, have an important impact on the final image quality and its computational complexity, so they must be specified before iteration begins. However, tuning of the parameters is not an easy task. In computer vision tasks, two stopping criteria were widely used: a fixed number of iterations or a predefined threshold, which the difference between adjacent iterations becoming smaller. However, these two strategies both have to choose the threshold or the number of iterations by experience, so they cannot be done completely automatically for all images.

Assuming a known noise distribution in synthetic image denoising, Roth experimentally determined the appropriate  $\lambda$  using a complicated training procedure with synthetic data and specified a fixed number of iterations. The Roth-FoE can generate satisfied results in the synthetic experiment with known noise variance.

We simulated noisy data by adding WGN to the same test image Castle from the Berkeley segmentation dataset, as in Roth's work. The true noise level we added was  $\sigma = 15$ , and the noise levels while using Roth-FoE are four different hypotheses with  $\hat{\sigma} \in \{10, 15, 30, 75\}$ . Other parameters were selected as recommended in Ref. 2. The peak signal-to-noise ratio (PSNR) curves of four experiences with 5000 iterations are given in Fig. 1. It shows that the noise variance estimation is vital in the performance of the final results using the stopping criterion for fixed numbers of iterations. However, if we can find an effective automatic stopping criterion to stop the algorithm at optimal times, such as using  $\hat{\sigma} \in \{30, 75\}$ , we can get the satisfied results, or even better than the results by using the perfect noise variance estimation with  $\hat{\sigma}=15$ . Unfortunately, in real image denoising, there is no ground truth and the accurate noise variance is unavailable, even though there are many types of noise variance estimate algorithms.<sup>4</sup> So the Roth-FoE is restricted to simulated experiments, and cannot be directly applied to real image denoising applications.

### 4 Automatic Stopping Criterion

According to Eq. (3), Roth-FoE can be understood as a regularization algorithm in inverse problems, and  $\lambda/\sigma^2$  can be considered as the regularization factor to balance the impact between the two terms: a data fidelity term that measures the likelihood of the input image given the output, and a prior term that encodes prior assumptions about the output. In this work, motivated by Fig. 1, we considered the parameter  $\lambda/\sigma^2$  as fixed and adopted a no-reference image quality assessment based on singular value decomposition (SVD) to design an optimal automatic stopping criterion for FoE-based image denoising.

Considering an  $n \times n$  window  $\omega_k$  at point  $(i, j)$  of image  $X$ , the gradient matrix is defined as

$$\mathbf{G} = \begin{bmatrix} \vdots & \vdots \\ g_i(k) & g_j(k) \\ \vdots & \vdots \end{bmatrix}, \quad k \in \omega_k,$$

where  $[g_i(k), g_j(k)]^T$  denotes the gradient of the image at point  $(i, j)$ . Computing the SVD of  $\mathbf{G}$  and assuming  $s_1 \geq s_2 \geq 0$ , it can be obtained that

$$\mathbf{G} = \mathbf{U}\mathbf{S}\mathbf{V}^T = \mathbf{U} \begin{bmatrix} s_1 & 0 \\ 0 & s_2 \end{bmatrix} [\mathbf{v}_1 \quad \mathbf{v}_2]^T.$$

The singular values  $s_1$  and  $s_2$  represent the energy in the directions  $\mathbf{v}_1$  and  $\mathbf{v}_2$ , respectively, and they reflect the strength of the gradients along the dominant direction and its perpendicular direction. Image patches can be classified into four types of idealized patches: flat, linear, quadratic, and edged regions. Define the image content metric of image patch  $\omega_k$  as

$$Q_k = s_1 \cdot \frac{s_1 - s_2}{s_1 + s_2}.$$

Zhu and Milanfar<sup>5,6</sup> demonstrated that for anisotropic patches ( $s_1 \neq s_2$ ), including the linear, anisotropic quadratic, and edged regions, the proposed metric  $Q_k$  is able to detect both blur and random noise. So in practice, when measuring the true content of an image as a whole, we can calculate  $Q$  in all anisotropic areas as

$$Q = \frac{1}{K} \sum_k Q_k. \tag{4}$$

We distinguish between isotropic and anisotropic areas by employing significance testing<sup>6</sup> based on local coherence  $R = (s_1 - s_2) / (s_1 + s_2)$ . In our experiments, we set the significance level to be 0.001, and use it to determine image patch labels. In the iteration,  $Q$  is computed every time as a sign of whether to stop the iteration to get the optimal results. The FoE denoising method with automatic stopping criterion (ASC-FoE) can be described as follows.

1. Initialization: set  $\lambda/\sigma^2=0.001$ ,  $\eta=0.1$ , and the initial iteration value  $X^{(0)}=Y$ . Compute its image quality measurement  $Q^{(0)}$  according to Eq. (4). Set  $Q_{\max}=Q^{(0)}$  and optimal iteration time  $I_{\text{opt}}=0$ .

**Table 1** PSNR, SSIM, and iteration number for Roth-FoE and ASC-FoE.

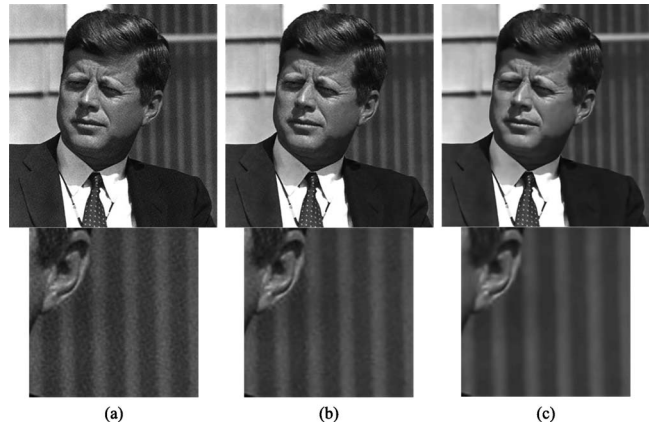
$\hat{\sigma}$	PSNR (dB)		SSIM		Iteration Number	
	ASC-FoE	Roth-FoE	ASC-FoE	Roth-FoE	ASC-FoE	Roth-FoE
25	23.2245	<b>23.2254</b>	0.5088	<b>0.5089</b>	4999	<b>5000</b>
30	<b>26.8926</b>	<b>26.8926</b>	<b>0.8041</b>	<b>0.8041</b>	<b>5000</b>	<b>5000</b>
35	<b>27.6546</b>	<b>27.6546</b>	<b>0.8047</b>	<b>0.8047</b>	<b>5000</b>	<b>5000</b>
40	<b>27.7328</b>	27.6133	<b>0.8050</b>	0.7916	4290	<b>5000</b>
50	<b>27.7814</b>	27.1281	<b>0.8047</b>	0.7916	3704	<b>5000</b>
75	<b>27.7977</b>	26.3426	<b>0.8067</b>	0.7455	3163	<b>5000</b>
100	<b>27.7933</b>	26.0431	<b>0.8066</b>	0.7364	3043	<b>5000</b>

- Iteration: update  $X^{(t+1)}$  and corresponding  $Q^{(t+1)}$  from Eqs. (3) and (4). If  $Q^{(t+1)} > Q_{\max}$ , let  $Q_{\max} = Q^{(t+1)}$ ,  $I_{\text{opt}} = t + 1$ .
- Stopping criterion: if  $Q^{(t+1+n)} < Q_{\max}$ , for  $n = 1, \dots, N$ , the iteration terminates, and let  $X^{(t+1)}$  be the denoised image.

### 5 Experimental Results

We design an experimental scheme to illustrate the image quality performance of the proposed method on both synthetic and real images. For the synthetic data, we simulated noisy data by adding WGN ( $\sigma=30$ ) to the test image Castle. We denoised it by Roth-FoE and the proposed ASC-FoE under seven different noise-level hypotheses with  $\hat{\sigma} \in \{25, 30, 35, 40, 50, 75, 100\}$ . We used the  $5 \times 5$  FoE model with 24 filters, and other parameters were selected as recommend in Ref. 2. Table 1 shows the performance comparison in terms of PSNR and structural similarity (SSIM) index as measured between the denoised image and the ground truth. Also, the iteration numbers are given out. It can be seen that the ASC-FoE outperforms the Roth-FoE and has less computational cost. It should be noted that, for fair comparison of the two methods in the synthetic ASC-FoE experiment, we used six different noise level hypotheses the same as the Roth-FoE instead of setting  $\lambda/\sigma^2 = 0.001$ , as described in ASC-FoE. In addition, using  $\lambda/\sigma^2 = 0.001$  is almost equal to using  $\hat{\sigma} = 75$ , and the result is also satisfied.

For the real noise image, the experimental results are shown in Fig. 2. We use the test image JFK ( $367 \times 343$ )<sup>7</sup> that suffers from the real noise shown in Fig. 2(a). The noise comes from film grain, scanning, and compression processes, and may not be pure Gaussian—indeed, it may be space variant. We estimated the standard deviation of the noise through the commonly used median absolute deviation (MAD) method for Roth-FoE. The measured value is  $\sigma = 4.2$ . The results of Roth-FoE and ASC-FoE are shown in Figs. 2(b) and 2(c), respectively. It can be seen that the



**Fig. 2** Denoised images and partially magnified images of Roth-FoE and ASC-FoE compared in this experiment on an image with real noise. (a) Real noisy image. (b) Roth-FoE ( $Q=27.3598$ ). (c) Proposed ASC-FoE ( $Q=28.5272$ ).

ASC-FoE outperforms the Roth-FoE in noise-suppressing capability, and obtains a more pleasant result. Moreover, the Roth-FoE requires 5000 iterations according to recommended parameter settings, while in this experiment we terminated the ASC-FoE at iteration 208 when the image quality achieved the best. Therefore, the iteration time reduction of ASC-FoE scores around 96% and would result in a speedup factor of more than 24.

### 6 Conclusion

For real image denoising applications, we develop an automatic stopping criterion in FoE image denoising. We demonstrate that our ASC-FoE can obtain a faster and more pleasant result than the original Roth-FoE without the explicit need to know the noise variance *a priori*.

### Acknowledgments

The authors are grateful to the associate editor Eddie Jacobs and the anonymous reviewers for their efforts, comments, and recommendations, which have led to a substantial improvement of this work. This study was partially supported by the National Natural Science Foundation of China (Grant No. 60902088) and NDTF Project of the ATR Laboratory (Grant No. 9140C8004011005).

### References

- C. Liu, R. Szeliski, S. B. Kang, C. L. Zitnick, and W. T. Freeman, "Automatic estimation and removal of noise from a single image," *IEEE Trans. Pattern Anal. Mach. Intell.* **30**, 299–314 (2008).
- S. Roth and M. J. Black, "Fields of experts," *Int. J. Comput. Vis.* **82**, 205–229 (2009).
- V. Katkovnik, A. Foi, K. Egiazarian, and J. Astola, "From local kernel to nonlocal multiple-model image denoising," *Int. J. Comput. Vis.* **86**, 1–32 (2010).
- D. Zoran and Y. Weiss, "Scale invariance and noise in natural images," in *Proc. IEEE Intl. Conf. on Computer Vision (ICCV09)*, Tokyo, Japan, pp. 64–69, IEEE, Piscataway, NJ (2009).
- X. Zhu and P. Milanfar, "A no-reference sharpness metric sensitive to blur and noise," presented at 1st Intl. Workshop on Quality of Multimedia Experience (QoMEX), San Diego, 2009.
- X. Zhu and P. Milanfar, "Automatic parameter selection for denoising algorithms using a no-reference measure of image content," *IEEE Trans. Image Process.* (in press).
- H. Takeda, S. Farsiu, and P. Milanfar, "Kernel regression for image processing and reconstruction," *IEEE Trans. Image Process.* **16**, 349–366 (2007).