Blur kernel estimation via salient edges and nonlocal regularization

Suil Son[1] and Suk I. Yoo[1]
[1] Department of Computer Science and Engineering, Seoul National University, Seoul, Republic of Korea
e-mail: likerivers12@snu.ac.kr, sukinyoo@snu.ac.kr

Abstract— **Blind image deblurring is a severely ill-posed inverse problem. To obtain a high quality latent image from a single blurred one, effective regularizations are required. In this paper, we propose a nonlocal regularization to improve blur kernel estimation. Under convolution operation, even similar patches could result in the quite different values. However, if the estimated kernel is correct, the nonlocal similar patches weighted by that kernel may result in the similar value by convolution. Therefore, the weighted nonlocal patches can improve the kernel estimation. We extract the nonlocal patches in terms of the weighted similarity by the kernel and then use them for regularization of the kernel estimation. Since the nonlocal regularization is a data-authentic prior, our approach not only mitigates the ill-posedness but also imposes the effective prior to kernel estimation. Experimental results show that our approach outperforms conventional blind deblurring algorithms.**

Keywords— **Blind image deblurring, nonlocal regularization, salient edges, blur kernel estimation, image restoration.**

I. Introduction

Image deblurring is a classical and challenging inverse problem that aims to recover a high-quality image from its associated blurry-noisy one. It has gained considerable attention in recent years because it involves many challenges in problem formulation, regularization, and optimization. The formation process of image blur is usually modeled as

$$B = k * I + n \tag{1}$$

where B, k, I and n represent the blurred image, blur kernel, latent image, and the additive noise, respectively. Here, $*$ denotes the convolution operator. Since obtaining the latent image is a severely ill-posed inverse problem, a good regularization is required in order to obtain a high quality latent image.

When the blur kernel is unknown, the image deblurring methods are considered blind. To make blind deblurring more tractable, various image priors have been proposed. Fergus *et al.* [1] adopted a zero mean mixture of Gaussian to fit for natural image gradients. Shan *et al.* [2] used a parametric model to approximate the heavy-tailed natural image prior. In addition, to avoid the delta kernel solution, they used a large regularization weight to suppress insignificant structures and preserve strong ones. Joshi *et al.* [3] and Cho *et al.* [4] directly restored sharp edges from blurred images and relied on them to estimate blur kernels. Levin *et al.* [5], [6] proposed a hyper-Laplacian prior to fit the distribution of natural image gradients. In the work of [5], they illustrated the limitation of the simple maximum a posteriori (*MAP*) approach, and proposed an efficient marginal likelihood approximation in [6]. To overcome the limitation of *MAP*, Krishnan *et al.* [7] proposed a normalized sparsity prior. In the formulation of [7], latent images were favored over blurred ones by the new regularizer. Goldstein and Fattal [8] estimated blur kernels by spectral irregularities.

An explicit edge prediction step for kernel estimation was also employed in various methods. In [3], Joshi *et al.* computed sharp edges by first locating step edges and then propagating the local intensity extrema toward the edge. Cho *et al.* [4] detected sharp edges from blurred images directly, and then they employed the Radon transform to estimate the blur kernel. However, these methods have difficulty in dealing with large blur. Cho and Lee [9] used bilateral filtering and shock filtering to predict sharp edges iteratively, and then they selected the salient edges for kernel estimation. Xu and Jia [10] adaptively selected useful edges for kernel estimation by using an effective mask computation algorithm. The kernel refinement was achieved by iterative support detection(ISD) method [11]. However, this method ignores the continuity of the motion blur kernel. Hu and Tang [12] learned good regions for kernel estimation and employed method [9] to estimate kernels. Although the performance is greatly improved, the sparsity and continuity of blur kernels still cannot be guaranteed.

Additional priors and regularization were also applied. Sun *et al.* [13] employed patch priors to restore useful edges for kernel estimation. To achieve better selection of edges, Xu *et al.* [14] and Pan and Su [15] proposed l_0-regularized kernel estimation. Due to the properties of l_0-norm, these methods are able to select large scale edges for kernel estimation. Sparse representation as a new powerful image prior has also been employed in blind deblurring [16], [17], [18]. Because sparse representation is able to provide data-authentic priors in the kernel estimation, these methods have also achieved state-of-the-art results.

Like the sparse representation, nonlocal information also provides data-authentic priors; therefore, we make use of this prior to improve the blur kernel estimation. Ordinary nonlocal similar patches have possibility of resulting in quite different values under convolution operation because the convolutional kernel amplifies or reduces the effect of each element depending on its values. In this paper, we select nonlocal similar patches based on the similarity weighted by the kernel. If the estimated kernel is correct, those nonlocal patches may result

978-1-4799-8637-8/15/$31.00 ©2015 IEEE

in the similar value by convolution. Therefore, the weighted nonlocal patches can be used to improve the kernel estimation. In addition, previous works [19], [20] show that the salient edges help kernel estimation. Therefore, we utilize the salient edges not only to estimate the blur kernel but also to extract candidate nonlocal patches.

The remainder of this paper is organized as follows. We firstly visit the background of blind deconvolution algorithm in Section II. Then we explain the method to extract salient structure in Section III. We show how to use the salient edges and the nonlocal regularization for improving the blur kernel estimation in Section IV. The final latent image estimation is explained in Section V. We illustrate the experimental results of our approach with comparison to other conventional methods in Section VI. Finally, we conclude in Section VII.

II. Background

When the formation process of a blurry image B is modeled as (1), blind deconvolution is to solve the following regularized minimization

$$\min_{I,k} \|B - k * I\|_2 + \lambda J(I) + \gamma G(k) \tag{2}$$

where the first term enforces the convolutional blur model, the functionals $J(I)$ and $G(k)$ are the priors for I and k, and λ and γ two nonnegative parameters that weigh their contribution. Additional constraints on k, such as positivity of its entries and integration to 1, are usually included.

Blind deconvolution can also be represented by the probabilistic formulation. In the Bayesian framework, this formulation corresponds to maximizing the posterior distribution, *MAP*,

$$\arg\max_{I,k} p(I,k|B) = \arg\max_{I,k} p(B|I,k)p(I)p(k), \tag{3}$$

where $p(B|I,k)$ models the noise affecting the blurry image; $p(I)$ models the distribution of typical sharp images; and $p(k)$ is the prior knowledge about the blur function. $p(B|I,k)$ was usually modeled by the Gaussian distribution [1], [6] or an exponential distribution [10]; $p(I)$ was by a heavy-tailed distribution of the image gradients; and $p(k)$ was by a Gaussian distribution [10], [9], a sparsity-including distribution [1], [2] or a uniform distribution [6].

Since under these assumptions the $MAP_{I,k}$ problem (3) is equivalent to problem (2), also the Bayesian approach suffers from local minima. Levin *et al.* [6] and Fergus *et al.* [1] propose to address the problem by marginalizing over all possible sharp images I and thus solve the following MAP_k problem

$$\arg\max_{k} p(k|B) = \arg\max_{k} \int p(B|I,k)p(I)p(k)dI. \tag{4}$$

Then they estimate I by solving a convex problem where k is given from the previous step. Levin *et al.* [5] showed that, using a sparsity-inducing prior for the image gradients and a uniform distribution for the blur, the $MAP_{I,k}$ approach favors the no-blur solution ($I = B, k = \delta$). In addition, they showed that, for sufficiently large images, the MAP_k approach converges to the true solution. Recently, despite the limitation of the $MAP_{I,K}$ approach, Perrone and Favaro [21] showed that the normalization of the kernel in the iterative step of the blur kernel estimation gives the desired solution to the kernel estimation problem. Therefore, iteratively obtaining I and k with normalization is an effective method to blind deconvolution.

For the kernel estimation, in spite of the formulation in Eq. (2), most blind deconvolution approaches solve the problem in the gradient domain as

$$\min_{I,k} \|\nabla B - k * \nabla I\|_2 + \lambda J(I) + \gamma G(k). \tag{5}$$

In this paper, we use both of the formulations. Specifically, we apply Eq. (5) for the estimation of k and Eq. (2) for the estimation of I.

III. Extracting salient structure

In order to avoid the delta kernel solution and get a better kernel estimate, most iterative deblurring methods [9], [10] are carefully designed to select sharp edges for kernel estimation. According to Pan *et al.* [19], image details caused by blur could adversely affect the kernel estimation. One effective way to remove these detail is to apply image denoising model based on the total variation (*TV*). To increase the robustness of kernel estimation, we adopt the edge selection method [19] to select the salient edges for kernel estimation. In addition, to mitigate the stair-casing effect in smooth area, we adjust the value of θ to be large in the smooth areas, and small near the edges as in [10], [19]. Thus, we use the following adaptive model to extract the main structure of an image I:

$$\min_{I_s} \|\nabla I_s\|_2 + \frac{1}{2\theta\omega(\mathbf{x})}\|I_s - I\|_2^2 \tag{6}$$

where $\omega(\mathbf{x}) = \exp(-\|r(\mathbf{x})\|^{0.8})$. Here, $r(\mathbf{x})$ is a criterion for selecting informative edges which is given by

$$r(\mathbf{x}) = \frac{\|\sum_{\mathbf{y}\in N_h(\mathbf{x})} \nabla B(\mathbf{y})\|_2}{\sum_{\mathbf{y}\in N_h(\mathbf{x})} \|\nabla B(\mathbf{y})\|_2 + 0.5} \tag{7}$$

where B is the blurred image, and $N_h(\mathbf{x})$ is an $h \times h$ window centered at pixel $\mathbf{x}$. The constant 0.5 is to prevent producing a large r in flat regions. If a local region is flat, then the associated r has a small value. On the other hand, if a local region contains strong image structures, the associated r has a large value. This has a strong penalty to these areas which are flat or contain narrow strips as well.

After computing I_s, we apply a gradient selection method to mitigate the possible adverse effect of salient edges. Using a shock filter [22], we compute the enhanced structure $\tilde{I}_s$ as

$$\frac{\partial \tilde{I}_s}{\partial t} = -\text{sign}(\Delta I_s \|\nabla I_s\|_2), \tag{8}$$

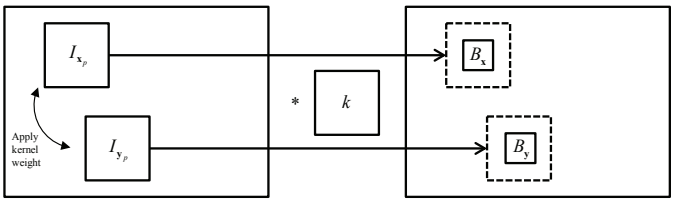


Fig. 1. Illustration of the nonlocal operation.

where $\Delta I_s = I_x^2 I_{xx} + 2I_x I_y I_{xy} + I_y^2 I_{yy}$. Finally, we compute salient edges ∇S as

$$\nabla S = \nabla \tilde{I}_s \odot H(\nabla \tilde{I}_s, t), \tag{9}$$

where $\odot$ is the element-wise multiplication, t is a threshold, and $H(\nabla \tilde{I_S}, t)$ is the unit binary mask function which is defined as

$$H(\nabla \tilde{I}_s, t) = \begin{cases} 1, & \|\nabla \tilde{I}_s\|_2 \geq t, \\ 0, & \text{otherwise.} \end{cases} \tag{10}$$

The application of Eq. (8) eliminates some noise in the $\|\nabla \tilde{I}_s\|_2$; thus, only the salient edges with large values influence the kernel estimation.

IV. Blur Kernel estimation

A. Estimating a blur kernel

The motion blur kernel describes the path of camera shake during the expose. This path usually looks like a narrow curve. Thus, most literature assumes that distributions of blur kernels can be modeled by a hyper-Laplacian. Under this assumption, the model for kernel estimation using the salient edges ∇S is

$$\begin{aligned} &\min_k \|\nabla B - k * \nabla S\|_2^2 + \gamma \|k\|_\alpha^\alpha, \\ &\text{s.t. } k \succeq 0, \ \|k\|_1 = 1, \end{aligned} \tag{11}$$

where $0 < \alpha < 1$.

Since a camera shake is a continuous movement, the shape of blur kernel is also continuous. To preserve the continuity of kernel, we apply an additional regularization by the following spatial term as in [19]:

$$C(k) = \{(x, y) \mid |\partial_x k(x, y)| + |\partial_y k(x, y)| \neq 0\} \tag{12}$$

where $0 < \alpha < 1$. $C(k)$ counts the number of pixels whose gradients are non-zero. This help the kernel estimation to keep the structure of kernel effectively and to remove some noise.

We additionally introduce the nonlocal prior such that patches after blurring are similar if their latent patches are similar. For the application of this prior, we firstly select nonlocal similar patches based on the similarity weighted by the estimating kernel. Ordinary nonlocal similar patches could have quite different value under convolution operation because the convolution amplifies or reduces the effect of each element depending on the kernel. If the kernel is correct, the nonlocal similar patches by kernel weight results in the similar value after convolution. Therefore, this prior knowledge helps the estimation of the correct kernel. We propose the nonlocal regularization as

$$\begin{aligned} J_{nl}(I, B, k) = \sum_{\{\mathbf{x} \in \Omega\}} \sum_{\{\mathbf{y} \in NL_K(\mathbf{x})\}} &\{((I_{\mathbf{x}_p} * k)_c - B_{\mathbf{x}}) \\ &-((I_{\mathbf{y}_p} * k)_c - B_{\mathbf{y}})\}, \end{aligned} \tag{13}$$

where Ω is a set of pixel locations that are counted for the nonlocal cost, $NL_K(\mathbf{x})$ is a set of nonlocal similar locations to $\mathbf{x}$, $\mathbf{x}_p$ is the patch whose center is $\mathbf{x}$, $\mathbf{y}_p$ is the nonlocal similar patch to $\mathbf{x}_p$ whose center is $\mathbf{y}$. Thus, $I_{\mathbf{x}_p}$ is the patch on I whose center is $\mathbf{x}$, and the $B_{\mathbf{x}}$ is the pixel $\mathbf{x}$ on B. The subscript c means extracting the pixel of the center of the patch. We extracted only the center of the patch to avoid the boundary problem of convolution. In Fig. 1, the simple nonlocal operation is illustrated for example. The set of nonlocal similar locations is defined as

$$NL_K(\mathbf{x}) = \{ \mathbf{y} \mid d(I_{\mathbf{x}_p}, I_{\mathbf{y}_p}) \leq K\text{-}dist(I_{\mathbf{x}_p}), \mathbf{y} \in \Omega\} \tag{14}$$

where $d(I_{\mathbf{x}_p}, I_{\mathbf{y}_p})$ is the distance between patches, and $K\text{-}dist(I_{\mathbf{x}_p})$ is the distance of the K^{th}-nearest patch from the patch $I_{\mathbf{x}_p}$. The distance measure is given by

$$\begin{aligned} d(I_{\mathbf{x}_p}, I_{\mathbf{y}_p}) = \| \ (I_{\mathbf{x}_p} \odot (k_- + \frac{1}{h \cdot w})) \\ -(I_{\mathbf{y}_p} \odot (k_- + \frac{1}{h \cdot w})) \ \|_F, \end{aligned} \tag{15}$$

where k_- is the flipped version of k, h and w are the height and the width of the kernel, respectively, $\odot$ is the element-wise multiplication, and $\|\cdot\|_F$ means the Frobenius norm. Here, k_- is applied because convolution operation measures correlation between an image patch and the flipped version of the kernel. The operation $\odot$ applies the kernel weight for each element of a patch, and $\frac{1}{h \cdot w}$ applies a minimal contribution for every element of the patch.

To effectively impose the nonlocal regularization, some details should be considered. First, we use the size of nonlocal patch as the same to the size of kernel. Since J_{nl} is computed from the convolved value of the patch, the size of patch should be at least the size of the kernel. Nonetheless, larger size patches may deteriorate the selection of similar patches. Therefore, we use the same size for the patch and the kernel and use only the center of the convolved patch for regularization. Second, we extract the candidate patches only from the salient edges for the same reason of avoiding the noise effect. For this candidate locations, Ω in Eq. (13) consists of the locations at which the value of H in Eq. (10) is non-zero.

Based on the above considerations, our kernel estimation model is defined as

$$\begin{aligned} &\min_k \|\nabla B - k * \nabla S\|_2^2 + \gamma \|k\|_\alpha^\alpha + \mu C(k) + \lambda J_{nl}(I, B, k), \\ &\text{s.t. } k \succeq 0, \ \|k\|_1 = 1, \end{aligned} \tag{16}$$

where γ, μ, and λ controls the contribution of each term.

Model (16) is difficult to be minimized directly because of the discrete counting metric. Thus, we approximate it by alternately minimizing

$$\min_k \|\nabla B - k * \nabla S\|_2^2 + \gamma \|k\|_\alpha^\alpha + \lambda J_{nl}(I, B, k), \quad \text{s.t. } k \succeq 0, \ \|k\|_1 = 1, \tag{17}$$

and

$$\min_{\hat{k}} \|\hat{k} - k\|_2^2 + \mu C(\hat{k}). \tag{18}$$

Model (17) can be optimized by using the constrained iterative reweighted least square (*IRLS*) method [23]. For model (18), we employ the alternating optimization method in [24].

B. Estimating an intermediate latent image

In this stage, we focus on the sharp edges restoration from the blurred image. Thus, we employ the anisotropic *TV* model because it has been proven an effective method which can be efficiently solved by many fast algorithms.

$$\min_I \|B - k * I\|_2^2 + \beta \|\nabla I\|_1. \tag{19}$$

We use the *IRLS* method to solve this model.

C. Multiscale implementation

For large kernels, an excessive number of I and k updates may be required to converge to a reasonable solution. To mitigate this problem, we perform multiscale estimation of the kernel using a coarse-to-fine pyramid of image resolutions, in a similar manner as in [1]. We use levels with a size ratio of $\sqrt{2}/2$ between them. The number of levels is determined by the size of the kernel k such that the kernel size at the coarsest level is 3×3. We downsample the input blurry image and then take discrete gradient to form the input ∇I each level. Once a kernel estimate k and the latent image I are computed, they are upsampled to act as the initialization of the kernel and sharp image at the next finer level. The resizing operations are done by using bilinear interpolation.

V. Final Latent Image Estimation

Once the kernel k has been estimated, we can use a variety of non-blind deconvolution methods to recover the latent image I from blurred image B. Model (19) can be used for this, but it may lead to the stair-casing effect and destroy textures. To overcome this problem, some adaptive regularization terms have been proposed and proved to be effective in edge-preserving [25]. Therefore, we utilize the predicted structure to guide the latent image restoration. The final latent image is obtained by solving

$$\min_I \|B - k * I\|_2^2 + \eta(\exp(-\|\partial_x S\|^{0.8}) \cdot \|\partial_x I\|_1 + \exp(-\|\partial_y S\|^{0.8}) \cdot \|\partial_y I\|_1). \tag{20}$$

This model can also be solved by the *IRLS* method efficiently.

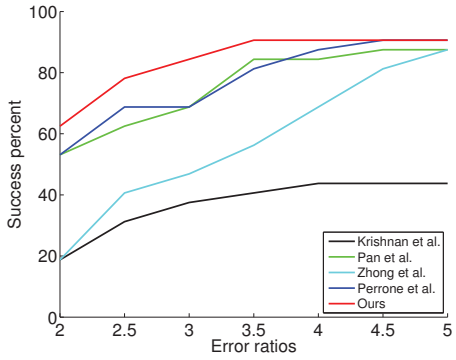


Fig. 2. Cumulative error ratio across test examples.

VI. Experimental Results

In this section, we evaluated our method and compared it with other conventional methods quantitatively and qualitatively. For the experiments, we set the parameters as follows. The value of θ in Eq. (6) was set to 1. In Eq. (16), γ was set to 0.01; μ to 0.01; and λ to from 0.01 to 0.1. In Eq. (19), β was set to 0.005. In Eq. (20), η was set to 0.003. In Eq. (14), K was set to 5.

We performed quantitative evaluation of our method using the data set from [5]. The test data consists of 32 examples with 8 kernels. The error ratio proposed by [5] was used to measure the quality of estimated results. Our method was compared with those of Krishnan *et al.* [7], Pan *et al.* [19], Zhong *et al.* [26], and Perrone *et al.* [21]. The comparison results are shown in Fig. 2. As can be seen, our method achieved the highest performance. In addition, Fig. 3 shows one example from the test dataset. As shown in the figure, The kernel estimates of Perrone *et al.* [21] and ours look more similar to the ground truth than others. These results demonstrate the effectiveness of the proposed method.

We also conducted experiments on various images and qualitatively examined the visual quality of the deblurred images. We tested on the image with some noise and listed the results in Fig. 4. As shown in the figure, the results of Zhong *et al.* [26] and Perrone *et al.* [21] contains some coarse texture which seem to be the effect of the noise. The results of Krishnan *et al.* [7], Pan *et al.* [19], and ours look smoother, and our result shows more contrastive visualization. Also we tested on the images with more structures and listed the results in Figs. 5 and 6. In Fig. 5, most of the methods show quite good results, but the result of Perrone *et al.* [21] has much of stair-casing effects. In Fig. 6, the results of Krishnan *et al.* [7] and Zhong *et al.* [26] failed to restore the sharp image. The result of Perrone *et al.* [21] shows a sharp visualization but it also contains some stair-casing effects. The results of Pan *et al.* [19] and ours show better visualization than others, and our result is more contrastive. These results show that our approach is effective in deblurring and restore more contrastive sharp images.

(a) Blurred image

(b) Ground truth

(c) Krishnan *et al.* [7]

(d) Pan *et al.* [19]

(e) Zhong *et al.* [26]

(f) Perrone *et al.* [21]

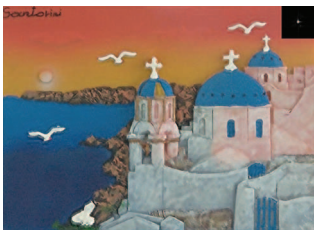
(g) Ours

Fig. 3. Comparison with other methods on a synthetic example

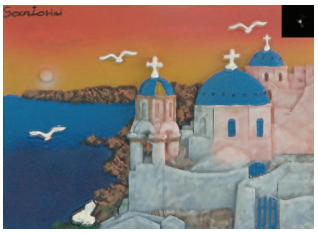
VII. Conclusion

In this paper, we have presented an effective method for blind image deblurring. The proposed method combines salient edges and nonlocal regularization. The salient edges provide reliable edge information, and the nonlocal prior offers data-authentic information. For the nonlocal regularization, we selected the nonlocal similar patches in terms of weighted similarity by the blur kernel. This nonlocal regularization improved the estimation of the blur kernel. Both quantitative and qualitative evaluations demonstrate that the proposed method performs favorably against several conventional algorithms.

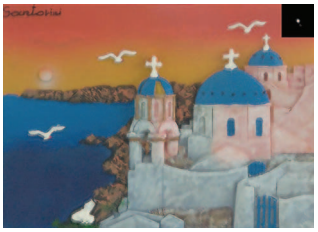
(a) Blurred image



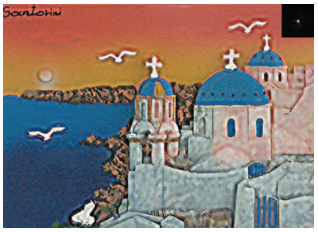
(b) Krishnan *et al.* [7]



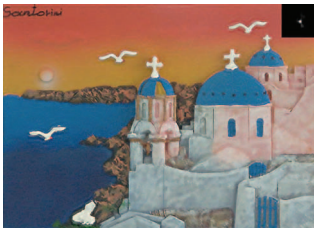
(c) Pan *et al.* [19]



(d) Zhong *et al.* [26]



(e) Perrone *et al.* [21]



(f) Ours

Fig. 4. Deblurring examples

References

[1] R. Fergus, B. Singh, A. Hertzmann, S. T. Roweis, and W. T. Freeman, "Removing camera shake from a single photograph," *ACM Transactions on Graphics (TOG)*, vol. 25, no. 3, pp. 787–794, 2006.

[2] Q. Shan, J. Jia, and A. Agarwala, "High-quality motion deblurring from a single image," in *ACM Transactions on Graphics (TOG)*, vol. 27, no. 3. ACM, 2008, p. 73.

[3] N. Joshi, R. Szeliski, and D. Kriegman, "Psf estimation using sharp edge prediction," in *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*. IEEE, 2008, pp. 1–8.

[4] T. S. Cho, S. Paris, B. K. Horn, and W. T. Freeman, "Blur kernel estimation using the radon transform," in *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*. IEEE, 2011, pp. 241–248.

[5] A. Levin, Y. Weiss, F. Durand, and W. T. Freeman, "Understanding and evaluating blind deconvolution algorithms," in *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*. IEEE, 2009, pp. 1964–1971.

[6] ——, "Efficient marginal likelihood optimization in blind deconvolution," in *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*. IEEE, 2011, pp. 2657–2664.

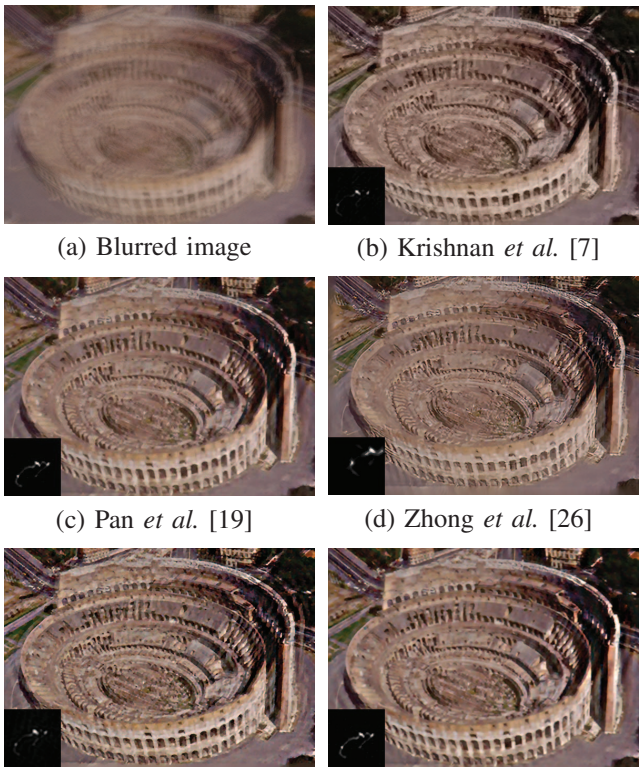
[7] D. Krishnan, T. Tay, and R. Fergus, "Blind deconvolution using a normalized sparsity measure," in *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*. IEEE, 2011, pp. 233–240.

[8] A. Goldstein and R. Fattal, "Blur-kernel estimation from spectral irregularities," in *Computer Vision–ECCV 2012*. Springer, 2012, pp. 622–635.

[9] S. Cho and S. Lee, "Fast motion deblurring," in *ACM Transactions on Graphics (TOG)*, vol. 28, no. 5. ACM, 2009, p. 145.

[10] L. Xu and J. Jia, "Two-phase kernel estimation for robust motion deblurring," in *Computer Vision–ECCV 2010*. Springer, 2010, pp. 157–170.

[11] Y. Wang and W. Yin, "Compressed sensing via iterative support de-

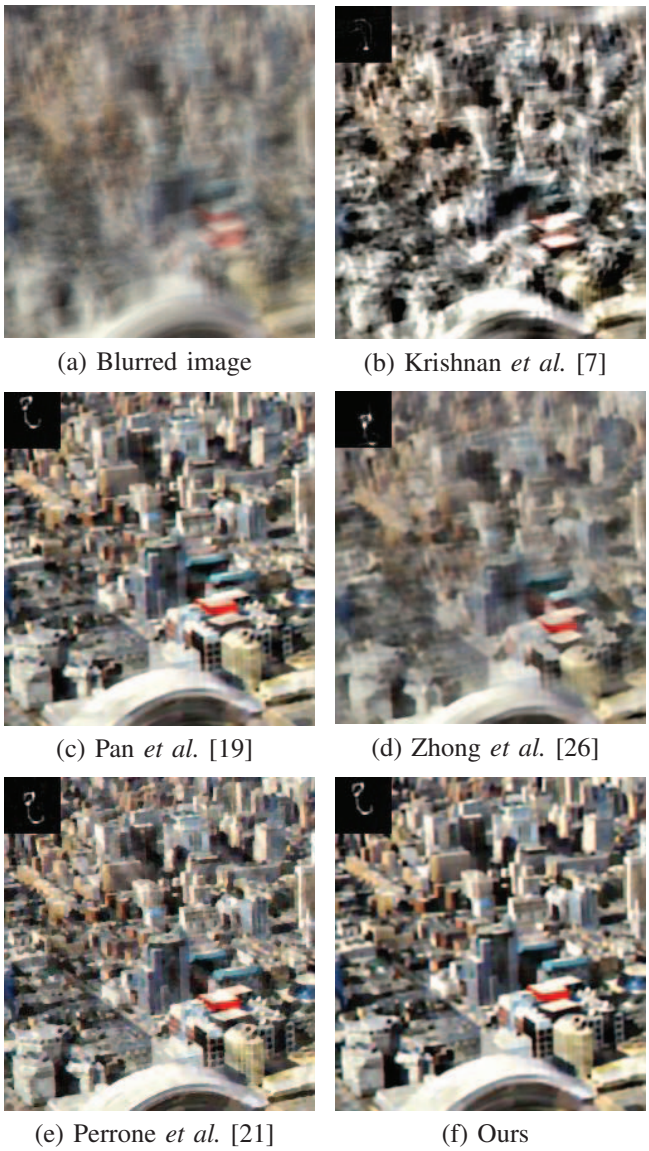


(a) Blurred image (b) Krishnan *et al.* [7]

(c) Pan *et al.* [19] (d) Zhong *et al.* [26]

(e) Perrone *et al.* [21] (f) Ours

Fig. 5. Deblurring examples



(a) Blurred image (b) Krishnan *et al.* [7]

(c) Pan *et al.* [19] (d) Zhong *et al.* [26]

(e) Perrone *et al.* [21] (f) Ours

Fig. 6. Deblurring examples

tection," *Rice University CAAM Technical Report TR09-30, http://www. caam. rice. edu/˜ wy1/paperfiles/Rice_ CAAM_ TR09-30. pdf*, 2009.

[12] Z. Hu and M.-H. Yang, "Good regions to deblur," in *Computer Vision–ECCV 2012*. Springer, 2012, pp. 59–72.

[13] L. Sun, S. Cho, J. Wang, and J. Hays, "Edge-based blur kernel estimation using patch priors," in *Computational Photography (ICCP), 2013 IEEE International Conference on*. IEEE, 2013, pp. 1–8.

[14] L. Xu, S. Zheng, and J. Jia, "Unnatural l0 sparse representation for natural image deblurring," in *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*. IEEE, 2013, pp. 1107–1114.

[15] J. Pan and Z. Su, "Fast l0-regularized kernel estimation for robust motion deblurring," *Signal Processing Letters, IEEE*, vol. 20, no. 9, pp. 841–844, 2013.

[16] Z. Hu, J.-B. Huang, and M.-H. Yang, "Single image deblurring with adaptive dictionary learning," in *Image Processing (ICIP), 2010 17th IEEE International Conference on*. IEEE, 2010, pp. 1169–1172.

[17] H. Zhang, J. Yang, Y. Zhang, and S. T. Huang, "Sparse representation based blind image deblurring," in *Advances in Neural Information Processing Systems*, 2014, pp. 3005–3013.

[18] J.-F. Cai, H. Ji, C. Liu, and Z. Shen, "Framelet-based blind motion deblurring from a single image," *Image Processing, IEEE Transactions on*, vol. 21, no. 2, pp. 562–572, 2012.

[19] J. Pan, R. Liu, Z. Su, and X. Gu, "Kernel estimation from salient structure for robust motion deblurring," *Signal Processing: Image Communication*, vol. 28, no. 9, pp. 1156–1170, 2013.

[20] J. Pan, R. Liu, Z. Su, and G. Liu, "Motion blur kernel estimation via salient edges and low rank prior," in *Multimedia and Expo (ICME), 2014 IEEE International Conference on*. IEEE, 2014, pp. 1–6.

[21] D. Perrone and P. Favaro, "Total variation blind deconvolution: The devil is in the details," in *Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on*. IEEE, 2014, pp. 2909–2916.

[22] S. Osher and L. I. Rudin, "Feature-oriented image enhancement using shock filters," *SIAM Journal on Numerical Analysis*, vol. 27, no. 4, pp. 919–940, 1990.

[23] A. Levin, R. Fergus, F. Durand, and W. T. Freeman, "Image and depth from a conventional camera with a coded aperture," in *ACM Transactions on Graphics (TOG)*, vol. 26, no. 3. ACM, 2007, p. 70.

[24] L. Xu, C. Lu, Y. Xu, and J. Jia, "Image smoothing via l 0 gradient minimization," in *ACM Transactions on Graphics (TOG)*, vol. 30, no. 6. ACM, 2011, p. 174.

[25] Z. Farbman, R. Fattal, D. Lischinski, and R. Szeliski, "Edge-preserving decompositions for multi-scale tone and detail manipulation," in *ACM Transactions on Graphics (TOG)*, vol. 27, no. 3. ACM, 2008, p. 67.

[26] L. Zhong, S. Cho, D. Metaxas, S. Paris, and J. Wang, "Handling noise in single image deblurring using directional filters," in *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*. IEEE, 2013, pp. 612–619.