

# IMAGE DEBLURRING USING ROBUST SPARSITY PRIORS

Xinxin Zhang, Ronggang Wang, Yonghong Tian, Wenmin Wang, Wen Gao

School of Electronic and Computer Engineering  
Peking University Shenzhen Graduate School  
Lishui Road 2199, Nanshan District, Shenzhen, China 518055

## ABSTRACT

In this paper, we propose a robust method to remove motion blur from a single photograph. We find that an inaccurate kernel and an unreliable final latent image reconstruction method are two main factors leading to low-quality restored images. To improve image quality, we do the following technical contributions. For robust blur kernel estimation, first, an edge mask and a smooth constraint are used to provide reliable intermediate latent images for salient structure extraction; second, we adopt an effective salient structure selection method to remove detrimental edges for kernel estimation; third, we use a gradient sparsity prior to remove kernel noise and ensure the continuity of blur kernels. For final latent image reconstruction, we combine the merits of both the TV- $l_2$  model and the hyper-Laplacian model to preserve tiny details and eliminate noise. Experimental results on synthetically blurred images and real photographs demonstrate that the proposed algorithm performs better than state-of-the-art approaches.

**Index Terms**— Motion blur, deblurring, kernel estimation, image reconstruction

## 1. INTRODUCTION

Motion blur caused by a relative motion between a camera and a scene is a common phenomenon in the process of imaging. It is usually modeled as a convolution process:

$$I = L \otimes k + N, \quad (1)$$

where  $I$ ,  $L$ ,  $k$  and  $N$  represent the motion blurred image, latent image, blur kernel and unknown sensor noise respectively,  $\otimes$  denotes the convolution operator. The objective is to recover  $L$  from  $I$  without specific knowledge of  $k$ , so single image blind deconvolution is an ill-posed problem.

Blind image deblurring has gained much attention and made significant progress in recent years. Although many approaches [1, 2, 3, 4, 5] perform well in some particular scenarios, e.g., low-light images, text and face images, none of them is robust enough in all cases.

Thanks to National Science Foundation of China 61370115, China 863 project of 2015AA015905, and Shenzhen Peacock Plan for funding.

The success of most existing algorithms can be attributed to the use of a wide range of parametric image priors. These constraints are used to avoid local minima, dense kernels and visual artifacts in the restored image. Fergus *et al.* [6] used the Gaussian mixture model as the natural image prior and combined it with ensemble learning [7] to get clear images. Similarly, Shan *et al.* [8] proposed a two piece-wise continuous function to fit the natural image gradient distribution. The hyper-Laplacian distribution proposed by Krishnan *et al.* [9] usually outperformed the former two, and this sparsity prior was used extensively in the subsequent works. Pan *et al.* [10] used a low rank prior to provide a data-authentic prior for the latent image and applied Gaussian regularization to kernel estimation. This constraint on the blur kernel ensured kernel sparsity, however it made the kernel noisy. When a hysteresis threshold is used, several subtle structures of blur kernels are lost. To solve this problem, Xu and Jia [11] proposed a kernel refinement method via iterative support detection. Pan and Liu [12] introduced a spatial prior to preserve the sparsity and continuity of the blur kernel.

Moreover, salient edge selection is also an important factor which affects reliable kernel estimation a lot. Cho and Lee [13] employed a bilateral filter together with a shock filter to predict sharp edges. Sharp edges with scales smaller than the kernel size, however, could damage kernel estimation. Xu and Jia [11] observed a new connection between image edges and the quality of kernel estimation and proposed a metric to select useful edges for kernel estimation. Pan and Liu [12] developed an adaptive edge selection method to choose reliable edges. Instead of using extra salient edge selection steps,  $l_0$  regularization was used in [14] and [1] to retain salient structures only. Although these methods have made tremendous progress, the results are still far from perfect.

In this paper, we propose a robust motion deblurring method for a single image. In the robust blur kernel estimation step, first, a mask and a smoothness prior are used to remove noise and ringing artifacts in the interim image; then an effective salient edge selection method is adopted to make the kernel estimation more accurate; finally, we use a gradient sparsity prior to eliminate kernel noise. To address the problem of ringing artifacts, we adopt a total variation (TV) regularization and a hyper-Laplacian prior in the final

latent image reconstruction step to get two latent images and average them to obtain the final result. Experimental results demonstrate the effectiveness of the proposed method.

## 2. ROBUST BLUR KERNEL ESTIMATION

In the proposed deblurring method, the blur kernel is estimated with robust sparse priors, and the proposed robust kernel estimation consists of two steps: intermediate image estimation and kernel estimation. We iterate these two steps for several times (5 times used in our experiments) at each level so as to solve this ill-posed problem. The estimation process is performed in a coarse-to-fine fashion with an initial rough kernel at the coarsest level in order to avoid local minima.

### 2.1. Intermediate latent image estimation

In many traditional methods, the restored interim images contain noise in large smooth areas and ringing artifacts near edges, which leads to unreliable salient edge selection. We introduce a mask to encode edge regions and use a smoothness prior to eliminate noise and ringing artifacts in locally smooth regions. What's more, a hyper-Laplacian image prior is adopted to make gradients in near-edge clear image regions obey a heavy-tailed distribution. A smooth interim image is obtained with these constraints which also helps to improve salient edge selection in kernel estimation. Our interim image estimation model is defined as

$$\begin{aligned} \min_L \|L \otimes k - I\|_2^2 + \lambda_1 \|\nabla L\|^{0.5} \circ M + \\ \lambda_2 \|\nabla L\|_2^2 \circ (J - M), \end{aligned} \quad (2)$$

where  $\nabla L = (\partial_x L, \partial_y L)^T$  is the gradient of the image  $L$ ,  $\lambda_1$  and  $\lambda_2$  are the weights and  $\circ$  represents the element-wise multiplication operator,  $J$  is an all-ones matrix and  $M$  is a 2D binary mask. We detect edges in an image with the Canny edge detector and utilize mathematical morphological operations to dilate the edges with a disk model. The radius of the disk model is equal to the kernel size in the current iteration. Then we set an element in  $M$  to 1 if its corresponding pixel belongs to the dilated edge region and set it to 0 otherwise. To solve (2), we fix  $k$  and optimize  $L$  and use the half-quadratic penalty method [15] to solve the highly non-convex function. We introduce an auxiliary variable  $u$  to substitute  $\nabla L$  and change (2) to the following optimization problem:

$$\begin{aligned} \min_{L,u} \|L \otimes k - I\|_2^2 + \lambda_1 \|u\|^{0.5} \circ M + \\ \lambda_2 \|u\|_2^2 \circ (J - M) + \lambda_3 \|\nabla L - u\|_2^2, \end{aligned} \quad (3)$$

where  $\lambda_3$  is a weight and when it is close to  $\infty$ , the solution of (3) converges to that of (2). We solve (3) by updating  $u$  and  $L$  iteratively.

By fixing all the variables except  $u$ , (3) is simplified to

$$\min_u \lambda_1 \|u\|^{0.5} \circ M + \lambda_2 \|u\|_2^2 \circ (J - M) + \lambda_3 \|\nabla L - u\|_2^2. \quad (4)$$

And  $u$  can be obtained with Newton-Paphson method.

We then fix  $u$ , and  $L$  can be optimized by minimizing

$$\min_L \|L \otimes k - I\|_2^2 + \lambda_3 \|\nabla L - u\|_2^2, \quad (5)$$

where fast Fourier transforms are used to compute  $L$ . Based on the Paseval theorem, the solution of  $L$  is expressed as

$$L = \mathcal{F}^{-1} \left( \frac{\mathcal{F}(I) \circ \overline{\mathcal{F}(k)} + \lambda_3 \mathcal{F}(u) \circ \overline{\mathcal{F}(\nabla)}}{\mathcal{F}(k) \circ \overline{\mathcal{F}(k)} + \lambda_3 \mathcal{F}(\nabla) \circ \overline{\mathcal{F}(\nabla)}} \right), \quad (6)$$

where  $\mathcal{F}$  and  $\mathcal{F}^{-1}$  represent the fast Fourier transform and its inverse operation respectively,  $(\cdot)$  denotes the conjugate operator. Iterating the above two steps (4) and (5) for twenty times is enough to get a satisfactory intermediate latent image.

### 2.2. Kernel estimation via salient edges

#### 2.2.1. Salient edge selection

Edge information could damage kernel estimation when the scale of an object is smaller than that of the blur kernel, while effective salient edges can avoid the delta kernel and get an accurate blur kernel. Therefore, we use reliable salient structures instead of all the image edges for kernel estimation. The results of previous salient edge selection methods [13, 11] often contain many small scale edges, which leads to inaccurate kernels. To increase the robustness of kernel estimation, we adopt a strong structure extraction method [8] to select useful edges for kernel estimation. It makes main structures stand out via Relative Total Variation (RTV). Compared with conventional methods, [8] is more effective and can control the magnitude and quantity of salient edges by two adjustable parameters  $\lambda$  and  $\sigma$ .

Suppose  $X$  is the salient edges image generated from  $L$ , tiny structures and noise are removed by setting thresholds for gradient map of  $X$  [13]. The final salient gradient map  $\nabla S = (\partial_x S, \partial_y S)^T$  used for kernel estimation is obtained by

$$\nabla S = \nabla X \circ H(\|\nabla X\|_2 - t), \quad (7)$$

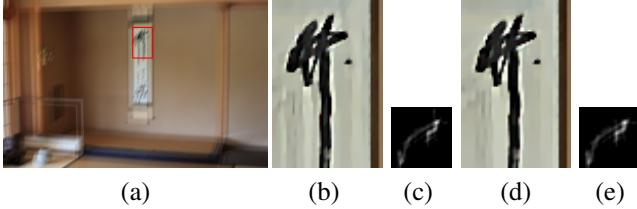
where  $H(\cdot)$  is the Heaviside function and  $t$  is a threshold.

#### 2.2.2. Kernel estimation

Most state-of-the-art works [1, 13] adopted a single Gaussian regularizer to guarantee the sparseness of blur kernels, which results in a lot of noise in the kernel sometimes. However, they ignore continuity by using a simple lagging threshold to remove kernel noise. To address this problem, a gradient sparsity prior is utilized to suppress noise in the kernel matrix and ensure the continuity at the same time.

The salient gradient map  $\nabla S$  is used to estimate the kernel by minimizing the energy function:

$$\min_k \|\nabla S \otimes k - \nabla I\|_2^2 + \gamma_1 \|k\|_2^2 + \gamma_2 \|\nabla k\|_2^2, \quad (8)$$



**Fig. 1.** Effect of prior  $\|\nabla k\|_2^2$ . (a) Blurred image. (b) and (c) are the deblurred result and the corresponding kernel when  $\gamma_2 = 0$ . (d) and (e) are the deblurred result and the corresponding kernel when  $\gamma_2 = 50$ .

where  $\gamma_1$  and  $\gamma_2$  are the adjustable weights controlling the sparseness and smoothness of the blur kernel respectively,  $x$  denotes the pixel. The function (8) can be efficiently solved by fast Fourier transform:

$$k = \mathcal{F}^{-1}\left(\frac{\mathcal{F}(\nabla I) \circ \overline{\mathcal{F}(\nabla S)}}{\mathcal{F}(\nabla S) \circ \overline{\mathcal{F}(\nabla S)} + \gamma_2 \mathcal{F}(\nabla) \circ \overline{\mathcal{F}(\nabla)} + \gamma_1}\right). \quad (9)$$

Then we set negative elements in  $k$  to 0 and normalize  $k$ . The optimized kernel contains little noise and preserves tiny structures. Fig. 1 shows the effect of this gradient sparsity prior. We can see that the kernel in Fig. 1(e) is clearer and smoother than that in Fig. 1(c) and the corresponding deblurred image in Fig. 1(d) is smoother than that in Fig. 1(b).

### 3. FINAL LATENT IMAGE RECONSTRUCTION

After the blur kernel is estimated, there is a lack of texture details in the intermediate latent image, so it is necessary to make use of additional non-blind deconvolution to get the final latent image. We notice that image deconvolution using a hyper-Laplacian prior [9] can obtain a clear image with main structures and few artifacts. Sometimes, however, it is too smooth to preserve some fine details. The latent image estimation method with an isotropic TV norm [16] can preserve abundant tiny textures, but noise and ringing artifacts are retained. Inspired by the artifacts removing algorithm proposed in [1], we combine these two algorithms so as to maximize their merits and minimize the weaknesses.

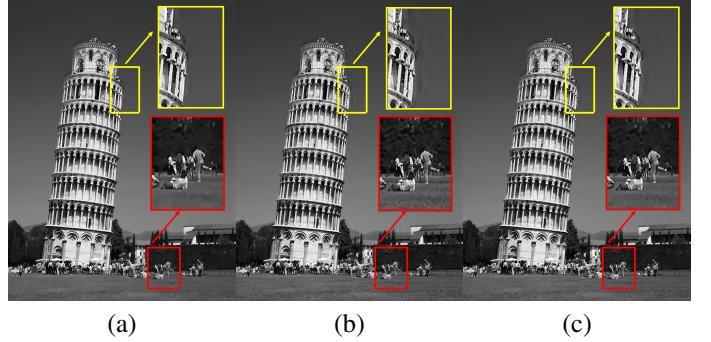
We get latent image  $L_1$  by minimizing the objective function:

$$\min_{L_1} \|L_1 \otimes k - \tilde{I}\|_2^2 + \rho \|\nabla L_1\|^\alpha, \quad (10)$$

where  $0.5 \leq \alpha \leq 0.8$ ,  $\rho$  is a weight. Smaller  $\alpha$  leads to a smoother result. We can adjust  $\alpha$  to get a satisfactory result.

Likewise, the latent image  $L_2$  can be obtained by solving the TV- $l_2$  minimization:

$$\begin{aligned} & \min_{L_2} \|L_2 \otimes k - \tilde{I}\|_2^2 + \mu \|L_2\|_{TV}, \\ & \|L_2\|_{TV} = \sqrt{(\partial_x L_2)^2 + (\partial_y L_2)^2}, \end{aligned} \quad (11)$$



**Fig. 2.** (a) The result of using hyper-Laplacian prior. (b) The result of using TV- $l_2$  prior. (c) The final result.

where  $\mu$  is a regularization parameter and  $\|L_2\|_{TV}$  is the isotropic TV norm,  $\tilde{I}$  is the blurred image enhanced by a shock filter. Since blurred images, especially heavily blurred ones, lose massive details and induce detrimental effects to the image reconstruction, so it is essential to highlight the meaningful edges.

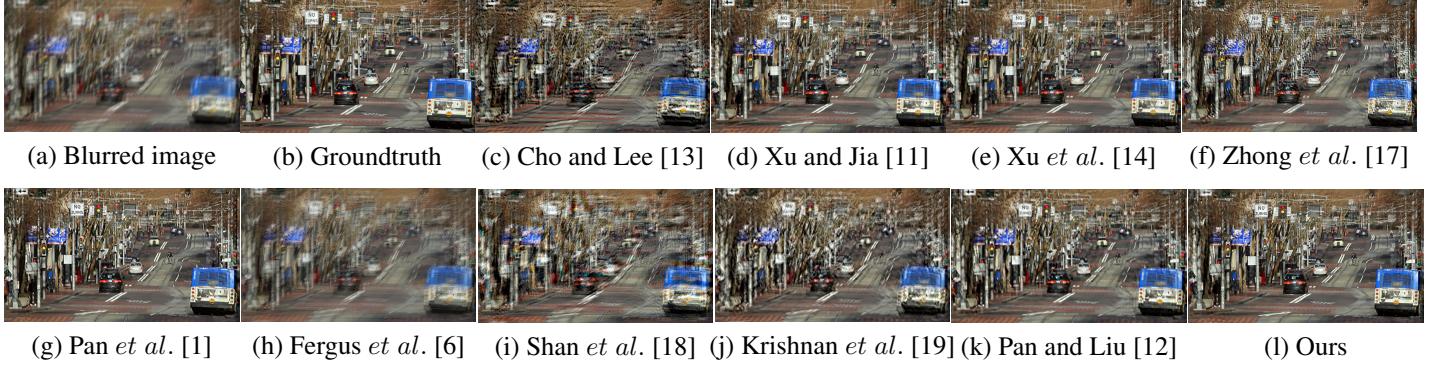
As tiny details will be lost during the process of denoising, rather than removing noise on the difference map between  $L_1$  and  $L_2$ , it would be better to take the average of these two images as the final restored image. An example is shown in Fig. 2. As can be seen there are no ringing artifacts and visible noise in Fig. 2(a), but the grassland is too smooth. Fig. 2(b) contains a few ringing artifacts on the right side of tower. The grassland, however, is rich in details. Fig. 2(c) has plentiful details as well as few ringing artifacts.

### 4. EXPERIMENTAL RESULTS

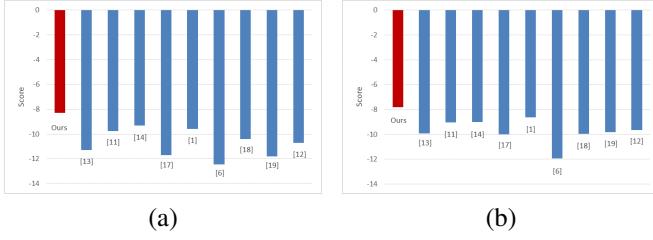
In this section, extensive experiments are performed on both synthetic motion blur database and real blurred images with our proposed method. We compare our method with nine state-of-the-art methods and do our utmost to fine tune the parameters of these methods. In the salient edge selection step, we set  $\sigma = 1.5$  and  $\lambda = 0.01$  as the initial values and gradually increase the values by multiplying them by 1.1 in each iteration. The parameter  $\gamma_1$  in model (8) is set to 2,  $\rho$  in Equation (10) is set to 0.001 and  $\mu$  in Equation (11) is set to 0.001 empirically.

**Objective evaluation on the dataset of [20].** We run our algorithm on a publicly available dataset [20]. It contains 40 sharp images of various scenes, each is blurred synthetically with two different kernels and added Gaussian noise of three different levels ( $\sigma = 0.0, 0.01, 0.02$ ). Particularly, these images contain different amounts of structural edges. Thus, various levels of deblurring difficulties exist for the methods under testing, especially those relying on salient edge selection.

Since most state-of-the-art methods do not handle noisy images particularly, so we just use the images with no ad-



**Fig. 3.** Comparison with state-of-the-art methods on a synthetic example from the dataset of [20]. The quality scores of restored images (c)-(l) are -13.473, -10.708, -10.634, -11.124, -10.266, -13.326, -12.514, -13.353, -11.874 and -9.418 respectively.

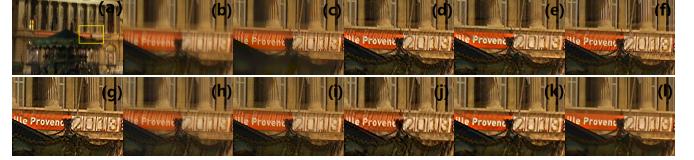


**Fig. 4.** Quantitative comparisons with other 9 state-of-the-art deblurring methods. (a) Average scores of results from images blurred with kernel 1. (b) Average scores of results from images blurred with kernel 2.

ditional noise for testing. For the objective evaluation, we adopt the no-reference quality assessment method introduced in [20] to quantitatively compare our scores against those of Cho and Lee [13], Xu and Jia [11], Xu et al. [14], Zhong et al. [17], Pan et al. [1], Fergus et al. [6], Shan et al. [18], Krishnan et al. [19] and Pan and Liu [12]. This quality evaluation method does not require access to the original images and better matches user's subjective perception than conventional approaches (e.g. PSNR, SSIM, FSIM). A larger score value means higher quality. For example, if the scores of image A and B are -8 and -10 respectively, it means A has higher quality than B. The average quality scores of different methods are illustrated in Fig. 4. We can see that the proposed method performs the best.

Fig. 3 provides one example from the test dataset. As shown, the results with other methods contain various degrees of visual artifacts, while the result with our method has richer details and fewer artifacts.

**Subjective evaluation on real blurred images.** To demonstrate the effectiveness of our algorithm, we test our method on several real blurred photographs. Fig. 5 shows one example of comparing the result of our method with other methods. We can see that Fig. 5(c), (h), (i) and (j) are still



**Fig. 5.** Comparison of state-of-the-art deblurring methods on a real blurred image. (a) Blurred image. (b) Blurred patch extracted from (a) corresponds to the yellow rectangle. (c)~(l) Deblurred image patches of [13], [11], [14], [17], [1], [6], [18], [19], [12] and ours respectively.

quite blurry. Fig. 5(d), (e) and (f) contain serious ringing artifacts. Although Fig. 5(g) and (k) are better than others, the annoying artifacts around the text exert a negative effect on visual quality. We can see the deblurred result (Fig. 5(l)) with our method is the best.

## 5. CONCLUSION

A robust algorithm for single blind image deblurring was proposed in this paper. In the process of kernel estimation, first, we used a mask to distinguish between the edge and smooth regions, then we used a smoothness prior to suppress ringing artifacts and noise in smooth areas; second, we adopted an adaptive salient edge selection method to provide reliable edge information for kernel estimation; third, a gradient sparsity prior in kernel estimation model was used to remove kernel noise and keep kernel sparse. We combined the hyper-Laplacian model with the  $TV-l_2$  model in the final latent image reconstruction step so as to preserve main and tiny edges and remove noise. Extensive experiments showed that the proposed algorithm performed better than state-of-the-art methods in both objective and subjective evaluations.

## 6. REFERENCES

- [1] J.Pan, H.Zhe, and Z.Su, “Deblurring text images via  $l_0$ -regularized intensity and gradient prior,” in *CVPR*, 2014, pp. 2901–2908.
- [2] H.Cho, J.Wang, and S.Lee, “Text image deblurring using text-specific properties,” in *ECCV*, 2012, pp. 524–537.
- [3] S.Cho, J.Wang, and S.Lee, “Handling outliers in non-blind image deconvolution,” in *CVPR*, 2011, pp. 495–502.
- [4] Z.Hu, S.Cho, and J.Wang, “Deblurring low-light images with light streaks,” in *CVPR*, 2014, pp. 3382–3389.
- [5] J.Pan, Z.Hu, and Z.Su, “Deblurring face images with exemplars,” in *ECCV*, 2014, pp. 47–62.
- [6] R.Fergus, B.Singh, and A.Hertzmann, “Removing camera shake from a single photograph,” *ACM Trans. Graph*, vol. 25, pp. 787–794, 2006.
- [7] J.Miskin, D.J.C, and Mackay, “Ensemble learning for blind image separation and deconvolution,” 2000.
- [8] L.Xu, Q.Yan, and Y.Xia, “Structure extraction from texture via relative total variation,” *ACM Trans. Graph*, vol. 31(6), pp. 139, 2010.
- [9] D.Krishnan and R.Fergus, “Fast image deconvolution using hyper-laplacian priors,” in *NIPS*, 2009, pp. 157–170.
- [10] J.Pan, R.Liu, and Z.Su, “Motion blur kernel estimation via salient edges and low rank prior,” in *ICME*, 2014, pp. 1–6.
- [11] L.Xu and J.Jia, “Two-phase kernel estimation for robust motion deblurring,” in *ECCV*, 2010, pp. 157–170.
- [12] J.Pan, R.Liu, and Z.Su, “Kernel estimation from salient structure for robust motion deblurring,” *Signal Processing: Image Communication*, vol. 28(9), pp. 1156–1170, 2013.
- [13] S.Cho and S.Lee, “Fast motion deblurring,” *ACM Trans. Graph*, vol. 28(5), pp. 145, 2009.
- [14] L.Xu, S.Zheng, and J.Jia, “Unnatural  $l_0$  sparse representation for natural image deblurring,” in *CVPR*, 2013, pp. 1107–1114.
- [15] D.Geman and C.Yang, “Nonlinear image recovery with half-quadratic regularization,” *PAMI*, vol. 4, pp. 932–946, 1995.
- [16] S.H.Chan, R.Khoshabeh, and K.B.Gibson, “An augmented lagrangian method for total variation video restoration,” *IEEE Transactions on Image Processing*, vol. 20(11), pp. 3097–3111, 2011.
- [17] L.Zhongu, S.Cho, and D.Metaxas, “Handling noise in single image deblurring using directional filters,” in *CVPR*, 2013, pp. 612–619.
- [18] Q.Shan, J.Jia, and A.Agarwala, “High-quality motion deblurring from a single image,” *ACM Trans. Graph*, vol. 27(3), pp. 73, 2008.
- [19] D.Krishnan, T.Tay, and R.Fergus, “Blind deconvolution using a normalized sparsity measure,” in *CVPR*, 2011, pp. 233–240.
- [20] Y.Liu, J.Wang, and S.Cho, “A no-reference metric for evaluating the quality of motion deblurring,” *ACM Trans. Graph*, vol. 32(175), pp. 1–12, 2013.