Supplementary Material of Blind Deblurring for Saturated Images

Liang Chen *[1], Jiawei Zhang †[2], Songnan Lin[2], Faming Fang[1], Jimmy S. Ren[2,3]
[1] Shanghai Key Laboratory of Multidimensional Information Processing,
School of Computer Science and Technology, East China Normal University
[2] SenseTime Research
[3] Qing Yuan Research Institute, Shanghai Jiao Tong University, Shanghai, China

In this supplementary material, we provide,

1. Detailed deviation steps including solving problems referring to the latent image I (*i.e*. Eq. (7) in the manuscript) and the blur kernel K (Eq. (12) in the manuscript), and the reason why Richardson-Lucy (RL) is inappropriate to update the blur kernel;
2. Detailed experimental settings in our ablation studies;
3. More experimental results.

1. Detailed Deviation to Update Latent Image I and Blur Kernel K

1.1. Solving the problem referring to I

According to Eq. (4) of our manuscript, the fidelity term can be presented as:

$$\begin{aligned} &\mathcal{L}(B, M \circ (I \otimes K)) \\ &= -\log \prod_i \mathcal{P}oisson(B_i; M_i(I \otimes K)_i) \\ &= \sum_i (M_i(I \otimes K)_i - \log(M_i(I \otimes K)_i)B_i) + C \end{aligned} \tag{1}$$

where M, B, and $P(I)$ are the latent map, blurry image, and the prior term enforced on the latent image; C is an irrelevant variant w.r.t. I and can be ignored during the optimization.

Then, we can get Eq. (5) of our manuscript, the problem referring to I is given as following:

$$\min_I \bar{\mathbf{1}}^T (M \circ (I \otimes K) - \log(M \circ (I \otimes K)) \circ B + \lambda P(I))\bar{\mathbf{1}}, \tag{2}$$

where $\circ$ is Hadamard product and λ is the weight parameter. $\bar{\mathbf{1}}$ is the all-one vector and $\bar{\mathbf{1}}^T(\cdot)\bar{\mathbf{1}}$ is actually the summation of every element of the matrix $(\cdot)$.

We can solve Eq. (2) by setting its derivative to zero. Before getting its derivative, we rewrite Eq. (2) into vectorized form:

$$\min_{\mathbf{I}} \mathbf{M}^T\mathbf{K}\mathbf{I} - \mathbf{B}^T \log(diag(\mathbf{M})\mathbf{K}\mathbf{I}) + \lambda \bar{\mathbf{1}}^T P(\mathbf{I})\bar{\mathbf{1}}, \tag{3}$$

where $\mathbf{M}$, $\mathbf{B}$, and $\mathbf{I}$ are M, B and I in their vectorized form, $\mathbf{K}$ is the Toeplitz for of K w.r.t. I. For the second term of Eq. (3), we denote it as A and its derivative w.r.t. $\mathbf{I}$ is:

$$\begin{aligned} \frac{\partial A}{\partial \mathbf{I}} &= \frac{\partial diag(\mathbf{M})\mathbf{K}\mathbf{I}}{\partial \mathbf{I}} \frac{\partial log(diag(\mathbf{M})\mathbf{K}\mathbf{I})}{\partial diag(\mathbf{M})\mathbf{K}\mathbf{I}} \frac{\partial \mathbf{B}^T \log(diag(\mathbf{M})\mathbf{K}\mathbf{I})}{\partial log(diag(\mathbf{M})\mathbf{K}\mathbf{I})}, \\ &= (\mathbf{K}^T diag(\mathbf{M})) diag(\frac{1}{diag(\mathbf{M})\mathbf{K}\mathbf{I}})\mathbf{B}, \\ &= \mathbf{K}^T diag(\frac{1}{diag(\mathbf{M})\mathbf{K}\mathbf{I}})(diag(\mathbf{M})\mathbf{B}), \end{aligned} \tag{4}$$

*This work was done when Liang Chen was an intern at SenseTime.
†Corresponding author

where the divide operation is element-wise. Then we can solve Eq. (3) by setting its derivative to zero as:

$$\mathbf{K}^T\mathbf{M} - \mathbf{K}^T diag(\frac{1}{diag(\mathbf{M})\mathbf{KI}})(diag(\mathbf{M})\mathbf{B}) + \lambda P'(\mathbf{I}) = 0. \tag{5}$$

Reformulate the above formation into its matrix form, we have:

$$M \otimes \widetilde{K} - \frac{M \circ B}{M \circ (I \otimes K)} \otimes \widetilde{K} + \lambda P_I'(I) = 0. \tag{6}$$

where $\widetilde{K}$ is the transpose of K that flips the shape of K upside down and left-to-right, $P_I'(I)$ is the first order derivative of $P_I(I)$ w.r.t. I. Recall that the sum of the kernel equals to 1, *i.e.* $\bar{\mathbf{1}}^T\widetilde{\mathbf{K}} = 1$, where $\widetilde{\mathbf{K}}$ is the vectorized form of $\widetilde{K}$, and $\bar{\mathbf{1}}$ is the all-one vector. Thus, we further have,

$$M \otimes \widetilde{K} - \frac{M \circ B}{M \circ (I \otimes K)} \otimes \widetilde{K} + \lambda P_I'(I) + \mathbf{1} - \mathbf{1} \otimes \widetilde{K} = 0, \tag{7}$$

$$(\frac{B}{I \otimes K} - M + \mathbf{1}) \otimes \widetilde{K} = \mathbf{1} + \lambda P_I'(I), \tag{8}$$

$$I \circ (\frac{B}{I \otimes K} - M + \mathbf{1}) \otimes \widetilde{K} = I \circ (\mathbf{1} + \lambda P_I'(I)), \tag{9}$$

where $\mathbf{1}$ is the all-one matrix.

In order to solve Eq. (9), we use the fixed point iteration scheme and rewrite it as:

$$\frac{I^{t+1}}{I^t} = \mathbf{1} = \frac{(\frac{B}{I^t \otimes K} - M + \mathbf{1}) \otimes \widetilde{K}}{\mathbf{1} + \lambda P_I'(I^t)}. \tag{10}$$

Thus, we can finally get Eq. (7) in our manuscript:

$$I^{t+1} = \frac{I^t \circ ((\frac{B}{I^t \otimes K} - M + \mathbf{1}) \otimes \widetilde{K})}{\mathbf{1} + \lambda P_I'(I^t)}. \tag{11}$$

1.2. Solving the problem referring to K

Recall that using the RL method to update I is based on the fact that $\bar{\mathbf{1}}^T\widetilde{\mathbf{K}} = 1$. However, this requirement is not satisfied when updating K *i.e.* $\bar{\mathbf{1}}^T\mathbf{I} \neq 1$. And we cannot use RL scheme to update the blur kernel.

To estimate the blur kernel, we use Gaussian distribution to replace the Poisson one and the fidelity term in Eq. (10) of our manuscript is:

$$\begin{aligned}
&\mathcal{L}(B, M \circ (I \otimes K)) \\
&= -\log \prod_i \mathcal{N}(B_i; M_i(I \otimes K)_i, M_i(I \otimes K)_i) \\
&= -\log \prod_i \mathcal{N}(B_i - M_i(I \otimes K)_i; 0, W_{i,i}^{-1}) \\
&= \sum_i (B_i - M_i(I \otimes K)_i)^2 W_{i,i} + C \\
&= \|B - M \circ (I \otimes K)\|_W^2 + C,
\end{aligned} \tag{12}$$

where $\| \bullet \|_W^2$ is the norm under metric W, and for $\forall\, x$, $\|x\|_W^2$ is computed as $x^T W x$; C is an irrelevant variant w.r.t. K and can be ignored during optimization. $W^{-1} = diag(M \circ (I \otimes K))$ which is fixed when updating K.

As the blur kernel estimation based on image gradients is more stable and accurate, we can then estimate it by:

$$\min_K \|\nabla B - M \circ (\nabla I \otimes K)\|_W^2 + \beta\|K\|^2, \tag{13}$$

Algorithm 1 Solving Eq. (15) using the conjugate gradient (CG) method

Input: W, ∇B, ∇I, M, β, and the maximum iteration step s_{max}
Output: $K_{s_{max}}$
$\mathbf{b} = W(\mathbf{M}\nabla\mathbf{I})^T\nabla\mathbf{B}$
$\mathbf{A} = W(\mathbf{M}\nabla\mathbf{I})^T\mathbf{M}\nabla\mathbf{I} + \beta$
$\mathbf{P}_0 = \mathbf{b} - \mathbf{A}\mathbf{K}_0$
$\mathbf{r}_0 = \mathbf{P}_0$
for $s = 0$ to s_{max} **do**
  $\alpha_s = (\mathbf{r}_s^T\mathbf{r}_s)/(\mathbf{P}_s^T\mathbf{A}\mathbf{P}_s)$
  $\mathbf{K}_{s+1} = \mathbf{K}_s + \alpha_s\mathbf{P}_s$
  $\mathbf{r}_{s+1} = \mathbf{r}_s - \alpha_s\mathbf{A}\mathbf{P}_s$
  $\beta_s = (\mathbf{r}_{s+1}^T\mathbf{r}_{s+1})/(\mathbf{r}_s^T\mathbf{r}_s)$
  $\mathbf{P}_{s+1} = \mathbf{r}_{s+1} + \beta_s\mathbf{P}_s$
end for

where ∇ is the gradient operator in horizontal and vertical dimensions (*i.e.* $\nabla = \{\nabla_h, \nabla_v\}$). In order to solve Eq. (13), we rewrite it to the vectorized form:

$$\min_{\mathbf{K}}(\nabla\mathbf{B} - \mathbf{M}\nabla\mathbf{I}\mathbf{K})^T W(\nabla\mathbf{B} - \mathbf{M}\nabla\mathbf{I}\mathbf{K}) + \beta\mathbf{K}^T\mathbf{K}, \tag{14}$$

where **B** and **K** are the vectorized form of B and K, **M** and **I** are the Toeplitze form of M and I w.r.t. K. Taking the derivative of Eq. (14) w.r.t. **K** and setting it to zero, we have,

$$W(\mathbf{M}\nabla\mathbf{I})^T\nabla\mathbf{B} = (W(\mathbf{M}\nabla\mathbf{I})^T\mathbf{M}\nabla\mathbf{I} + \beta)\mathbf{K}. \tag{15}$$

We can use a conjugate gradient method to solve Eq. (15). The overall algorithm is shown in Algorithm 1.

2. Detailed Experimental Settings

In Section 5.2 of our manuscript, we conduct ablation studies to compare our method with existing ones. In this section, we give detailed experimental settings of these ablation studies.

2.1. Relation with Chen *et al.*[3]

In [3], the deblurring process is conducted by minimizing the following equation,

$$\min_{I,K}\sum_i \overline{M}_i(B_i - (I\otimes K)_i)^2 + \lambda P(I) + \beta\|K\|^2. \tag{16}$$

Note the prior terms for the latent image (*i.e.* $P(I)$) and blur kernel (*i.e.* $\|K\|^2$) are the same in our work and that from [3]. In [3], the latent image I, blur kernel K, and latent map $\overline{M}$ are computed as:

$$\begin{cases} I = arg\min_I \sum_i \overline{M}_i(B_i - (I\otimes K)_i)^2 + \lambda P(I), & (17)\\ K = arg\min_K \sum_i \overline{M}_i(\nabla B - \nabla I\otimes K)_i^2 + \beta\|K\|^2, & (18)\\ \overline{M}_i = \left(\exp\left(\dfrac{(B_i - (I\otimes K)_i)^2 - \alpha}{\beta}\right) + 1\right)^{-1}. & (19)\end{cases}$$

Please refer to [3] for more details.

In comparison, the map M in our setting is given by :

$$M_i = \begin{cases} 1, & \text{if } (I\otimes K)_i \leq 1 \\ \frac{1}{(I\otimes K)_i}, & \text{Otherwise}\end{cases} \tag{20}$$

Algorithm 2 Deblurring process in [3] with our map setting

Input: blurred image B, parameters λ, β and initial kernel $K^{0,0}$.
Output: blur kernel K and intermediate latent image I.

```
Initialize M^0=1, I^{0,0} = B.
t=1, x=1, j=0.
while j < j_max do
    while t < t_max do
        Compute I^{t,j} using Eq. (17) given M^{t-1} and K;
        Update M^t using Eq. (20) given I^{t,j} and K;
        t ← t + 1
    end while
    while Stopping criterion is not satisfied do
        Compute K^{x,j} using Eq. (18) given M^x;
        Update M^x using Eq. (20) given K^{x-1,j} and I;
        x ← x + 1;
    end while
    j ← j + 1
end while
```

Algorithm 3 Our deblurring process with map setting from [3]

Input: blurred image B, parameters λ, β and initial kernel $K^{0,0}$.
Output: blur kernel K and intermediate latent image I.

```
Initialize M^0=1, I^{0,0} = B.
t=1, x=1, j=0.
while j < j_max do
    while t < t_max do
        Compute I^{t,j} using Eq. (11) given M^{t-1} and K;
        Update M^t using Eq. (19) given I^{t,j} and K;
        t ← t + 1
    end while
    while Stopping criterion is not satisfied do
        Update W^x using Eq. (13) given K^{x-1,j} and M^x;
        Compute K^{x,j} using Eq. (13) given W^x and M^x;
        Update M^x using Eq. (19) given K^{x-1,j} and I;
        x ← x + 1;
    end while
    j ← j + 1
end while
```

To evaluate the effectiveness of these two map settings, we first replace the map setting in [3] (*i.e.* $\overline{M}$ in Eq. (16)) with our map setting (*i.e.* M in Eq. (20)) and update I and K according to Eq. 16 and Eq. 17. We denote this experiment as 'Chen et al. [3] with our M' in Figure 8 (b) of our manuscript. The overall algorithm is given in Algorithm 2. Then, we replace our map setting with that from [3] and update update I and K according to Eq. 2 and Eq. 13. We denote this experiment as 'Ours with saturate map of [3]' in Figure 8 (b) of our manuscript. The overall algorithm is given in Algorithm 3. Deblurring results from the above-mentioned strategies in dataset [9] are shown in Figure 8 (b) of our manuscript.

2.2. Relation with Whyte *et al*. [13]

In the non-blind deblurring method [13], the imaging process is modeled as:

$$B = \mathcal{P}oisson(R(I \otimes K)), \tag{21}$$

Algorithm 4 Extend the non-blind deblurring method [13] into our optimization framework

Input: blurred image B, parameters λ, β and initial kernel $K^{0,0}$.
Output: blur kernel K and intermediate latent image I.

```
I^{0,0} = B.
t=1, x=1, j=0.
while j < j_max do
  while t < t_max do
    Compute I^{t,j} using Eq. (23) given K;
    t ← t + 1
  end while
  while Stopping criterion is not satisfied do
    Compute K^{x,j} using Eq. (13) given W^x = diag(R(I^j ⊗ K^{x,j}))^{-1};
    x ← x + 1;
  end while
  j ← j + 1
end while
```

where R is a smooth function from [1], $R(x) = x - \frac{1}{a}\log(1 + \exp(a(x-1)))$, and a is fixed as 50 in their implementation. Their fidelity term can be presented as:

$$L(B, I \otimes K) = R(I \otimes K) - \log(R(I \otimes K)) \circ B. \tag{22}$$

To compare their function-based model with the proposed latent map-based model, we replace our fidelity term Eq. (1) with theirs Eq. (22) in our updating scheme. Specifically, the latent image is updated via the following formation:

$$I^{t+1} = \frac{I^t \circ ((\frac{B \circ R'(I^t \otimes K)}{R(I^t \otimes K)} - M + \mathbf{1}) \otimes \widetilde{K})}{\mathbf{1} + \lambda P_I'(I^t)}, \tag{23}$$

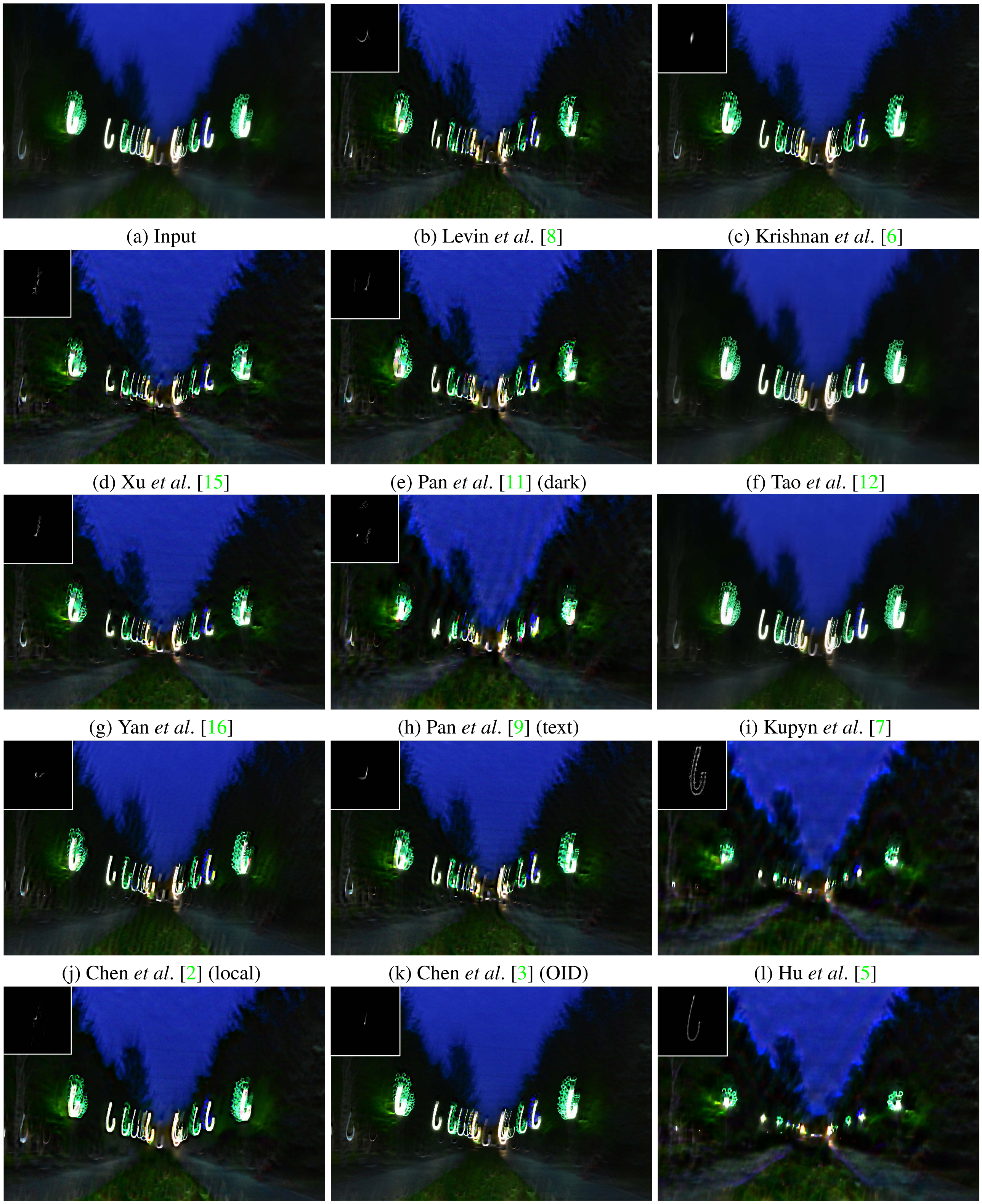
where $R'(\bullet)$ is the derivative of $R(\bullet)$ w.r.t. I, and $R'(x) = 1/(1+\exp(a(x-1)))$. The blur kernel is updated by minimizing the following equation:

$$\min_K \|B - R(I \otimes K)\|_W^2, \ \text{s.t.} \ W^{-1} = diag(R(I \otimes K)) \tag{24}$$

This updating process is similar to that of Eq. (13), and we omit it here. The overall deblurring process is illustrated in Algorithm 4. The result, which is denoted as 'Extension of Whyte et al.', is shown in Figure 8 (a) of our manuscript.

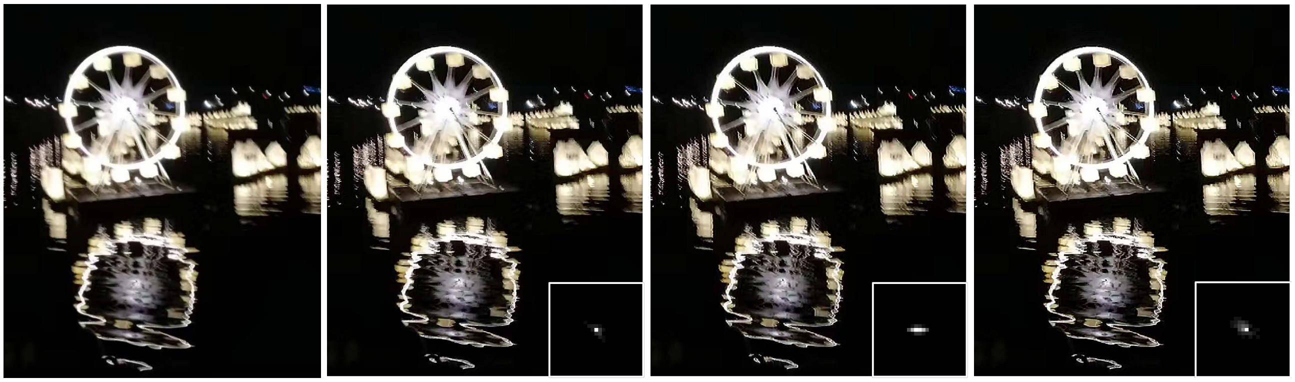
3. More Experimental Results

In this section, we show more deblurring results of images with large saturated regions. The results from state-of-the-art methods are directly provided by the authors or generated with given codes after tuning the hyper-parameters for better results.

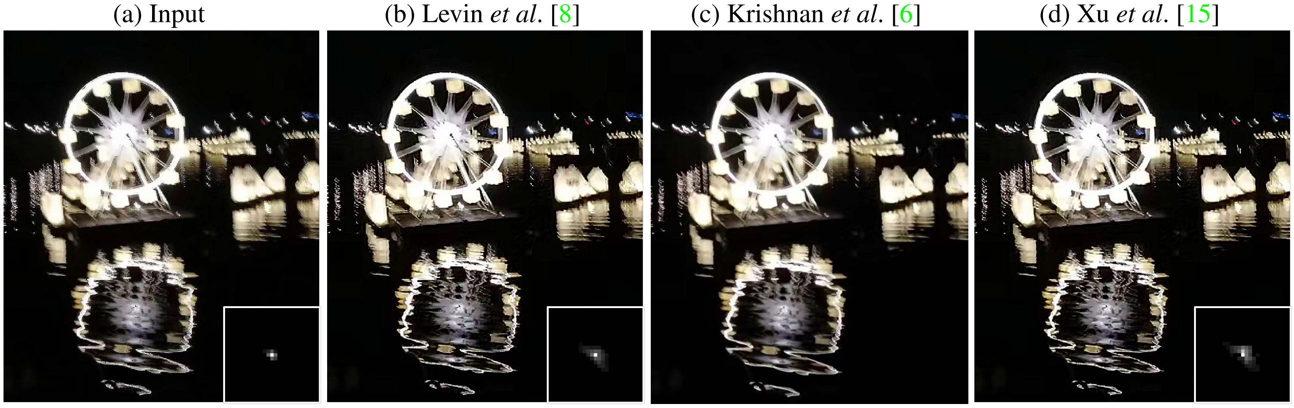


(a) Input (b) Levin *et al*. [8] (c) Krishnan *et al*. [6]
(d) Xu *et al*. [15] (e) Pan *et al*. [11] (dark) (f) Tao *et al*. [12]
(g) Yan *et al*. [16] (h) Pan *et al*. [9] (text) (i) Kupyn *et al*. [7]
(j) Chen *et al*. [2] (local) (k) Chen *et al*. [3] (OID) (l) Hu *et al*. [5]
(m) Pan *et al*. [10] (robust) (n) Dong *et al*. [4] (o) Ours

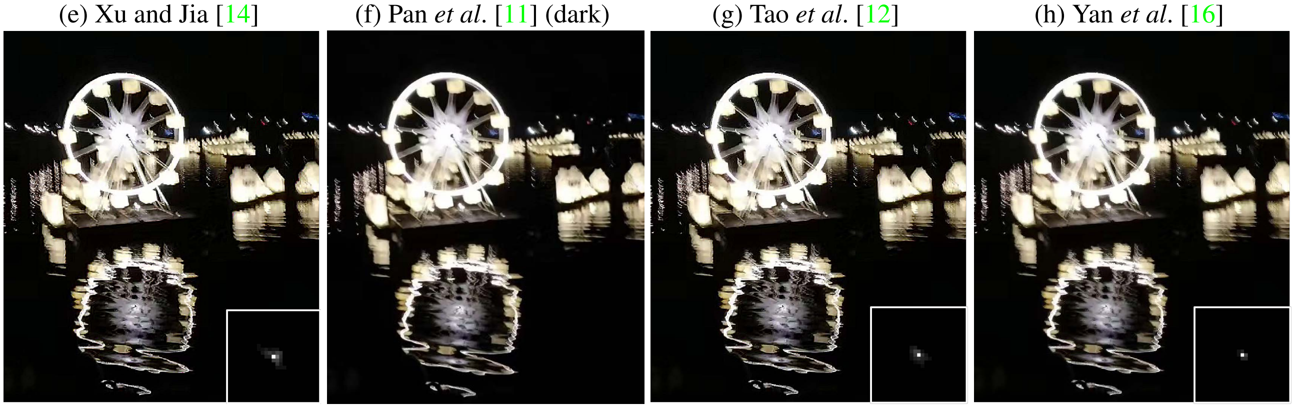
Figure 1. Deblurring results of a real-world blurry example. Our method generates results with fewer artifacts.



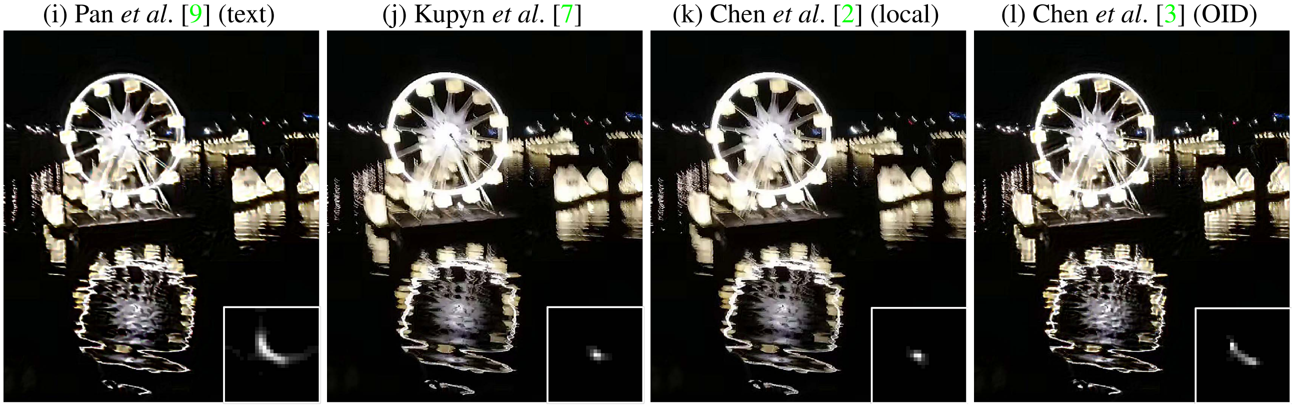
(a) Input (b) Levin *et al*. [8] (c) Krishnan *et al*. [6] (d) Xu *et al*. [15]



(e) Xu and Jia [14] (f) Pan *et al*. [11] (dark) (g) Tao *et al*. [12] (h) Yan *et al*. [16]

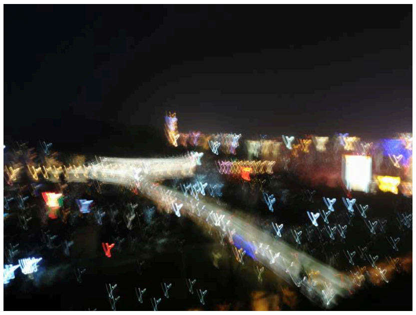


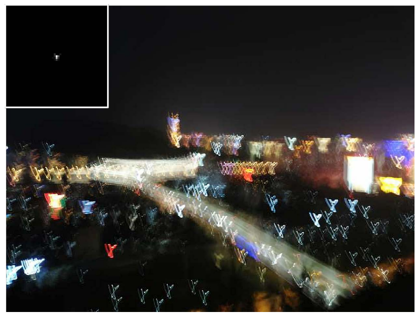
(i) Pan *et al*. [9] (text) (j) Kupyn *et al*. [7] (k) Chen *et al*. [2] (local) (l) Chen *et al*. [3] (OID)

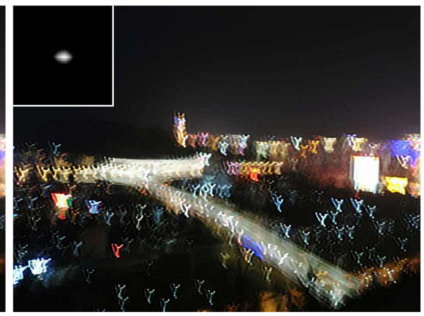


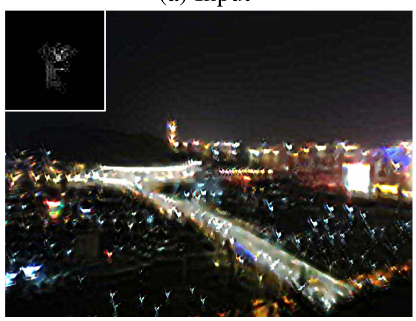
(m) Hu *et al*. [5] (n) Pan *et al*. [10] (robust) (o) Dong *et al*. [4] (p) Ours

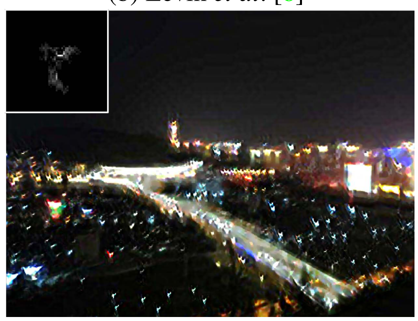
Figure 2. Deblurring results of a real-world blurry example. Our method generates results with fewer artifacts.

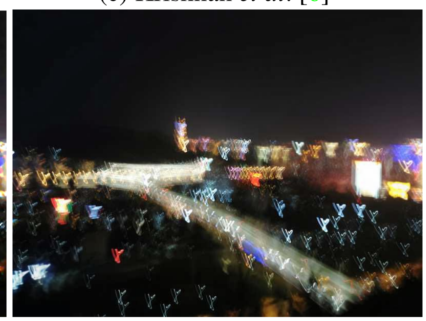

(a) Input

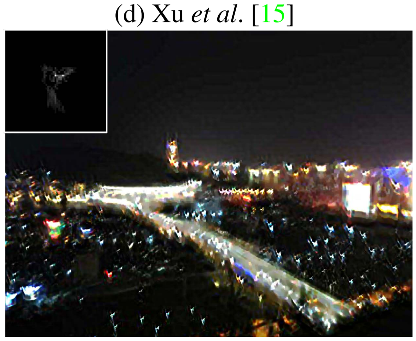

(b) Levin *et al*. [8]

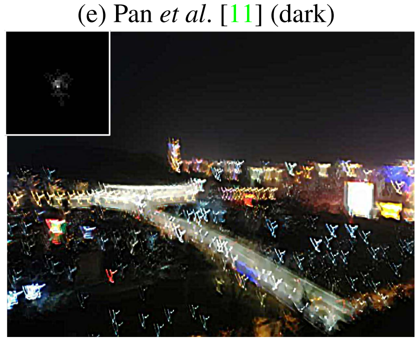

(c) Krishnan *et al*. [6]

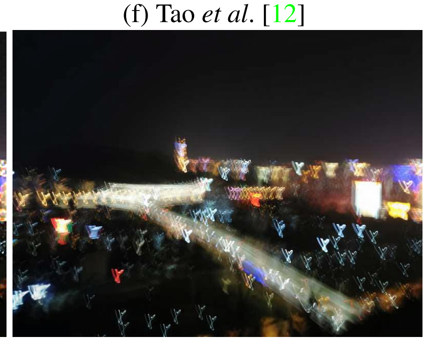

(d) Xu *et al*. [15]


(e) Pan *et al*. [11] (dark)

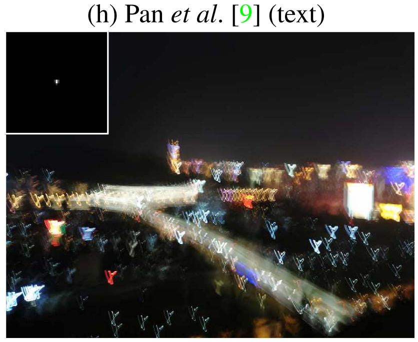

(f) Tao *et al*. [12]

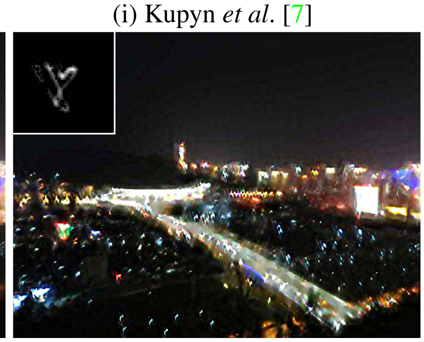

(g) Yan *et al*. [16]

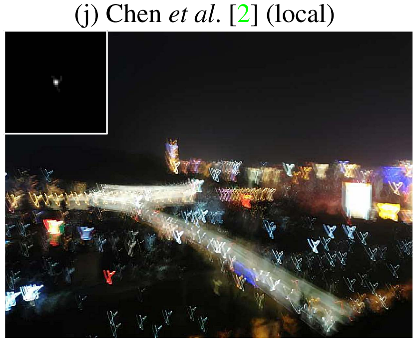

(h) Pan *et al*. [9] (text)

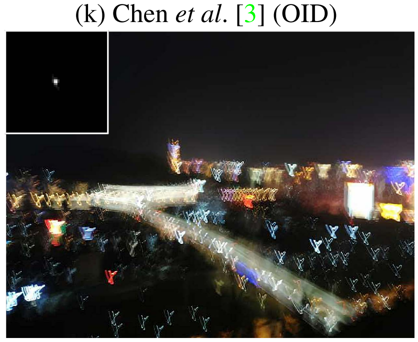

(i) Kupyn *et al*. [7]


(j) Chen *et al*. [2] (local)


(k) Chen *et al*. [3] (OID)


(l) Hu *et al*. [5]


(m) Pan *et al*. [10] (robust)


(n) Dong *et al*. [4]

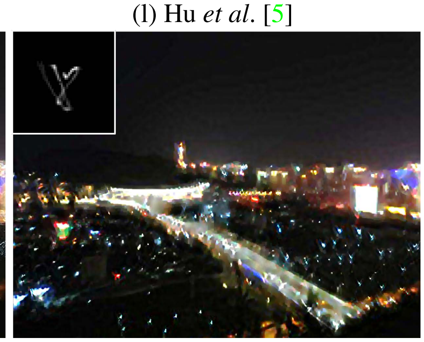
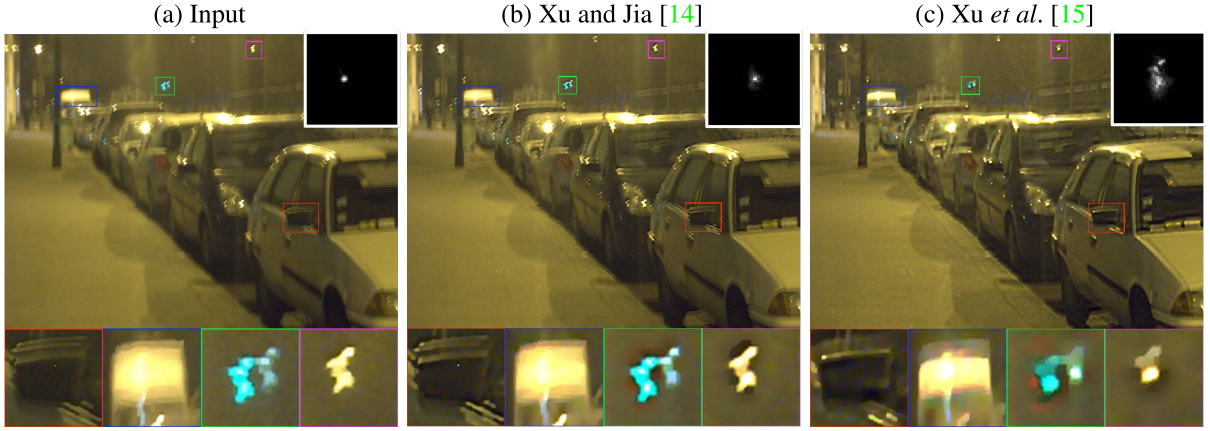

(o) Ours

Figure 3. Deblurring results of a real-world blurry example. Our method generates results with fewer artifacts.

(a) Input (b) Xu and Jia [14] (c) Xu *et al*. [15]



(d) Zhong *et al*. [17] (e) Yan *et al*. [16] (f) Pan *et al*. [9] (text)

(g) Tao *et al*. [12] (h) Kupyn *et al*. [7] (i) Chen *et al*. [3] (OID)

(j) Hu *et al*. [5] (k) Pan *et al*. [10] (robust) (l) Ours

Figure 4. Deblurring results of a real-world blurry example. Our method generates results with fewer artifacts.

References

[1] Chunhui Chen and Olvi L Mangasarian. A class of smoothing functions for nonlinear and mixed complementarity problems. *Computational Optimization and Applications*, 1996. 5

[2] Liang Chen, Faming Fang, Tingting Wang, and Guixu Zhang. Blind image deblurring with local maximum gradient prior. In *CVPR*, 2019. 6, 7, 8

[3] Liang Chen, Faming Fang, Jiawei Zhang, Jun Liu, and Guixu Zhang. Oid: Outlier identifying and discarding in blind image deblurring. In *ECCV*, 2020. 3, 4, 6, 7, 8, 9

[4] Jiangxin Dong, Jinshan Pan, Zhixun Su, and Ming-Hsuan Yang. Blind image deblurring with outlier handling. In *ICCV*, 2017. 6, 7, 8

[5] Zhe Hu, Sunghyun Cho, Jue Wang, and Ming-Hsuan Yang. Deblurring low-light images with light streaks. In *CVPR*, 2014. 6, 7, 8, 9

[6] Dilip Krishnan, Terence Tay, and Rob Fergus. Blind deconvolution using a normalized sparsity measure. In *CVPR*, 2011. 6, 7, 8

[7] Orest Kupyn, Tetiana Martyniuk, Junru Wu, and Zhangyang Wang. Deblurgan-v2: Deblurring (orders-of-magnitude) faster and better. In *ICCV*, 2019. 6, 7, 8, 9

[8] Anat Levin, Yair Weiss, Fredo Durand, and William T Freeman. Efficient marginal likelihood optimization in blind deconvolution. In *CVPR*, 2011. 6, 7, 8

[9] Jinshan Pan, Zhe Hu, Zhixun Su, and Ming-Hsuan Yang. l_0-regularized intensity and gradient prior for deblurring text images and beyond. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(2):342–355, 2017. 4, 6, 7, 8, 9

[10] Jinshan Pan, Zhouchen Lin, Zhixun Su, and Ming-Hsuan Yang. Robust kernel estimation with outliers handling for image deblurring. In *CVPR*, 2016. 6, 7, 8, 9

[11] Jinshan Pan, Deqing Sun, Hanspeter Pfister, and Ming-Hsuan Yang. Blind image deblurring using dark channel prior. In *CVPR*, 2016. 6, 7, 8

[12] Xin Tao, Hongyun Gao, Xiaoyong Shen, Jue Wang, and Jiaya Jia. Scale-recurrent network for deep image deblurring. In *CVPR*, 2018. 6, 7, 8, 9

[13] Oliver Whyte, Josef Sivic, and Andrew Zisserman. Deblurring shaken and partially saturated images. *IJCV*, 2014. 4, 5

[14] Li Xu and Jiaya Jia. Two-phase kernel estimation for robust motion deblurring. In *ECCV*, 2010. 7, 9

[15] Li Xu, Shicheng Zheng, and Jiaya Jia. Unnatural l_0 sparse representation for natural image deblurring. In *CVPR*, 2013. 6, 7, 8, 9

[16] Yanyang Yan, Wenqi Ren, Yuanfang Guo, Rui Wang, and Xiaochun Cao. Image deblurring via extreme channels prior. In *CVPR*, 2017. 6, 7, 8, 9

[17] Lin Zhong, Sunghyun Cho, Dimitris Metaxas, Sylvain Paris, and Jue Wang. Handling noise in single image deblurring using directional filters. In *CVPR*, 2013. 9