Learning Discriminative Data Fitting Functions for Blind Image Deblurring Supplemental Material

Jinshan Pan[1] Jiangxin Dong[2] Yu-Wing Tai[3] Zhixun Su[2] Ming-Hsuan Yang[4]
[1]Nanjing University of Science and Technology [2]Dalian University of Technology
[3]Tencent Youtu Lab [4]UC Merced

Overview

In this supplemental material, we give the derivation details of important equations of the main paper in Sections 1 and 2, Section 3 presents the algorithm details of the discriminative non-blind deconvolution. We demonstrate the flexibility of the proposed method in Section 4 and provide more experimental results of the proposed algorithm against the state-of-the-art deblurring methods in Section 5.

1. Derivations of (13) in the Manuscript

Let $\mathcal{L}_j = \frac{1}{2}\|\mathbf{k}_j(\omega) - \mathbf{k}_j^{gt}\|_2^2$. Based on the chain rule, the gradient of $\mathcal{L}_j$ with respect to ω_i is

$$\begin{aligned}\frac{\partial \mathcal{L}_j}{\partial \omega_i} =& \frac{\partial \mathcal{L}_j}{\partial \mathbf{k}_j}\frac{\partial \mathbf{k}_j}{\partial \omega_i} \\ =& (\mathbf{k}_j - \mathbf{k}_j^{gt})^\top \frac{\partial \mathbf{k}_j}{\partial \omega_i}.\end{aligned} \tag{1}$$

In order to compute $\frac{\partial \mathbf{k}_j}{\partial \omega_i}$, we first introduce an important proposition.

Proposition 1: *Let* $\mathbf{A}$ *denote a matrix. The gradient with respect to its inverse form is:*

$$\partial \mathbf{A}^{-1} = -\mathbf{A}^{-1}(\partial \mathbf{A})\mathbf{A}^{-1}, \tag{2}$$

where ∂ denotes the gradient operator. The detailed proof can be found in [5].

Based on **Proposition 1** and (8) in the manuscript, we have

$$\frac{\partial \mathbf{k}_j}{\partial \omega_i} = -\left(\sum_i \omega_i \mathbf{A}_{ij}^\top \mathbf{A}_{ij} + \gamma\right)^{-1} \mathbf{A}_{ij}^\top \mathbf{A}_{ij}\mathbf{k}_j + \left(\sum_i \omega_i \mathbf{A}_{ij}^\top \mathbf{A}_{ij} + \gamma\right)^{-1} \left(\mathbf{A}_{ij}^\top \mathbf{b}_{ij}\right) \tag{3}$$

2. Derivations of (16) in the Manuscript

As stated in Section 2.5 of the manuscript, the weights for the non-blind deconvolution is obtained by solving (15) in the manuscript. Let $\mathcal{L}_j^I$ denote the loss function:

$$\mathcal{L}_j^I = \frac{1}{2}\sum_j \|\mathbf{I}_j - \mathbf{I}_j^{gt}\|^2, \tag{4}$$

where $\mathbf{I}_j$ is obtained according to (12) in the manuscript, which is defined as

$$I_j = \mathcal{F}^{-1}\left(\frac{\sum_i \omega_i \overline{\mathcal{F}(k_j)}\mathcal{F}(f_i * B_j) + \beta\overline{\mathcal{F}(\nabla_h)}\mathcal{F}(g_j^h) + \overline{\mathcal{F}(\nabla_v)}\mathcal{F}(g_j^v)}{\sum_i \omega_i \overline{\mathcal{F}(f_i)}\overline{\mathcal{F}(k_j)}\mathcal{F}(k_j)\mathcal{F}(f_i) + \beta(\sum_{i\in\{h,v\}}\overline{\mathcal{F}(\nabla_i)}\mathcal{F}(\nabla_i))}\right). \tag{5}$$

As we use the total variation regularization in the non-blind deconvolution step, the solutions of g_j^h and g_j^v are $g_j^h = \mathcal{S}_{\frac{\lambda}{2\beta}}(\nabla_v I_j)$ and $g_j^v = \mathcal{S}_{\frac{\lambda}{2\beta}}(\nabla_v I_j)$, where

$$\mathcal{S}_{\lambda/(2\beta)}(x) = \begin{cases} x - \lambda/(2\beta), & x > \lambda/(2\beta), \\ x + \lambda/(2\beta), & x < -\lambda/(2\beta), \\ 0, & \text{otherwise.} \end{cases} \tag{6}$$

Based on the chain rule, the gradient of $\mathcal{L}_I$ with respect to ω_i is

$$\begin{aligned} \frac{\partial \mathcal{L}_j^I}{\partial \omega_i} =& \frac{\partial \mathcal{L}_j^I}{\partial \mathbf{I}_j} \frac{\partial \mathbf{I}_j}{\partial \omega_i} \\ =& (\mathbf{I}_j - \mathbf{I}_j^{gt})^\top \frac{\partial \mathbf{I}_j}{\partial \omega_i}. \end{aligned} \tag{7}$$

We note that (5) gives the relationship between I_j and w_i in the frequency domain. We can easily get its equivalent matrix-vector form according to convolution property:

$$\mathbf{I}_j = \left(\sum_i \omega_i \mathbf{K}_{ij}^\top \mathbf{K}_{ij} + \beta(\mathbf{D}_h^\top \mathbf{D}_h + \mathbf{D}_v^\top \mathbf{D}_v) \right)^{-1} \left(\sum_i \omega_i \mathbf{K}_{ij}^\top \mathbf{B}_{ij} + \beta(\mathbf{D}_h^\top \mathbf{g}_j^h + \mathbf{D}_v^\top \mathbf{g}_j^v) \right), \tag{8}$$

where $\mathbf{K}_{ij}$ is the matrix of $f_i * k_j$ with respect to I, $\mathbf{D}_h$ and $\mathbf{D}_v$ are the matrix forms of ∇_h and ∇_v, and $\mathbf{B}_{ij}$, $\mathbf{g}_j^h$, and $\mathbf{g}_j^v$ are the vector forms of $f_i * B_j$, g_j^h, and g_j^v, respectively.

Based on **Proposition 1**, we can get

$$\begin{aligned} \frac{\partial \mathbf{I}_j}{\partial \omega_i} =& - \left(\sum_i \omega_i \mathbf{K}_{ij}^\top \mathbf{K}_{ij} + \beta(\mathbf{D}_h^\top \mathbf{D}_h + \mathbf{D}_v^\top \mathbf{D}_v) \right)^{-1} \mathbf{K}_{ij}^\top \mathbf{K}_{ij} \mathbf{I}_j \\ &+ \left(\sum_i \omega_i \mathbf{K}_{ij}^\top \mathbf{K}_{ij} + \beta(\mathbf{D}_h^\top \mathbf{D}_h + \mathbf{D}_v^\top \mathbf{D}_v) \right)^{-1} \mathbf{K}_{ij}^\top \mathbf{B}_{ij}. \end{aligned} \tag{9}$$

We use FFT to compute (9). Thus, $\frac{\partial \mathcal{L}_j^I}{\partial \omega_i}$ can be obtained by

$$\frac{\partial \mathcal{L}_j^I}{\partial \omega_i} = (\mathbf{I}_j - \mathbf{I}_j^{gt})^\top \mathbf{W}_i. \tag{10}$$

Each W_i is defined by

$$W_i = \mathcal{F}^{-1} \left(\frac{\Delta_b}{\Delta_d} - \frac{\Delta_f \Delta_n}{\Delta_d^2} \right), \tag{11}$$

where $\Delta_d = \sum_i \omega_i \overline{\mathcal{F}(f_i)\mathcal{F}(k_j)} \mathcal{F}(k_j)\mathcal{F}(f_i) + \beta \left(\sum_{i \in \{h,v\}} \overline{\mathcal{F}(\nabla_i)} \mathcal{F}(\nabla_i) \right)$, $\Delta_f = \overline{\mathcal{F}(f_i)\mathcal{F}(k_j)} \mathcal{F}(k_j)\mathcal{F}(f_i)$, $\Delta_b = \overline{\mathcal{F}(f_i)\mathcal{F}(k_j)}$ $\mathcal{F}(f_i * B)$, and $\Delta_n = \sum_i \omega_i \overline{\mathcal{F}(k_j)} \mathcal{F}(f_i * B_j) + \beta \left(\overline{\mathcal{F}(\nabla_h)} \mathcal{F}(g_j^h) + \overline{\mathcal{F}(\nabla_v)} \mathcal{F}(g_j^v) \right)$.

3. Algorithm of Learning Discriminative Non-blind Deconvolution

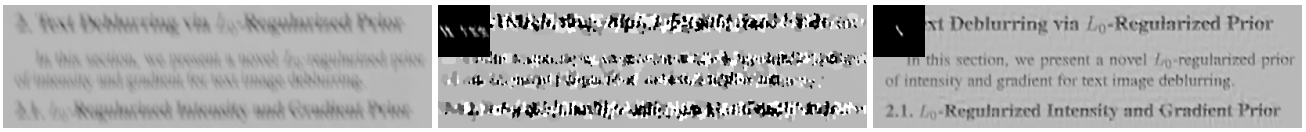
In the manuscript, we only give some main steps of the algorithm for learning discriminative non-blind deconvolution due to space limit. In this supplemental material, we summarize the algorithm for learning discriminative non-blind deconvolution in Algorithm 1.

4. Extensions of Proposed Method

As stated in the manuscript, our method can be applied to other deblurring tasks with specific image priors. In this supplemental material, we show more results on the text deblurring [4] to demonstrate the flexibility of the proposed method.

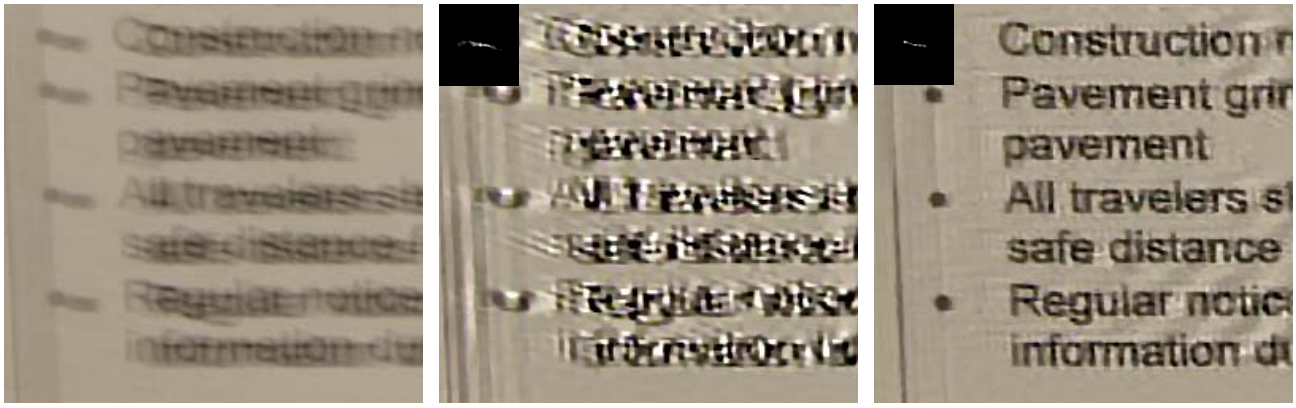
Algorithm 1 Algorithm for Learning Discriminative Non-blind Deconvolution

Input: Blurred image B_j and blur kernel k_j^{gt} and ground truth clear image I_j^{gt}.
$I_j \leftarrow B_j$, $\beta \leftarrow 2\lambda$.
while $l \leq \max\_iter1$ **do**
  repeat
    solve for g_j^h and g_j^v using one dimension shrinkage operator.
    solve for I_j using (5).
    $\beta \leftarrow 2\beta$.
  until $\beta > \beta_{\max}$
  $\omega_i = \omega_i - \alpha_I \sum_j \frac{\partial \mathcal{L}_j^I}{\partial \omega_i}$.
end while
Output: The weight ω_i.



(a) Blurred image (b) Pan et al. [4] (c) Ours

Figure 1. Extension of text image deblurring algorithm [4]. The results generated by the proposed method contain clearer characters and fewer ringing artifacts.



(a) Blurred image (b) Pan et al. [4] (c) Ours

Figure 2. Extension of text image deblurring algorithm [4]. The results generated by the proposed method contain clearer characters and fewer ringing artifacts.

5. More Experimental Results

In this section, we show more experimental results in Figures 3, 4, 5, 6 and 7.

References

[1] S. Cho and S. Lee. Fast motion deblurring. In *SIGGRAPH Asia*, volume 28, page 145, 2009. 4, 5

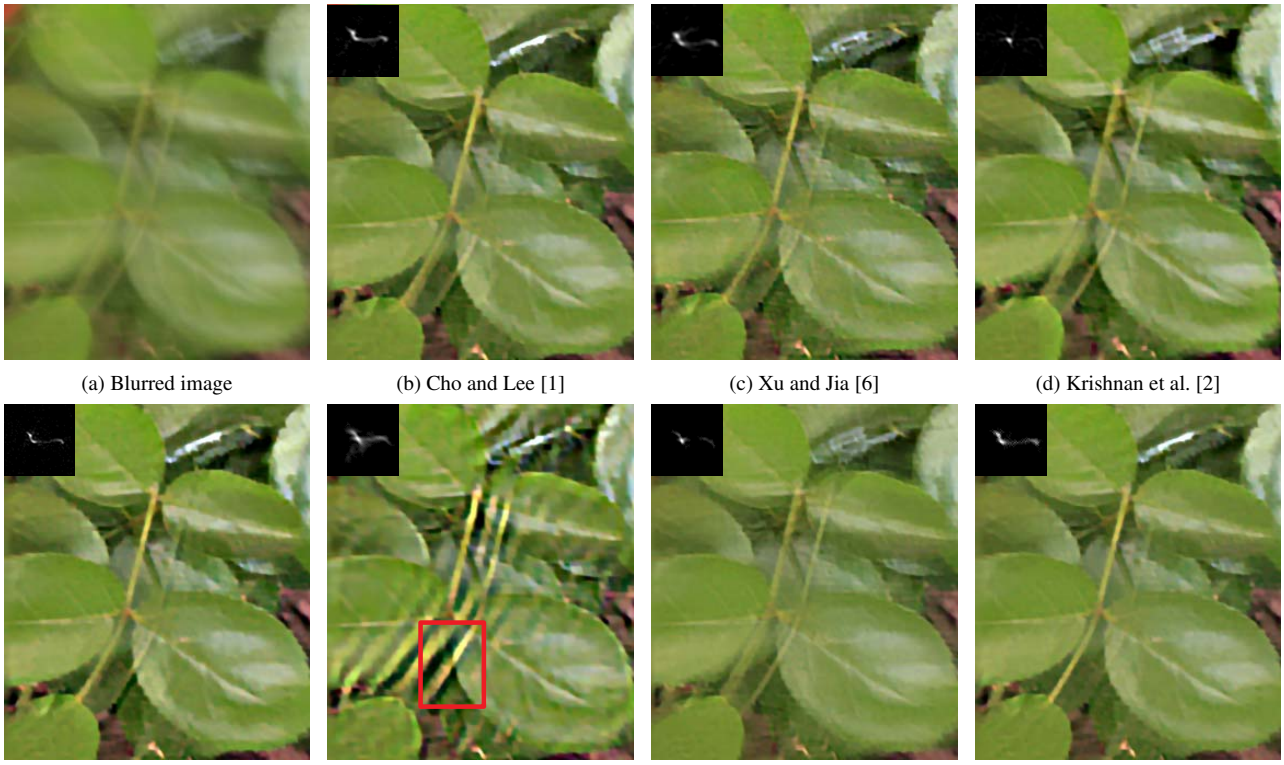
[2] D. Krishnan, T. Tay, and R. Fergus. Blind deconvolution using a normalized sparsity measure. In *CVPR*, pages 2657–2664, 2011. 4

[3] A. Levin, Y. Weiss, F. Durand, and W. T. Freeman. Understanding and evaluating blind deconvolution algorithms. In *CVPR*, pages 1964–1971, 2009. 4

[4] J. Pan, Z. Hu, Z. Su, and M.-H. Yang. Deblurring text images via L_0-regularized intensity and gradient prior. In *CVPR*, pages 2901–2908, 2014. 2, 3, 4

[5] K. B. Petersen and M. S. Pedersen. *The Matrix Cookbook*. http://matrixcookbook.com. 1

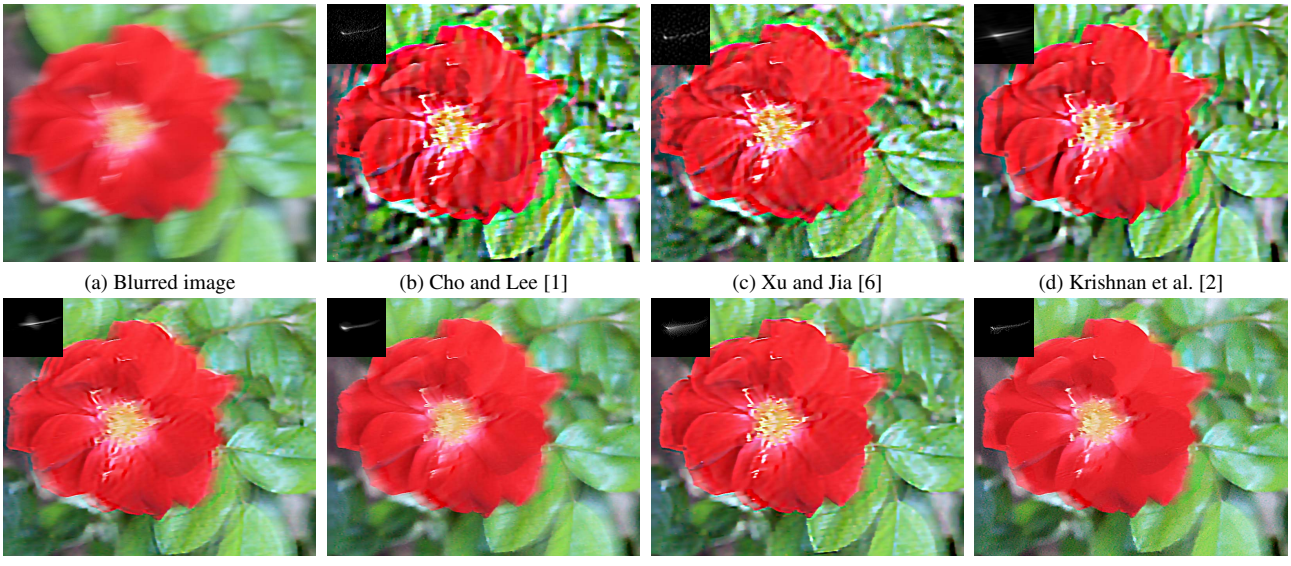
[6] L. Xu and J. Jia. Two-phase kernel estimation for robust motion deblurring. In *ECCV*, pages 157–170, 2010. 4, 5



(a) Blurred image (b) Cho and Lee [1] (c) Xu and Jia [6] (d) Krishnan et al. [2]

(e) Xu et al. [3] (f) Pan et al. [4] (g) With fixed data fitting functions (h) Ours

Figure 3. A real captured image. Comparisons on a natural image. The results in (b)-(g) still contain significant blur residual. Compared with results using fixed data fitting functions in (g), our method generates better deblurred results. (Best viewed on high-resolution display with zoom-in.)



(a) Blurred image (b) Cho and Lee [1] (c) Xu and Jia [6] (d) Krishnan et al. [2]

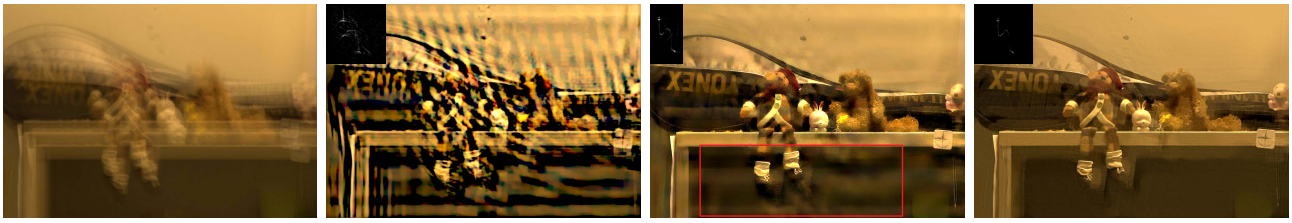
(e) Levin et al. [3] (f) Zhong et al. [7] (g) Pan et al. [4] (h) Ours

Figure 4. Another real captured image. Our method generates visually comparable or even better deblurring results. (Best viewed on high-resolution display with zoom-in.)

[7] L. Zhong, S. Cho, D. Metaxas, S. Paris, and J. Wang. Handling noise in single image deblurring using directional filters. In *CVPR*, pages 612–619, 2013. 4

(a) Blurred image (b) Cho and Lee [1] (c) Xu and Jia [6] (d) Ours

Figure 5. Real captured image from [6]. Our method generates clear results with fewer noise. (Best viewed on high-resolution display with zoom-in.)



(a) Blurred image (b) Cho and Lee [1] (c) Xu and Jia [6] (d) Ours

Figure 6. Real captured image from [6]. Our method generates clear results with fewer ringing artifacts. The part in the red box in (c) contain some ringing artifacts. (Best viewed on high-resolution display with zoom-in.)

(a) Blurred image (b) Cho and Lee [1] (c) Xu and Jia [6] (d) Ours

Figure 7. Real captured image from [6]. Our method generates the results with much clearer characters. (Best viewed on high-resolution display with zoom-in.)