Article

Effective Alternating Direction Optimization Methods for Sparsity-Constrained Blind Image Deblurring

Naixue Xiong [1,2], Ryan Wen Liu [1,*], Maohan Liang [1], Di Wu [3], Zhao Liu [1,*] and Huisi Wu [4]

[1] Hubei Key Laboratory of Inland Shipping Technology, School of Navigation, Wuhan University of Technology, Wuhan 430063, China; xiongnaixue@gmail.com (N.X.); mhliang@whut.edu.cn (M.L.)

[2] Department of Business and Computer Science, Southwestern Oklahoma State University, Oklahoma, OK 73096, USA

[3] School of Computer, Wuhan University, Wuhan 430072, China; wudi2012@whu.edu.cn

[4] College of Computer Science and Software Engineering, Shenzhen University, Shenzhen 518060, China; hswu@szu.edu.cn

* Correspondence: wenliu@whut.edu.cn (R.W.L.); zhaoliu@whut.edu.cn (Z.L.);
  Tel.: +86-27-8658-7900 (R.W.L.); +86-27-8658-5810 (Z.L.)

Abstract: Single-image blind deblurring for imaging sensors in the Internet of Things (IoT) is a challenging ill-conditioned inverse problem, which requires regularization techniques to stabilize the image restoration process. The purpose is to recover the underlying blur kernel and latent sharp image from only one blurred image. Under many degraded imaging conditions, the blur kernel could be considered not only spatially sparse, but also piecewise smooth with the support of a continuous curve. By taking advantage of the hybrid sparse properties of the blur kernel, a hybrid regularization method is proposed in this paper to robustly and accurately estimate the blur kernel. The effectiveness of the proposed blur kernel estimation method is enhanced by incorporating both the L_1-norm of kernel intensity and the squared L_2-norm of the intensity derivative. Once the accurate estimation of the blur kernel is obtained, the original blind deblurring can be simplified to the direct deconvolution of blurred images. To guarantee robust non-blind deconvolution, a variational image restoration model is presented based on the L_1-norm data-fidelity term and the total generalized variation (TGV) regularizer of second-order. All non-smooth optimization problems related to blur kernel estimation and non-blind deconvolution are effectively handled by using the alternating direction method of multipliers (ADMM)-based numerical methods. Comprehensive experiments on both synthetic and realistic datasets have been implemented to compare the proposed method with several state-of-the-art methods. The experimental comparisons have illustrated the satisfactory imaging performance of the proposed method in terms of quantitative and qualitative evaluations.

Keywords: imaging sensors; blind deblurring; image restoration; total variation; total generalized variation; alternating direction method of multipliers

1. Introduction

1.1. Background and Related Work

Single-image blind deblurring for imaging sensors has recently received increasing attention in modern imaging applications, e.g., the Internet of Things (IoT), astronomical imaging, biomedical imaging, computational photography and microscopy [1–3]. It is well known that the image pixel

intensity can be determined by the total incoming light sensed by the imaging sensor over the exposure time. As shown in Figure 1, the discrete image degradation model can be written as follows:

$$B\left(x,y\right) = \sum_{m=1}^{M} w_m \left(H_m L\right)\left(x,y\right) + \xi\left(x,y\right), \tag{1}$$

where $B\left(x,y\right)$ is the observed image after camera exposure, $L\left(x,y\right)$ denotes the latent sharp image to be restored, $\xi\left(x,y\right)$ denotes the additive white Gaussian noise, w_m is a weight, which essentially represents the length of exposure time at camera pose m, and H_m is a transformation matrix related to the camera rotation or translation at pose m during exposure. In this work, we only consider the case of uniform (i.e., spatially invariant) image deblurring. Thus, the matrix H_m only corresponds to the camera translation along both the X and Y axes. For the sake of simplicity, the original image degradation model (1) can be rewritten as a convolution version as follows:

$$B = L \otimes k + \xi, \tag{2}$$

where $\otimes$ is the mathematical operation of convolution, k denotes the blur kernel related to the weight w and transformation matrix H in (1). The purpose of single-image blind deblurring is to recover both k and L from only one blurred image B. It is a challenging ill-conditioned inverse problem, since many different pairs k and L can lead to the same B [4]. The constraints on both the blur kernel and the latent sharp image should be exploited to select the optimal pairs k and L for enhancing imaging performance.
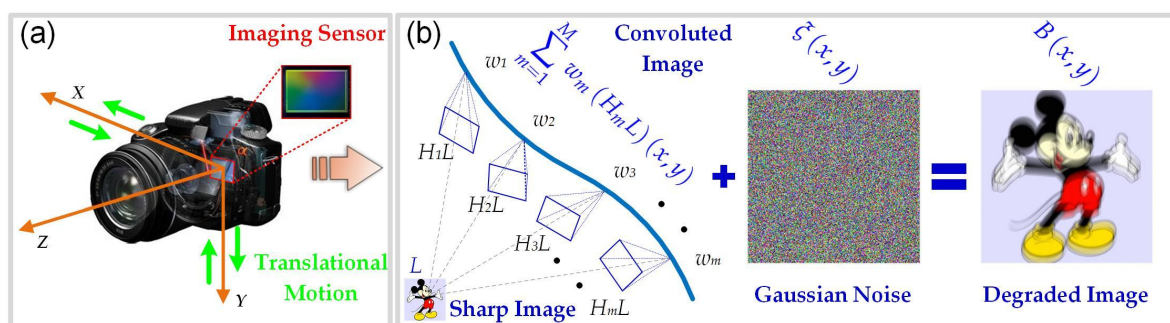

Figure 1. Diagram of the image degradation model for the motion blur of the imaging sensor in the Internet of Things (IoT). (**a**) Camera translation along both the X and Y axes considered in this work. (**b**) Discrete image degradation model with the curve being the sensor motion trajectory over the exposure time.

To cope with the ill-conditioned nature of blind deblurring, many statistical priors learned from blur kernels and latent sharp images have been developed to regularize the restoration process. In the literature [4,5], current single-image blind deblurring methods are widely divided into two categories: (1) methods that simultaneously estimate both the blur kernel and latent sharp image; (2) methods that first estimate the blur kernel, then recover the latent sharp image. It is well known that the support size of the blur kernel is often extremely smaller compared to the image size. Therefore, the joint maximum a posteriori (MAP) estimation of k and L often fails since the number of unknowns is larger than the number of known variables in B. In contrast, the estimation of the blur kernel can be obtained accurately through the MAP estimation of k alone [4]. To guarantee the high-quality blind deblurring, this paper mainly focuses on the second type of method, i.e., estimating the blur kernel first and then dealing with the corresponding non-blind deconvolution problem.

The pioneering work [6] mainly focused on the estimation of simple and small blur kernels, which are very rare in many practical scenarios. To make blind deblurring more practical, most current state-of-the-art methods were usually proposed by exploiting the prior knowledge from the statistics

of blur kernels and sharp images. In 2006, Fergus et al. [7] contributed the original work on practical blind deblurring where the blur kernels were quite large and complex. In particular, the authors proposed a variational Bayesian image deblurring model by combining the mixture-of-Gaussian image prior with the mixture-of-exponential kernel prior. In [8], the blur kernel prior was assumed to follow an exponential distribution. Under the MAP framework, the exponential distribution results in an L_1-norm constraint on kernel intensity, which has a good interpretation on the sparsity of the blur kernel [9,10]. Under some imaging conditions, the blur kernel can also be assumed as a (piecewise) sufficiently smooth function. As a consequence, many researchers proposed to replace the L_1-norm of kernel intensity with its squared L_2-norm version [11–13]. Current experiments have shown that both L_1- and squared L_2-regularized optimization methods could achieve accurate kernel estimation on the benchmark dataset introduced by [4]. In the case of large blur kernels, however, it is difficult to robustly and accurately estimate the blur kernels using these methods mentioned. Many efforts [4,5] have been devoted to theoretically explain the reason why it is difficult for accurately estimating the blur kernels, especially for the large ones in practice. More recently, hybrid sparsity priors on blur kernels [14,15] have been considered and achieved robust image restoration results.

Once the blur kernel is estimated, the blind deblurring problem (2) essentially becomes a non-blind image deconvolution. During the past several decades, numerous numerical methods have been developed to handle non-blind deconvolution. One of the most popular methods is the Tikhonov regularization [16,17], followed by its various extensions [18,19]. These methods can be easily implemented, but commonly generate over-smoothing effects on the restored images. Other widely-used methods, such as the Richardson–Lucy method [20] and Wiener filter [21], easily suffer from noise amplification and ringing-like artifacts. To overcome the undesirable artifacts, Yuan et al. [22] proposed to develop a progressive inter-scale and intra-scale image deconvolution approach based on the bilateral Richardson-Lucy method. Current research illustrates that images have the properties of sparse gradients. Many efforts [23–25] were made to enhance non-blind deconvolution by imposing the total variation (TV) regularizer. From a statistical point of view, the TV regularizer corresponds to an assumption of a Laplacian sparse prior on image gradients. Recently, the extended TV regularizers, such as non-convex TV (NCTV) [11,13,26] and higher-order TV (HOTV) [27,28], have been attracting increasing attention for improving non-blind deconvolution. Both TV and HOTV regularizers have also been combined to overcome the potential disadvantages existing in these two regularizers [29–31]. The newly-developed total generalized variation (TGV) regularizer, originally proposed by Bredies et al. [32] in 2010, achieved great success for the restoration of blurred images [33–35]. Motivated by the concepts of non-local means (NLM) and graph Laplacian, the non-local TV (NLTV) regularizer has significantly improved the deconvolution quality [36–39]. The NLTV-regularized variational models can guarantee the highest-quality deconvolution because they take full advantage of the high degree of geometrical self-similarity that is inherent in natural images.

1.2. Motivation and Contributions

In the current literature [7–13], most existing blur kernel estimation methods were proposed based on the assumptions that the blur kernel was spatially sparse or piecewise smooth within the support of a continuous curve. As a consequence, the proposed methods could not always guarantee high-accuracy estimation under certain degradation conditions. In recent years, more attention has been paid to the sparse image priors for improving the estimation accuracy. To further enhance the estimation quality, in our opinion, it is still necessary to investigate advanced sparsity constraints on the blur kernel. The robust estimation method will be proposed in this paper by taking into account the sparsity and smoothing properties of the blur kernel. In particular, the sparsity property is promoted using the L_1-norm of kernel intensity; the smoothing property is utilized through the introduction of the squared L_2-norm of the intensity derivative. By making full use of the advantages of both the L_1-norm and the squared L_2-norm on kernel prior representation, the proposed method could potentially generate satisfactory estimation under more different degradation conditions. Essentially,

most of the previous works [8,9,11–13] on blur kernel estimation can be considered as a special case of our proposed hybrid regularization method. If we only use the L_1-norm term, it can take full advantage of the property of spatial sparsity, but the resulting estimated blur kernel easily suffers from the isolated points [15]. If we only use the squared L_2-norm term, the continuous and smoothing properties of blur kernels under certain imaging conditions could be well preserved. However, the potential spatial sparsity property may be ignored, leading to inaccurate estimation of the blur kernel. Therefore, to guarantee the accuracy of the estimated blur kernel, it is necessary to combine the L_1-norm of kernel intensity with the squared L_2-norm of the intensity derivative in our work. If images are degraded by Gaussian, average or pillbox (disc) blur kernels, which have the weak spatial sparsity properties, but high smoothing properties, the proposed hybrid regularization method could theoretically generate higher estimation accuracy compared with traditional single regularization methods. It is worth mentioning that the hybrid blur kernel prior proposed in this work is extremely different from the current hybrid versions [14,15]. Current work [13,40–42] has illustrated that the L_0 quasi-norm has a good natural interpretation of the sparsity property of the image gradient and benefits for image detail enhancement. In particular, it performs well in penalizing small gradient magnitudes and encouraging large ones to preserve fine details. To improve the accuracy of blur kernel estimation, the L_0 quasi-norm of the image gradient is also incorporated into our blur kernel estimation method. Owing to the non-convex nature of the L_0 quasi-norm and the non-smooth nature of the L_1-norm, the commonly-used numerical methods could not be effectively adopted to solve the blur kernel estimation problem. To guarantee a feasible solution, the resulting non-convex non-smooth optimization problem will be effectively dealt with by developing an alternating direction method of multipliers (ADMM)-based numerical method [43]. The preliminary results on blur kernel estimation can be found in our previous short-version conference paper [44].

Existing work has illustrated that the TV regularizer, first proposed by Rudin et al. [23] in 1992, has the capacity of preserving edges and smoothing flat regions. TV-regularized variational image restoration models with the L_1-norm [11] or the squared L_2-norm [24] data-fidelity terms have gained considerable attention. However, the image quality could be degraded because the results often suffer from undesirable staircase-like artifacts in regions with gradual intensity variations [45]. The reason behind this phenomenon is that the TV regularizer favors solutions that are piecewise constant. To effectively suppress the artifacts, many extensions of TV [11,13,26–28,36–39] could be used to improve the image quality. For example, the patch-based NLTV regularizer has the capacity of guaranteeing the highest-quality image restoration. However, the NLTV-regularized variational model is practically limited due to the high computational cost. To make it easier to implement blind deblurring in practice, it is necessary to balance the trade-off between computational cost and imaging performance. Motivated by the success of the TGV regularizer, we tend to propose an effective non-blind deconvolution method based on the TGV regularizer of second-order (i.e., TGV^2) [32]. The TGV^2 regularizer is able to suppress the undesirable artifacts, while preserving the image edges since it favors piecewise polynomial intensities [34]. The quality of restored images could be correspondingly enhanced. From an optimization point of view, the resulting image deconvolution model could not be directly solved using traditional numerical methods because of the non-smooth nature of the TGV^2 regularizer. To achieve a robust and effective solution, an ADMM-based optimization method will be developed to solve the resulting non-smooth minimization problem. In particular, the original complex minimization problem can be decomposed into several simple subproblems by introducing some auxiliary variables. Each of these subproblems has a closed-form solution or can be efficiently solved using the current numerical method. The effectiveness of the proposed method will be demonstrated using comprehensive experiments on both synthetic and realistic blurred images.

In conclusion, the main contributions of this paper, given the state-of-the-art research work, are mainly summarized by the following three aspects:

- To accurately estimate the blur kernel, a hybrid regularization method was proposed by combining the L_1-norm of kernel intensity with the squared L_2-norm of the intensity derivative. An alternating direction method was presented to effectively solve the resulting blur kernel estimation problem.

- The TGV^2-regularized variational model with an L_1-norm data-fidelity term was proposed for enhancing the non-blind deconvolution result. To guarantee the stability and effectiveness of the solution, an ADMM-based numerical method was developed to solve the resulting non-smooth optimization problem.

- The satisfactory blind deblurring performance of the proposed method has been illustrated using comprehensive experiments on both synthetic and realistic blurred images (with large blur kernels). The proposed method has also been successfully exploited for single-image deblurring in the field of ocean engineering.

The main benefit of the proposed method is that it takes full advantage of the hybrid constraints for blur kernel estimation and the TGV^2 regularizer for non-blind deconvolution. Therefore, it can accurately estimate the blur kernel and guarantee high-quality image deconvolution. Experiments using synthetic, as well as realistic blurred images will be implemented to verify the effectiveness of our proposed method in practical applications.

2. Hybrid Regularized Blur Kernel Estimation

As discussed in Section 1.2, our robust two-step framework for single-image blind deblurring is illustrated in Figure 2. This section mainly focuses on the first blur kernel estimation step, which is separated into the following two aspects: (1) sharp edges restoration; and (2) blur kernel estimation. In order to enhance the deblurring performance, we exploit the following statistical priors for blur kernel estimation: an L_0-sparsity prior on the latent gradient image x and a hybrid sparsity prior on the blur kernel k. Under these sparsity-constrained priors, blur kernel estimation in this paper is equivalent to solving the following minimization problem:

$$(x^*, k^*) = \underset{x,k}{\mathrm{argmin}} \left\{ \frac{1}{2} \|x \otimes k - y\|_2^2 + \gamma \|x\|_0 + \eta_1 \|k\|_1 + \eta_2 \|\nabla k\|_2^2 \right\}, \tag{3}$$

where γ, η_1, η_2 are predefined positive regularization parameters, the L_0 quasi-norm $\|\circ\|_0$ counts the number of nonzero elements, x denotes $\nabla L = (\partial_h L, \partial_v L)^T$ and y denotes $\nabla B = (\partial_h B, \partial_v B)^T$ with ∂_h and ∂_v being the finite differences along the horizontal and vertical directions, respectively. The proposed blur kernel estimation model (3) is mainly composed of four terms: The first term, called the squared L_2-norm data-fidelity term, denotes a measure of the distance between the restored data and the observed version. The second term is the L_0 quasi-norm regularization term, which can preserve the sparsity of natural image gradients. The third and fourth terms are, respectively, the L_1-norm and the squared L_2-norm constraints on the blur kernel, which can stabilize the final estimation result. The accurate estimation of the blur kernel is beneficial for generating high-quality non-blind image deconvolution.
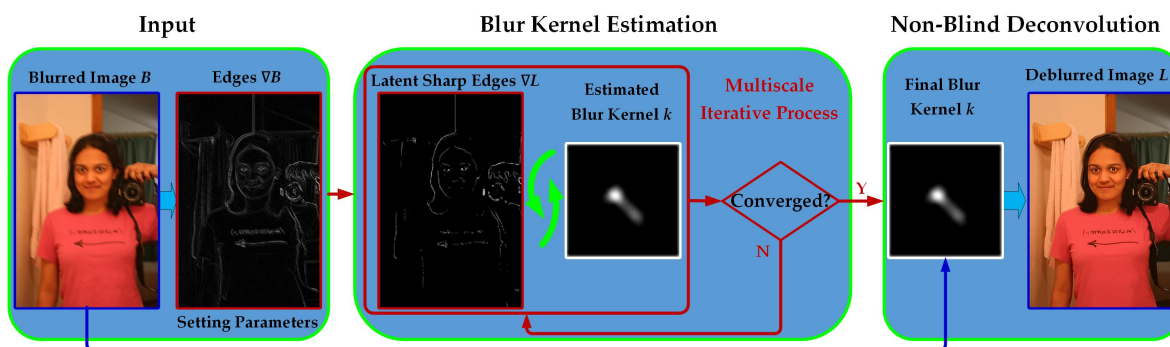
Figure 2. The illustration of our proposed robust regularization method for single-image blind deblurring. The proposed method first estimates the blur kernel, then recovers the latent sharp image.

2.1. Sharp Edge Restoration

Since x and k are independent in (3), in the first step, the latent sharp edges x at the $(m+1)$-th outer iteration can be recovered by solving the following minimization problem:

$$x_{m+1} = \underset{x}{\arg\min} \left\{ \frac{1}{2} \|x \otimes k_m - y\|_2^2 + \gamma \|x\|_0 \right\},$$ (4)

for $m = 0, 1, \cdots, M_{\max}$ with $M_{\max}$ denoting the maximum number of outer iterations. As discussed in [11,13], strong edges are not always beneficial for accurate estimation of the blur kernel. To select the informative edges, an effective method used to measure the usefulness of gradients is given by:

$$r(p) = \frac{\left\| \sum_{q \in \mathcal{N}_h(p)} \nabla B(q) \right\|_2}{\sum_{q \in \mathcal{N}_h(p)} \|\nabla B(q)\|_2 + 0.5},$$ (5)

where B is the blurred image and $\mathcal{N}_h(p)$ denotes an $h \times h$ window centered at pixel $p \in \Omega$ (image domain). The measure metric (5), first proposed by Xu and Jia [11], enables accurate estimation of the blur kernel by removing some narrow strips. To incorporate the measure metric into our kernel estimation framework, the problem (4) can be reformulated as:

$$\mathbf{x}_{m+1} = \underset{\mathbf{x}}{\arg\min} \left\{ \frac{1}{2} \|\mathcal{K}_m \mathbf{x} - \mathbf{y}\|_2^2 + \gamma \|\kappa \circ \mathbf{x}\|_0 \right\},$$ (6)

where $\circ$ represents the pointwise product and $\kappa(p) = \exp(-|r(p)|^{0.8})$ for $p \in \Omega$ with the variable $r(p)$ being defined in (5). For the sake of simplicity, the convolution version in (6) is expressed in a matrix-vector multiplication form. In (6), $\mathcal{K}_m$ is a block Toeplitz matrix with Toeplitz blocks transformed from the blur kernel k at the m-th outer iteration. $\mathbf{x}$ and $\mathbf{y}$ represent the vector version of x and y, respectively. It is well known that the model (6) is difficult to solve directly because of the non-smooth and non-convex natures of the L_0 quasi-norm $\|\kappa \circ \mathbf{x}\|_0$. To guarantee solution efficiency and stability, we propose to develop an ADMM-based numerical method [43,46,47] to solve the unconstrained optimization problem (6). To apply ADMM, we first replace $\mathbf{x}$ by $\mathbf{v}$ and then transform (6) into the following constrained optimization problem:

$$\min_{\mathbf{v},\mathbf{x}} \left\{ \frac{1}{2} \|\mathcal{K}_m \mathbf{x} - \mathbf{y}\|_2^2 + \gamma \|\kappa \circ \mathbf{v}\|_0 \right\}, \quad \text{s.t. } \mathbf{v} = \mathbf{x}.$$ (7)

Note that the updates of $\mathbf{v}$ and $\mathbf{x}$ are independent of each other. Let $\mathcal{L}_{\mathcal{A}}\left(\mathbf{v}, \mathbf{x}; \varphi^{\mathbf{x}}\right)$ represent the augmented Lagrangian function of (7), which is defined as follows:

$$\mathcal{L}_{\mathcal{A}}\left(\mathbf{v}, \mathbf{x}; \varphi^{\mathbf{x}}\right) = \frac{1}{2}\left\|\mathcal{K}_m \mathbf{x} - \mathbf{y}\right\|_2^2 + \gamma\left\|\kappa \circ \mathbf{v}\right\|_0 + \frac{\beta_1}{2}\left\|\mathbf{v} - \mathbf{x}\right\|_2^2 - \left\langle\varphi^{\mathbf{x}}, \mathbf{v} - \mathbf{x}\right\rangle, \tag{8}$$

where β_1 is a pre-defined penalty parameter and $\varphi^{\mathbf{x}}$ denotes the Lagrangian multiplier. In particular, ADMM solves problem (6) by minimizing $\mathcal{L}_{\mathcal{A}}\left(\mathbf{v}, \mathbf{x}; \varphi^{\mathbf{x}}\right)$ with respect to $\mathbf{v}$ and $\mathbf{x}$ alternatively given the other fixed, followed by an update of the Lagrangian multiplier $\varphi^{\mathbf{x}}$, i.e.,

$$\begin{cases} \mathbf{v}_{i+1} = \underset{\mathbf{v}}{\arg\min}\left\{\gamma\left\|\kappa \circ \mathbf{v}\right\|_0 + \frac{\beta_1}{2}\left\|\mathbf{v} - \left(\mathbf{x}_{m,i} + \frac{\varphi_i^{\mathbf{x}}}{\beta_1}\right)\right\|_2^2\right\} \\ \mathbf{x}_{m,i+1} = \underset{\mathbf{x}}{\arg\min}\left\{\frac{1}{2}\left\|\mathcal{K}_m \mathbf{x} - \mathbf{y}\right\|_2^2 + \frac{\beta_1}{2}\left\|\mathbf{x} - \left(\mathbf{v}_{i+1} - \frac{\varphi_i^{\mathbf{x}}}{\beta_1}\right)\right\|_2^2\right\} \end{cases} \tag{9}$$

with $\mathbf{x}_{m,0} = \mathbf{x}_m$. At each iteration, the Lagrangian multiplier $\varphi^{\mathbf{x}}$ can be updated through $\varphi_{i+1}^{\mathbf{x}} = \varphi_i^{\mathbf{x}} - \tau\beta_1\left(\mathbf{v}_{i+1} - \mathbf{x}_{m,i+1}\right)$ for $i = 0, 1, \cdots, I_{\max}$ with $I_{\max}$ denoting the maximum number of inner iterations. Here, $\tau \in \left(0, (\sqrt{5}+1)/2\right)$ denotes the step length. In order to accelerate the convergence of numerical solution, during each iteration, the penalty parameter β_1 can be updated as $\beta_1 \leftarrow \rho\beta_1$ with ρ being a positive step length. The explicit solution of the $\mathbf{v}$-subproblem in (9) can be directly obtained using the element-wise hard thresholding operator formulated in [48], i.e.,

$$\mathbf{v}_{i+1} = \mathcal{H}_{\kappa\gamma, \beta_1}\left(\mathbf{x}_{m,i} + \frac{\varphi_i^{\mathbf{x}}}{\beta_1}\right), \tag{10}$$

where $\mathcal{H}_{a,b}\left(\cdot\right)$ is defined as:

$$\mathcal{H}_{a,b}\left(\mathbf{s}\right) = \begin{cases} 0, & \text{if } |\mathbf{s}| < \sqrt{2a/b}, \\ \mathbf{s}, & \text{otherwise.} \end{cases}$$

where both a and b are intermediate variables.

Essentially, the $\mathbf{x}$-subproblem in (9) is a least-squares optimization problem; the corresponding normal equation can be readily obtained as follows:

$$\left(\mathcal{K}_m^\top \mathcal{K}_m + \beta_1 \mathbf{I}\right)\mathbf{x} = \mathcal{K}_m^\top \mathbf{y} + \beta_1\left(\mathbf{v}_{i+1} - \frac{\varphi_i^{\mathbf{x}}}{\beta_1}\right), \tag{11}$$

where superscript $\top$ denotes the transpose operator for real matrices or vectors and $\mathbf{I}$ denotes an identity matrix. Under the periodic boundary condition, $\mathcal{K}_m^\top \mathcal{K}_m$ is a block circulant matrix with circulant blocks. It can be diagonalized using the two-dimensional discrete Fourier transform. Let $\mathcal{F}$ denote the forward fast Fourier transform (FFT) operator. Applying by $\mathcal{F}$ on both sides of (11) and yields:

$$\left(\overline{\mathcal{F}(\mathcal{K}_m)}\mathcal{F}(\mathcal{K}_m) + \beta_1 \mathcal{F}(\mathbf{I})\right)\mathcal{F}(\mathbf{x}) = \overline{\mathcal{F}(\mathcal{K}_m)}\mathcal{F}(\mathbf{y}) + \beta_1 \mathcal{F}\left(\mathbf{v}_{i+1} - \frac{\varphi_i^{\mathbf{x}}}{\beta_1}\right). \tag{12}$$

To decrease the computational cost, both $\overline{\mathcal{F}(\mathcal{K}_m)}\mathcal{F}(\mathcal{K}_m) + \beta_1\mathcal{F}(\mathbf{I})$ and $\overline{\mathcal{F}(\mathcal{K}_m)}\mathcal{F}(\mathbf{y})$ can be computed only once at the beginning of the iterative algorithm. Thus, solving (12) is straightforward,

which means that it is relatively easy to achieve the solution of (11). As a consequence, the solution of the least-squares optimization problem (11) is given by:

$$\mathbf{x}_{m,i+1} = \mathcal{F}^{-1} \left(\frac{\overline{\mathcal{F}(\mathcal{K}_m)} \mathcal{F}(\mathbf{y}) + \beta_1 \mathcal{F} \left(\mathbf{v}_{i+1} - \varphi_i^{\mathbf{x}} / \beta_1 \right)}{\overline{\mathcal{F}(\mathcal{K}_m)} \mathcal{F}(\mathcal{K}_m) + \beta_1 \mathcal{F}(\mathbf{I})} \right), \tag{13}$$

where $\mathcal{F}^{-1}(\cdot)$ denotes the inverse FFT operator and $\overline{\mathcal{F}(\cdot)}$ represents the complex conjugate operator. The minimization process (9) is implemented alternately until the solution converges to the optimal one. Finally, the recovered sharp edge $\mathbf{x}_{m+1} = \mathbf{x}_{m,I_{\max}}$ is achieved to enhance the accuracy of blur kernel estimation in the next step. The whole optimization procedure of ADMM for Subproblem (6) is summarized in Algorithm 1.

Algorithm 1 ADMM for Subproblem (6).

1: **Input**: Blur kernel $\mathbf{k}_m$, blurred image gradient $\mathbf{y}$, $\tau = 1.618$, $\rho = 3$, $I_{\max} = 5$ and $\epsilon = 5 \times 10^{-6}$.

2: **Initialize**: $\varphi_0^{\mathbf{x}} = 0$, $\beta_1 = 0.03$ and $i = 0$.

3: $\mathbf{x}_{m,0} = \mathbf{x}_m$.

4: **while** (not converged and $i \leq I_{\max}$) **do**

5:     Compute $\mathbf{v}_{i+1}$ according to (10).

6:     Compute $\mathbf{x}_{m,i+1}$ according to (13).

7:     Update Lagrangian multiplier and parameter: $\varphi_{i+1}^{\mathbf{x}} = \varphi_i^{\mathbf{x}} - \tau\beta_1 \left(\mathbf{v}_{i+1} - \mathbf{x}_{m,i+1} \right)$ and $\beta_1 \leftarrow \rho\beta_1$.

8:     Check convergence condition: $\|\mathbf{v}_{i+1} - \mathbf{x}_{m,i+1}\|_\infty < \epsilon$

9: **end while**

10: $\mathbf{x}_{m+1} = \mathbf{x}_{m,I_{\max}}$.

2.2. Blur Kernel Estimation

In the blur kernel estimation step, given the recovered sharp edge $\mathbf{x}_{m+1}$, the blur kernel $\mathbf{k}$ in (3) at the $(m+1)$-th outer iteration can be estimated by considering the following minimization problem:

$$\mathbf{k}_{m+1} = \underset{\mathbf{k}}{\operatorname{argmin}} \left\{ \frac{1}{2} \|\mathcal{X}_{m+1}\mathbf{k} - \mathbf{y}\|_2^2 + \eta_1 \|\mathbf{k}\|_1 + \eta_2 \|\nabla\mathbf{k}\|_2^2 \right\}. \tag{14}$$

where the optimal parameters η_1 and η_2 are manually selected by extensive experiments. It is obvious that Model (14) is essentially a convex optimization problem. Analogous to the optimization of problem (6), ADMM can also be adopted to efficiently solve (14) in our experiments. We first replace $\mathbf{k}$ by $\mathbf{h}$ and then obtain the corresponding augmented Lagrangian function $\mathcal{L}_{\mathcal{A}} \left(\mathbf{h}, \mathbf{k}; \varphi^{\mathbf{k}} \right)$ as follows:

$$\mathcal{L}_{\mathcal{A}} \left(\mathbf{h}, \mathbf{k}; \varphi^{\mathbf{k}} \right) = \frac{1}{2} \|\mathcal{X}_{m+1}\mathbf{k} - \mathbf{y}\|_2^2 + \eta_1 \|\mathbf{h}\|_1 + \eta_2 \|\nabla\mathbf{k}\|_2^2 + \frac{\beta_2}{2} \|\mathbf{h} - \mathbf{k}\|_2^2 - \langle \varphi^{\mathbf{k}}, \mathbf{h} - \mathbf{k} \rangle, \tag{15}$$

where β_2 is a pre-defined penalty parameter and $\varphi^{\mathbf{k}}$ denotes the Lagrangian multiplier. Given the fixed $\mathbf{k}_{m,j}$, the minimization of $\mathcal{L}_{\mathcal{A}} \left(\mathbf{h}, \mathbf{k}; \varphi^{\mathbf{k}} \right)$ with respect to $\mathbf{h}$ could be easily handled through the widely-used **shrinkage** operator [49–51], which operates pointwise on scalars or matrices. The solution $\mathbf{h}_{j+1}$ at the $(j+1)$-th inner iteration is given by:

$$\begin{aligned} \mathbf{h}_{j+1} &= \underset{\mathbf{h}}{\operatorname{argmin}} \left\{ \eta_1 \|\mathbf{h}\|_1 + \frac{\beta_2}{2} \left\| \mathbf{h} - \left(\mathbf{k}_{m,j} + \frac{\varphi_j^{\mathbf{k}}}{\beta_2} \right) \right\|_2^2 \right\} \\ &= \max \left\{ \left| \mathbf{k}_{m,j} + \frac{\varphi_j^{\mathbf{k}}}{\beta_2} \right| - \frac{\eta_1}{\beta_2}, 0 \right\} \circ \mathbf{sign} \left(\mathbf{k}_{m,j} + \frac{\varphi_j^{\mathbf{k}}}{\beta_2} \right), \end{aligned} \tag{16}$$

with $\mathbf{k}_{m,0} = \mathbf{k}_m$. Here, the sign function $\mathbf{sign}\,(\cdot)$ is defined as:

$$\mathbf{sign}\,(s) = \begin{cases} 1 & : s > 0, \\ 0 & : s = 0, \\ -1 & : s < 0. \end{cases}$$

Given the fixed $\mathbf{h}_{j+1}$, the minimization of $\mathcal{L}_{\mathcal{A}}\left(\mathbf{h}, \mathbf{k}; \varphi^{\mathbf{k}}\right)$ with respect to $\mathbf{k}$ is equivalent to solving a least-squares optimization problem. Analogous to solving the **x**-subproblem in (9), the corresponding solution $\mathbf{k}_{m,j+1}$ is obtained as follows:

$$\mathbf{k}_{m,j+1} = \mathcal{F}^{-1}\left(\frac{\overline{\mathcal{F}(\mathcal{X}_{m+1})}\mathcal{F}(\mathbf{y}) + \beta_2\mathcal{F}\left(\mathbf{h}_{j+1} - \varphi_j^{\mathbf{k}}/\beta_2\right)}{2\eta_2\overline{\mathcal{F}(\nabla)}\mathcal{F}(\nabla) + \overline{\mathcal{F}(\mathcal{X}_{m+1})}\mathcal{F}(\mathcal{X}_{m+1}) + \beta_2\mathcal{F}(\mathbf{I})}\right). \tag{17}$$

It is tractable to obtain the efficient solution $\mathbf{k}_{m,j+1}$ in (17) using one forward and one inverse FFT. At each iteration, the Lagrangian multiplier $\varphi^{\mathbf{k}}$ can be updated as follows:

$$\varphi_{j+1}^{\mathbf{k}} = \varphi_j^{\mathbf{k}} - \tau\beta_2\left(\mathbf{h}_{j+1} - \mathbf{k}_{m,j+1}\right), \tag{18}$$

for $j = 0, 1, \cdots, J_{\max}$ with $J_{\max}$ denoting the maximum number of inner iterations. The estimated blur kernel $\mathbf{k}_{m+1} = \mathbf{k}_{m,J_{\max}}$ can be achieved for sharp edge restoration in the next step. Note that the convergence of our proposed numerical algorithm can be guaranteed according to the existing convergence results for ADMM in the literature [43,52,53]. Finally, our proposed hybrid regularized variational model for blur kernel estimation is summarized in Algorithm 2.

Algorithm 2 Hybrid regularized blur kernel estimation.

1: **Input**: Blurred image gradient $\mathbf{y}$, $\tau = 1.618$, $\gamma = 5 \times 10^{-2}$, $\eta_1 = \eta_2 = 10^{-3}$, and $M_{\max} = 15$.
2: **Initialize**: $\mathbf{k}_0 = \texttt{uniform}$ and $m = 0$.
3: **while** (not converged and $m \leq M_{\max}$) **do**
　　// Step 1 : *Sharp Edges Restoration* $\mathbf{x}_{m+1}$
4:　　　Update $\mathbf{x}_{m+1}$ by Algorithm 1.
　　// Step 2 : *Blur Kernel Estimation* $\mathbf{k}_{m+1}$
5:　　　$\mathbf{k}_{m,0} = \mathbf{k}_m$.
6:　　　**for** $j = 0$ to $J_{\max}$ **do**
7:　　　　　Compute $\mathbf{h}_{j+1}$ according to (16).
8:　　　　　Compute $\mathbf{k}_{m,j+1}$ according to (17).
9:　　　　　Update Lagrangian multiplier: $\varphi_{j+1}^{\mathbf{k}} = \varphi_j^{\mathbf{k}} - \tau\beta_2\left(\mathbf{h}_{j+1} - \mathbf{k}_{m,j+1}\right)$.
10:　　　**end for**
11:　　　$\mathbf{k}_{m+1} = \mathbf{k}_{m,J_{\max}}$, $\gamma \leftarrow 0.5\gamma$.
12: **end while**
13: **Output**: Estimated blur kernel $\mathbf{k}$.

---

3. Robust Non-Blind Deconvolution

This section mainly focuses on developing a high-order variational model for robust non-blind deconvolution. For the sake of better writing, the original image degradation model $B = L \otimes k + \xi$ in (2) can be rewritten as follows:

$$\mathbf{B} = \mathcal{K}\mathbf{L} + \xi. \tag{19}$$

Once the blur kernel $\mathcal{K}$ (i.e., $\mathbf{k}$ in Algorithm 2) is estimated accurately, blind deblurring can be simplified to the non-blind deconvolution problem. As a consequence, this problem could be handled through the commonly-used regularization methods. One of the most famous methods is the TV-regularized deconvolution method [23–25]. However, the undesirable staircase-like artifacts generated in restored images often lead to significant degradation of visual image quality. In order to enhance the imaging performance, a robust non-blind image deconvolution method will be proposed based on the second-order regularizer TGV^2 [32]. Recently, TGV^2 has been successfully utilized as a regularization scheme in various practical applications [33,54–56] and outperforms the popular TV regularizer. In particular, TGV^2 is capable of preserving image edges and suppressing undesirable artifacts.

Inspired by the work in [11], the L_1-norm data-fidelity term will be incorporated into our TGV^2-regularized non-blind deconvolution to suppress the potential outliers. The used assumption behind the squared L_2-norm data-fidelity term is that the data-fidelity costs follow a Gaussian distribution. This assumption often fails because the squared L_2-norm could make the restored images vulnerable to undesirable outliers. In contrast, the L_1-norm introduced in this work is much more robust to the presence of outliers compared with the squared L_2-norm. The proposed non-blind deconvolution model L_1-TGV^2 is given by:

$$\mathbf{L}^* = \underset{\mathbf{L}}{\operatorname{argmin}} \left\{ \|\mathcal{K}\mathbf{L} - \mathbf{B}\|_1 + \lambda \text{TGV}_\alpha^2 (\mathbf{L}) \right\}, \tag{20}$$

where $\lambda > 0$ denotes a predefined regularization parameter. For a scalar field $\mathbf{L} \in \mathcal{L}^1(\Omega)$, the discretized $\text{TGV}_\alpha^2(L)$ [32] is defined as follows:

$$\text{TGV}_\alpha^2(\mathbf{L}) = \underset{\mathbf{V} \in \mathcal{C}_c^2(\Omega, R^2)}{\operatorname{argmin}} \left\{ \alpha_1 \|\nabla \mathbf{L} - \mathbf{V}\|_1 + \alpha_0 \|\mathcal{E}(\mathbf{V})\|_1 \right\}, \tag{21}$$

where α_1 and α_0 are positive tuning parameters, $\mathcal{C}_c^2(\Omega, R^2)$ denotes the space of the vector field and $\mathcal{E}(\mathbf{V}) = \frac{1}{2}(\nabla \mathbf{V} + \nabla \mathbf{V}^T)$ with $\mathbf{V} = [\mathbf{V}_1 \, \mathbf{V}_2]^T$ being the symmetrized gradient of a complex-valued vector field $\mathbf{V}$. Due to the non-smooth nature of the TGV^2 regularizer, in this paper, we propose to develop an ADMM-based numerical algorithm to effectively solve the non-smooth optimization problem (20). Three auxiliary variables $\mathbf{W}$, $\mathbf{Y}$ and $\mathbf{Z}$ are first introduced, and (20) is then transformed into the following constrained minimization problem:

$$\min_{\mathbf{W}, \mathbf{Y}, \mathbf{Z}, \mathbf{L}, \mathbf{V}} \left\{ \|\mathbf{W}\|_1 + \lambda \left(\alpha_1 \|\mathbf{Y}\|_1 + \alpha_0 \|\mathbf{Z}\|_1 \right) \right\}$$
$$\text{s.t.} \quad \mathbf{W} = \mathcal{K}\mathbf{L} - \mathbf{B}, \ \mathbf{Y} = \nabla \mathbf{L} - \mathbf{V}, \ \mathbf{Z} = \mathcal{E}(\mathbf{V}). \tag{22}$$

It is obvious that $\mathbf{L}$ and $\mathbf{V}$ are coupled together. For the fixed values of $\mathbf{L}$ and $\mathbf{V}$, the updates of $\mathbf{W}$, $\mathbf{Y}$ and $\mathbf{Z}$ are independent of each other. Thus, the variables $\mathbf{W}$, $\mathbf{Y}$, $\mathbf{Z}$, $\mathbf{L}$ and $\mathbf{V}$ can be decomposed into two blocks, i.e., $(\mathbf{W}, \mathbf{Y}, \mathbf{Z})$ and $(\mathbf{L}, \mathbf{V})$. Let $\mathcal{L}_\mathcal{A}(\mathbf{W}, \mathbf{Y}, \mathbf{Z}, \mathbf{L}, \mathbf{V}; \xi, \zeta, \eta)$ denote the augmented Lagrangian function of (22), which can be defined as follows:

$$\begin{aligned}
\mathcal{L}_\mathcal{A}(\mathbf{W}, \mathbf{Y}, \mathbf{Z}, \mathbf{L}, \mathbf{V}; \xi, \zeta, \eta) \ &= \|\mathbf{W}\|_1 + \frac{\rho_1}{2} \left\| \mathbf{W} - (\mathcal{K}\mathbf{L} - \mathbf{B}) - \frac{\xi}{\rho_1} \right\|_2^2 \\
&+ \lambda \alpha_1 \|\mathbf{Y}\|_1 + \frac{\rho_2}{2} \left\| \mathbf{Y} - (\nabla \mathbf{L} - \mathbf{V}) - \frac{\zeta}{\rho_2} \right\|_2^2 + \lambda \alpha_0 \|\mathbf{Z}\|_1 + \frac{\rho_3}{2} \left\| \mathbf{Z} - \mathcal{E}(\mathbf{V}) - \frac{\eta}{\rho_3} \right\|_2^2
\end{aligned} \tag{23}$$

where $\xi \in \mathcal{R}^{mn}$, $\zeta \in \mathcal{R}^{2mn}$ and $\eta \in \mathcal{R}^{4mn}$ denote the Lagrange multipliers and ρ_1, ρ_2 and ρ_3 represent the positive penalty parameters that control the weights of penalty terms. It is numerically intractable to directly obtain the solutions of (22) through commonly-used methods [57]. In order to guarantee a stable solution, it is necessary to alternatively solve the $\mathbf{W}$-, $\mathbf{Y}$-, $\mathbf{Z}$-, $\mathbf{L}$- and $\mathbf{V}$-subproblems and then update the Lagrange multipliers (i.e., ξ, ζ and η) until the obtained solution meets the predefined

threshold. In particular, each of these subproblems has a closed-form solution or can be efficiently solved using the existing simple numerical method.

3.1. $\{\mathbf{W}, \mathbf{Y}, \mathbf{Z}\}$-Subproblems

Given the fixed values of variables $\mathbf{L}^t$, $\mathbf{V}^t$, ξ^t, ζ^t and η^t, the $\mathbf{W}$-subproblem, the $\mathbf{Y}$-subproblem and the $\mathbf{Z}$-subproblem in Equation (23) can be efficiently solved by considering the following L_1-regularized least-squares minimization problems:

$$\mathbf{W}^{t+1} = \underset{\mathbf{W}}{\operatorname{argmin}} \left\{ \|\mathbf{W}\|_1 + \frac{\rho_1}{2} \left\| \mathbf{W} - \left(\mathcal{K} \mathbf{L}^t - \mathbf{B} + \frac{\xi^t}{\rho_1} \right) \right\|_2^2 \right\}, \tag{24}$$

$$\mathbf{Y}^{t+1} = \underset{\mathbf{Y}}{\operatorname{argmin}} \left\{ \lambda \alpha_1 \|\mathbf{Y}\|_1 + \frac{\rho_2}{2} \left\| \mathbf{Y} - \left(\nabla \mathbf{L}^t - \mathbf{V}^t + \frac{\zeta^t}{\rho_2} \right) \right\|_2^2 \right\}, \tag{25}$$

$$\mathbf{Z}^{t+1} = \underset{\mathbf{Z}}{\operatorname{argmin}} \left\{ \lambda \alpha_0 \|\mathbf{Z}\|_1 + \frac{\rho_3}{2} \left\| \mathbf{Z} - \left(\mathcal{E}\left(\mathbf{V}^t\right) + \frac{\eta^t}{\rho_3} \right) \right\|_2^2 \right\}. \tag{26}$$

Note that the unknown variables $\mathbf{W}$, $\mathbf{Y}$ and $\mathbf{Z}$ are componentwise separable in the $\{\mathbf{W}, \mathbf{Y}, \mathbf{Z}\}$-subproblems (24)–(26). These subproblem can be effectively dealt with through the commonly-used **shrinkage** operator [49,50]. This operator is fast and easy to implement in practice. The solutions $\mathbf{W}^{t+1}$, $\mathbf{Y}^{t+1}$ and $\mathbf{Z}^{t+1}$ are obtained as follows:

$$\mathbf{W}^{t+1} = \max \left\{ \left| \mathcal{K} \mathbf{L}^t - \mathbf{B} + \frac{\xi^t}{\rho_1} \right| - \frac{1}{\rho_1}, 0 \right\} \circ \mathbf{sign}\left(\mathcal{K} \mathbf{L}^t - \mathbf{B} + \frac{\xi^t}{\rho_1} \right), \tag{27}$$

$$\mathbf{Y}^{t+1} = \max \left\{ \left| \nabla \mathbf{L}^t - \mathbf{V}^t + \frac{\zeta^t}{\rho_2} \right| - \frac{\lambda \alpha_1}{\rho_2}, 0 \right\} \circ \mathbf{sign}\left(\nabla \mathbf{L}^t - \mathbf{V}^t + \frac{\zeta^t}{\rho_2} \right), \tag{28}$$

$$\mathbf{Z}^{t+1} = \max \left\{ \left| \mathcal{E}\left(\mathbf{V}^t\right) + \frac{\eta^t}{\rho_3} \right| - \frac{\lambda \alpha_0}{\rho_3}, 0 \right\} \circ \mathbf{sign}\left(\mathcal{E}\left(\mathbf{V}^t\right) + \frac{\eta^t}{\rho_3} \right), \tag{29}$$

3.2. $(\mathbf{L}, \mathbf{V})$-Subproblems

The minimization with respect to $(\mathbf{L}, \mathbf{V})$ in (23) is essentially a least-squares optimization problem. However, it is impossible to directly obtain the solutions $\mathbf{L}^{t+1}$ and $\mathbf{V}^{t+1}$ through the forward and inverse FFT operators because the updates of $\mathbf{L}$ and $\mathbf{V}$ are coupled to each other. To guarantee the solution stability, the minimizations with respect to both $\mathbf{L}^{t+1}$ and $\mathbf{V}^{t+1}$ should be simultaneously implemented. Given the fixed values of $\mathbf{W}^{t+1}$, $\mathbf{Y}^{t+1}$, $\mathbf{Z}^{t+1}$, ξ^t, ζ^t and η^t, the coupled $(\mathbf{L}, \mathbf{V})$-subproblem in the augmented Lagrangian function (23) is quadratic, resulting in the following system of linear equations:

$$\begin{cases} \mathbf{L}^{t+1} = \underset{\mathbf{L}}{\operatorname{argmin}} \left\{ \frac{\rho_1}{2} \left\| \mathcal{K} \mathbf{L} - \left(\mathbf{W}^{t+1} + \mathbf{B} - \frac{\xi^t}{\rho_1} \right) \right\|_2^2 + \frac{\rho_2}{2} \left\| \nabla \mathbf{L} - \left(\mathbf{Y}^{t+1} + \mathbf{V} - \frac{\zeta^t}{\rho_2} \right) \right\|_2^2 \right\}, \\ \mathbf{V}^{t+1} = \underset{\mathbf{V}}{\operatorname{argmin}} \left\{ \frac{\rho_2}{2} \left\| \mathbf{V} - \left(\nabla \mathbf{L} - \mathbf{Y}^{t+1} + \frac{\zeta^t}{\rho_2} \right) \right\|_2^2 + \frac{\rho_3}{2} \left\| \mathcal{E}\left(\mathbf{V}\right) - \left(\mathbf{Z}^{t+1} - \frac{\eta^t}{\rho_3} \right) \right\|_2^2 \right\}. \end{cases} \tag{30}$$

Instead of directly solving the system of linear equations (30), we tend to solve the corresponding first-order necessary optimality conditions as follows:

$$
\begin{cases}
\left(\rho_1 \mathcal{K}^\top \mathcal{K} + \rho_2 \nabla^\top \nabla \right) \mathbf{L} - \rho_2 \partial_x^\top \mathbf{V}_1 - \rho_2 \partial_y^\top \mathbf{V}_2 \\
\quad - \left[\rho_1 \mathcal{K}^\top \left(\mathbf{W}^{t+1} + \mathbf{B} - \dfrac{\xi^t}{\rho_1} \right) + \rho_2 \nabla^\top \left(\mathbf{Y}^{t+1} - \dfrac{\zeta^t}{\rho_2} \right) \right] = 0, \\
- \rho_2 \partial_x \mathbf{L} + \left(\rho_2 \mathbf{I} + \rho_3 \partial_x^\top \partial_x + \dfrac{\rho_3}{2} \partial_y^\top \partial_y \right) \mathbf{V}_1 + \dfrac{\rho_3}{2} \partial_y^\top \partial_x \mathbf{V}_2 \\
\quad - \left[\rho_2 \left(\dfrac{\zeta_1^t}{\rho_2} - \mathbf{Y}_1^{t+1} \right) + \rho_3 \left(\partial_x^\top \left(\mathbf{Z}_1^{t+1} - \dfrac{\eta_1^t}{\rho_3} \right) + \partial_y^\top \left(\mathbf{Z}_3^{t+1} - \dfrac{\eta_3^t}{\rho_3} \right) \right) \right] = 0, \\
- \rho_2 \partial_y \mathbf{L} + \dfrac{\rho_3}{2} \partial_x^\top \partial_y \mathbf{V}_1 + \left(\rho_2 \mathbf{I} + \rho_3 \partial_y^\top \partial_y + \dfrac{\rho_3}{2} \partial_x^\top \partial_x \right) \mathbf{V}_2 \\
\quad - \left[\rho_2 \left(\dfrac{\zeta_2^t}{\rho_2} - \mathbf{Y}_2^{t+1} \right) + \rho_3 \left(\partial_y^\top \left(\mathbf{Z}_2^{t+1} - \dfrac{\eta_2^t}{\rho_3} \right) + \partial_x^\top \left(\mathbf{Z}_3^{t+1} - \dfrac{\eta_3^t}{\rho_3} \right) \right) \right] = 0.
\end{cases}
\tag{31}
$$

For the sake of better reading, the original system of linear equations (31) can be rewritten as follows:

$$
\begin{bmatrix}
\mathbf{R}_1 & \mathbf{R}_4^\top & \mathbf{R}_5^\top \\
\mathbf{R}_4 & \mathbf{R}_2 & \mathbf{R}_6^\top \\
\mathbf{R}_5 & \mathbf{R}_6 & \mathbf{R}_3
\end{bmatrix}
\begin{bmatrix}
\mathbf{L} \\
\mathbf{V}_1 \\
\mathbf{V}_2
\end{bmatrix}
=
\begin{bmatrix}
\mathbf{D}_1 \\
\mathbf{D}_2 \\
\mathbf{D}_3
\end{bmatrix},
\tag{32}
$$

with:

$$
\begin{cases}
\mathbf{R}_1 = \rho_1 \mathcal{K}^\top \mathcal{K} + \rho_2 \nabla^\top \nabla, \\
\mathbf{R}_2 = \rho_2 \mathbf{I} + \rho_3 \partial_x^\top \partial_x + \dfrac{1}{2} \rho_3 \partial_y^\top \partial_y, \\
\mathbf{R}_3 = \rho_2 \mathbf{I} + \rho_3 \partial_y^\top \partial_y + \dfrac{1}{2} \rho_3 \partial_x^\top \partial_x, \\
(\mathbf{R}_4, \mathbf{R}_5, \mathbf{R}_6) = \left(-\rho_2 \partial_x, -\rho_2 \partial_y, \dfrac{1}{2} \rho_3 \partial_x^\top \partial_y \right),
\end{cases}
$$

and:

$$
\begin{cases}
\mathbf{D}_1 = \rho_1 \mathcal{K}^\top \left(\mathbf{W}^{t+1} + \mathbf{B} - \dfrac{\xi^t}{\rho_1} \right) + \rho_2 \nabla^\top \left(\mathbf{Y}^{t+1} - \dfrac{\zeta^t}{\rho_2} \right), \\
\mathbf{D}_2 = \rho_2 \left(\dfrac{\zeta_1^t}{\rho_2} - \mathbf{Y}_1^{t+1} \right) + \rho_3 \left(\partial_x^\top \left(\mathbf{Z}_1^{t+1} - \dfrac{\eta_1^t}{\rho_3} \right) + \partial_y^\top \left(\mathbf{Z}_3^{t+1} - \dfrac{\eta_3^t}{\rho_3} \right) \right), \\
\mathbf{D}_3 = \rho_2 \left(\dfrac{\zeta_2^t}{\rho_2} - \mathbf{Y}_2^{t+1} \right) + \rho_3 \left(\partial_y^\top \left(\mathbf{Z}_2^{t+1} - \dfrac{\eta_2^t}{\rho_3} \right) + \partial_x^\top \left(\mathbf{Z}_3^{t+1} - \dfrac{\eta_3^t}{\rho_3} \right) \right).
\end{cases}
$$

Let $\mathcal{F}$ denote the discrete Fourier transform operator for real (complex) matrices or vectors. To efficiently solve the system of linear equations (32), we multiply both sides of (32) by $\mathcal{F}$, such that the coefficient matrix will be blockwise diagonal, i.e.,

$$
\begin{bmatrix}
\mathcal{F}(\mathbf{R}_1) & \mathcal{F}(\mathbf{R}_4)^\top & \mathcal{F}(\mathbf{R}_5)^\top \\
\mathcal{F}(\mathbf{R}_4) & \mathcal{F}(\mathbf{R}_2) & \mathcal{F}(\mathbf{R}_6)^\top \\
\mathcal{F}(\mathbf{R}_5) & \mathcal{F}(\mathbf{R}_6) & \mathcal{F}(\mathbf{R}_3)
\end{bmatrix}
\begin{bmatrix}
\mathcal{F}(\mathbf{L}) \\
\mathcal{F}(\mathbf{V}_1) \\
\mathcal{F}(\mathbf{V}_2)
\end{bmatrix}
=
\begin{bmatrix}
\mathcal{F}(\mathbf{D}_1) \\
\mathcal{F}(\mathbf{D}_2) \\
\mathcal{F}(\mathbf{D}_3)
\end{bmatrix}.
\tag{33}
$$

Essentially, (33) is a linear system with three equations and three variables. We propose to directly use Cramer's rule to effectively yield the closed-form solutions as follows:

$$
\mathbf{L}^{t+1} = \mathcal{F}^{-1} \left(\dfrac{\mathbf{det_L}}{\mathbf{det_T}} \right), \quad \mathbf{V}_1^{t+1} = \mathcal{F}^{-1} \left(\dfrac{\mathbf{det_{V_1}}}{\mathbf{det_T}} \right) \text{ and } \mathbf{V}_2^{t+1} = \mathcal{F}^{-1} \left(\dfrac{\mathbf{det_{V_2}}}{\mathbf{det_T}} \right),
\tag{34}
$$

where $\mathcal{F}^{-1}(\circ)$ represents the inverse Fourier transform operator. In particular, we have the determinants $\mathbf{det_L} = \begin{vmatrix} \mathbf{S} & \mathbf{T}_2 & \mathbf{T}_3 \end{vmatrix}_*$, $\mathbf{det_{V_1}} = \begin{vmatrix} \mathbf{T}_1 & \mathbf{S} & \mathbf{T}_3 \end{vmatrix}_*$, $\mathbf{det_{V_2}} = \begin{vmatrix} \mathbf{T}_1 & \mathbf{T}_2 & \mathbf{S} \end{vmatrix}_*$ and $\mathbf{det_T} = \begin{vmatrix} \mathbf{T}_1 & \mathbf{T}_2 & \mathbf{T}_3 \end{vmatrix}_*$ with $\mathbf{T}_1 = \left[\mathcal{F}(\mathbf{R}_1)^\top \ \mathcal{F}(\mathbf{R}_4)^\top \ \mathcal{F}(\mathbf{R}_5)^\top \right]^\top$, $\mathbf{T}_2 = \left[\mathcal{F}(\mathbf{R}_4) \ \mathcal{F}(\mathbf{R}_2)^\top \ \mathcal{F}(\mathbf{R}_6)^\top \right]^\top$,

$\mathbf{T}_3 = \left[\mathcal{F}(\mathbf{R}_5) \; \mathcal{F}(\mathbf{R}_6) \; \mathcal{F}(\mathbf{R}_3)^\top \right]^\top$ and $\mathbf{S} = \left[\mathcal{F}(\mathbf{D}_1)^\top \; \mathcal{F}(\mathbf{D}_2)^\top \; \mathcal{F}(\mathbf{D}_3)^\top \right]^\top$. The definition of determinant $| \cdot |_*$ in this work is briefly introduced as follows:

$$\begin{vmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{vmatrix}_* = \begin{bmatrix} r_{11} \circ r_{22} \circ r_{33} + r_{12} \circ r_{23} \circ r_{31} \\ + r_{13} \circ r_{21} \circ r_{32} - r_{13} \circ r_{22} \circ r_{31} \\ - r_{12} \circ r_{21} \circ r_{33} - r_{11} \circ r_{23} \circ r_{32} \end{bmatrix}.$$

To shorten the computational time, all matrix elements in $\mathbf{det_T}$ for (34) should be calculated before the execution of our ADMM-based numerical algorithm. At each iteration, $\mathbf{D}_1$, $\mathbf{D}_2$ and $\mathbf{D}_3$ are first computed; $\mathbf{S} = \left[\mathcal{F}(\mathbf{D}_1)^\top \; \mathcal{F}(\mathbf{D}_2)^\top \; \mathcal{F}(\mathbf{D}_3)^\top \right]^\top$ can be easily calculated correspondingly. The final solutions $\mathbf{L}^{t+1}$, $\mathbf{V}_1^{t+1}$ and $\mathbf{V}_2^{t+1}$ can be naturally obtained using Cramer's rule (34).

3.3. Update the Lagrange Multipliers

At each iteration of our proposed numerical method, the Lagrange multipliers (ξ, ζ, η) should be updated as follows:

$$\xi^{t+1} = \xi^t - \tau \rho_1 \left(\mathbf{W}^{t+1} - \left(\mathcal{K} \mathbf{L}^{t+1} - \mathbf{B} \right) \right), \tag{35}$$

$$\zeta^{t+1} = \zeta^t - \tau \rho_2 \left(\mathbf{Y}^{t+1} - \left(\nabla \mathbf{L}^{t+1} - \mathbf{V}^{t+1} \right) \right), \tag{36}$$

$$\eta^{t+1} = \eta^t - \tau \rho_3 \left(\mathbf{Z}^{t+1} - \mathcal{E}\left(\mathbf{V}^{t+1} \right) \right), \tag{37}$$

where the step length $\tau = 1.618$ is used throughout this paper. In conclusion, an ADMM-based numerical method was proposed to decompose the original complex optimization problem (20) into several simpler subproblems. Each subproblem has a closed-form solution or can be efficiently solved using the existing numerical method. In particular, the $\{\mathbf{W}, \mathbf{Y}, \mathbf{Z}\}$-subproblems (24–26) could be easily solved using the **shrinkage** operator. The solutions of $\mathbf{L}$ and $\mathbf{V}$ were simultaneously obtained through Cramer's rule (34). The optimization procedure of our proposed method for non-blind image deconvolution is summarized in Algorithm 3.

Algorithm 3 ADMM for the L_1-TGV^2 Model (20).

1: **Input**: Blurred image $\mathbf{B}$, blur kernel k (i.e., $\mathcal{K}$ in Section 3), $\rho_1 = 50$, $\rho_2 = 0.5$, $\rho_3 = 5$, $\alpha_1 = 1$, $\alpha_2 = 1.5$, $\tau = 1.618$, $T_{\max} = 10$ and $\epsilon = 5 \times 10^{-5}$.

2: **Initialize**: $\mathbf{L}^0 = \mathbf{B}$, $\mathbf{V}^0 = 0$, $\xi^0 = 0$, $\zeta^0 = 0$, $\eta^0 = 0$ and $t = 0$.

3: **while** (not converged and $t \leq T_{\max}$) **do**

4:　　Compute $\mathbf{W}^{t+1}$ according to $\mathbf{W}^{t+1} = \mathbf{shrinkage}\left(\mathcal{K}\mathbf{L}^t - \mathbf{B} + \xi^t/\rho_1, 1/\rho_1 \right)$.

5:　　Compute $\mathbf{Y}^{t+1}$ according to $\mathbf{Y}^{t+1} = \mathbf{shrinkage}\left(\nabla \mathbf{L}^t - \mathbf{V}^t + \zeta^t/\rho_2, \lambda \alpha_1/\rho_2 \right)$.

6:　　Compute $\mathbf{Z}^{t+1}$ according to $\mathbf{Z}^{t+1} = \mathbf{shrinkage}\left(\mathcal{E}\left(\mathbf{V}^t \right) + \eta^t/\rho_3, \lambda \alpha_0/\rho_3 \right)$.

7:　　Compute $\left(\mathbf{L}^{t+1}, \mathbf{V}^{t+1} \right)$ according to
　　　$\mathbf{L}^{t+1} = \mathcal{F}^{-1}\left(\frac{\mathbf{det_L}}{\mathbf{det_T}} \right)$, $\mathbf{V}_1^{t+1} = \mathcal{F}^{-1}\left(\frac{\mathbf{det_{V_1}}}{\mathbf{det_T}} \right)$ and $\mathbf{V}_2^{t+1} = \mathcal{F}^{-1}\left(\frac{\mathbf{det_{V_2}}}{\mathbf{det_T}} \right)$.

8:　　Update Lagrangian multipliers (ξ, ζ, η):
　　　$\xi^{t+1} = \xi^t - \tau \rho_1 \left(\mathbf{W}^{t+1} - \left(\mathcal{K}\mathbf{L}^{t+1} - \mathbf{B} \right) \right)$.
　　　$\zeta^{t+1} = \zeta^t - \tau \rho_2 \left(\mathbf{Y}^{t+1} - \left(\nabla \mathbf{L}^{t+1} - \mathbf{V}^{t+1} \right) \right)$.
　　　$\eta^{t+1} = \eta^t - \tau \rho_3 \left(\mathbf{Z}^{t+1} - \mathcal{E}\left(\mathbf{V}^{t+1} \right) \right)$.

9:　　Check convergence condition: $\left\| \mathbf{L}^{t+1} - \mathbf{L}^t \right\|_\infty < \epsilon$

10: **end while**

11: Output: Deblurred image $\mathbf{L}$.

---

4. Experimental Results and Discussion

Comprehensive blind deblurring experiments on both synthetic and realistic blurred images will be performed to verify the effectiveness of our proposed method in this section.

4.1. Experimental Settings

The proposed blind deconvolution framework was evaluated on a synthetic blur-image dataset [4] and realistic blurred images. The synthetic dataset has been widely exploited as a benchmark dataset to evaluate the performance of blur kernel estimation. Our numerical experiments were implemented using MATLAB R2011a (The MathWorks, Natick, Inc., MA) on a machine with a 3.30-GHz Intel(R) Pentium(R) G3260 CPU and 4 GB RAM. For both synthetic and realistic datasets, the parameter values for blur kernel estimation in Section 2 were set as follows: $\tau = 1.618$, $\rho = 3$, $I_{max} = 5$, $\gamma = 5 \times 10^{-2}$, $\eta_1 = \eta_2 = 10^{-3}$ and $M_{max} = 15$. The resulting optimal parameters for non-blind deconvolution in Section 3 were set empirically, i.e., $\rho_1 = 50$, $\rho_2 = 0.5$, $\rho_3 = 5$, $\alpha_1 = 1$, $\alpha_2 = 1.5$ and $T_{max} = 10$. The deblurring results have illustrated the satisfactory performance of the manually-selected parameters in our experiments. For the sake of better comparison, the competing blind deblurring methods yield the restoration results with the input parameters manually optimized by the authors. To further improve deblurring performance, there is a great potential to develop automatic estimation methods to adaptively select the optimal parameters in our future work. Similar to [4], the sum of squared differences (SSD) and SSD ratio were used simultaneously to quantitatively evaluate the performance of blur kernel estimation in Section 4.2. In particular, the SSD ratio is measured between the deconvolution error with the estimated blur kernel and the deconvolution error with the ground truth kernel [4].

4.2. Experiments on Synthetically-Blurred Images

Numerous experiments are implemented in this subsection to evaluate the performance of our proposed method on one widely-used synthetic blur image dataset [4], which can be downloaded from the link: www.wisdom.weizmann.ac.il/~levina/papers/LevinEtalCVPR2011Code.zip. In Figure 3, it could be found that the dataset is composed of four grayscale images of the size 256×256 and eight different uniform blur kernels, resulting in a total of 32 synthetic blurred images in our experiments. The proposed method will be compared with several state-of-the-art blind deblurring methods [7,11–13,58] in terms of SSD and the SSD ratio. In order to guarantee an unbiased comparison, the final deblurred results are all generated using the sparse non-blind deconvolution method proposed in [58]. Furthermore, to enhance the robustness of blur kernel estimation, a widely-used multiscale scheme [7] was introduced in Algorithm 2. Experimental results on the SSD ratio for different kernel estimation methods are summarized in Figure 4. It can be found that our proposed method is able to generate more robust estimation results on this synthetic dataset under consideration in most of the cases. In contrast, the accuracy of blur kernel estimation for other competing methods is limited due to the simple assumption of the blur kernel prior.
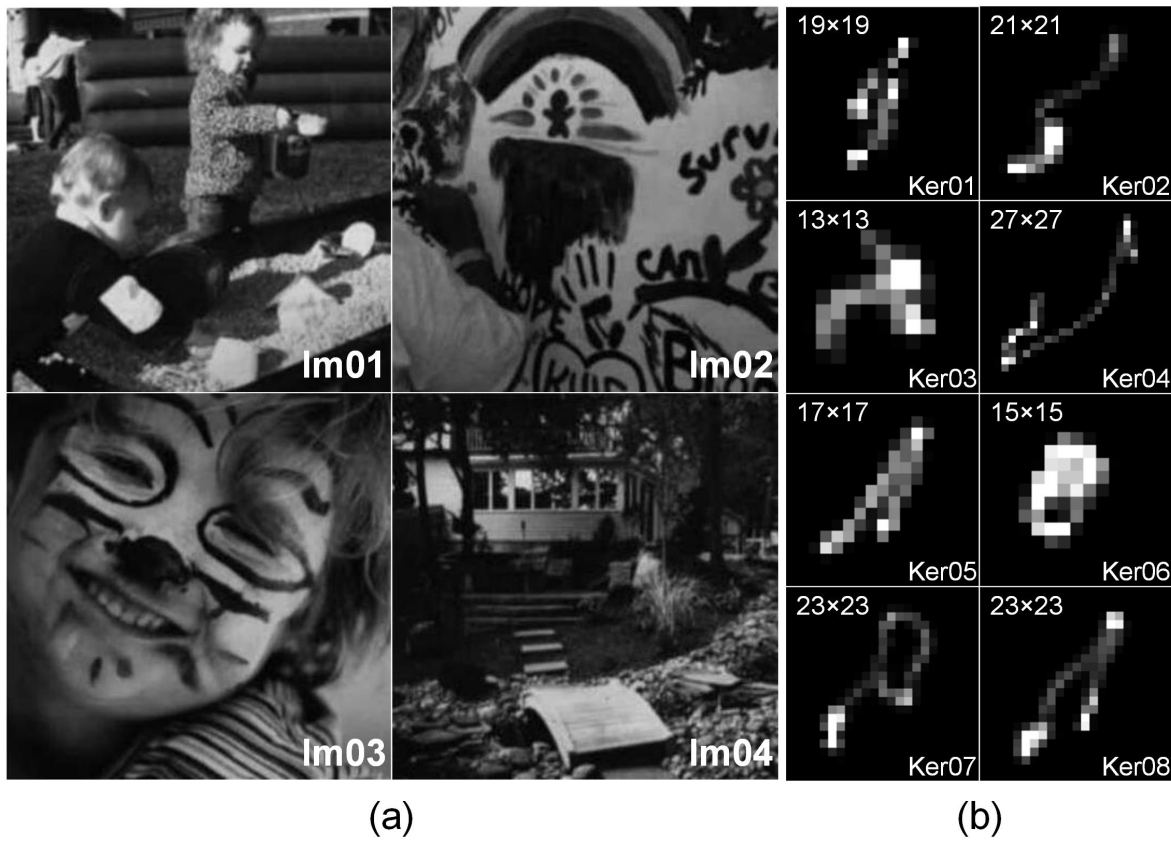
Figure 3. The experimental dataset of [4]. From left to right: (**a**) four gray-scale test images of the size 256 × 256 and (**b**) eight uniform blur kernels of different sizes (the blur kernel sizes are illustrated in the upper-left panels), resulting in 32 test images in our synthetic experiments.
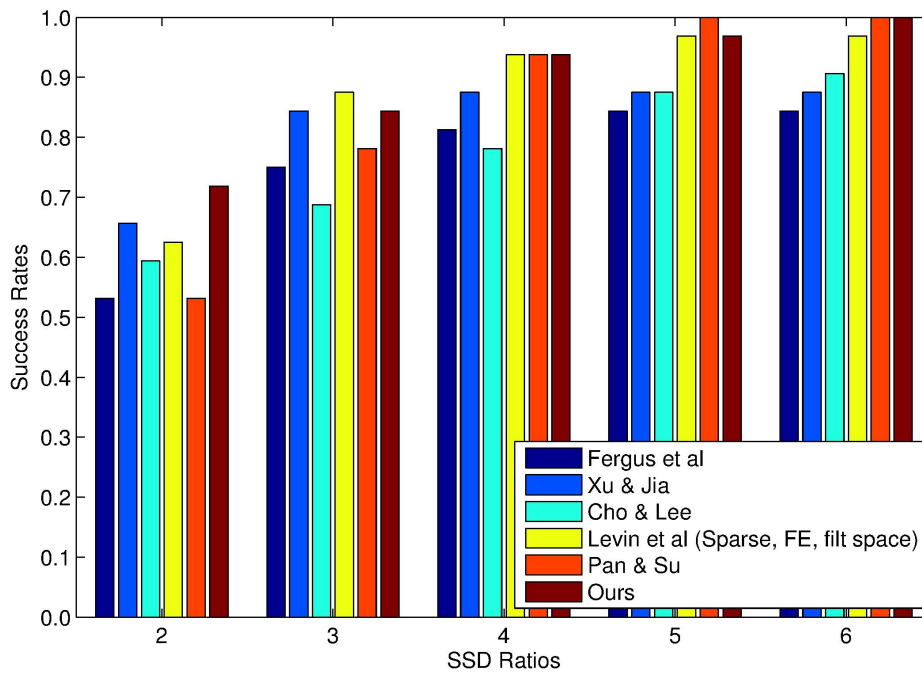

Figure 4. Cumulative histograms of the sum of squared differences (SSD) ratios on the blur-image dataset of Levin et al. [4].

The objective function Equation (3), comprising the L_0 quasi-norm and the L_1-norm regularization terms, is both non-convex and non-smooth. Thus, it is much more difficult to estimate the complexity of the proposed method from a theoretical point of view. It is well known that computational cost is highly dependent on the algorithm complexity. For the sake of simplicity, we only tend to compare the computational time for different blur kernel estimation methods under the same imaging conditions. The methods of Xu and Jia [11] and Cho and Lee [12] are efficiently implemented using C code in our experiments. In contrast, the other competing methods are performed in MATLAB. Since all test images in Figure 3 are mainly composed of fine details of different textures, we take images Im02 and Im04 as examples to evaluate the computational efficiency. The computational time of different competing blur kernel estimation methods is summarized in Table 1. It could be found that the methods of Xu and Jia [11] and Cho and Lee [12] generate the lowest computational cost due to the C implementation. Our proposed method yields significantly faster computational speeds compared with Fergus et al. [7] and Levin et al. [58]. However, the method of Pan and Su [13] achieves the highest computational efficiency due to the fast alternating direction optimization method. As shown in Table 1, the proposed method yields the best evaluation results in terms of the SSD metric under consideration in most of the cases. The satisfactory performance of our proposed method benefits from the combination of the L_1-norm of kernel intensity with the squared L_2-norm of the intensity derivative. It could be further observed that both computational cost and image quality highly depend on the size of the blur kernel. As the size becomes larger, the computational cost obviously becomes higher, and the image quality could becomes lower for all competing methods under different imaging conditions. In addition, there is no significant difference in blind deblurring performance between different test images for the same blurring degradation.

Table 1. Sum of squared differences (SSD)/CPU computational time (unit: seconds) of different blind deblurring methods on one synthetic test image from [4].

Methods	Ker01	Ker02	Ker03	Ker04	Ker05	Ker06	Ker07	Ker08
	Im02							
Fergus et al. [7]	156.1/449.6	181.8/403.7	44.28/409.3	88.89/311.0	69.13/350.5	44.92/377.8	120.6/407.8	77.30/466.8
Xu and Jia [11]	44.45/1.264	82.76/1.263	76.92/1.061	38.64/1.263	138.2/1.248	67.12/1.279	231.6/1.280	81.19/1.264
Cho and Lee [12]	75.82/1.159	103.2/1.163	74.06/1.021	112.1/1.250	70.55/1.140	41.28/1.133	75.42/1.236	130.6/1.245
Pan and Su [13]	42.11/1.474	102.2/1.564	56.28/1.328	132.5/1.559	46.36/1.428	37.04/1.452	99.73/1.673	57.67/1.572
Levin et al. [58]	85.76/115.9	121.8/138.7	37.92/68.14	73.36/299.1	102.7/94.09	83.74/83.95	106.9/185.6	57.93/179.0
Ours	35.20/2.041	74.64/2.008	36.85/1.823	95.50/2.159	34.26/1.884	34.20/2.052	73.17/2.136	42.25/2.128
	Im04							
Fergus et al. [7]	99.55/436.2	162.4/389.7	51.64/418.6	72.67/299.2	47.70/338.7	45.65/374.6	90.47/388.4	65.20/472.4
Xu and Jia [11]	38.43/1.294	93.93/1.248	70.88/1.092	195.7/1.280	60.74/1.263	36.03/1.123	125.3/1.264	69.44/1.263
Cho and Lee [12]	112.0/1.243	113.5/1.130	69.76/1.046	135.2/1.269	116.0/1.223	43.65/1.086	269.0/1.238	123.1/1.242
Pan and Su [13]	45.42/1.494	131.4/1.560	54.48/1.325	127.5/1.572	55.22/1.433	36.64/1.457	125.5/1.579	62.00/1.570
Levin et al. [58]	88.21/108.5	123.3/135.6	39.84/79.05	122.4/284.9	63.71/91.66	63.30/79.67	114.0/169.6	57.70/168.3
Ours	36.43/1.995	93.36/1.970	36.58/1.895	90.51/2.129	39.91/1.867	35.20/1.923	88.96/2.171	45.18/2.163

For a better comparison, Figure 5 visually displays a test image from the synthetic blur-image dataset [4] recovered by different deblurring methods. As can be observed, our proposed method estimates a more accurate blur kernel. The proposed method generates the highest quality of deblurred image, since it achieves a more natural-looking appearance. We can conclude that our proposed method has superior performance compared with other competing methods.
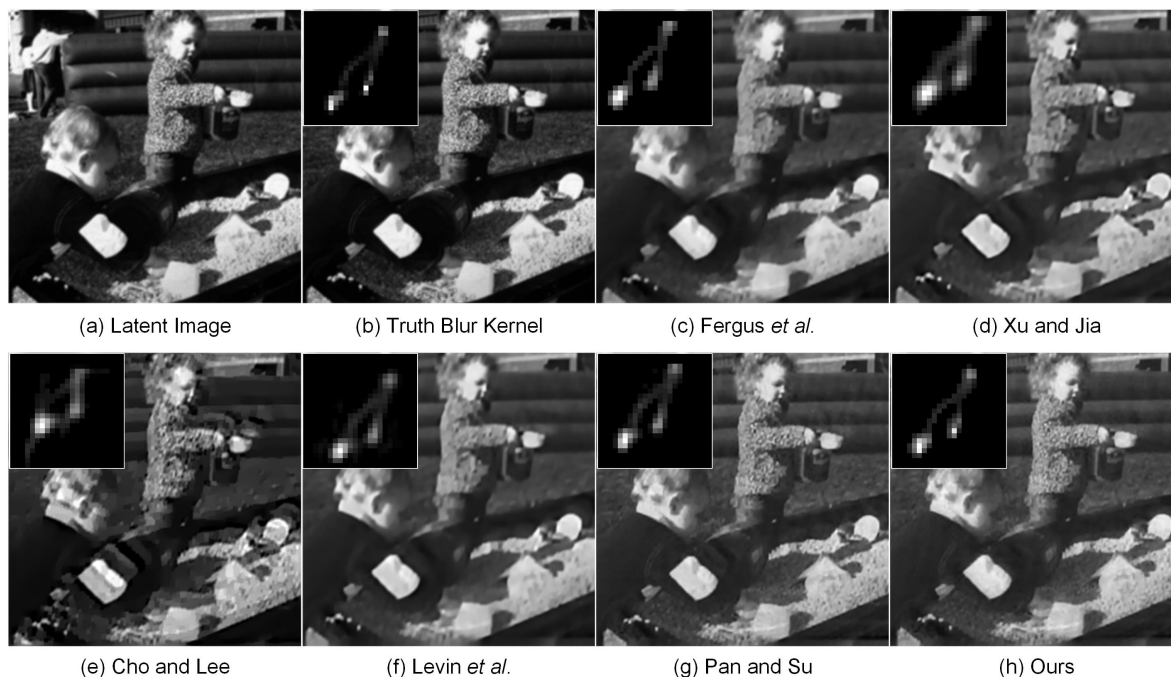
Figure 5. Comparison of results on one test image from [4]. From top-left to bottom-right: (**a**) latent sharp image, deblurred versions with (**b**) the truth blur kernel, estimated blur kernels generated by (**c**) Fergus et al. [7], (**d**) Xu and Jia [11], (**e**) Cho and Lee [12], (**f**) Levin et al. [58], (**g**) Pan and Su [13] and (**h**) our proposed method, respectively. The estimated blur kernels are illustrated in the upper-left panels.

4.3. Experiments on a Large Blur Kernel

Single-image blind deconvolution with a large blur kernel is an extremely challenging problem in practical application. It is necessary to investigate whether our proposed method is able to deal with the large blur kernel. In this subsection, our proposed method will be compared with three state-of-the-art deblurring methods, i.e., Xu and Jia [11] in ECCV-2010, Krishnan et al. [9] in CVPR-2011 and Pan et al. [59] in CVPR-2016. Once the blur kernel is estimated, the method of Xu and Jia [11] implements a robust non-blind deconvolution by integrating the L_1-norm data-fidelity term and the TV regularizer. Krishnan et al. [9] directly uses the hyper-Laplacian prior-based fast non-blind deconvolution method [26] to generate the final recovered image. The non-blind deconvolution results in our proposed method will be achieved using the combination of the L_1-norm data-fidelity term and the TGV^2 regularizer summarized in Algorithm 3. In contrast, the method in Pan et al. [59] simultaneously estimates the latent sharp image and blur kernel by introducing the assumption of the dark channel prior.

The blind deblurring results with large blur kernels are visually displayed in Figures 6 and 7. As shown in Figure 6, the method of Krishnan et al. [9] is unable to yield the satisfactory estimation of the large blur kernel of the size 159×159. The resulting deblurred image suffers from the loss of important geometrical structures, which significantly degrades the visual quality. In contrast, the methods of Xu and Jia [11], Pan et al. [59] and our proposed method have the capacity of guaranteeing the accurate estimation of the blur kernel. The main geometrical structures in the deblurred images could be reconstructed correspondingly. However, an excess smoothing could be observed in the methods of Xu and Jia [11] and Pan et al. [59], which leads to the loss of small structural features. Owing to the second-order regularizer TGV^2, our non-blind deconvolution method produces a much more natural-looking result than Xu and Jia [11] and Pan et al. [59]. As shown by the arrows in Figure 6, more details could be preserved in our proposed method; thus, the resulting deblurring performance outperforms other comparative methods. More blind deblurring results with large blur kernels are

visually illustrated in Figure 7. The sizes of the estimated blur kernels in our experiments are 101×101, 95×95 and 91×91, respectively. It could be found that the method of Krishnan et al. [9] fails to yield the accurate estimation of the blur kernel and generates unsatisfactory deblurring performance. In contrast, the proposed method is able to generate comparable results to the state-of-the-art blind deblurring methods, i.e., Xu and Jia [11] and Pan et al. [59]. The final high-quality recovered images could be achieved with more structures and details preserved. Therefore, there is a huge potential to use our proposed method to restore the blurred images with large blur kernels in practice.

(a) Blurred Image　　(b) Xu and Jia　　(c) Krishnan *et al.*　　(d) Pan *et al.*　　(e) Ours
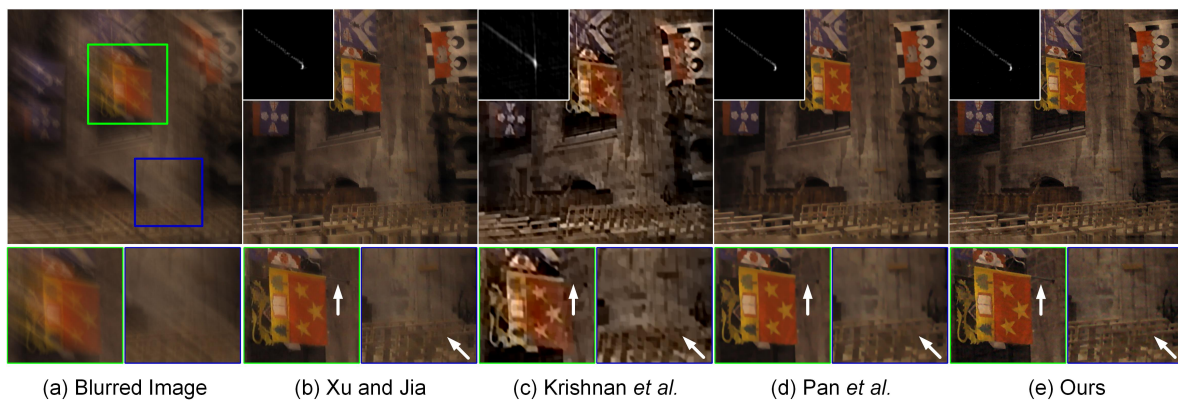
Figure 6. Restoration of a blurred image with a large motion kernel of the size 159×159. From left to right: (**a**) input blurred image, deblurred versions generated by (**b**) Xu and Jia [11], (**c**) Krishnan et al. [9], (**d**) Pan et al. [59] and (**e**) our proposed method, respectively. The estimated blur kernels and local magnification views respectively are illustrated in the upper-left and bottom panels.
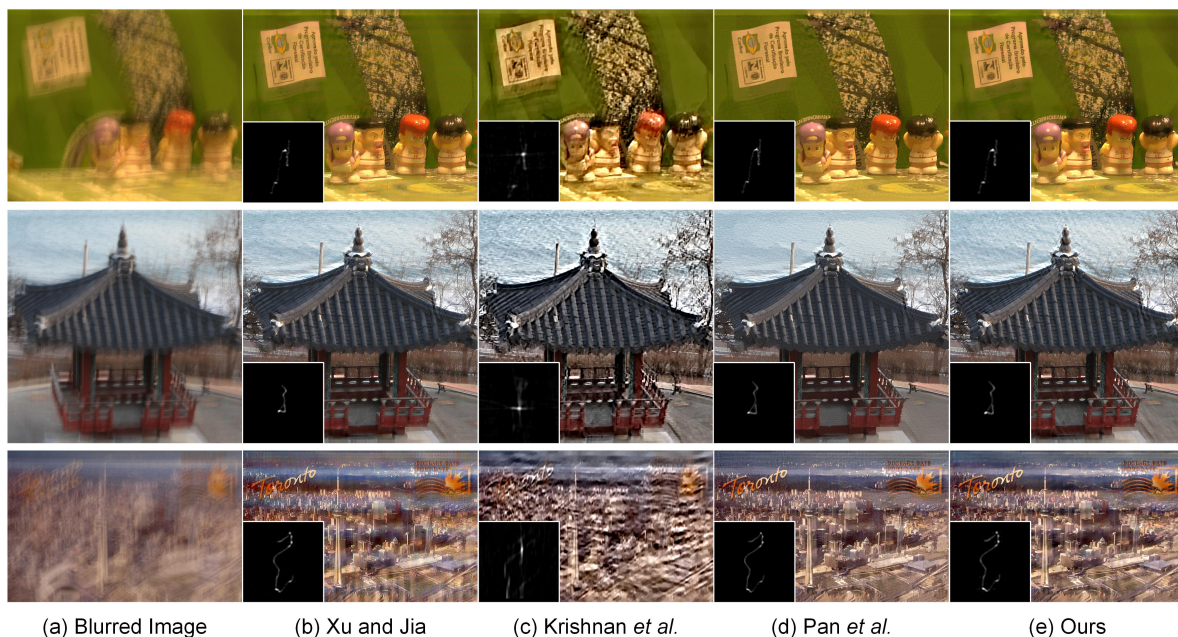

(a) Blurred Image　　(b) Xu and Jia　　(c) Krishnan *et al.*　　(d) Pan *et al.*　　(e) Ours

Figure 7. Blind deconvolution of three different realistic images. From left to right: (**a**) input blurred image, deblurred versions generated by (**b**) Xu and Jia [11], (**c**) Krishnan et al. [9], (**d**) Pan et al. [59] and (**e**) our proposed method, respectively. The sizes of the estimated blur kernels from top to bottom are 101×101, 95×95 and 91×91, respectively.

4.4. Experiments on Ocean Engineering

In the field of ocean engineering, computer vision-assisted automatic detection and tracking systems with airborne and shipborne imaging sensors have been widely used to improve maritime control, safety and rescue operations. However, the resulting imaging performance sometimes suffers from motion blur, noise, haze and sensor nonlinearities, which could significantly degrade the visual image quality under poor weather conditions. In this paper, we mainly focus our attention on the restoration of blurred images, since this degradation condition is more common than other conditions in ocean engineering. The experimental images captured with airborne and shipborne imaging sensor systems, as well as the corresponding blur kernel estimation and image deconvolution results are visually displayed in Figures 8 and 9. The sizes of the estimated blur kernels are 35×35 and 95×95, respectively. As shown in Figure 8, the method of Krishnan et al. [9] is unable to achieve high-quality blur kernel estimation resulting in unsatisfactory deblurring performance with ringing-like artifacts. In contrast, Pan et al. [59] and our proposed method have the capacity of producing accurate estimation of the blur kernel for this example. The final high-quality restored images could be guaranteed using non-blind deconvolution methods. However, due to the low-contrast structure shown in Figure 9, the latest dark channel prior-based method [59] could not accurately estimate the blur kernel in the case of shipborne imaging. The reason behind this phenomenon may be that the statistical properties between images captured by shipborne cameras and natural images are essentially different. The assumption of the dark channel prior is not always valid under different imaging conditions. The method of Xu and Jia [11] also fails to estimate the blur kernel and generates a low-quality restored image. Figure 9 visually illustrates that our proposed method is still able to guarantee accurate kernel estimation and image deconvolution. More geometric structures and fine details could be preserved in our recovered images, beneficial for detecting and tracking moving vessels in practice. The maritime control, safety and rescue operations could be correspondingly improved in the field of maritime management and ocean engineering.
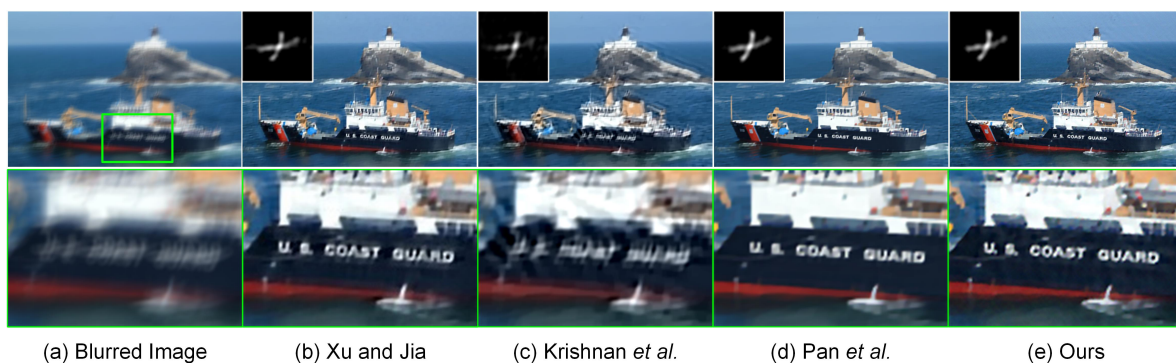

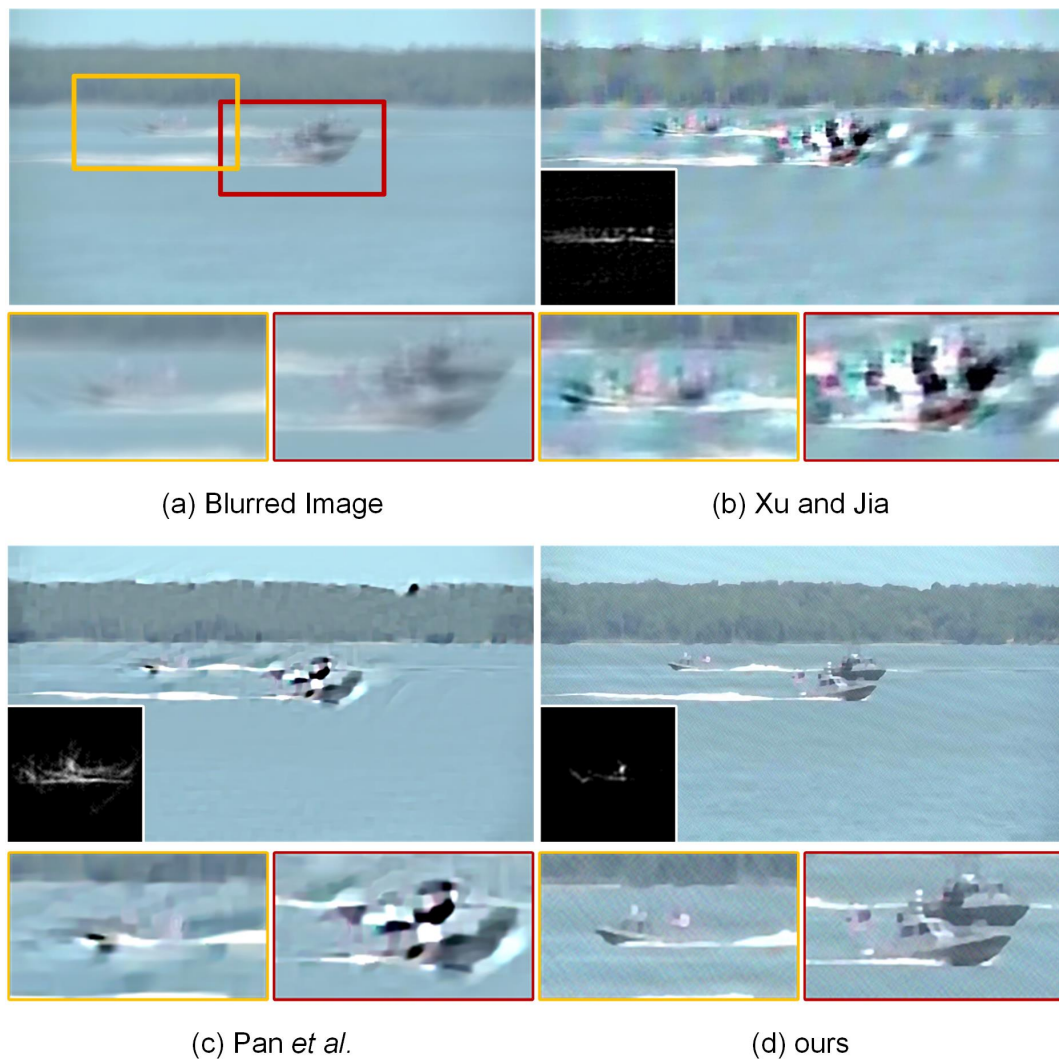| (a) Blurred Image | (b) Xu and Jia | (c) Krishnan *et al.* | (d) Pan *et al.* | (e) Ours |

Figure 8. Restoration of a blurred image captured with an airborne camera. From left to right: (**a**) input blurred image, deblurred versions generated by (**b**) Xu and Jia [11], (**c**) Krishnan et al. [9], (**d**) Pan et al. [59] and (**e**) our proposed method, respectively. The estimated blur kernels of the size 35×35 and local magnification views respectively are illustrated in the upper-left and bottom panels.

Figure 9. Restoration of a blurred image captured with a shipborne camera. From left to right: (**a**) input blurred image, deblurred versions generated by (**b**) Xu and Jia [11], (**c**) Pan et al. [59] and (**d**) our proposed method, respectively. The estimated blur kernels of the size 95×95 and local magnification views respectively are illustrated in the bottom-left and bottom panels.

4.5. Experiments on More Realistic Blurred Images

To better evaluate our proposed method, this subsection is concluded by testing blind deblurring on more realistic human and nature images. Our experimental results will be compared with the recovered results generated by the three above-mentioned methods, i.e., Xu and Jia [11], Krishnan et al. [9] and Pan et al. [59]. The estimation of the blur kernel for each method is implemented directly using the codes and parameter settings provided by the authors. As shown in Figure 10, the recovered result generated by Krishnan et al. [9] suffers from the over-smoothing of detailed texture structures due to the inaccurate estimation of the blur kernel. The loss of geometrical structures easily makes the deblurred image look less natural, resulting in significant visual quality degradation. The local magnification views shown in Figure 10 visually illustrate that the proposed method yields a more "natural-looking" performance compared with Xu and Jia [11] and Pan et al. [59]. In particular, more geometrical details in the face and hand regions could be preserved in our proposed method. The sharp edges, slightly over-smoothed by Xu and Jia [11] and Pan et al. [59], have been reconstructed accurately using our proposed method. Its superior performance benefits from the hybrid blur kernel constraints and edge-preserving TGV^2 regularizer.

(a) Blurred Image   (b) Xu and Jia   (c) Krishnan *et al.*   (d) Pan *et al.*   (e) Ours

Figure 10. Single-image blind deblurring results with a blur kernel of size 27×27 (top) and their local magnification views (middle and bottom). From left to right: (**a**) input blurred images, deblurred versions generated by (**b**) Xu and Jia [11], (**c**) Krishnan et al. [9], (**d**) Pan et al. [59] and (**e**) our proposed method, respectively.

The excellent deblurring performance of our proposed method could also be visually found in Figure 11. As shown in the local magnification views, the inaccurate estimation of the blur kernel for Xu and Jia [11] causes the degraded visual quality with a significant loss of fine details. Krishnan et al. [9] is able to guarantee the quality of blur kernel estimation in this case. However, the final deblurring result tends to be unsatisfactory because the restored "number" could not be preserved correctly, shown in the local magnification views. The main geometrical structures and fine details in the recovered images are preserved by Pan et al. [59] and our proposed method. The ringing-like artifacts could be visually found near the ear region in the deblurring result by Pan et al. [59]. Our proposed method is able to overcome this limitation, but still generates slight ringing-like artifacts in the jaw region. The reason may be that our non-blind deconvolution method summarized in Algorithm 3 is performed with a constant regularization parameter λ. To further enhance the image quality, the regularization parameter should be selected spatially variant to suppress the ringing-like artifacts. More realistic deblurring results on human images are visually displayed in Figure 12. It could be observed that our proposed method yields deblurring results that are visually comparable with the current state-of-the-art methods. Figure 13 illustrates the realistic deblurring results on five different natural images. The sizes of the estimated blur kernels from top to bottom are 35×35, 55×55, 41×41, 35×35 and 55×55, respectively. Since these realistic images contain sufficient textures and geometrical structures, all competing methods have the capacity of accurately estimating the blur kernels in these cases. Therefore, our experimental results are visually comparable to others. The quality of the deblurred images could be correspondingly guaranteed in practical applications.
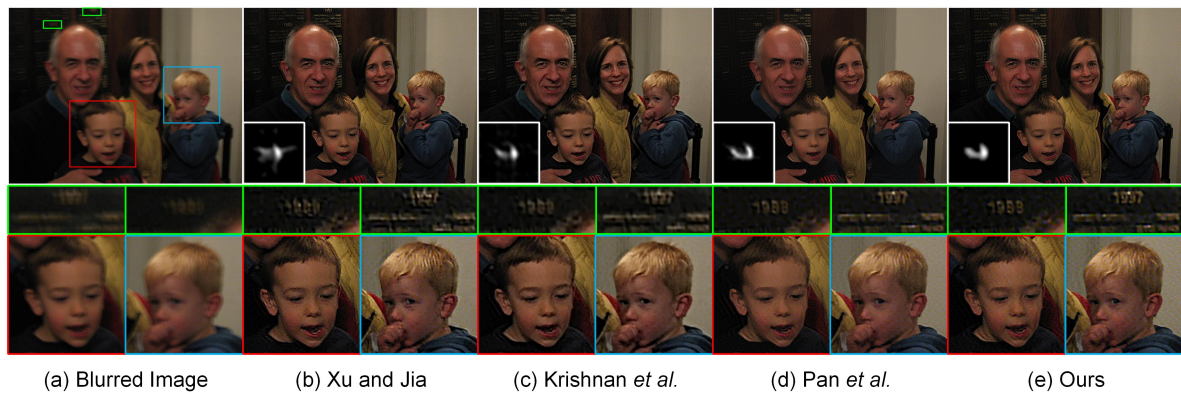
(a) Blurred Image　　　(b) Xu and Jia　　　(c) Krishnan *et al.*　　　(d) Pan *et al.*　　　(e) Ours

Figure 11. Restoration of a blurred image with a blur kernel of the size 27×27 (top) and the corresponding local magnification views (middle and bottom). From left to right: (**a**) input blurred images, deblurred versions generated by (**b**) Xu and Jia [11], (**c**) Krishnan et al. [9], (**d**) Pan et al. [59] and (**e**) our proposed method, respectively.
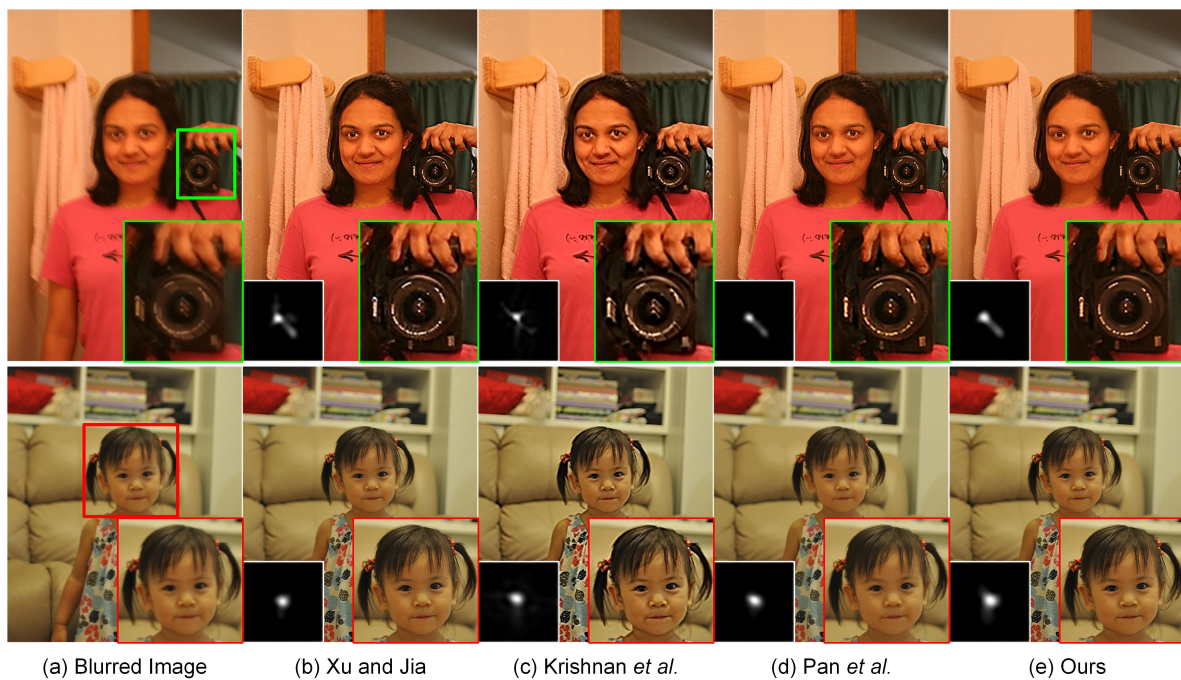

(a) Blurred Image　　　(b) Xu and Jia　　　(c) Krishnan *et al.*　　　(d) Pan *et al.*　　　(e) Ours

Figure 12. Blind deconvolution of two realistic human images. From left to right: (**a**) input blurred images, deblurred versions generated by (**b**) Xu and Jia [11], (**c**) Krishnan et al. [9], (**d**) Pan et al. [59] and (**e**) our proposed method, respectively. The sizes of the estimated blur kernels from top to bottom are 25×25 and 23×23, respectively.

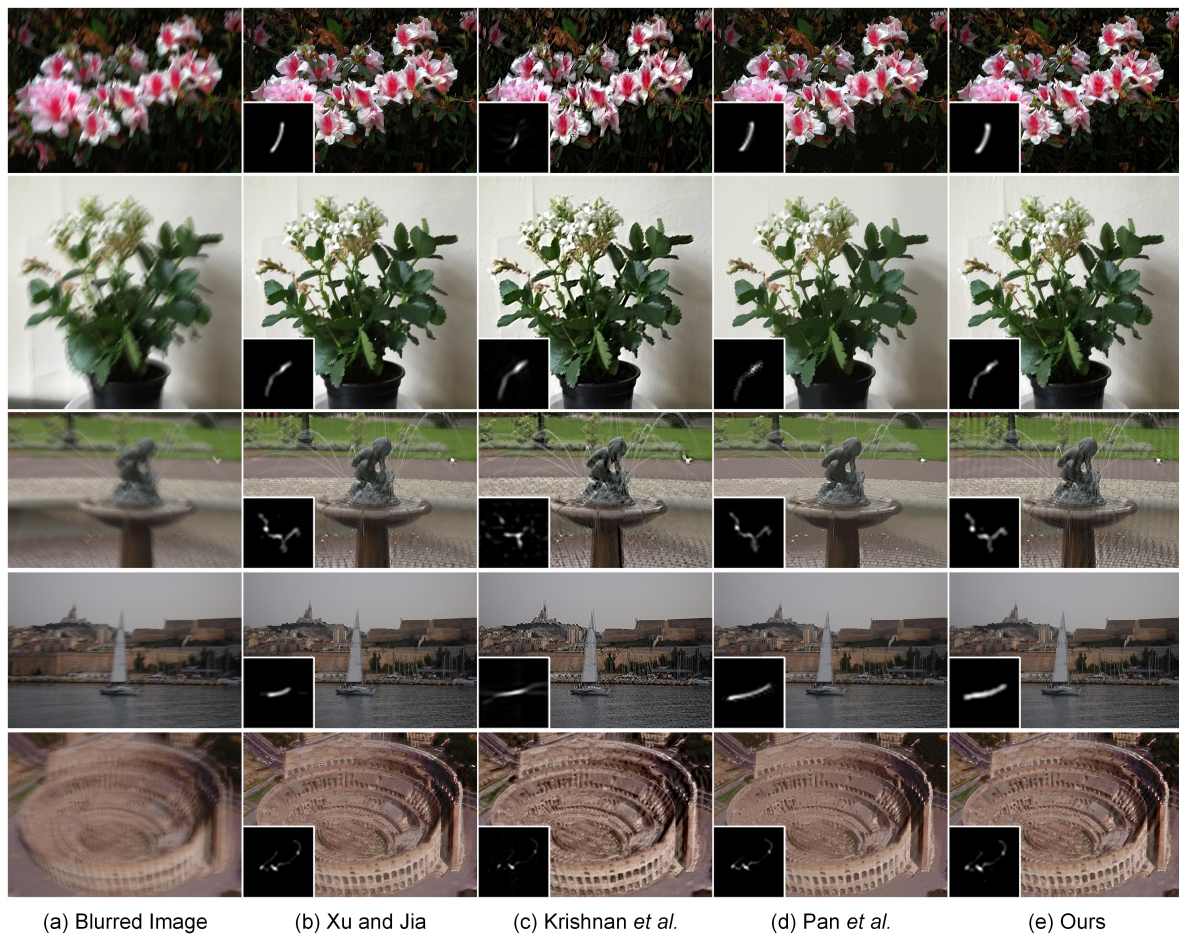(a) Blurred Image　　(b) Xu and Jia　　(c) Krishnan *et al.*　　(d) Pan *et al.*　　(e) Ours

Figure 13. Blind deconvolution of five different realistic natural images. From left to right: (**a**) input blurred images, deblurred versions generated by (**b**) Xu and Jia [11], (**c**) Krishnan et al. [9], (**d**) Pan et al. [59] and (**e**) our proposed method, respectively. The sizes of the estimated blur kernels from top to bottom are 35×35, 55×55, 41×41, 35×35 and 55×55, respectively (the images are best viewed in full-screen mode).

5. Conclusions and Future Work

The major contributions of this work are mainly two-fold. First, a hybrid regularization method was developed to robustly estimate the blur kernel by incorporating both the L_1-norm of kernel intensity and the squared L_2-norm of the intensity derivative. The underlying assumption behind the proposed method was that the blur kernel was not only spatially sparse, but also piecewise smooth within the support of a continuous curve. An alternating direction algorithm was then proposed to effectively solve the resulting problem of blur kernel estimation. Second, to guarantee high-quality non-blind deconvolution, the TGV²-regularized variational model with an L_1-norm data-fidelity term was presented to enhance the final image quality. The resulting optimization problem was then effectively solved using an ADMM-based numerical method. Comprehensive experiments implemented on both synthetic and realistic blurred images have illustrated the effectiveness of the proposed method. Given the recent progress in image deblurring, the proposed deblurring framework has several potential limitations in its current version. To further improve the blind deblurring performance, there is a huge potential to extend our future work along the following directions:

- The constant parameters (i.e., η_1 and η_2) for both the L_1-norm of kernel intensity and the squared L_2-norm of intensity derivative in (3) are manually selected in our current work. Essentially, it is necessary to automatically and adaptively select the parameters according to the statistical

properties of the blur kernel. For instance, if the blur kernel can be better sparsely represented in the spatial domain, η_1 should be larger; whereas η_2 plays a more important role if the blur kernel has a significant piecewise smooth structure. In our future work, an automatic estimation method should be developed to adaptively select the weighting parameters η_1 and η_2 in (3) to enhance the accuracy of blur kernel estimation.

- The single-image blind deblurring method proposed in this work is performed based on a common assumption that the blur kernel is uniform (i.e., spatially invariant) across the image plane. Recent work in the literature [2,60–65] has illustrated that the uniform simple assumption does not always hold in practice. To further enhance image quality, the assumption of the non-uniform (i.e., spatially variant) blur kernel has gained increasing attention in modern imaging sciences. In our opinion, the proposed hybrid regularized blur kernel estimation method discussed in Section 2 can be naturally extended to the case of non-uniform deblurring in future work.

As discussed beforehand, our proposed method suffers from some potential limitations (i.e., constant weighting parameters and uniform blurring assumption). Numerous experiments implemented on both synthetic and realistic blurred images have demonstrated its satisfactory deblurring performance. Therefore, it is still worthy of consideration since it could guarantee reliable performance compared to current state-of-the-art uniform blind deblurring methods. Recent research [9,41,59] has indirectly shown that our proposed method could be easily extended to the case of the non-uniform scenario. We believe there is a great potential for restoring blurred images using the proposed method in practical applications.

Author Contributions: The work presented in this paper corresponds to a collaborative development by all authors. R.W.L., Z.L. and N.X. conceived of and designed the experiments. R.W.L., M.L., D.W. and H.W. performed the numerical experiments. R.W.L., D.W. and H.W. analyzed the experimental results. R.W.L., M.L., Z.L. and N.X. wrote the manuscript and improved this manuscript's English language and style.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Yang, F.; Huang, Y.; Luo, Y.; Li, L.; Li, H. Robust image restoration for motion blur of image sensors. *Sensors* **2016**, *16*, 845.
2. Cheong, H.; Chae, E.; Lee, E.; Jo, G.; Paik, J. Fast image restoration for spatially varying defocus blur of imaging sensor. *Sensors* **2015**, *15*, 880–898.
3. Ruiz, P.; Zhou, X.; Mateos, J.; Molina, R.; Katsaggelos, A.K. Variational Bayesian blind image deconvolution: A review. *Digit. Signal Prog.* **2015**, *47*, 116–127.
4. Levin, A.; Weiss, Y.; Durand, F.; Freeman, W.T. Understanding blind deconvolution algorithms. *IEEE Trans. Pattern Anal. Mach. Intell.* **2011**, *33*, 2354–2367.
5. Wang, C.; Yue, Y.; Dong, F.; Tao, Y.; Ma, X.; Clapworthy, G.; Ye, X. Enhancing Bayesian estimators for removing camera shake. *Comput. Graph. Forum* **2013**, *32*, 113–125.
6. Kundur, D.; Hatzinakos, D. Blind image deconvolution. *IEEE Signal Process. Mag.* **1996** *13*, 43–64.
7. Fergus, R.; Singh, B.; Hertzmann, A.; Roweis, S.T.; Freeman, W.T. Removing camera shake from a single photograph. *ACM Trans. Graph.* **2006**, *25*, 787–794.
8. Shan, Q.; Jia, J.Y.; Agarwala, A. High-quality motion deblurring from a single image. *ACM Trans. Graph.* **2008**, *28*, 73.
9. Krishnan, D.; Tay, T.; Fergus, R. Blind deconvolution using a normalized sparsity measure. In Proceedings of the 24th IEEE Conference on Computer Vision and Pattern Recognition, Colorado Springs, CO, USA, 20–25 June 2011; pp. 233–240.

10. Kotera, J.; Šroubek, F.; Milanfar, P. Blind deconvolution using alternating maximum a posteriori estimation with heavy-tailed priors. In Proceedings of the 15th International Conference on Computer Analysis of Images and Patterns, York, UK, 27–29 August 2013; pp. 59–66.

11. Xu, L.; Jia, J.Y. Two-phase kernel estimation for robust motion deblurring. In Proceedings of the 11th European Conference on Computer Vision, Crete, Greece, 5–11 September 2010; pp. 157–170.

12. Cho, S.; Lee, S. Fast motion deblurring. *ACM Trans. Graph.* **2009**, *28*, 145.

13. Pan, J.S.; Su, Z.X. Fast l^0-regularized kernel estimation for robust motion deblurring. *IEEE Signal Process. Lett.* **2013**, *20*, 841–844.

14. Cai, J.F.; Ji, H.; Liu, C.; Shen, Z. Framelet-based blind motion deblurring from a single image. *IEEE Trans. Image Process.* **2012**, *21*, 562–572.

15. Shao, W.Z.; Li, H.B.; Elad, M. Bi-l_0-l_2-norm regularization for blind motion deblurring. *J. Vis. Commun. Image Represent.* **2015**, *33*, 42–59.

16. Engl, H.W.; Hanke, M.; Neubauer, A. *Regularization of Inverse Problems*; Kluwer Academic Publishers: Dordrecht, Netherlands, 1996.

17. Donatelli, M. A multigrid for image deblurring with Tikhonov regularization. *Numer. Linear Algebr. Appl.* **2015**, *12*, 715–729.

18. Hamarik, U.; Palm, R.; Raus, T. Extrapolation of Tikhonov regularization method. *Math. Model. Anal.* **2010**, *15*, 55–68.

19. Liu, W.; Wu, C.S. A predictor-corrector iterated Tikhonov regularization for linear ill-posed inverse problems. *Appl. Math. Comput.* **2013**, *221*, 802–818.

20. Lucy, L.B. An iterative technique for the rectification of observed distributions. *Astro. J.* **1974**, *79*, 745–754.

21. Wiener, N. *Extrapolation, Interpolation, and Smoothing of Stationary Time Series*; MIT Press: Cambridge, MA, USA, 1949.

22. Yuan, L.; Sun, J.; Quan, L.; Shum, H.Y. Progressive inter-scale and intra-scale non-blind image deconvolution. *ACM Trans. Graph.* **2008**, *27*, 74.

23. Rudin, L.I.; Osher, S.; Fatemi, E. Nonlinear total variation based noise removal algorithms. *Phys. D* **1992**, *60*, 259–268.

24. Chan, T.F.; Wong, C.K. Total variation blind deconvolution. *IEEE Trans. Image Process.* **1998**, *7*, 370–375.

25. Babacan, S.D.; Molina, R.; Katsaggelos, A.K. Variational Bayesian blind deconvolution using a total variation prior. *IEEE Trans. Image Process.* **2009**, *18*, 12–26.

26. Krishnan, D.; Fergus, R. Fast image deconvolution using hyper-Laplacian priors. In Proceedings of the 23rd Annual Conference on Neural Information Processing Systems, Vancouver, BC, Canada, 7–9 December 2009; pp. 1033–1041.

27. Chan, T.; Marquina, A.; Mulet, P. High-order total variation-based image restoration. *SIAM J. Sci. Comput.* **2000**, *22*, 503–516.

28. Stefan, W.; Renaut, R.A.; Gelb, A. Improved total variation-type regularization using higher order edge detectors. *SIAM J. Imaging Sci.* **2010**, *3*, 232–251.

29. Lysaker, M.; Tai, X.C. Iterative image restoration combining total variation minimization and a second-order functional. *Int. J. Comput. Vis.* **2006**, *66*, 5–18.

30. Liu, R.W.; Wu, D.; Wu, C.S.; Xu, T.; Xiong, N. Constrained nonconvex hybrid variational model for edge-preserving image restoration. In Proceedings of the 2015 IEEE International Conference on Systems, Man, and Cybernetics, Hong Kong, China, 9–12 October 2015; pp. 1809–1814.

31. Papafitsoros, K.; Schönlieb, C.B. A combined first and second order variational approach for image reconstruction. *J. Math. Imaging Vis.* **2014**, *48*, 308–338.

32. Bredies, K.; Kunisch, K.; Pock, T. Total generalized variation. *SIAM J. Imaging Sci.* **2010**, *3*, 492–526.

33. Bredies, K.; Dong, Y.; Hintermüller, M. Spatially dependent regularization parameter selection in total generalized variation models for image restoration. *Int. J. Comput. Math.* **2013**, *90*, 109–123.

34. Liu, R.W.; Shi, L.; Yu, S.C.H.; Wang, D. Box-constrained second-order total generalized variation minimization with a combined $L^{1,2}$ data-fidelity term for image reconstruction. *J. Electron. Imaging* **2015**, *24*, 033026.

35. Duan, J.; Lu, W.; Tench, C.; Gottlob, I.; Proudlock, F.; Samani, N.N.; Bai, L. Denoising optical coherence tomography using second order total generalized variation decomposition. *Biomed. Signal Process. Control* **2016**, *24*, 120–127.

36. Zhang, X.; Burger, M.; Bresson, X.; Osher, S. Bregmanized nonlocal regularization for deconvolution and sparse reconstruction. *SIAM J. Imaging Sci.* **2010**, *3*, 253–276.

37. Tang, S.; Gong, W.; Li, W.; Wang, W. Non-blind image deblurring method by local and nonlocal total variation models. *Signal Process.* **2014**, *94*, 339–349.

38. Liu, R.W.; Shi, L.; Yu, S.C.H.; Wang, D. A two-step optimization approach for nonlocal total variation-based Rician noise reduction in magnetic resonance images. *Med. Phys.* **2015**, *42*, 5167–5187.

39. Zhang, C.; Wu, D.; Liu, R.W.; Xiong, N. Non-local regularized variational model for image deblurring under mixed Gaussian-impulse noise. *J. Internet Technol.* **2015**, *16*, 1301–1320.

40. Xu, L.; Lu, C.; Xu, Y.; Jia, J. Image smoothing via l0 gradient minimization. *ACM Trans. Graph.* **2011**, *30*, 174.

41. Xu, L.; Zheng, S.; Jia, J. Unnatural l0 sparse representation for natural image deblurring. In Proceedings of the 26th IEEE Conference on Computer Vision and Pattern Recognition, Portland, OR, USA, 23–28 June 2013; pp. 1107–1114.

42. Pan, J.; Hu, Z.; Su, Z.; Yang, M.H. L0-regularized intensity and gradient prior for deblurring text images and beyond. *IEEE Trans. Pattern Anal. Mach. Intell.* **2016**, doi:10.1109/TPAMI.2016.2551244.

43. Boyd, S.; Parikh, N.; Chu, E.; Peleato, B.; Eckstein, J. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Found. Trends Mach. Learn.* **2010**, *3*, 1–122.

44. Liu, R.W.; Wu, D.; Wu, C.S.; Xiong, N. Hybrid regularized blur kernel estimation for single-image blind deconvolution. In Proceedings of the 2015 IEEE International Conference on Systems, Man, and Cybernetics, Hong Kong, China, 9–12 October 2015; pp. 1815–1820.

45. Liu, R.W.; Shi, L.; Huang, W.; Xu, J.; Yu, S.C.H.; Wang, D. Generalized total variation-based MRI Rician denoising model with spatially adaptive regularization parameters. *Magn. Reson. Imaging* **2014**, *32*, 702–720.

46. Liu, Q.; Lai, Z.; Zhou, Z.; Kuang, F.; Jin, Z. A truncated nuclear norm regularization method based on weighted residual error for matrix completion. *IEEE Trans. Image Process.* **2016**, *25*, 316–330.

47. Lu, Y.; Lai, Z.; Xu, Y.; Li, X.; Zhang, D.; Yuan, C. Low-rank preserving projections. *IEEE Trans. Cybern.* **2015**, *46*, 1900–1913.

48. Dong, B.; Zhang, Y. An efficient algorithm for l_0 minimization in wavelet frame based image restoration. *J. Sci. Comput.* **2013**, *54*, 350–368.

49. Chan, R.H.; Tao, M.; Yuan, X.M. Constrained total variation deblurring models and fast algorithms based on alternating direction method of multipliers. *SIAM J. Imaging Sci.* **2013**, *6*, 680–697.

50. Lai, Z.; Xu, Y.; Chen, Q.; Yang, J.; Zhang, D. Multilinear sparse principal component analysis. *IEEE Trans. Neural Netw. Learn. Syst.* **2014**, *25*, 1942–1950.

51. Duan, J.; Pan, Z.; Zhang, B.; Liu, W.; Tai, X.C. Fast algorithm for color texture image inpainting using the non-local CTV model. *J. Glob. Optim.* **2015**, *62*, 853–876.

52. Figueiredo, M.A.; Bioucas-Dias, J.M. Restoration of Poissonian images using alternating direction optimization. *IEEE Trans. Image Process.* **2010**, *19*, 3133–3145.

53. Han, D.; Yuan, X.M. A note on the alternating direction method of multipliers. *J. Optim. Theory Appl.* **2012**, *155*, 227–238.

54. Knoll, F.; Bredies, K.; Pock, T.; Stollberger, R. Second order total generalized variation (TGV) for MRI. *Magn. Reson. Med.* **2011**, *65*, 480–491.

55. Valkonen, T.; Bredies, K.; Knoll, F. Total generalized variation in diffusion tensor imaging. *SIAM J. Imaging Sci.* **2013**, *6*, 487–525.

56. Lu, W.; Duan, J.; Qiu, Z.; Pan, Z.; Liu, R.W.; Bai, L. Implementation of high-order variational models made easy for image processing. *Math. Meth. Appl. Sci.* **2016**, *39*, 4208–4233.

57. Duan, J.; Pan, Z.; Yin, X.; Wei, W.; Wang, G. Some fast projection methods based on Chan-Vese model for image segmentation. *EURASIP J. Image Video Process.* **2014**, *2014*, 1–16.

58. Levin, A.; Weiss, Y.; Durand, E.; Freeman, W.T. Efficient marginal likelihood optimization in blind deconvolution. In Proceedings of the 24th IEEE Conference on Computer Vision and Pattern Recognition, Colorado Springs, CO, USA, 20–25 June 2011; pp. 2657–2664.

59. Pan, J.; Sun, D.; Pfister, H.; Yang, M.H. Blind image deblurring using dark channel prior. In Proceedings of the 29th IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 26 June–1 July 2016; pp. 1–9.

60. Whyte, O.; Sivic, J.; Zisserman, A.; Ponce, J. Non-uniform deblurring for shaken images. *Int. J. Comput. Vis.* **2012**, *98*, 168–186.

61. Cho, S.; Cho, H.; Tai, Y.W.; Lee, S. Registration based non-uniform motion deblurring. *Comput. Graph. Forum* **2012**, *31*, 2183–2192.

62. Paramanand, C.; Rajagopalan, A.N. Non-uniform motion deblurring for bilayer scenes. In Proceedings of the 26th IEEE Conference on Computer Vision and Pattern Recognition, Portland, OR, USA, 23–28 June 2013; pp. 1115–1122.

63. Sun, J.; Cao, W.; Xu, Z.; Ponce, J. Learning a convolutional neural network for non-uniform motion blur removal. In Proceedings of the 28th IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 769–777.

64. Sroubek, F.; Kamenicky, J.; Lu, Y.M. Decomposition of space-variant blur in image deconvolution. *IEEE Signal Process. Lett.* **2016**, *23*, 346–350.

65. Zhang, X.; Wang, R.; Jiang, X.; Wang, W.; Gao, W. Spatially variant defocus blur map estimation and deblurring from a single image. *J. Vis. Commun. Image Represent.* **2016**, *35*, 257–264.