

Received June 1, 2017, accepted June 20, 2017, date of publication June 23, 2017, date of current version July 17, 2017.

Digital Object Identifier 10.1109/ACCESS.2017.2719119

# Iteratively Reweighted Blind Deconvolution With Adaptive Regularization Parameter Estimation

HOUZHANG FANG<sup>1</sup>, YI CHANG<sup>2</sup>, (Student Member, IEEE), GANG ZHOU<sup>3</sup>, AND LIZHEN DENG<sup>4</sup>

<sup>1</sup>National Laboratory of Radar Signal Processing, Xidian University, Xi'an 710071, China

<sup>2</sup>Science and Technology on Multispectral Information Processing Laboratory, School of Automation, Huazhong University of Science and Technology, Wuhan 430074, China

<sup>3</sup>State Key Laboratory of Material Processing and Die & Mould Technology, Huazhong University of Science and Technology, Wuhan 430074, China

<sup>4</sup>College of Telecommunication and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210003, China

Corresponding author: Houzhang Fang (houzhangfang@xidian.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 41501371, in part by the Fundamental Research Funds for the Central Universities under Grant JB170210, in part by the China Postdoctoral Science Foundation under Grant 2017M612453 and Grant 2016M591891, and in part by the Natural Science Foundation of Jiangsu Province under Grant BK20140874 and Grant BK20150864.

**ABSTRACT** In many realistic image processing applications, the acquired images often suffer from mixed noises and blurring, which greatly degrade the image quality. In this paper, we propose an iteratively reweighted blind deconvolution method with robust regression for obtaining high quality images with mixed noises present. First, we construct a variational regularization model, including a robust regression data term with an adaptive reweighted least square criterion, which is robust to the mixed noises. To preserve the sharp edges and suppress the noise, a total variation-based regularization term for the image is incorporated into the model. Moreover, a Laplacian regularization term is imposed on the point spread function (PSF) for better smoothness. The subsequent optimization problems for the image and the PSF are solved using the limited-memory BFGS-B algorithm suitable for the large-scale problems. In addition, to improve the practicality of the method, a variant of the generalized cross validation method is derived and adopted to automatically estimate the regularization parameters for the image and the PSF. Experiments on simulated and real images demonstrate that the proposed method is superior to the state-of-the-art methods in terms of both subjective measure and visual quality.

**INDEX TERMS** Iteratively reweighted, blind deconvolution, regularization parameter selection, image restoration, robust regression.

## I. INTRODUCTION

Deterministic and stochastic distortions, usually blurs and noise perturbation, are commonly found in images acquired by astronomical, medical, and microscopic imaging systems, and many others (see [1] for a review). The blur and noise are related to the process of image formation and recording, respectively. Blind image deconvolution (BID) [1] has been widely utilized to simultaneously estimate the original image and the point spread function (PSF) of the imaging system from the degraded observations. However, the BID problem is highly ill-posed. In order to obtain physically meaningful solutions, the variational regularization based blind image deconvolution methods with the appropriate prior knowledge incorporated has been introduced (see [2]–[4] and [1] for a review).

In traditional blind image restoration methods, single noise is assumed in the process of image recording due to its

simplicity (see [3]–[5]). However, many images are often affected by a mixture of noise sources. In the astronomical and microscopic imaging, the noises are generally regarded as a combination of the Gaussian noise and the Poissonian noise that are always present in digital images acquired using digital detectors such as CCD cameras [2], [6]–[8]. The Gaussian component is typically related to the readout noise present in the electronic part of the imaging system and usually cannot be neglected, whereas the Poisson noise component is from fluctuations in the number of counting photons at each pixel. A number of methods have been developed for coping with image deconvolution in the context of Poisson-Gaussian noises. Lantéri et al. considered the image deconvolution techniques under a mixture of Poissonian and Gaussian noises [9] and applied the split gradient method to obtain the maximum likelihood (ML) iterative algorithms with three classical regularization terms for imposing a smoothness

constraint on the solution. Li et al. proposed a reweighted  $\ell_2$  fidelity in the variational model with framelet regularization in [10] and adopted the split Bregman algorithm to convert the original minimization problem into three easily solvable subminimization problems that can be solved using quadratic minimization, soft shrinkage and matrix vector multiplications, and the proposed method is effective in removing mixed Poisson-Gaussian noises. Chouzenoux et al. [11] adopted the primal-dual splitting algorithm to solve the constructed variational model, which combines the Poisson-Gaussian negative log-likelihood in the presence of Poisson-Gaussian noises with a convex nonsmooth regularizer. The method presented in [11] is competitive in restoring images corrupted by mixed noises. Recently, Kubínová and Nagy [12] proposed combining robust regression with a weighted least-squares function with solution-dependent weights to approximate the true negative log-likelihood function and imposed the Tikhonov regularization for the image in order to achieve a reasonable solution. The results presented in [12] suggest that the proposed method can remove mixed Poisson-Gaussian noises and several types of outliers. Fang et al. [13] presented an iteratively reweighted blind deconvolution method where the data term weight is constructed based on the residual error and the median absolute deviation of the residual error, which is capable of coping with the mixed noises. However, the ability to remove the mixed noises for the algorithm in [13] is still limited.

The selection of the regularization parameter is crucial in variational regularization based blind deconvolution and plays an important role in controlling the contribution of the regularization term to the solution. A large number of existing methods, such as [3] and [4], choose the regularization parameters by trial and error tests which can be highly time-consuming. A considerable effort has been devoted to adaptively choosing the regularization parameter. Liao et al. in [14] employed the variable-splitting and penalty method [15] to solve total variation (TV) based blind deconvolution. The advantage of applying the variable-splitting and penalty technique to total variation based blind deconvolution is that the resulting sub-minimization problem about the image is a least square problem and thus the parameter associated with the image can be estimated adaptively using the generalized cross-validation (GCV) method [16]–[18]. Wang and Ng [19] also utilized the variable-splitting and penalty techniques [15] to solve total variation blind deconvolution in [20] and framelet based blind deconvolution in [21], where the regularization parameters for the image and PSF can be updated using the GCV method. In Poisson maximum likelihood estimation, Bardsley and Goldes [22] derived a weighted least-square function to approximate the negative-log Poisson likelihood function and made use of an approximated GCV function to estimate the regularization parameter. Then, Bardsley and Goldes [23] employed the GCV method to estimate the regularization parameter associated with total variation-regularized negative-log Poisson likelihood problems. Kubínová and Nagy [24] further

invoked the work in [22] and [23] by investigating the selection of the regularization parameter in [12] and incorporating the non-negative constraints into the modified GCV function. We also derived a formula to automatically update the regularization parameter associated with the proposed iteratively reweighted blind deconvolution [13] and applied the proposed method to restore the passive millimeter-wave image [25], [26]. Liu et al. [5] adopted the fast gradient projection algorithm [27] to convert the total variation blind deconvolution in [20] into two separate image denoising subproblems that can be solved fast.

In this work, we propose an iteratively reweighted least squares blind deconvolution algorithm, which is designed to simultaneously estimate the latent image and the degradation PSF. Motivated by the non-blind deconvolution work in [12], [22], and [24], the data fidelity term in our blind deconvolution model is a weighted least squares problem with the weights depending on the current image and PSF, which can handle the mixed Poisson-Gaussian noises. In order to preserve the sharp edges and guarantee restoration with minimal artifacts, total variation regularization is utilized for the image. The Laplacian regularization is adopted for the PSF, which is effective in imposing smoothness on the PSF and is typically employed as a model for the PSF that is expected to be smooth, such as Gaussian blur [5], [14], [26], Gaussian-like blur [13] and atmospheric turbulence [28]. Because of the highly nonlinear characteristic of the data fidelity term, a variant of the GCV method is derived and adopted to estimate the regularization parameters for the image and PSF.

The outline of this paper is as follows. In Section II, we present our algorithm. In Section III, we present the experimental results. Conclusions are drawn in Section IV.

## II. ROBUST REGRESSION FOR WEIGHTED LEAST SQUARES BLIND DECONVOLUTION

### A. THE DEGRADATION MODEL

In this study, we consider the degradation model with the matrix-vector form

$$\mathbf{y} = \mathbf{H}\mathbf{f} + \mathbf{n} = \mathbf{F}\mathbf{h} + \mathbf{n}, \quad (1)$$

where  $\mathbf{y}$ ,  $\mathbf{h}$ ,  $\mathbf{f}$ , and  $\mathbf{n}$  denote the observed image, PSF of the imaging system, high-resolution image, and noise, respectively.  $\mathbf{H} \in \mathbb{R}^{N \times N}$  represents the block-circulant blurring matrix notation of the convolution of the PSF  $\mathbf{h} \in \mathbb{R}^N$ . Similarly,  $\mathbf{F} \in \mathbb{R}^{N \times N}$  denotes the block-circulant matrix formed from the image  $\mathbf{f} \in \mathbb{R}^N$  ( $N$  is the number of pixels of the image).

### B. DISCRETE TOTAL VARIATION AND LAPLACIAN OPERATOR

It is known that the TV regularization model can well preserve the important image features, such as sharp edges or object boundaries. It is first proposed for image denoising by Rudin et al. [29] and then extended to image deconvolution in [30]. We denote the total finite difference operator

by  $\mathbf{D} = [\mathbf{D}_1; \mathbf{D}_2] \in \mathbb{R}^{2N \times N}$ , where  $\mathbf{D}_1, \mathbf{D}_2 \in \mathbb{R}^{N \times N}$  are discretizations of the corresponding horizontal and vertical first-order forward finite difference operators with appropriate boundary conditions. Thus, the discrete total variation of the image  $\mathbf{f}$  can be written as

$$\|\mathbf{D}\mathbf{f}\|_1 := \sum_{i=1}^N \sqrt{(\mathbf{D}_1\mathbf{f})_i^2 + (\mathbf{D}_2\mathbf{f})_i^2}. \quad (2)$$

Note that the TV regularization term  $\|\mathbf{D}\mathbf{f}\|_1$  over the image is non-differentiable. To address this obstacle, in our computations, we apply a continuous and differentiable approximation to (2)

$$T_\varepsilon(\mathbf{f}) = \|\mathbf{D}\mathbf{f}\|_{1,\varepsilon} := \sum_{i=1}^N \psi \left( (\mathbf{D}_1\mathbf{f})_i^2 + (\mathbf{D}_2\mathbf{f})_i^2 \right), \quad (3)$$

where  $\psi(t) = \sqrt{t + \varepsilon}$ ,  $\varepsilon$  is a small constant. This is a very standard approximation to the  $\ell_1$  norm regularization [13], [31].

The Laplacian operator  $\mathbf{L} \in \mathbb{R}^{N \times N}$  is the discrete matrix realization of the Laplacian kernel

$$\mathbf{L}_{\text{kernel}} = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix}. \quad (4)$$

### C. REWEIGHTED LEAST SQUARES BLIND DECONVOLUTION MODEL WITH ROBUST REGRESSION

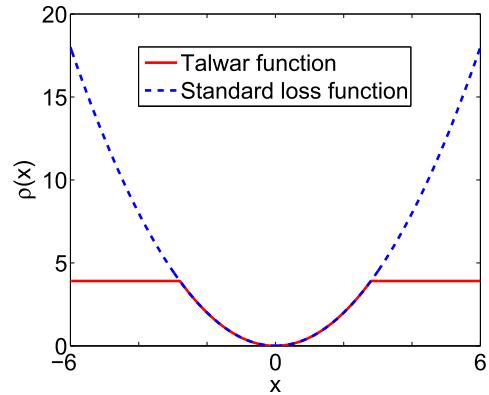
As shown in (1), we assume that the noise  $\mathbf{n}$  is a normally-distributed random variable with mean 0 and variance  $\mathbf{H}\mathbf{f} + \sigma^2$ , which aims at approximating a combination of Gaussian noise (with variance  $\sigma^2$ ) and Poisson noise (with variance  $\mathbf{H}\mathbf{f}$ ). We formulate the reweighted least squares blind image deconvolution problem as

$$J(\mathbf{f}, \mathbf{h}) = \sum_{i=1}^N \rho \left( \frac{[\mathbf{H}\mathbf{f}]_i - \mathbf{y}_i}{\sqrt{[\mathbf{H}\mathbf{f}]_i + \sigma^2}} \right) + \lambda \|\mathbf{D}\mathbf{f}\|_{1,\varepsilon} + \frac{\gamma}{2} \|\mathbf{L}\mathbf{h}\|_2^2, \quad (5)$$

where  $\lambda$  and  $\gamma$  are the regularization parameters for the image and PSF, respectively. The first term on the right-hand side (RHS) is the data fidelity term that is formulated using robust regression with the reweighted least squares criterion.  $\rho(\cdot)$  is the Talwar function (see Figure 1)

$$\rho(x) = \begin{cases} x^2/2, & |x| \leq \beta, \\ \beta^2/2, & |x| > \beta, \end{cases} \quad (6)$$

where the tuning constant  $\beta$  determines the trade-off between the robustness and efficiency. As suggested in [12] and [32], we use the Talwar function with the tuning parameter  $\beta = 2.795$  such that the asymptotic efficiency for the standard loss function  $x^2/2$  is 95% when the distribution of the disturbances is the unit normal Gaussian distribution. In the first term on the RHS, the random variable inside the robust



**FIGURE 1.** The Talwar loss function  $\rho(x)$  (solid line) for the tuning parameter  $\beta = 2.795$  corresponding to 95% efficiency with the standard least squares loss function  $x^2/2$  (dotted line) as reference.

loss function  $\rho(\cdot)$  in (5) is rescaled to obtain an approximately unit normal distribution.

In general, the regularization constraints should be chosen based on some prior assumptions about the unknown quantities. The total variation regularization term is adopted for the image for its excellent edge-preserving ability in the second term on the RHS. The Laplacian regularization term is utilized for the PSF. In comparison to that of the TV regularization constraint, the  $\ell_2$ -norm Laplacian regularization  $\|\mathbf{L}\mathbf{h}\|_2^2$  is much better and more appropriate for the smooth PSF, such as Gaussian blur [5], [14], [26], Gaussian-like blur [13] and atmospheric turbulence blur [28]. The two regularization terms for the image and the PSF play important roles on avoiding the convergence to the degenerate solutions.

### D. NUMERICAL OPTIMIZATION ALGORITHMS

It is well known that the variational regularization based blind image deconvolution usually leads to a challenging nonconvex joint minimization problem with respect to the image  $\mathbf{f}$  and the PSF  $\mathbf{h}$ . The most typical approach is an alternative minimization (AM) scheme (see [5], [13], [14], [20] for instance): in each step of the iterative procedure we minimize one variable and keep the other one fixed. Specifically, for a given initial guess of the image and the PSF  $(\mathbf{f}^{(0)}, \mathbf{h}^{(0)})$  for  $(\mathbf{f}, \mathbf{h})$ , we first minimize the functional (5) by solving  $\mathbf{f}^{(1)} = \arg \min_{\mathbf{f}} J(\mathbf{f}, \mathbf{h}^{(0)})$  and then solving  $\mathbf{h}_1 = \arg \min_{\mathbf{h}} J(\mathbf{f}^{(1)}, \mathbf{h})$ . The process is repeated until the given stopping criteria are met. The alternative minimization iteration procedure is described in Algorithm 1.

Next, we describe the numerical solution of our blind deconvolution algorithm outlined in Algorithm 1 in detail. Two steps are presented in Algorithm 1. Step 1 in (7) is a large-scale nonblind image deconvolution problem. The data term in (7) is a highly nonlinear term. Hence the typical numerical optimization methods, such as linear conjugate gradient, cannot be used. To solve the image in (7), the nonlinear optimization algorithms have to be used. The limited-memory BFGS-B (L-BFGS-B) algorithm [33] is a widely

**Algorithm 1** Outline of the alternative minimization iterations

- 1: **Input:** The observed image  $\mathbf{y}$ , the maximum number of iterations  $K_{\text{MaxIter}}$ , and iteration stopping errors  $\epsilon_f, \epsilon_h$ .
- 2: **Initialization:** an initial image  $\mathbf{f}^{(0)}$  and an initial PSF  $\mathbf{h}^{(0)}$ .
- 3: **while**  $k = 0$  to  $K_{\text{MaxIter}}$  and  $\|\mathbf{f}^{(k)} - \mathbf{f}^{(k-1)}\|_2 / \|\mathbf{f}^{(k-1)}\|_2 < \epsilon_f$  or  $\|\mathbf{h}^{(k)} - \mathbf{h}^{(k-1)}\|_2 / \|\mathbf{h}^{(k-1)}\|_2 < \epsilon_h$  **do**
- 4:   Update  $k := k + 1$ .
- 5:   Given the PSF  $\mathbf{h}^{(k)}$ , compute the latent image  $\mathbf{f}^{(k)}$ :
$$\begin{aligned}\mathbf{f}^{(k)} &= \arg \min_{\mathbf{f}} J(\mathbf{f}, \mathbf{h}^{(k-1)}) \\ &= \arg \min_{\mathbf{f}} \sum_{i=1}^N \rho \left( \frac{[\mathbf{H}^{(k-1)}\mathbf{f}]_i - \mathbf{y}_i}{\sqrt{[\mathbf{H}^{(k-1)}\mathbf{f}]_i + \sigma^2}} \right) + \lambda \|\mathbf{D}\mathbf{f}\|_{1,\varepsilon}.\end{aligned}\quad (7)$$
- 6:   Given the image  $\mathbf{f}^{(k)}$ , compute the PSF  $\mathbf{h}^{(k)}$ :
$$\begin{aligned}\mathbf{h}^{(k)} &= \arg \min_{\mathbf{h}} J(\mathbf{f}^{(k)}, \mathbf{h}) \\ &= \arg \min_{\mathbf{h}} \sum_{i=1}^N \rho \left( \frac{[\mathbf{F}^{(k)}\mathbf{h}]_i - \mathbf{y}_i}{\sqrt{[\mathbf{F}^{(k)}\mathbf{h}]_i + \sigma^2}} \right) + \frac{\gamma}{2} \|\mathbf{L}\mathbf{h}\|_2^2.\end{aligned}\quad (8)$$
- 7: **end while**
- 8: Output: the restored image  $\mathbf{f}^{(k)}$  and PSF  $\mathbf{h}^{(k)}$ ;

used quasi-Newton algorithm in engineering and scientific computation and has been proved suitable for large-scale problem. Thompson et al. [34] investigated the performance of the several nonlinear optimization methods that solve the blind image deconvolution maximum likelihood optimization problem and concluded that the L-BFGS-B algorithm is relatively optimal in restoring the image details and in running speed compared with other nonlinear methods, such as the nonlinear conjugate gradient method [35] and the truncated Newton method [36]. In this research, we employ the L-BFGS-B algorithm for numerical optimization of our nonblind deconvolution about the image in (7) and the PSF in (8).

A key step for the optimization algorithm with the L-BFGS-B is to calculate the gradient of the objective function. The gradient of the objective function with respect to the image  $\mathbf{f}$  in (7) by invoking the calculus of the variation has the form

$$\begin{aligned}\text{grad}J(\mathbf{f}, \mathbf{h}^{(k)})(\mathbf{f}) &= \left( \mathbf{H}^{(k)} \right)^T \mathbf{z} + \lambda \nabla T_\varepsilon(\mathbf{f}), \\ z_i &= \begin{cases} \frac{1}{2} - \frac{1}{2} \left( \frac{y_i + \sigma^2}{[\mathbf{H}^{(k)}\mathbf{f}]_i + \sigma^2} \right), & \left| \frac{[\mathbf{H}^{(k)}\mathbf{f}]_i - \mathbf{y}_i}{\sqrt{[\mathbf{H}^{(k)}\mathbf{f}]_i + \sigma^2}} \right| \leq \beta, \\ 0, & \text{otherwise,} \end{cases}\end{aligned}\quad (9)$$

where the superscript  $T$  denotes the transpose operation and the gradient of the total variation regularization term about

the image  $\nabla T_\varepsilon(\mathbf{f}) = L_1(\mathbf{f})\mathbf{f}$ ,  $L_1(\mathbf{f}) = \mathbf{D}_1^T \text{diag}(\psi'(\mathbf{D}\mathbf{f}^2))\mathbf{D}_1 + \mathbf{D}_2^T \text{diag}(\psi'(\mathbf{D}\mathbf{f}^2))\mathbf{D}_2$ ,  $\psi(t) = \sqrt{t + \varepsilon}$ ,  $\mathbf{D}\mathbf{f}^2 := (\mathbf{D}_1\mathbf{f})^2 + (\mathbf{D}_2\mathbf{f})^2$ .

Similarly, for the PSF  $\mathbf{h}$ , the corresponding gradient for the objective function in (8) has the form

$$\begin{aligned}\text{grad}J(\mathbf{f}^{(k)}, \mathbf{h})(\mathbf{h}) &= \left( \mathbf{F}^{(k)} \right)^T \mathbf{z} + \gamma \mathbf{L}^T \mathbf{L}\mathbf{h}, \\ z_i &= \begin{cases} \frac{1}{2} - \frac{1}{2} \left( \frac{y_i + \sigma^2}{[\mathbf{F}^{(k)}\mathbf{h}]_i + \sigma^2} \right)^2, & \left| \frac{[\mathbf{F}^{(k)}\mathbf{h}]_i - \mathbf{y}_i}{\sqrt{[\mathbf{F}^{(k)}\mathbf{h}]_i + \sigma^2}} \right| \leq \beta, \\ 0, & \text{otherwise.} \end{cases}\end{aligned}\quad (10)$$

The derivation of (9) and (10) can be found in the appendix.

### E. CONSTRAINTS FOR THE IMAGE AND THE PSF

The existing research has demonstrated that the AM algorithm does not always obtain the physically reasonable solutions. The reason is that the objective functional (5) is a nonconvex and thus the joint minimization problem for (5) over the image and the PSF may not have a unique solution. Recent studies on joint iterative blind deconvolution [37] have indicated that blind deconvolution algorithms can converge to a satisfactory solution. A key step to success for blind deconvolution is to separate the normalization and the nonnegative constraints on the PSF from the minimization step. Specifically, after updating the image  $\mathbf{f}$  and the PSF  $\mathbf{h}$  at each AM iteration, we impose the nonnegativeness, normalization, and the centralization constraints on the PSF

$$\mathbf{h} \geq 0, \quad \sum_i \mathbf{h}_i = 1, \quad \mathbf{h} = \text{centralize}(\mathbf{h}), \quad (11)$$

where  $\text{centralize}(\mathbf{h})$  means that we shift the PSF to the location of the centroid. The observed image  $\mathbf{y}$  in our experiments is also set to  $[0, 1]$ .

## III. THE SELECTION OF THE REGULARIZATION PARAMETERS

The regularization parameter is important in striking a balance between the data fidelity term and regularization constraint term. Thus, it is necessary to develop a method to adaptively select the regularization parameters. In the case of least squares estimation, such methods have been very well-developed, such as the discrepancy principle, the unbiased predictive risk estimator, L-curve and generalized cross validation [38]. However, these methods cannot be directly applied to the minimization problems in which the data term is a reweighted nonlinear term. In this study, we will consider a variant of the generalized cross validation method, which is an approximation of the leave-one-out cross validation function for large-scale problem [16]–[18], [38]. Once the image and PSF are obtained by algorithm 1, the regularization parameters  $\lambda$  and  $\gamma$  are estimated in algorithm 1.

### A. THE CONSTRUCTION OF THE GCV FUNCTION FOR $\lambda$

As pointed out in [38, Ch. 7], suppose that we solve a linear least squares problem with Tikhonov regularization

(with penalty Laplacian operator  $\mathbf{L}$ )

$$\|\mathbf{A}\mathbf{f} - \mathbf{y}\|_2^2 + \alpha\|\mathbf{L}\mathbf{f}\|_2^2, \quad (12)$$

where  $\mathbf{A}$  is the discretized blurring operator, and  $\mathbf{f}$  and  $\mathbf{y}$  are the latent image and the observed image, respectively.  $\alpha$  is the regularization parameter. The corresponding GCV functional can be easily written as

$$\text{GCV}(\alpha) = \frac{N\|\mathbf{r}_\alpha\|_2^2}{[\text{trace}(\mathbf{I} - \mathbf{A}_\alpha)]^2}, \quad (13)$$

where  $\mathbf{r}_\alpha = \mathbf{A}\mathbf{f}_\alpha - \mathbf{y} = (\mathbf{A}_\alpha - \mathbf{I})\mathbf{y}$  is the regularized residual, and  $\text{trace}(\cdot)$  is the matrix trace operation. The influence matrix  $\mathbf{A}_\alpha$  has the form

$$\mathbf{A}_\alpha = \mathbf{A}(\mathbf{A}^T\mathbf{A} + \alpha\mathbf{L}^T\mathbf{L})^{-1}\mathbf{A}^T. \quad (14)$$

The derivation of the GCV functional for regularized least squares problems takes advantage of the fact that the regularization operator in those problems is linear. However, the data fidelity term and regularization term have a more complicated form in (7) compared with that of (12) and thus it is not easy to directly write out the residual matrix and the influence matrix for the GCV function based on (7). For (7), the influence matrix  $\mathbf{A}_\lambda$  is nonlinear and hence a linear approximation is necessary.

Following the work in [38, Ch. 7] and [23], if a non-quadratic regularized functional is used, the regularized operator  $\mathbf{L}^T\mathbf{L}$  for the influence matrix in (14) can be replaced by the Hessian matrix of the regularized functional. In this study, the TV regularization is adopted for the image and thus the Hessian matrix of the regularization function  $T_\varepsilon(\mathbf{f}) = \|\mathbf{D}\mathbf{f}\|_{1,\varepsilon}$  can be written as

$$\nabla^2 T_\varepsilon(\mathbf{f}) = L_1(\mathbf{f}) + 2L_2(\mathbf{f}), \quad (15)$$

where

$$\begin{aligned} L_2(\mathbf{f}) &= \mathbf{D}_1^T \text{diag}((\mathbf{D}_1\mathbf{f})^2 \odot \psi''(\mathbf{D}\mathbf{f}^2))\mathbf{D}_1 \\ &\quad + \mathbf{D}_1^T \text{diag}((\mathbf{D}_{12}\mathbf{f})^2 \odot \psi''(\mathbf{D}\mathbf{f}^2))\mathbf{D}_2 \\ &\quad + \mathbf{D}_2^T \text{diag}(\mathbf{D}_{12}\mathbf{f} \odot \psi''(\mathbf{D}\mathbf{f}^2))\mathbf{D}_1 \\ &\quad + \mathbf{D}_2^T \text{diag}((\mathbf{D}_2\mathbf{f})^2 \odot \psi''(\mathbf{D}\mathbf{f}^2))\mathbf{D}_2. \end{aligned} \quad (16)$$

For  $\psi(t) = \sqrt{t + \varepsilon}$ ,  $\psi'(t) = (t + \varepsilon)^{-1/2}$ ,  $\psi''(t) = -(t + \varepsilon)^{-3/2}/4$ , and  $\mathbf{D}_{12}\mathbf{f} := \mathbf{D}_1\mathbf{f} \odot \mathbf{D}_2\mathbf{f}$ . ‘ $\odot$ ’ indicates the component-wise multiplication.

Motivated by the work in [23], [24], and [38], we replace the least squares form of the numerator in (13) by a weighted least squares form  $N\|\mathbf{W}\mathbf{r}_\lambda\|_2^2$  due to the introduction of the reweighted techniques and the Talwar loss function. The construction of the weight  $\mathbf{W}$  is to satisfy the original functional form

$$\|\mathbf{W}\mathbf{r}_\lambda\|^2 = \sum_{i=1}^N \rho \left( \frac{[\mathbf{H}\mathbf{f}]_i - \mathbf{y}_i}{\sqrt{[\mathbf{H}\mathbf{f}]_i + \sigma^2}} \right). \quad (17)$$

We compare the numerator (13) and  $\|\mathbf{W}\mathbf{r}_\lambda\|_2^2$  in (17) and consider the definition of the Tawlar function in (6), and

hence write out the diagonal component  $[\mathbf{W}]_{ii}$  of the weight diagonal matrix  $\mathbf{W}$  as follows

$$[\mathbf{W}]_{ii} = \begin{cases} \frac{1}{\sqrt{[\mathbf{H}\mathbf{f}_\lambda]_i + \sigma^2}}, & \left| \frac{[\mathbf{H}\mathbf{f}_\lambda]_i - \mathbf{y}_i}{\sqrt{[\mathbf{H}\mathbf{f}_\lambda]_i + \sigma^2}} \right| \leq \beta, \\ \frac{\beta}{[\mathbf{H}\mathbf{f}_\lambda]_i - \mathbf{y}_i}, & \text{otherwise.} \end{cases} \quad (18)$$

Based on the work in [23] and [24], the aim is to minimize the functional with respect to the regularization parameter  $\lambda$

$$\text{GCV}(\lambda) = \frac{N\|\mathbf{W}\mathbf{r}_\lambda\|_2^2}{[\text{trace}(\mathbf{I} - \mathbf{A}_\lambda)]^2}, \quad (19)$$

where the influence matrix  $\mathbf{A}_\lambda$  in the GCV function (19) can be approximated as

$$\mathbf{A}_\lambda = \mathbf{W}\mathbf{H} \left[ \mathbf{D}_\lambda \left( \mathbf{H}^T \mathbf{W}^2 \mathbf{H} + \lambda \nabla^2 T_\varepsilon(\mathbf{f}_\lambda) \right) \mathbf{D}_\lambda \right]^\dagger \mathbf{D}_\lambda \mathbf{H}^T \mathbf{W}, \quad (20)$$

where  $\dagger$  denotes “pseudo-inverse” and  $\mathbf{D}_\lambda$  is a diagonal matrix  $[\mathbf{D}_\lambda]_{ii} = 1$  for  $[\mathbf{f}_\lambda]_i > 0$  and is 0 otherwise.

The GCV method selects the value of  $\lambda$  in (7) that minimizes the GCV functional

$$\lambda_{\text{GCV}} = \arg \min_\lambda \text{GCV}(\lambda). \quad (21)$$

## B. RANDOMIZED TRACE ESTIMATION

The presence of the matrix  $\mathbf{D}_\lambda$  in the expression  $\mathbf{A}_\lambda$  and the size of the  $\mathbf{A}_\lambda$  make the evaluation of the term  $\text{trace}(\mathbf{I} - \mathbf{A}_\lambda)$  impractical. A computationally cheaper random trace estimation is utilized to replace the direct calculation of the trace in the denominator of (19) by resorting to the fact  $\text{trace}(\mathbf{B}) \approx \mathbf{v}^T \mathbf{B} \mathbf{v}$  for  $\mathbf{B} \in \mathbb{R}^{N \times N}$ , where  $\mathbf{v}$  is a realization of a  $N \times 1$  random vector  $\mathbf{V}$  whose components take the values of 1 and  $-1$  each with probability 0.5 [38]. This fact motivates the trace approximation

$$[\text{trace}(\mathbf{I} - \mathbf{A}_\lambda)]^2 \approx [\mathbf{v}^T \mathbf{v} - \mathbf{v}^T \mathbf{A}_\lambda \mathbf{v}]^2. \quad (22)$$

Note that practically  $\mathbf{A}_\lambda$  is not directly computed due to the large-scale nature of the problem.  $\mathbf{A}_\lambda \mathbf{v}$  in (22) is approximated with  $\mathbf{W}\mathbf{H}\mathbf{u}_1$  and the vector  $\mathbf{u}_1$  is obtained by solving the following linear system with a truncated conjugate gradient technique

$$\left[ \mathbf{D}_\lambda \left( \mathbf{H}^T \mathbf{W}^2 \mathbf{H} + \lambda \nabla^2 T_\varepsilon(\mathbf{f}_\lambda) \right) \mathbf{D}_\lambda \right] \mathbf{u}_1 = \mathbf{D}_\lambda \mathbf{H}^T \mathbf{W} \mathbf{v}. \quad (23)$$

## C. THE CONSTRUCTION OF THE GCV FUNCTION FOR THE PARAMETER $\gamma$

Because the image and the PSF in the data fidelity term of (8) is symmetrical and the regularization operator in (8) is linear, the GCV function for the  $\lambda$  can be easily derived

$$\text{GCV}(\gamma) = \frac{N\|\mathbf{W}\mathbf{r}_\gamma\|_2^2}{[\text{trace}(\mathbf{I} - \mathbf{A}_\gamma)]^2}, \quad (24)$$

where the numerator  $\|\mathbf{W}\mathbf{r}_\gamma\|_2^2$  in (24) can be obtained similar to (17)

$$\|\mathbf{W}\mathbf{r}_\gamma\|_2^2 = \sum_{i=1}^N \rho \left( \frac{[\mathbf{F}\mathbf{h}]_i - \mathbf{y}_i}{\sqrt{[\mathbf{F}\mathbf{h}]_i + \sigma^2}} \right). \quad (25)$$

The weight  $\mathbf{W}$  is a diagonal matrix:

$$[\mathbf{W}]_{ii} = \begin{cases} \frac{1}{\sqrt{[\mathbf{F}\mathbf{h}]_i + \sigma^2}}, & \left| \frac{[\mathbf{F}\mathbf{h}]_i - \mathbf{y}_i}{\sqrt{[\mathbf{F}\mathbf{h}]_i + \sigma^2}} \right| \leq \beta, \\ \frac{\beta}{[\mathbf{F}\mathbf{h}]_i - \mathbf{y}_i}, & \text{otherwise.} \end{cases} \quad (26)$$

The influence matrix is given by

$$\mathbf{A}_\gamma = \mathbf{W}\mathbf{F} \left[ \mathbf{D}_\gamma (\mathbf{F}^T \mathbf{W}^2 \mathbf{F} + \gamma \mathbf{L}^T \mathbf{L}) \mathbf{D}_\gamma \right]^\dagger \mathbf{D}_\gamma \mathbf{F}^T \mathbf{W}, \quad (27)$$

where  $\dagger$  denotes “pseudo-inverse” and  $\mathbf{D}_\gamma$  is a diagonal matrix  $[\mathbf{D}_\gamma]_{ii} = 1$  for  $[\mathbf{h}_\gamma]_i > 0$  and is 0 otherwise.

In order to overcome the computational difficulty of large-scale problem in (24), the random trace estimation is again employed to approximate the denominator of (24) by

$$[\text{trace}(\mathbf{I} - \mathbf{A}_\gamma)]^2 \approx [\mathbf{v}^T \mathbf{v} - \mathbf{v}^T \mathbf{A}_\gamma \mathbf{v}]^2, \quad (28)$$

where  $\mathbf{v}$  is a realization of a  $N \times 1$  random vector  $\mathbf{V}$  whose components take on the values of 1 and  $-1$  each with probability 0.5 [38]. Note that  $\mathbf{A}_\gamma \mathbf{v}$  in (28) is approximated by  $\mathbf{W}\mathbf{F}\mathbf{u}_2$  and the vector  $\mathbf{u}_2$  is similarly obtained by solving the following linear system with a truncated conjugate gradient technique

$$\left[ \mathbf{D}_\gamma (\mathbf{F}^T \mathbf{W}^2 \mathbf{F} + \gamma \mathbf{L}^T \mathbf{L}) \mathbf{D}_\gamma \right] \mathbf{u}_2 = \mathbf{D}_\gamma \mathbf{F}^T \mathbf{W} \mathbf{v}. \quad (29)$$

In our implementation, we use Matlab’s fminbnd function to solve the minimization problem of the GCV function about the parameter. A lower bound of 0 and an upper bound of 1 are used in each case.

Taking all the above into account, we get the complete iteration algorithm for iteratively reweighted blind image deconvolution and it is summarized below in Algorithm 2.

#### IV. THE CONVERGENCE ANALYSIS

Several authors have discussed the convergence of the alternative minimization (AM) method for blind deconvolution, see [21], [37], [39], [40] for details. The variational regularization based blind image deconvolution usually leads to a challenging joint minimization problem with respect to the image  $\mathbf{f}$  and the PSF  $\mathbf{h}$ . Generally, the objective functional (5) is not jointly convex in the image variable  $\mathbf{f}$  and the PSF  $\mathbf{h}$ , so convergence of the AM method to a global minimum is not guaranteed and we may compute only a local minimizer of (5). However, it was proved in [39] that for the total variation (TV) based blind deconvolution model the AM procedure converges globally, but the solution depends on the initialization. More recently, Perrone and Favaro in [37] and [40] confirmed from a mathematical point of view that a variant of the total variation based blind deconvolution algorithm [20] can successfully converge to the desired solution, even when

---

#### Algorithm 2 Iteratively Reweighted Blind Deconvolution Algorithm

---

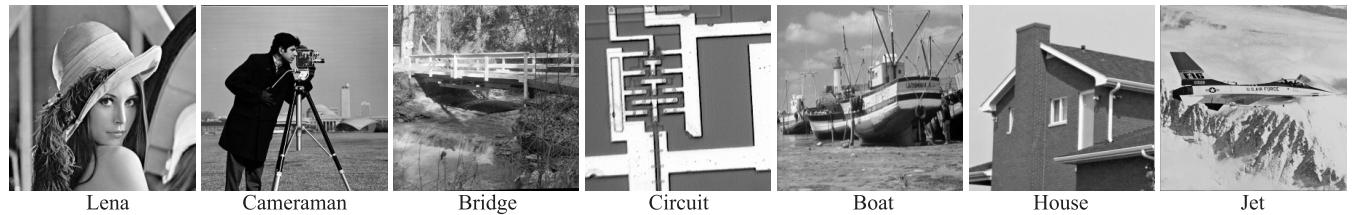
- 1: **Input:** The observed image  $\mathbf{y}$ , the maximum number of iterations  $K_{\text{MaxIter}}$ , and iteration stopping errors  $\epsilon_f$ ,  $\epsilon_h$ .
  - 2: **Initialization:**  $\mathbf{f}^{(0)} = \mathbf{y}$ .
  - 3: **while**  $k = 0$  to  $K_{\text{MaxIter}}$  and  $\|\mathbf{f}^{(k)} - \mathbf{f}^{(k-1)}\|_2 / \|\mathbf{f}^{(k-1)}\|_2 < \epsilon_f$  or  $\|\mathbf{h}^{(k)} - \mathbf{h}^{(k-1)}\|_2 / \|\mathbf{h}^{(k-1)}\|_2 < \epsilon_h$  **do**
  - 4:     Update  $k := k + 1$ ;
  - 5:     **% Updating procedure for f**
  - 6:     Restore the image  $\mathbf{f}$  with (7) and (9) using the L-BFGS-B method;
  - 7:     Compute the parameter  $\lambda^{(k)}$  based on Eq. (21);
  - 8:     **% Updating procedure for h**
  - 9:     Estimate the PSF  $\mathbf{h}$  with (8) and (10) using the L-BFGS-B method;
  - 10:    Compute the parameter  $\gamma^{(k)}$  based on Eq. (24);
  - 11:    **% Constraints on h<sup>(k)</sup>**
  - 12:    Impose the constraints on  $\mathbf{h}^{(k)}$  based on (11);
  - 13: **end while**
- 

starting at the no-blur solution. They found that the core element that makes the TV-like priors based blind deconvolution work is to separate the normalization and the positivity constraints from the minimization step with respect to the PSF. To further demonstrate and stress their finds, they implemented a variant of TV-based blind deconvolution by removing all the unnecessary recent improvements for blind deconvolution, such as filtering, blur kernel prior, edge enhancement, and optimized the subproblems with respect to the image and the PSF using the gradient descent method.

For the minimization subproblem with respect to the image in our model (7), we can show that the objective functional in (7) has positive semidefinite Hessian matrix, and hence it is a convex functional about the image, see [24] for similar convex analysis. This implies that the minimization of the objective functional (7) with respect to the image can converge to a global minimizer. A similar conclusion is also valid for the PSF in (8). Therefore, our algorithm framework can converge to the desired solution from a theoretical point of view. In the experiments, we observed that the algorithm consistently converged to visually high quality results.

#### V. EXPERIMENTAL RESULTS AND ANALYSIS

In this section, we present the performance of the proposed algorithm with both simulated and real data sets. As shown in Fig. 2, the seven test images Lena, Cameraman, Bridge, Circuit, House, Boat, and Jet are used for simulation. The size of the Jet image is  $512 \times 512$  and the size of each of other images is  $256 \times 256$ . All of the test images were normalized to between 0 and 1. In Algorithm 2, the algorithm is terminated when the convergence criterion  $\|\mathbf{f}^{(k)} - \mathbf{f}^{(k-1)}\|_2 / \|\mathbf{f}^{(k-1)}\|_2 < \epsilon_f$  or  $\|\mathbf{h}^{(k)} - \mathbf{h}^{(k-1)}\|_2 / \|\mathbf{h}^{(k-1)}\|_2 < \epsilon_h$  is satisfied where we set  $\epsilon_f = 1 \times 10^{-4}$  and  $\epsilon_h = 1 \times 10^{-4}$ . In our implementation, the maximum iteration number  $K_{\text{MaxIter}}$  of Algorithm 2 is set to be 100. The small constant  $\varepsilon$  is set to  $10^{-10}$ . For the



**FIGURE 2.** Seven test images for simulation. From left to right: Lena, Cameraman, Bridge, Circuit, Boat, House, and Jet.

L-BFGS-B method, the bounds on the image and the PSF are  $[0, +\infty]$  and  $[0, 1]$ , respectively. The inner number of iterations for each of the image and the PSF in the L-BFGS-B method is 100. The proposed algorithm is implemented in MATLAB.

We compare the proposed algorithm with the methods of Liao and Ng [14] and Wang and Ng [19], both of which are performed by using the GCV method to estimate the regularization parameter in a blind deconvolution problem. Liao et al.'s method (called as GCV-TVL2BD) is a very representative method in the total variation based blind image deconvolution, which has been proved to be superior to the variational Bayesian blind deconvolution methods in [41] and [42]. Wang et al.'s method (called as GCV-TVTVBD) [19] adopts the total variation regularization for the image and the PSF and is competitive in restoring real images.

In practical applications, the Gaussian noise standard deviation  $\sigma$  in (5) is unknown. Thus we adopt a robust estimator of the standard deviation in [43]

$$s = c \operatorname{med}_i \{\operatorname{med}_j \|\mathbf{x}_i - \mathbf{x}_j\| \}, \quad (30)$$

where  $\mathbf{x} \in \mathbb{R}^N$  is a vector and  $c = 1.1926$  is a constant factor. We can understand the formula (30) as follows. For each  $i$  we compute the median of  $\{\|\mathbf{x}_i - \mathbf{x}_j\|; j = 1, 2, \dots, N\}$ . This produces  $N$  values, the median of which gives our final estimate  $s$ . This estimator is more efficient than the median of the absolute deviations (MAD) rule and is still valid at asymmetric distributions (see [43]).

#### A. SIMULATED EXPERIMENTS FOR THE GAUSSIAN BLUR AND DISCUSSION

The images were degraded by the size  $11 \times 11$  Gaussian blur kernel with the standard deviation of 2.1. Then, the blurred images were contaminated by Gaussian noise with three noise standard deviations: 0.001, 0.003, and 0.005. Then, the resulting images were contaminated by Poisson noise with three noise levels: 500, 1500, and 5500. We employ the MATLAB function `fspecial()` to simulate the PSF. For the improvement in the restored image quality to be measured, the peak signal-to-noise ratio (PSNR) and Structural Similarity Index Measure (SSIM) [44] are adopted to quantitatively evaluate the quality of the observed degraded image  $\mathbf{y}$  and of the restored image  $\mathbf{f}$ . PSNR is calculated by  $10 \log(N/\|\mathbf{f} - \hat{\mathbf{f}}\|^2)$  where  $\mathbf{f}$  and  $\hat{\mathbf{f}}$  denote the original and the restored image, respectively.

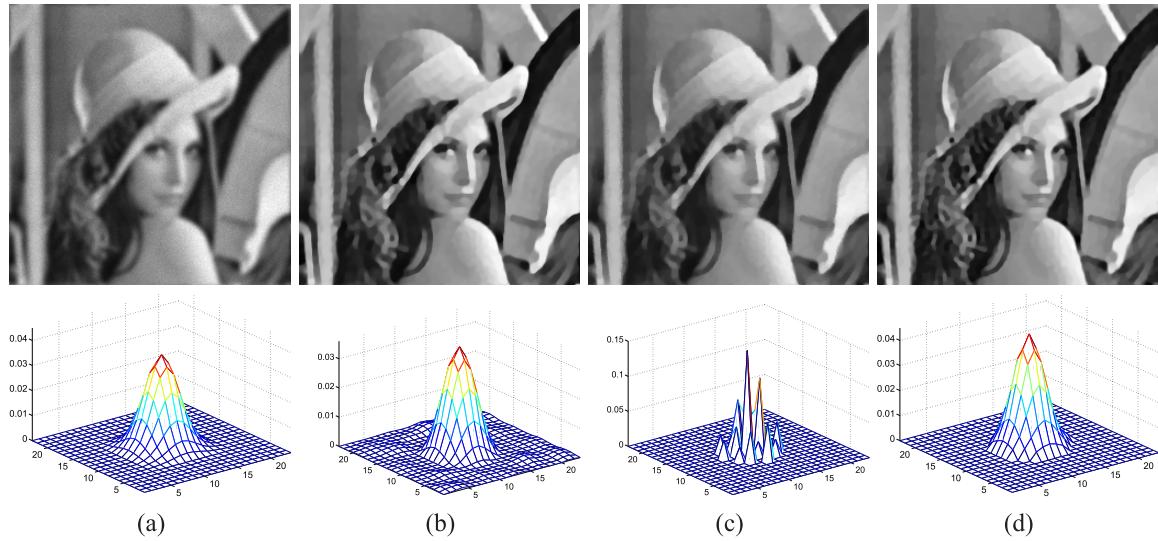
$N$  denotes the total number of image pixels. In all of the simulated experiments, the observed image is used as the initial estimation of the image  $\mathbf{f}^{(0)}$  and a Gaussian function is used as the initial estimation of the PSF  $\mathbf{h}^{(0)}$ .

Table 1 lists the PSNR and SSIM values of the degraded images and of the seven test images calculated by the three methods. It can be seen that the proposed method has achieved the highest PSNR and SSIM values among the three methods in Table 1.

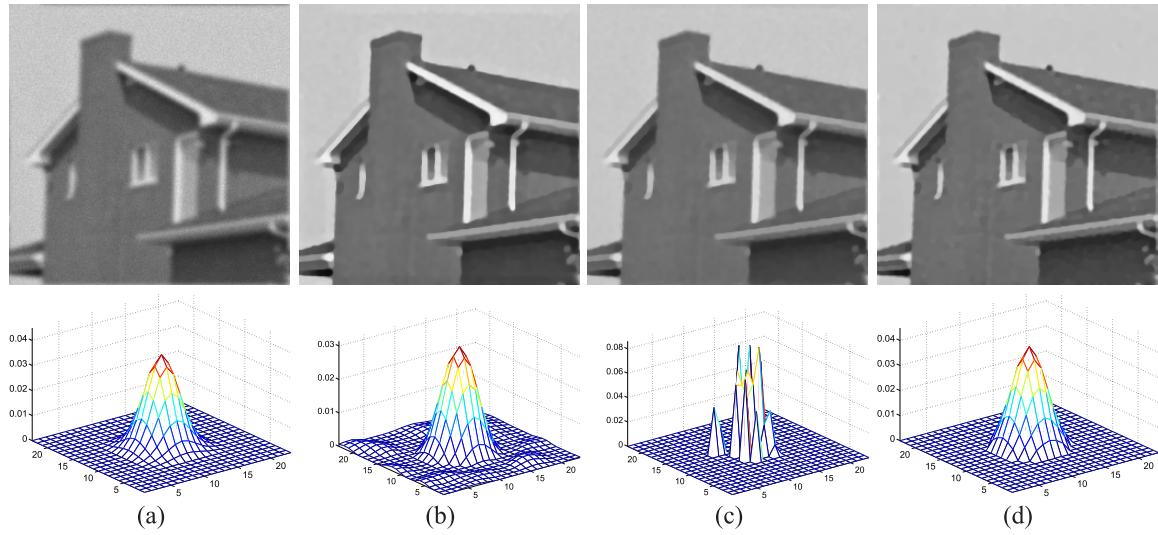
Figures 3 and 4 show the image deconvolution results under the same Gaussian noise and Poisson noise levels for the Lena image and the House image, respectively. The blurred and noisy images shown in Figs. 3(a) and 4(a) were obtained by adding Gaussian blur with `fspecial('gaussian', 11, 2.1)` and they are then contaminated by Gaussian noise with standard deviations 0.001 and Poisson noise levels  $I_{\max} = 1500$ . From Figs. 3(b) and 4(b), it is seen that the algorithm of Liao et al. can restore some details and suppress the mixed noises with a low noise levels, but there are still errors in the estimated PSFs. The results shown in Figs. 3(c) and 4(c) are obtained by the algorithm of Wang et al. Although the noise is suppressed well, we can still observe the residual blur and the estimated PSFs in Figs. 3(c) and 4(c) have very significant estimation errors. The main reason is that the total variation may not be a suitable prior for the Gaussian blur and thus the resulting estimated PSFs obviously deviated from the true PSFs. It is clear that the mixed noises are effectively removed in Figs. 3(d) and 4(d) by the proposed algorithm compared with those by the algorithms of Liao et al. and Wang et al. and the important structural details are preserved. The introduction of the reweighted strategy offers the proposed algorithm the ability to deal with the mixed noises, thus a better deconvolution performance. Figures 5 demonstrates the image deconvolution results for the Cameraman image. The blurred and noisy image shown in Fig. 5(a) was obtained by adding Gaussian blur with `fspecial('gaussian', 11, 2.1)` and it is then contaminated by Gaussian noise with standard deviations 0.003 and Poisson noise levels  $I_{\max} = 1500$ . It is seen that the recovered image by Liao et al.'s method in Fig. 5(b) degraded a little bit with a ringing effect that occurs along the edges of the image. The estimated PSF as shown in Fig. 5(b) has an obvious error compared with the true PSF in Fig. 5(a). It is shown in Fig. 5(c) that it tends to underestimate the PSF such that the result still looks blurry. It is seen from Fig. 5(d) that Algorithm 2 estimated the PSF very accurately

**TABLE 1.** PSNR (dB)/SSIM comparisons of three methods using seven test images for Gaussian blur with size  $11 \times 11$  and standard derivation of 2.1 and three different Gaussian and three different Poisson noise levels.

Images	Gaussian noise std.	$I_{\max}$	Observed	GCV-TVL2BD [14]	GCV-TVTVB[19]	Proposed
Lena	0.001	5500	22.51/0.6972	26.27/0.8158	25.41/0.7996	<b>26.75/0.8227</b>
		1500	22.63/0.6402	25.50/0.7844	24.99/0.7785	<b>26.26/0.8007</b>
		500	22.68/0.5437	24.83/0.7616	24.33/0.7555	<b>25.36/0.7773</b>
	0.003	5500	22.47/0.6950	26.17/0.8128	25.42/0.7984	<b>26.86/0.8247</b>
		1500	22.60/0.6385	25.62/0.7890	24.98/0.7779	<b>26.21/0.7991</b>
		500	22.72/0.5413	25.30/0.7636	24.54/0.7551	<b>26.13/0.7731</b>
	0.005	5500	22.46/0.6905	26.11/0.8108	25.36/0.7928	<b>26.68/0.8188</b>
		1500	22.56/0.6349	25.56/0.7866	24.94/0.7724	<b>26.26/0.7992</b>
		500	22.59/0.5387	24.92/0.7627	24.12/0.7491	<b>25.57/0.7755</b>
	0.001	5500	20.77/0.6742	23.16/0.7712	21.96/0.7380	<b>23.58/0.7910</b>
		1500	20.76/0.6022	22.68/0.7490	21.87/0.7346	<b>23.24/0.7783</b>
		500	20.62/0.4914	21.66/0.7240	21.18/0.7164	<b>22.35/0.7592</b>
Cameraman	0.003	5500	20.75/0.6706	22.59/0.7631	22.18/0.7509	<b>23.70/0.7929</b>
		1500	20.65/0.5994	22.51/0.7122	21.69/0.7345	<b>23.08/0.7750</b>
		500	20.72/0.4903	22.10/0.7258	21.51/0.7176	<b>22.85/0.7574</b>
	0.005	5500	20.73/0.6634	22.49/0.7607	22.18/0.7473	<b>23.57/0.7881</b>
		1500	21.03/0.5935	22.32/0.7442	21.95/0.7305	<b>23.19/0.7740</b>
		500	20.48/0.4878	21.96/0.7254	21.17/0.7129	<b>23.01/0.7559</b>
Circuit	0.001	5500	20.63/0.6694	24.68/0.8317	23.83/0.8052	<b>25.60/0.8444</b>
		1500	20.88/0.5946	23.80/0.8000	23.38/0.7893	<b>25.01/0.8260</b>
		500	21.09/0.4820	23.30/0.7779	22.90/0.7711	<b>24.12/0.8015</b>
	0.003	5500	20.63/0.6665	24.63/0.8282	23.80/0.8043	<b>25.52/0.8429</b>
		1500	20.88/0.5922	24.01/0.8075	23.37/0.7778	<b>24.92/0.8248</b>
		500	21.09/0.4804	23.32/0.7790	22.91/0.7708	<b>24.41/0.8041</b>
House	0.005	5500	20.62/0.6612	24.55/0.8264	23.72/0.7996	<b>25.38/0.8402</b>
		1500	20.88/0.5881	23.91/0.8050	23.25/0.7820	<b>24.96/0.8246</b>
		500	21.09/0.4778	23.18/0.7294	22.71/0.7200	<b>24.63/0.7988</b>
	0.001	5500	24.85/0.7308	29.48/0.8385	27.47/0.7994	<b>29.61/0.8420</b>
		1500	24.65/0.6428	28.22/0.8182	22.74/0.8121	<b>28.70/0.8277</b>
		500	24.27/0.5077	27.51/0.8040	27.46/0.8036	<b>28.00/0.8130</b>
Boat	0.003	5500	24.87/0.7273	29.53/0.8380	28.36/0.8246	<b>29.61/0.8418</b>
		1500	24.68/0.6400	28.46/0.8220	27.73/0.8120	<b>28.84/0.8275</b>
		500	24.27/0.5059	27.56/0.8044	26.99/0.7971	<b>28.10/0.8054</b>
	0.005	5500	24.89/0.7204	29.36/0.8361	28.00/0.8155	<b>29.46/0.8396</b>
		1500	24.69/0.6348	28.34/0.8206	27.39/0.8024	<b>28.82/0.8261</b>
		500	24.24/0.5024	27.52/0.8039	26.63/0.7863	<b>28.08/0.8094</b>
Bridge	0.001	5500	23.21/0.6523	25.64/0.7618	25.08/0.7446	<b>25.83/0.7624</b>
		1500	23.20/0.5993	25.04/0.7222	24.59/0.7147	<b>25.31/0.7304</b>
		500	22.92/0.5043	24.67/0.6913	24.17/0.6854	<b>25.25/0.7105</b>
	0.003	5500	23.24/0.6499	25.51/0.7561	24.48/0.7134	<b>25.78/0.7599</b>
		1500	23.21/0.5973	25.08/0.7241	24.58/0.7138	<b>25.45/0.7362</b>
		500	22.93/0.5022	24.65/0.6901	24.15/0.6841	<b>25.17/0.7073</b>
Jet	0.005	5500	23.25/0.6452	25.44/0.7518	24.94/0.7368	<b>25.77/0.7591</b>
		1500	23.21/0.5935	25.03/0.7219	24.52/0.7100	<b>25.44/0.7354</b>
		500	22.93/0.5000	24.64/0.6907	24.11/0.6807	<b>25.24/0.7076</b>
	0.001	5500	21.55/0.4891	23.76/0.6061	23.40/0.5898	<b>23.86/0.6307</b>
		1500	22.98/0.7359	23.30/0.5737	22.89/0.5536	<b>24.42/0.8500</b>
		500	21.96/0.4318	22.86/0.5322	22.44/0.5146	<b>23.06/0.5785</b>
	0.003	5500	21.58/0.4885	23.72/0.6062	22.83/0.5526	<b>23.87/0.6323</b>
		1500	21.79/0.4702	23.30/0.5749	22.89/0.5526	<b>23.51/0.6044</b>
		500	21.97/0.4312	22.92/0.5330	22.44/0.5137	<b>22.92/0.5655</b>
Bridge	0.005	5500	21.60/0.4870	23.62/0.6012	23.14/0.5760	<b>23.99/0.6408</b>
		1500	21.80/0.4689	23.22/0.5702	22.77/0.5459	<b>23.53/0.6070</b>
		500	21.99/0.4318	22.87/0.5325	22.29/0.5059	<b>23.19/0.5770</b>
	0.001	5500	24.25/0.8841	26.72/0.8453	26.67/0.8450	<b>27.45/0.9515</b>
		1500	24.26/0.8499	26.81/0.8351	25.99/0.8255	<b>27.05/0.9263</b>
		500	23.95/0.7750	25.94/0.8090	25.41/0.8047	<b>26.47/0.8731</b>
Jet	0.003	5500	24.24/0.8830	27.09/0.8498	26.62/0.8439	<b>27.59/0.9512</b>
		1500	24.25/0.8490	26.83/0.8351	26.03/0.8256	<b>27.04/0.9352</b>
		500	23.94/0.7740	26.02/0.8125	25.43/0.8048	<b>26.45/0.8943</b>
	0.005	5500	24.22/0.8811	27.15/0.8492	26.31/0.8420	<b>27.47/0.9495</b>
		1500	24.26/0.8478	26.49/0.8305	25.72/0.8246	<b>27.09/0.9320</b>
		500	23.87/0.7718	25.92/0.8087	25.09/0.8026	<b>26.49/0.8842</b>



**FIGURE 3.** Blind deconvolution results for the Lena image. (a) Image blurred by `fspecial('gaussian',11,2.1)` and contaminated by Gaussian noise (standard deviation  $\sigma = 0.001$ ) and Poisson noise ( $I_{\max} = 1500$ ) (PSNR = 22.63 dB, SSIM = 0.6402). Deconvolution results achieved using (b) GCV-TVL2BD [14] (PSNR = 25.50 dB, SSIM = 0.7844), (c) GCV-TVTVBD [19] (PSNR = 24.99 dB, SSIM = 0.7785), (d) our method (PSNR = 26.26 dB, SSIM = 0.8007). The bottom line: the original PSF and the estimated PSFs by three methods.

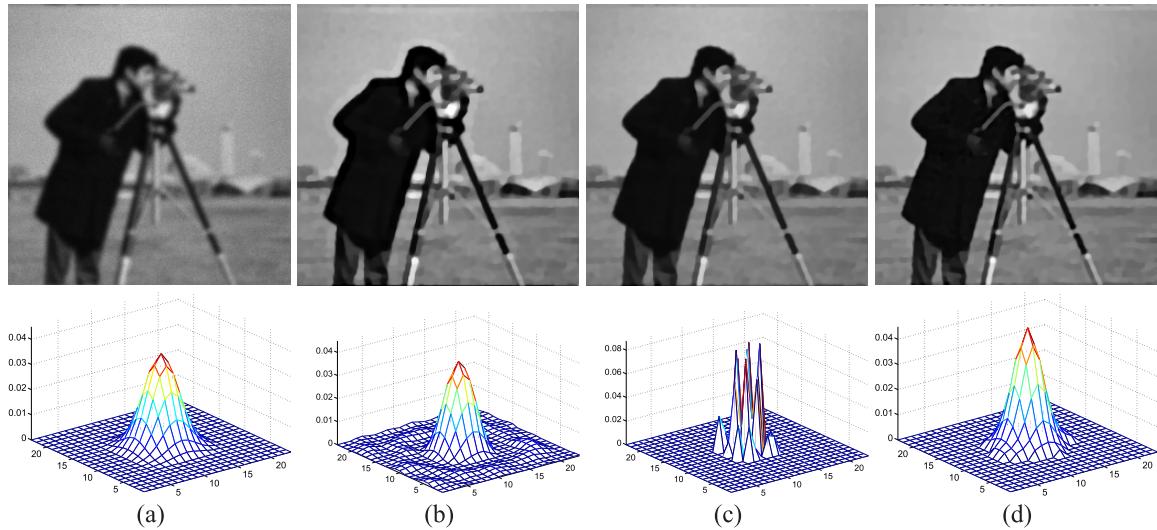


**FIGURE 4.** Blind deconvolution results for the House image. (a) Image blurred by `fspecial('gaussian',11,2.1)` and contaminated by Gaussian noise (standard deviation  $\sigma = 0.001$ ) and Poisson noise ( $I_{\max} = 1500$ ) (PSNR = 24.65 dB, SSIM = 0.6428). Deconvolution results achieved using (b) GCV-TVL2BD [14] (PSNR = 28.22 dB, SSIM = 0.8182), (c) GCV-TVTVBD [19] (PSNR = 22.74 dB, SSIM = 0.8121), (d) our method (PSNR = 28.70 dB, SSIM = 0.8277). The bottom line: the original PSF and the estimated PSFs by three methods.

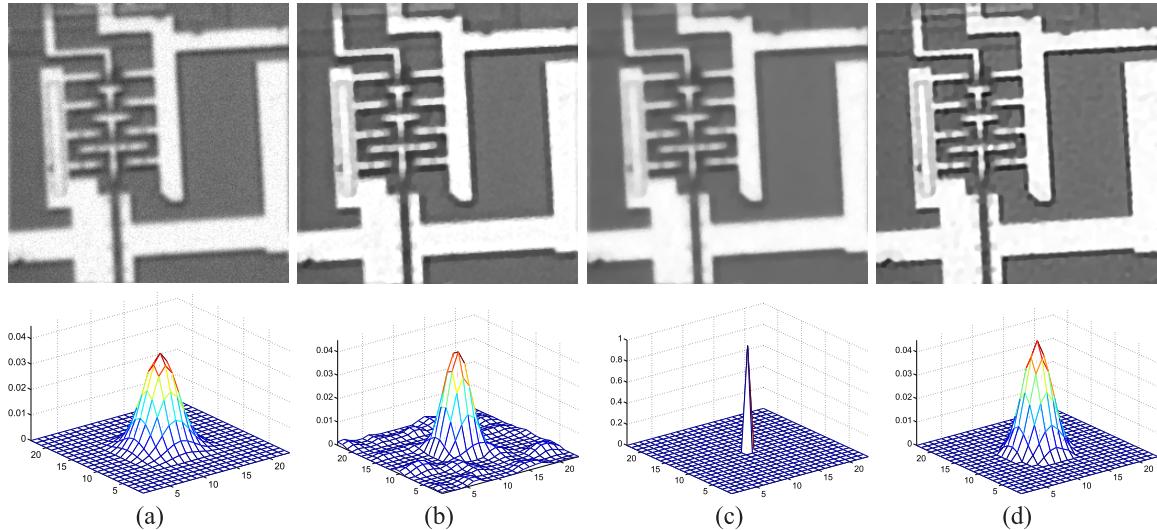
and yielded a high-quality deconvolution image with little artifacts.

Figures 6 shows the image deconvolution results for the Circuit image when the mixed noises are strong. The blurred and noisy images shown in Figs. 6(a) was obtained by adding Gaussian blur with `fspecial('gaussian',11,2.1)` and it is then contaminated by Gaussian noise with standard deviations 0.005 and Poisson noise levels  $I_{\max} = 500$ . It is seen that there exists remarkable residual noise in Fig. 6(b) recovered by the Liao et al.’s method. The resulting estimated PSF

from Wang et al.’s method deviated from the true PSF. Thus, the result on tested image from Wang et al.’s method is not satisfactory and there exists significant blurring effects. In contrast, the result from the proposed method has the least artifacts among all. The proposed method combines robust regression and the reweighted technique and thus the mixed strong noises can be removed well and the estimated PSF is very close to the original true PSF. Both of Liao et al.’s method and the proposed method imposed the smoothness constraint of PSFs. This is sound because the simulated PSF



**FIGURE 5.** Blind deconvolution results for the Cameraman image. (a) Image blurred by `fspecial('gaussian',11,2.1)` and contaminated by Gaussian noise (standard deviation  $\sigma = 0.003$ ) and Poisson noise ( $I_{\max} = 1500$ ) (PSNR = 20.65 dB, SSIM = 0.5994). Deconvolution results achieved using (b) GCV-TVL2BD [14] (PSNR = 22.51 dB, SSIM = 0.7122), (c) GCV-TTVB [19] (PSNR = 21.69 dB, SSIM = 0.7345), (d) our method (PSNR = 23.08 dB, SSIM = 0.7750). The bottom line: the original PSF and the estimated PSFs by three methods.



**FIGURE 6.** Blind deconvolution results for the Circuit image. (a) Image blurred by `fspecial('gaussian',11,2.1)` and contaminated by Gaussian noise (standard deviation  $\sigma = 0.005$ ) and Poisson noise ( $I_{\max} = 500$ ) (PSNR = 21.09 dB, SSIM = 0.4778). Deconvolution results achieved using (b) GCV-TVL2BD [14] (PSNR = 23.18 dB, SSIM = 0.7294), (c) GCV-TTVB [19] (PSNR = 22.71 dB, SSIM = 0.7200), (d) our method (PSNR = 24.63 dB, SSIM = 0.7988). The bottom line: the original PSF and the estimated PSFs by three methods.

is a Gaussian kernel. Therefore, the shape of the estimated PSFs obtained by the Liao et al.'s method and the proposed method is close to the original. However, the mixed strong noises still lead to an obvious noise error in the estimated PSF by Liao et al.'s method. The proposed method can provide better deconvolution results and is more visually pleasant with fewer artifacts than those by the other two methods.

Overall, compared with other methods, Algorithm 2 consistently performed well over these images, and the results are of good quality with few noticeable image artifacts. It is

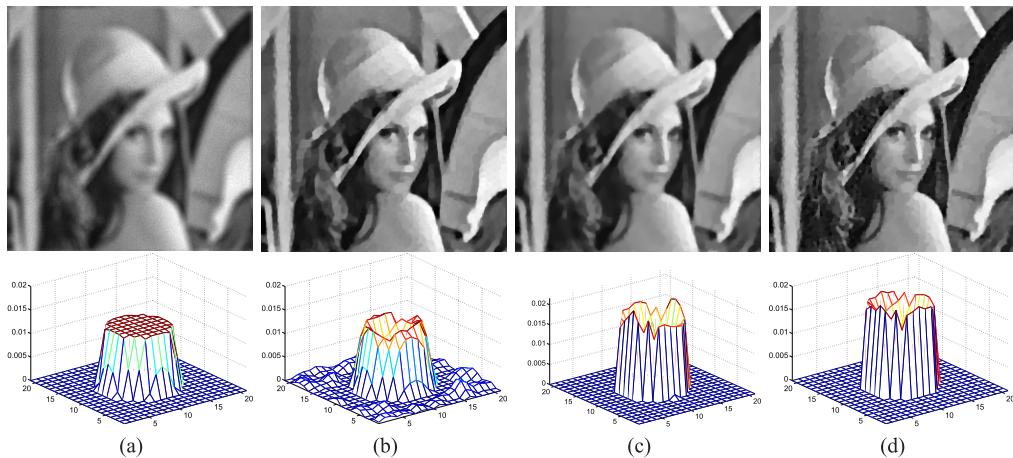
seen that Algorithm 2 has notable advantages over the other two methods in terms of PSNR values and visual quality.

## B. BLIND DECONVOLUTION FOR THE OUT-OF-FOCUS BLUR

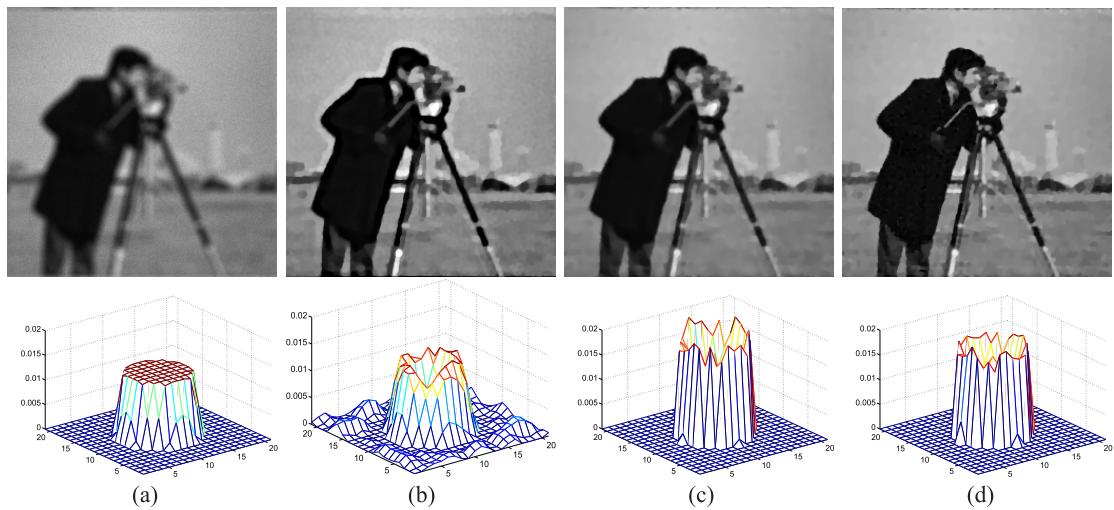
In this section, the performance of the proposed blind deconvolution method is demonstrated for the images which are blurred by the out-of-focus blur with the size  $5 \times 5$  and are then polluted by the Gaussian noise with the standard deviation  $\sigma = 0.001$  and Poisson noise with the maximum

**TABLE 2.** PSNR (dB)/SSIM comparisons of three methods using seven test images for Gaussian blur with size  $11 \times 11$  and standard derivation of 2.1 and the standard deviation  $\sigma = 0.001$  and Poisson noise with the maximum intensity  $I_{\max} = 1500$ .

Images	Observed	GCV-TVL2BD [14]	GCV-TVTVBD [19]	Proposed
Lena	20.93/0.5432	25.22/0.7460	24.87/0.7504	<b>25.98/0.7744</b>
Cameraman	19.31/0.5315	22.97/0.7069	22.62/0.7445	<b>23.36/0.7592</b>
Bridge	20.37/0.3435	23.09/0.5316	22.96/0.5304	<b>23.91/0.5711</b>
Circuit	18.99/0.4801	23.76/0.7444	24.43/0.7733	<b>25.10/0.8008</b>
Boat	21.64/0.5004	24.44/0.6847	24.45/0.6862	<b>25.06/0.7047</b>
House	22.84/0.5785	27.98/0.7876	27.27/0.7923	<b>28.64/0.8046</b>
Jet	22.61/0.7777	26.31/0.8533	25.35/0.8557	<b>26.62/0.8650</b>



**FIGURE 7.** Blind deconvolution results for the Lena image. (a) Image blurred by `fspecial('disk',5)` and contaminated by Gaussian noise (standard deviation  $\sigma = 0.001$ ) and Poisson noise ( $I_{\max} = 1500$ ) (PSNR = 20.93 dB, SSIM = 0.5412). Deconvolution results achieved using (b) GCV-TVL2BD [14] (PSNR = 25.22 dB, SSIM = 0.7460), (c) GCV-TVTVBD [19] (PSNR = 24.87 dB, SSIM = 0.7504), (d) our method (PSNR = 25.98 dB, SSIM = 0.7744). The bottom line: the original PSF and the estimated PSFs by three methods.

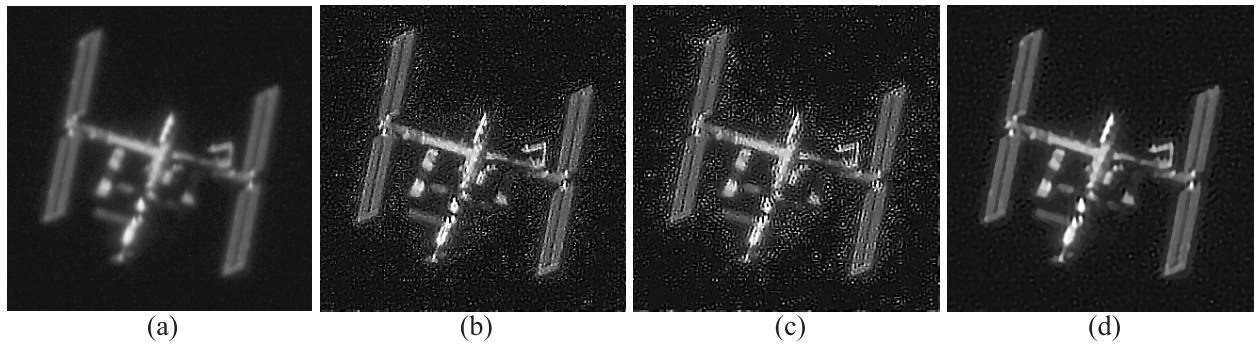


**FIGURE 8.** Blind deconvolution results for the Cameraman image. (a) Image blurred by `fspecial('disk',5)` and contaminated by Gaussian noise (standard deviation  $\sigma = 0.001$ ) and Poisson noise ( $I_{\max} = 1500$ ) (PSNR = 19.31 dB, SSIM = 0.5315). Deconvolution results achieved using (b) GCV-TVL2BD [14] (PSNR = 22.97 dB, SSIM = 0.7069), (c) GCV-TVTVBD [19] (PSNR = 22.62 dB, SSIM = 0.7445), (d) our method (PSNR = 23.36 dB, SSIM = 0.7592). The bottom line: the original PSF and the estimated PSFs by three methods.

intensity  $I_{\max} = 1500$  and compared to the methods of Liao and Ng [14] and Wang and Ng [19].

The quantitative assessments in terms of PSNR and SSIM are listed in Table 2. It can be seen that the proposed method

consistently provides the best PSNR and SSIM values for seven different images. Figs. 7 and 8 are the simulated results achieved by three methods. Figures 7(b) and 8(b) show the restored images by GCV-TVL2BD, from which



**FIGURE 9.** Blind deconvolution results for the real observed image. (a) Real degraded international space station image. Deconvolution results achieved using (b) GCV-TVL2BD [14], (c) GCV-TVTBVD [19], (d) our method.

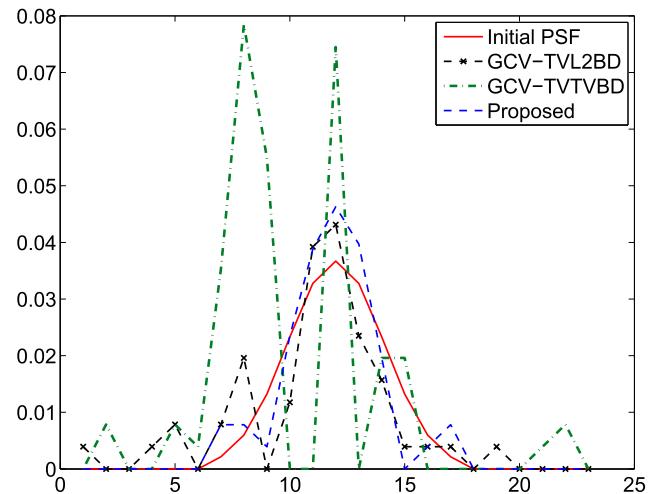
we observe that the GCV-TVL2BD method can suppress the mixed noises to a certain extent, but at the cost that the image details are also removed. The PSFs in Figs. 7(b) and 8(b) also contain a slight amount error, which can be affected by the mixed noises. Figures 7(c) and 8(c) show the restored images by GCV-TVTBVD, where the total variation regularization is imposed for the PSF in this method, we can see that the shape of the estimated PSF is close to the true PSF. However, it is seen that the image detail information of recovered images from GCV-TVTBVD is unexpectedly degraded. Figures 7(d) and 8(d) are the results of the proposed method, from which we note that the mixed noises are removed to a large extent and the edges and details information are restored and preserved well. Compared to other evaluated methods, the proposed method performs consistently over these images and the results have better visual quality.

### C. REAL DATA EXPERIMENTS

In the second part of the experiments, the proposed method is applied to a real observed image of the international space station, which was taken with the ground based telescope in Munich, shown in Fig. 9(a).<sup>1</sup> The main difficulties behind the blind deconvolution for real images are two-fold. The first is that there is no prior knowledge about the noise and its statistical distribution. The second is that there is no exact mathematical expression for the shape of the PSF for this kind of image. The astronomical observation is commonly degraded by atmospheric turbulence. According to our previous studies in restoring this kind of astronomical observation images [3], [4], [28], an approximation Gaussian-like or Gaussian blur kernel for the initial PSF of the blind deconvolution algorithm can be applied. We adopt a  $23 \times 23$  Gaussian blur kernel with the standard deviation of 2.1 as the initial estimate of PSF for the three methods.

Fig. 9(b) and 9(c) represent the restored results of GCV-TVL2BD in [14] and GCV-TVTBVD in [19]. We can see from Fig. 9(b) and 9(c) that the mixed noises are not removed well and are amplified remarkably. The reason may be that the type of the mixed noises are more complex in real astronomical observation image that may be contaminated

<sup>1</sup><http://www.iss-tracking.de/images/stationpic.html>



**FIGURE 10.** One-dimensional slices through the origin of the theoretical and initial PSFs in the restoration of the international space station image for the three methods.

not only by Poisson or Gaussian noise but also by other unknown type of noises. The proposed method yields a reasonably good restoration result, as shown in Fig. 9(d). The reason is that the proposed method incorporates a robust regression with the reweighted mechanism into the data term that makes the proposed blind deconvolution robust to multiple mixed noises, and thus the multiple mixed noises can be removed to a large extent. Therefore, the proposed method can produce a high quality restoration. The estimated PSFs by the three methods and the initial PSF are also shown in Fig. 10. It is clear that the estimated PSF by the proposed method is relatively smoother than that of GCV-TVL2BD and GCV-TVTBVD.

### VI. CONCLUSION

In this paper, we present a blind deconvolution method by incorporating a robust regression data term with the reweighted least square criterion and a TV-based regularization term for the image and a Laplacian regularization term for the PSF into a variational regularization framework. The image and the PSF are estimated using the L-BFGS-B optimization algorithm within an alternating minimization

framework. We have extended the variant of the GCV functional to the total variation regularization case for estimating the regularization parameters. Experimental results demonstrate that the GCV criterion-based regularization parameter choice can yield satisfactory restorations for a broader class of images under the mixed noises.

The proposed method also has some limitations. In the aspect of the computational complexity, the introduced algorithm is more computationally intensive than the GCV-TVL2BD and GCV-TVTVBD methods since (7) and (8) cannot be solved directly in the frequency domain or by simply employing the existing methods and a numerical optimization approach to a large scale problem is needed. The regularization parameters estimation by the nonlinear GCV function requires computing the minimizer of (21) and (24). We also need to find the image  $\mathbf{f}$  and the PSF  $\mathbf{h}$  inside the GCV function using the iterative optimization technique. It is time-consuming and unsuited for real-time applications for the moment. A more efficient parameter search method for minimizing (21) and (24) than fminbnd in MATLAB can be also developed.

## APPENDIX I.

### DERIVATION OF THE GRADIENT EQUATIONS

In this Appendix we derive the gradient formulas (9) and (10) that correspond to the objective functional (7) and (8), respectively.

(1) The derivation of the gradient of objective functional  $J(\mathbf{f}, \mathbf{h}^{(k-1)})$  about the image  $\mathbf{f}$ .

When  $\left| \frac{(\mathbf{H}\mathbf{f})_i - b_i}{\sqrt{(\mathbf{H}\mathbf{f})_i + \sigma^2}} \right| \leq \beta$ , the energy functional  $J(\mathbf{f}, \mathbf{h}^{(k-1)})$  can be written as follows:

$$J(\mathbf{f}, \mathbf{h}^{(k-1)}) = \sum_{i=1}^N \rho \left( \frac{[\mathbf{H}^{(k-1)}\mathbf{f}]_i - y_i}{\sqrt{[\mathbf{H}^{(k-1)}\mathbf{f}]_i + \sigma^2}} \right) + \lambda \|\mathbf{D}\mathbf{f}\|_{1,\varepsilon}, \quad (31)$$

To conveniently derive the formula, we remove the superscript  $(k-1)$  and rewrite the formula (31) into:

$$J_1(\mathbf{f}) = \frac{1}{2} \sum_{i=1}^N \left( \frac{\mathbf{H}\mathbf{f} - \mathbf{y}}{\sqrt{\mathbf{H}\mathbf{f} + \sigma^2}} \right)_i^2 + \frac{\lambda}{2} \sum_{i=1}^N \psi \left( [\mathbf{D}_1\mathbf{f}]_i^2 + [\mathbf{D}_2\mathbf{f}]_i^2 \right). \quad (32)$$

Given  $\mathbf{f} \in \mathbb{R}^N$ , to compute the gradient of  $J_1$  at  $\mathbf{f}$  note that for any  $\eta \in \mathbb{R}^N$

$$\frac{d}{d\tau} J_1(\mathbf{f} + \tau\eta) \Big|_{\tau=0} = \frac{d}{d\tau} \left[ J_1^I + J_1^{II} \right] \Big|_{\tau=0}, \quad (33)$$

where  $\tau$  is a small constant and

$$J_1^I = \frac{1}{2} \sum_{i=1}^N \frac{(\mathbf{H}\mathbf{f} + \tau\mathbf{H}\eta - \mathbf{y})_i^2}{[\mathbf{H}\mathbf{f} + \tau\mathbf{H}\eta]_i + \sigma^2}, \quad (34)$$

$$J_1^{II} = \lambda \sum_{i=1}^N \psi \left( [\mathbf{D}_1\mathbf{f} + \tau\mathbf{D}_1\eta]_i^2 + [\mathbf{D}_2\mathbf{f} + \tau\mathbf{D}_2\eta]_i^2 \right). \quad (35)$$

For the first part

$$\begin{aligned} \frac{d}{d\tau} J_1^I \Big|_{\tau=0} &= \frac{d}{d\tau} \left( \frac{1}{2} \sum_{i=1}^N \frac{(\mathbf{H}\mathbf{f} + \tau\mathbf{H}\eta - \mathbf{y})_i^2}{[\mathbf{H}\mathbf{f} + \tau\mathbf{H}\eta]_i + \sigma^2} \right) \Big|_{\tau=0} \\ &= \frac{1}{2} \sum_{i=1}^N \frac{d}{d\tau} \left( \frac{(\mathbf{H}\mathbf{f} + \tau\mathbf{H}\eta - \mathbf{y})_i^2}{[\mathbf{H}\mathbf{f} + \tau\mathbf{H}\eta]_i + \sigma^2} \right) \Big|_{\tau=0} \\ &= \frac{1}{2} \sum_{i=1}^N \frac{(\mathbf{H}\mathbf{f} - \mathbf{y})_i (\mathbf{H}\mathbf{f} + \mathbf{y} + 2\sigma^2)_i}{([\mathbf{H}\mathbf{f}]_i + \sigma^2)^2} [\mathbf{H}\eta]_i \\ &= \frac{1}{2} \sum_{i=1}^N \frac{([\mathbf{H}\mathbf{f}]_i + \sigma^2)^2 - (y_i + \sigma^2)^2}{([\mathbf{H}\mathbf{f}]_i + \sigma^2)^2} [\mathbf{H}\eta]_i \\ &= \frac{1}{2} \sum_{i=1}^N \left[ 1 - \left( \frac{y_i + \sigma^2}{[\mathbf{H}\mathbf{f}]_i + \sigma^2} \right)^2 \right] [\mathbf{H}\eta]_i \\ &= \frac{1}{2} \left\langle \left[ 1 - \left( \frac{y_i + \sigma^2}{[\mathbf{H}\mathbf{f}]_i + \sigma^2} \right)^2 \right], \mathbf{H}\eta \right\rangle \\ &= \frac{1}{2} \left\langle \mathbf{H}^T \left[ 1 - \left( \frac{y_i + \sigma^2}{[\mathbf{H}\mathbf{f}]_i + \sigma^2} \right)^2 \right], \eta \right\rangle, \end{aligned} \quad (36)$$

where  $\langle \cdot, \cdot \rangle$  denotes the inner product.

For the second part

$$\begin{aligned} \frac{d}{d\tau} J_1^{II} \Big|_{\tau=0} &= \lambda \sum_{i=1}^N \psi' \left( [\mathbf{D}_1\mathbf{f} + \tau\mathbf{D}_1\eta]_i^2 + [\mathbf{D}_2\mathbf{f} + \tau\mathbf{D}_2\eta]_i^2 \right) \Big|_{\tau=0} \\ &= \sum_{i=1}^N \psi' \left( [\mathbf{D}_1\mathbf{f}]_i^2 + [\mathbf{D}_2\mathbf{f}]_i^2 \right) ([\mathbf{D}_1\mathbf{f}]_i [\mathbf{D}_1\eta]_i + [\mathbf{D}_2\mathbf{f}]_i [\mathbf{D}_2\eta]_i) \\ &= \left\langle \text{diag}(\psi'(\mathbf{D}\mathbf{f}^2)) \mathbf{D}_1\mathbf{f}, \mathbf{D}_1\eta \right\rangle + \left\langle \text{diag}(\psi'(\mathbf{D}\mathbf{f}^2)) \mathbf{D}_2\mathbf{f}, \mathbf{D}_2\eta \right\rangle \\ &= \left\langle \mathbf{D}_1^T \text{diag}(\psi'(\mathbf{D}\mathbf{f}^2)) \mathbf{D}_1\mathbf{f} + \mathbf{D}_2^T \text{diag}(\psi'(\mathbf{D}\mathbf{f}^2)) \mathbf{D}_2\mathbf{f}, \eta \right\rangle. \end{aligned} \quad (37)$$

Adding (36) and (37) and using the fundamental lemma of calculus of variations [38], we obtain the gradient of the objective functional  $J_1(\mathbf{f})$  about the image  $\mathbf{f}$

$$\begin{aligned} \text{grad}J_1(\mathbf{f})(\mathbf{f}) &= \frac{1}{2} \mathbf{H}^T \left[ 1 - \left( \frac{y_i + \sigma^2}{[\mathbf{H}\mathbf{f}]_i + \sigma^2} \right)^2 \right] \\ &\quad + \mathbf{D}_1^T \text{diag}(\psi'(\mathbf{D}\mathbf{f}^2)) \mathbf{D}_1\mathbf{f} \\ &\quad + \mathbf{D}_2^T \text{diag}(\psi'(\mathbf{D}\mathbf{f}^2)) \mathbf{D}_2\mathbf{f}, \end{aligned} \quad (38)$$

when  $\left| \frac{\mathbf{H}\mathbf{f} - \mathbf{y}}{\sqrt{\mathbf{H}\mathbf{f} + \sigma^2}} \right| \leq \beta$ .

For  $\left| \frac{\mathbf{H}\mathbf{f} - \mathbf{y}}{\sqrt{\mathbf{H}\mathbf{f} + \sigma^2}} \right| > \beta$ , the gradient in the first part of the functional (7) is 0.

Hence, the final gradient of objective functional  $J_1(\mathbf{f})$  about the image  $\mathbf{f}$  is given by

$$\begin{aligned} \text{grad}J_1(\mathbf{f})(\mathbf{f}) &= \mathbf{H}^T \mathbf{z} + \lambda \nabla T_\varepsilon(\mathbf{f}), \\ z_i &= \begin{cases} \frac{1}{2} - \frac{1}{2} \left( \frac{y_i + \sigma^2}{[\mathbf{H}\mathbf{f}]_i + \sigma^2} \right), & \left| \frac{[\mathbf{H}\mathbf{f}]_i - y_i}{\sqrt{[\mathbf{H}\mathbf{f}]_i + \sigma^2}} \right| \leq \beta, \\ 0, & \text{otherwise,} \end{cases} \end{aligned} \quad (39)$$

where the superscript  $T$  denotes the transpose operation and the gradient of the total variation regularization term over the image  $\nabla T_\varepsilon(\mathbf{f}) = L_1(\mathbf{f})\mathbf{f}$ ,  $L_1(\mathbf{f}) = \mathbf{D}_1^T \text{diag}(\psi'(\mathbf{D}\mathbf{f}^2))\mathbf{D}_1 + \mathbf{D}_2^T \text{diag}(\psi'(\mathbf{D}\mathbf{f}^2))\mathbf{D}_2$ ,  $\psi(t) = \sqrt{t + \varepsilon}$ ,  $\mathbf{D}\mathbf{f}^2 := (\mathbf{D}_1\mathbf{f})^2 + (\mathbf{D}_2\mathbf{f})^2$ .

(2) The derivation of the gradient of objective functional  $J(\mathbf{f}^{(k)}, \mathbf{h})$  about the image  $\mathbf{h}$ . When  $\left| \frac{(\mathbf{F}\mathbf{h})_i - \mathbf{y}_i}{\sqrt{(\mathbf{F}\mathbf{h})_i + \sigma^2}} \right| \leq \beta$ , the objective functional  $J(\mathbf{f}^{(k)}, \mathbf{h})$  can be written as follows:

$$J(\mathbf{f}^{(k)}, \mathbf{h}) = \sum_{i=1}^N \rho \left( \frac{[\mathbf{F}^{(k)}\mathbf{h}]_i - \mathbf{y}_i}{\sqrt{[\mathbf{F}^{(k)}\mathbf{h}]_i + \sigma^2}} \right) + \frac{\gamma}{2} \|\mathbf{L}\mathbf{h}\|_2^2, \quad (40)$$

To conveniently derive the formula, we remove the superscript  $(k)$  and rewrite the formula (40):

$$J_2(\mathbf{h}) = \frac{1}{2} \sum_{i=1}^N \left( \frac{\mathbf{F}\mathbf{h} - \mathbf{y}}{\sqrt{\mathbf{F}\mathbf{h} + \sigma^2}} \right)_i^2 + \frac{\gamma}{2} \sum_{i=1}^N (\mathbf{L}\mathbf{h})_i^2. \quad (41)$$

Given  $\mathbf{h} \in \mathbb{R}^N$ , to compute the gradient of  $J_2$  at  $\mathbf{h}$  note that for any  $\xi \in \mathbb{R}^N$

$$\frac{d}{d\mu} J_2(\mathbf{f} + \mu\xi) \Big|_{\mu=0} = \frac{d}{d\mu} \left[ J_2^I + J_2^{II} \right] \Big|_{\mu=0}, \quad (42)$$

where  $\mu$  is a small constant and

$$J_2^I = \frac{1}{2} \sum_{i=1}^N \frac{(\mathbf{F}\mathbf{h} + \mu\mathbf{F}\xi - \mathbf{y})_i^2}{[\mathbf{F}\mathbf{h} + \mu\mathbf{F}\xi]_i + \sigma^2}, \quad (43)$$

$$J_2^{II} = \frac{\gamma}{2} \sum_{i=1}^N (\mathbf{L}\mathbf{h} + \mu\mathbf{L}\xi)_i^2. \quad (44)$$

For the first part, similar to (1), we can easily obtain the gradient as follows

$$\text{grad}J_2^I(\mathbf{h}) = \mathbf{F}^T \left( \frac{1}{2} - \frac{1}{2} \left( \frac{\mathbf{y} + \sigma^2}{\mathbf{F}\mathbf{h} + \sigma^2} \right)^2 \right). \quad (45)$$

For the second part,  $J_2^{II}$  is a quadratic functional and hence its gradient can be easily obtained by

$$\text{grad}J_2^{II}(\mathbf{h}) = \gamma \mathbf{L}^T \mathbf{L}\mathbf{h}. \quad (46)$$

Adding (45) and (46) and we get the gradient of the functional  $J_2$  about the PSF  $\mathbf{h}$

$$\text{grad}J_2(\mathbf{h}) = \mathbf{F}^T \left( \frac{1}{2} - \frac{1}{2} \left( \frac{\mathbf{y} + \sigma^2}{\mathbf{F}\mathbf{h} + \sigma^2} \right)^2 \right) + \gamma \mathbf{L}^T \mathbf{L}\mathbf{h}. \quad (47)$$

For  $\left| \frac{\mathbf{F}\mathbf{h} - \mathbf{y}}{\sqrt{\mathbf{F}\mathbf{h} + \sigma^2}} \right| > \beta$ , the gradient in the first part of the functional (8) is  $\mathbf{0}$ .

Hence, the final gradient of objective functional  $J_2(\mathbf{h})$  about the PSF  $\mathbf{h}$  is given by

$$\begin{aligned} \text{grad}J_2(\mathbf{h})(\mathbf{h}) &= \mathbf{F}^T \mathbf{z} + \gamma \mathbf{L}^T \mathbf{L}\mathbf{h}, \\ z_i &= \begin{cases} \frac{1}{2} - \frac{1}{2} \left( \frac{\mathbf{y}_i + \sigma^2}{[\mathbf{F}\mathbf{h}]_i + \sigma^2} \right)^2, & \left| \frac{[\mathbf{F}\mathbf{h}]_i - \mathbf{y}_i}{\sqrt{[\mathbf{F}\mathbf{h}]_i + \sigma^2}} \right| \leq \beta, \\ 0, & \text{otherwise.} \end{cases} \end{aligned} \quad (48)$$

## ACKNOWLEDGMENT

The authors are grateful to the Associate Editor Gustavo Callico and the reviewers for their valuable time and the comments and Dr. Ming Zhang for improving the grammar. The authors would like to thank Marie Kubínová for her discussions pertaining to the work of [24].

## REFERENCES

- [1] P. Campisi and K. Egiazarian, Eds., *Blind Image Deconvolution: Theory and Applications*. New York, NY, USA: CRC Press, 2007.
- [2] C. L. Matson, K. Borelli, S. Jefferies, J. C. C. Beckner, E. K. Hege, and M. Lloyd-Hart, “Fast and optimal multiframe blind deconvolution algorithm for high-resolution ground-based imaging of space objects,” *Appl. Opt.*, vol. 48, no. 1, pp. A75–A92, 2009.
- [3] L. Yan, H. Fang, and S. Zhong, “Blind image deconvolution with spatially adaptive total variation regularization,” *Opt. Lett.*, vol. 37, no. 14, pp. 2778–2780, 2012.
- [4] H. Fang, L. Yan, H. Liu, and Y. Chang, “Blind Poissonian images deconvolution with framelet regularization,” *Opt. Lett.*, vol. 38, no. 4, pp. 389–391, 2013.
- [5] H. Liu, J. Gu, M. Q.-H. Meng, and W.-S. Lu, “Fast weighted total variation regularization algorithm for blur identification and image restoration,” *IEEE Access*, vol. 4, pp. 6792–6801, 2016.
- [6] D. L. Snyder, A. M. Hammoud, and R. L. White, “Image recovery from data acquired with a charge-coupled-device camera,” *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 10, no. 5, pp. 1014–1023, 1993.
- [7] D. L. Snyder, C. W. Helstrom, A. D. Lanterman, M. Faisal, and R. L. White, “Compensation for readout noise in CCD images,” *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 12, no. 2, pp. 272–283, 1995.
- [8] F. Benvenuto, A. La Camera, C. Theys, A. Ferrari, H. Lantéri, and M. Bertero, “The study of an iterative method for the reconstruction of images corrupted by Poisson and Gaussian noise,” *Inverse Problems*, vol. 24, no. 3, pp. 035016-1–035016-20, 2008.
- [9] H. Lantéri and C. Theys, “Restoration of astrophysical images—The case of Poisson data with additive Gaussian noise,” *EURASIP J. Appl. Signal Process.*, vol. 2005, no. 15, pp. 2500–2513, 2005.
- [10] J. Li, Z. Shen, R. Jin, and X. Zhang, “A reweighted  $\ell_2$  method for image restoration with Poisson and mixed Poisson–Gaussian noise,” *Inverse Problems Imag.*, vol. 9, no. 3, pp. 875–894, 2015.
- [11] E. Chouzenoux, A. Jezierska, J.-C. Pesquet, and H. Talbot, “A convex approach for image restoration with exact Poisson–Gaussian likelihood,” *SIAM J. Imag. Sci.*, vol. 8, no. 4, pp. 2662–2682, 2015.
- [12] M. Kubínová and J. G. Nagy, “Iteratively reweighted deconvolution and robust regression,” in *Proc. 16th AMOSTECH Conf.*, 2015, pp. 1–7.
- [13] H. Fang, Y. Shi, D. Pan, and G. Zhou, “Iteratively reweighted blind deconvolution for passive millimeter-wave images,” *Signal Process.*, vol. 138, pp. 182–194, Sep. 2017.
- [14] H. Liao and M. K. Ng, “Blind deconvolution using generalized cross-validation approach to regularization parameter estimation,” *IEEE Trans. Image Process.*, vol. 20, no. 3, pp. 670–680, Mar. 2011.
- [15] Y. Huang, M. K. Ng, and Y.-W. Wen, “A fast total variation minimization method for image restoration,” *SIAM J. Multiscale Model. Simul.*, vol. 7, no. 2, pp. 774–795, 2008.
- [16] G. Wahba, “Practical approximate solutions to linear operator equations when the data are noisy,” *SIAM J. Numer. Anal.*, vol. 14, no. 4, pp. 651–667, 1977.
- [17] G. H. Golub, M. Heath, and G. Wahba, “Generalized cross-validation as a method for choosing a good ridge parameter,” *Technometrics*, vol. 21, no. 2, pp. 215–223, 1979.
- [18] G. H. Golub and U. von Matt, “Generalized cross-validation for large-scale problems,” *J. Comput. Graph. Stat.*, vol. 6, no. 1, pp. 1–34, 1997.
- [19] W. Wang and M. K. Ng, “On algorithms for automatic deblurring from a single image,” *J. Comput. Math.*, vol. 30, no. 1, pp. 80–100, 2012.
- [20] T. F. Chan and C.-K. Wong, “Total variation blind deconvolution,” *IEEE Trans. Image Process.*, vol. 7, no. 3, pp. 370–375, Mar. 1998.
- [21] J.-F. Cai, H. Ji, C. Liu, and Z. Shen, “Framelet-based blind motion deblurring from a single image,” *IEEE Trans. Image Process.*, vol. 21, no. 2, pp. 562–572, Feb. 2012.
- [22] J. M. Bardsley and J. Goldes, “Regularization parameter selection methods for ill-posed Poisson maximum likelihood estimation,” *Inverse Problems*, vol. 25, no. 1, pp. 095005-1–095005-18, 2009.

- [23] J. M. Bardsley and J. Goldes, "Regularization parameter selection and an efficient algorithm for total variation-regularized positron emission tomography," *Numer. Algorithms*, vol. 57, no. 2, pp. 255–271, 2011.
- [24] M. Kubíková and J. G. Nagy. (2017). "Robust regression for mixed Poisson-Gaussian model." [Online]. Available: <https://arxiv.org/abs/1611.07774v2>
- [25] L. Yujiri, M. Shoucri, and P. Moffa, "Passive millimeter wave imaging," *IEEE Microw. Mag.*, vol. 4, no. 3, pp. 39–50, Sep. 2003.
- [26] H. Fang and L. Yan, "Parametric blind deconvolution for passive millimeter wave images with framelet regularization," *Optik*, vol. 125, no. 3, pp. 1460–1545, 2014.
- [27] A. Beck and M. Teboulle, "Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems," *IEEE Trans. Image Process.*, vol. 18, no. 11, pp. 2419–2434, Nov. 2009.
- [28] L. Yan, M. Jin, H. Fang, H. Liu, and T. Zhang, "Atmospheric-turbulence-degraded astronomical image restoration by minimizing second-order central moment," *IEEE Geosci. Remote Sens. Lett.*, vol. 9, no. 4, pp. 672–676, Jul. 2012.
- [29] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Phys. D, Nonlinear Phenomena*, vol. 60, nos. 1–4, pp. 259–268, 1992.
- [30] L. I. Rudin and S. Osher, "Total variation based image restoration with free local constraints," in *Proc. IEEE Int. Conf. Image Process.*, vol. 1, Nov. 1994, pp. 31–35.
- [31] T. F. Chan and S. Esedoglu, "Aspects of total variation regularized  $L^1$  function approximation," *SIAM J. Appl. Math.*, vol. 35, no. 5, pp. 1817–1837, 2005.
- [32] D. E. Coleman, P. W. Holland, N. Kaden, V. Klema, and S. C. Peters, "A system of subroutines for iteratively reweighted least squares computations," *ACM Trans. Math. Softw.*, vol. 6, no. 3, pp. 327–336, 1980.
- [33] C. Zhu, R. H. Byrd, P. Lu, and J. Nocedal, "Algorithm 778: L-BFGS-B: Fortran subroutines for large-scale bound-constrained optimization," *ACM Trans. Math. Softw.*, vol. 23, no. 4, pp. 550–560, 1997.
- [34] D. Thompson, B. Calef, and M. Werth, "Performance comparison of optimization methods for blind deconvolution," in *Proc. Adv. Maui Opt. Space Surveill. Technol. Conf. (AMOS)*, 2016, pp. 1–11.
- [35] J. Nocedal and S. J. Wright, *Numerical Optimization*, 2nd ed. New York, NY, USA: Springer, 2006.
- [36] R. S. Dembo and T. Steihaug, "Truncated-Newton algorithms for large-scale unconstrained optimization," *Math. Programm.*, vol. 26, no. 2, pp. 190–212, 1983.
- [37] D. Perrone and P. Favaro, "Total variation blind deconvolution: The devil is in the details," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 2909–2916.
- [38] C. R. Vogel, *Computational Methods for Inverse Problems*. Philadelphia, PA, USA: SIAM, 2002.
- [39] T. F. Chan and C. K. Wong, "Convergence of the alternating minimization algorithm for blind deconvolution," *Linear Algebra Appl.*, vol. 316, nos. 1–3, pp. 259–285, Sep. 2000.
- [40] D. Perrone and P. Favaro, "A clearer picture of total variation blind deconvolution," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 38, no. 6, pp. 1041–1055, Jun. 2016.
- [41] S. D. Babacan, R. Molina, and A. K. Katsaggelos, "Variational Bayesian blind deconvolution using a total variation prior," *IEEE Trans. Image Process.*, vol. 18, no. 1, pp. 12–26, Jan. 2009.
- [42] D. G. Tzirkas, A. C. Likas, and N. P. Galatsanos, "Variational Bayesian sparse kernel-based blind image deconvolution with student's-t priors," *IEEE Trans. Image Process.*, vol. 18, no. 4, pp. 753–764, Apr. 2009.
- [43] P. J. Rousseeuw and C. Croux, "Alternatives to the median absolute deviation," *J. Amer. Statist. Assoc.*, vol. 88, no. 424, pp. 1273–1283, 1993.
- [44] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.

**HOUZHANG FANG** received the M.S. degree in applied mathematics and the Ph.D. degree in control science and engineering from the Huazhong University of Science and Technology, in 2010 and 2014, respectively. He is currently a Lecturer with Xidian University. His current research interests include image restoration, blind deconvolution, super resolution, and object detection and recognition.



**YI CHANG** received the B.S. degree in automation from the University of Electronic Science and Technology of China, Chengdu, China, in 2011, and the M.S. degree in pattern recognition and intelligent systems from the Huazhong University of Science and Technology, China, in 2014, where he is currently pursuing the Ph.D. degree with the School of Automation.

From 2014 to 2015, he was a Research Assistant with Peking University, Beijing, China. His current research interests include image processing, computer vision, and machine learning.



**GANG ZHOU** was born in Hunan Province, China, in 1984. He received the degree from the Changsha University of Science and Technology, the B.S. degree in 2006, the M.S. degree in statistic and signals process from Guizhou Nationalities University in 2010, and the Ph.D. degree in pattern recognition and intelligent system from the Huazhong University of Science and Technology in 2016. He is currently a Research Fellow with the State Key Laboratory of Material Processing and Die & Mould Technology, Huazhong University of Science and Technology, Wuhan, China. His main researches include camera calibration, structured light 3-D measuring, super-resolution, moving detection, and pattern recognition.

**LIZHEN DENG** received the B.S. degree in electronic information science and technology from Huaibei Coal Industry Teachers College, Huaibei, China, in 2007, the M.S. degree in communication and information system from the Nanjing University of Aeronautics and Astronautics, Nanjing, China, in 2010, and the Ph.D. degree in electrical engineering from the Huazhong University of Science and Technology, China, in 2014. In 2014, she joined the Nanjing University of Posts and Telecommunications, Nanjing. Her current research interests include image processing, computer vision, pattern recognition, and spectral data processing.