Blind multi-Poissonian image deconvolution with sparse log-step gradient prior

WENDE DONG,[1,2,†] **QIXIANG WANG,**[1,†] **SHUYIN TAO,**[3] **AND CHAO TIAN**[4,5,*]

[1]*College of Automation Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China*
[2]*Key Laboratory of Space Photoelectric Detection and Perception (Nanjing University of Aeronautics and Astronautics), Ministry of Industry and Information Technology, Nanjing 211106, China*
[3]*School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing 210094, China*
[4]*Institute of Artificial Intelligence, Hefei Comprehensive National Science Center, Hefei, Anhui, 230088, China*
[5]*School of Engineering Science, University of Science and Technology of China, Hefei, Anhui, 230026, China*
[†]*These authors contributed equally to this work*
[*]*ctian@ustc.edu.cn*

Abstract: Blind image deconvolution plays a very important role in the fields such as astronomical observation and fluorescence microscopy imaging, in which the noise follows Poisson distribution. However, due to the ill-posedness, it is a very challenging task to reach a satisfactory result from a single blurred image especially when the power of the Poisson noise is at a high level. Therefore, in this paper, we try to achieve high-quality restoration results with multi-blurred images which are contaminated by Poisson noise. Firstly, we design a novel sparse log-step gradient prior which adopts a mixture of logarithm and step functions to regularize the image gradients and combine it with the Poisson distribution to formulate the blind multi-image deconvolution problem. Secondly, we incorporate the methods of variable splitting and Lagrange multiplier to convert the original problem into sub-problems, then we alternately solve them to achieve the estimation of all the blur kernels of corresponding blurred images. Besides, we also design a non-blind multi-image deconvolution algorithm which is based on the log-step gradient prior to reach the final restored image. Experimental results on both synthetic and real-world blurred images show that the proposed prior is very capable of suppressing negative artifacts caused by ill-posedness. The algorithm can achieve restored image of very high quality which is competitive with some state-of-the-art methods.

1. Introduction

Image blur is usually modeled by convolving the latent image with a blur kernel plus different types of noise. Astronomical observation and fluorescence microscopy imaging are two typical fields in which the Poisson noise dominates. In astronomical observation, because of low illumination, long exposure time has to be taken, the obtained image tends to be blurred due to mechanical vibration or atmospheric disturbance. While in fluorescence microscopy imaging, since the depth of field is very small, blur caused by defocus occurs frequently. Besides, in order to reduce the possibility of photo-damage to the sample caused by excitation laser, the exposure time must be controlled very short which means the Poisson noise in the obtained image may be of high level.

Image deconvolution is the inverse estimation of the latent image from the blurred. According to whether the blur kernel is known, it is divided into two categories, i.e., non-blind deconvolution

and blind deconvolution. In non-blind deconvolution, the blur kernel is measured by some auxiliary means and only the latent image needs to be restored, while in blind deconvolution, both the blur kernel and latent image needs to be estimated. However, image deconvolution is inherently ill-posed which means that the solution tends to be contaminated by amplified noise and ringing artifacts [1]. The type and level of the noise, the energy distribution of the blur kernel significantly affect the quality of the result, thus it is necessary to introduce some extra constraints to form new regularized problems whose solutions are more stable.

Maximum a-posteriori (MAP) estimation is one of the most commonly used methods to construct regularized deconvolution problem models. Take blind image deconvolution as example, it combines the distribution of the noise, the prior distributions of the latent image and blur kernel to construct an optimization problem. In practice, researchers usually adopt Gaussian distribution to model the noise and have proposed many approaches [2–6], but in some specific environments such as astronomical observation and fluorescence microscopy imaging, the noise evidently follows Poisson distribution, it is unreasonable to model the problem with Gaussian distribution, this is the reason why we focus on Poissonian image deconvolution in this paper (for simplicity, we use the word "Poissonian" to denote "Poisson noise contaminated" or similar expressions in this paper).

In the related work, almost all the researches focus on non-blind or blind single Poissonian image deconvolution. Nowadays, with the development of modern imaging systems, it is very easy to capture multiple images of the same scene. Therefore, under poor imaging conditions where image blur and heavy Poisson noise is inevitable, we can obtain multi-blurred images of the target. Because multi-blurred images contain complementary information, we may use them to achieve better results than using single blurred image. Therefore, we aim to obtain restored images of high quality from multi-Poissonian blurred images in this paper. The main contributions are as follows:

Firstly, we construct a sparse prior which combines the logarithm and step functions to penalize the square of image gradients (for simplicity, we call it log-step gradient prior in the following parts). In contrast to the classic priors such as total variation (TV) and L_0-norm based priors, it is more capable of suppressing negative artifacts under heavy Poisson noise.

Secondly, we prove that the core problem accompanied with the log-step gradient prior has a closed-form solution. Based on which, we convert the original problem into easier sub-problems and design an efficient optimization algorithm to achieve the result.

We use synthetic and real-world blurred images to evaluate the proposed algorithm, experimental results show that it outperforms some state-of-the-art methods.

The rest of this paper is organized as follows. In Section 2, we make a detailed introduction of the related work. In Section 3, we present the sparse log-step gradient prior and formulate the problem. In Section 4, we demonstrate the optimization approach. In Section 5, we test our approach on both synthetic and real-world blurred images. Finally, we make a conclusion in Section 6.

2. Related work

Regularization is very important for image deconvolution. In MAP estimation, constructing regularization terms is equivalent to designing priors which can formulate the probabilistic distributions of the latent image and blur kernel. Markov Random Fields theory and Hammersley-Clifford theorem [7] provide an exercisable way for constructing priors. Roughly, there are mainly two kinds of priors, i.e., Gaussian prior and sparse prior. The typical Gaussian priors such as Tikhonov regularization [8,9] introduces quadric term in the problem model which is easy to solve, but the restored image tends to be ambiguous on edges and degraded by negative artifacts, thus Gaussian priors are not very suitable for image deconvolution in the condition of heavy noise. In contrast, sparse priors are very capable of suppressing negative artifacts, e.g., in recent

years, the L_1-norm and L_p-norm ($0 < p < 1$) based priors have been widely used in non-blind image deconvolution and proven to be very effective [10,11]. The L_0-norm based prior has been successfully adopted in blind image deconvolution because it can result in restored image with sharp edges and smooth regions which is very beneficial for blur kernel estimation [12,13]. Besides, there are also some robust sparse priors of other types, e.g., the Fields of Experts [14,15] and Gaussian Scale Mixture Fields of Experts [16] which are trained from natural images.

How to design efficient algorithms to solve the sparse prior regularized deconvolution problem is another important research field. Specific to non-blind Poissonian image deconvolution, since the Poissonian fidelity term is non-quadratic, it is not an easy task to solve the regularized Poissonian deconvolution problem. In [17] and [18], the authors propose a one-step-late framework which is derived from the Expectation-Maximization (EM) method, many subsequent researches take it as reference, e.g., in [19] and [20], the one-step-late method is used to solve the deconvolution problem regularized by TV and natural image gradient prior, respectively. In [21] and [22], it is used to solve the Tikhonov and a sparse representation regularized deconvolution problem to improve the quality of the 4Pi-microscopy image. In [23] and [24], the authors propose the residual RL (RL is short for Richardson-Lucy algorithm [25,26]), gain controlled RL and joint bilateral RL which are also based on the one-step-late framework. However, the problem is that if the noise power is at a high level, the one-step-late method is of low efficiency and may not converge to satisfactory result. Another classic framework for solving the non-blind Poissonian deconvolution problem is the alternating direction method of multipliers (ADMM) [27], which converts the original problem into simpler sub-problems and alternately solve them to reach the final result. Similarly, the authors of [28] propose to use the split Bregman approach to simplify the original problem. In contrast to the one-step-late scheme, the two methods are more efficient and stable, therefore, just like the one-step-late approach, the ADMM and split Bregman methods are also widely used as guidance to solve various regularized Poissonian deconvolution problems [29–31]. Furthermore, they are also improved for better performance, e.g., the authors of [32] propose a fast algorithm based on the variable splitting approach to solve the TV regularized deconvolution problem, while [33] utilizes the second-order Taylor expansion to linearize the quadratic term and design a general inverse operator to accelerate the iterative steps.

In blind Poissonian image deconvolution, the above two frameworks are also very widely used, e.g., in [34], the Huber-Markov random field is used to regularize the latent image, the original problem is first divided into two non-blind deconvolution problems, one is with respect to the latent image and the other is with respect to the blur kernel, they are alternately solved with the one-step-late approach to reach the result. Similarly, the methods in [35] and [36] propose to use the sparse L_1-norm constraint on wavelet or framelet coefficients of the latent image, the authors adopt the split Bregman method to solve for the latent image, the blur kernel is also estimated with the one-step-late method. In [37], the authors use the L_0-norm to model the redundant analysis (e.g., curvelet or framelet) of the latent image and TV to regularize the blur kernel, the latent image is estimated with the Greedy Analysis Pursuit (GAP) algorithm but the blur kernel is still estimated with the one-step-late method. In [38], the authors incorporate the L_0-norm operator and TV to regularize the latent image and blur kernel, respectively. The main contribution is that they completely abandon the one-step-late scheme. Instead, they design an approach which combines the methods of variable splitting and Lagrange multiplier to convert the original problem into three sub-problems, then propose an alternating minimization algorithm which can estimate the blur kernel, latent image and Lagrange multiplier simultaneously. However, although there have been many single blind image deconvolution algorithms, it shows that when the blur kernel is large or the noise level is high, it is very difficult for them to reach acceptable result due to ill-posedness.

In multi-image deconvolution, more than one blurred versions of the same latent image are provided. Since the blur kernels for generating multi-blurred images may be different one from the other and the noise is random, the information degraded in one blurred image may be retained in others, the restored image may be of better quality than that of the single blind image deconvolution. Until now, most of the blind multi-image deconvolution methods consider Gaussian noise, e.g., the method in [39] takes Laplacian operator to construct Tikhonov regularization terms to estimate the latent image and blur kernels. The authors of [40] propose a method which takes TV for regularization, and [41] uses the results of [40] as initial input and propose a blur kernel refinement method. In [42], the Huber-Markov random field is used for regularization. In [43] and [44], a coupled penalty function for the latent image and blur kernels is proposed, it links the unknown latent image along with the blur kernel and noise variance of each observation, all the variables are estimated jointly under a marginalization framework. In [45], the variational Bayesian approach is used to solve the multi-image deconvolution problem, which can adjust the parameters automatically and achieve higher image quality. In [46], the latent image and blur kernels are represented by groups of sparse domains to exploit the local and nonlocal information, the sparse solution is promoted by using the split-Bregman algorithm to solve the L_1-norm optimization problem. In [47], a blind multi-image deconvolution algorithm which is based on the low-rank representation is presented. In [48], a new method relies on a novel sparse prior named smoothed normal with unknown variance (NUV) is proposed, the prior promotes piecewise smooth image with sharp edges which is very useful for blur kernel estimation. In [49], the authors improve the blind multi-image deconvolution method proposed in [50], which uses the hyper-Laplacian prior (L_p-norm of image gradients) and L_1-norm to regularize the latent image and blur kernels, respectively. The method has been used very successfully in restoring the real multi-blurred images captured by an imaging system with rotational rectangular aperture. Besides the above mentioned, there are also some multi-image deconvolution methods which take different types of degraded images as input, e.g., the approach with noisy blurred pairs [23,51].

3. Log-step gradient prior and problem model

3.1. *Log-step gradient prior*

Just as mentioned in Section 2, sparse priors play a very important role in improving the ill-posedness of deconvolution problem. In the following paragraphs, we will make a detailed analysis of some classic sparse priors and demonstrate the characteristics of the proposed prior.

In Fig. 1, the black dotted line denotes the real probabilistic distribution of horizontal gradients of the natural image "cameraman.tif" which is commonly used in image processing. The green, blue and cyan curves denote the L_0-norm (Eq. (1)), L_1-norm (Eq. (2)), and L_p-norm (Eq. (3)) gradient priors, respectively. To make a clear comparison, the image gradient is normalized to [-1, 1] and the logarithms of the real distribution and different priors are normalized to [0, 1]. Meanwhile, in Fig. 2, we present a blurred image contaminated by heavy noise and the corresponding restored images with these sparse priors.

$$L_0(\mathbf{o}) = \sum_{i=1}^{N} Step(|\mathbf{d}_x\mathbf{o}|_i + |\mathbf{d}_y\mathbf{o}|_i), \qquad (1)$$

$$L_1(\mathbf{o}) = ||\mathbf{d}_x\mathbf{o}||_1 + ||\mathbf{d}_y\mathbf{o}||_1, \qquad (2)$$

$$L_p(\mathbf{o}) = ||\mathbf{d}_x\mathbf{o}||_p^p + ||\mathbf{d}_y\mathbf{o}||_p^p \quad (0<p<1), \qquad (3)$$

where $\mathbf{o}$ is the latent image, $\mathbf{d}_x$ and $\mathbf{d}_y$ represent the convolution matrices of the horizontal and vertical difference operators, respectively. N is the number of pixels in $\mathbf{o}$ and i is the pixel index.

The Step function is defined as follow.

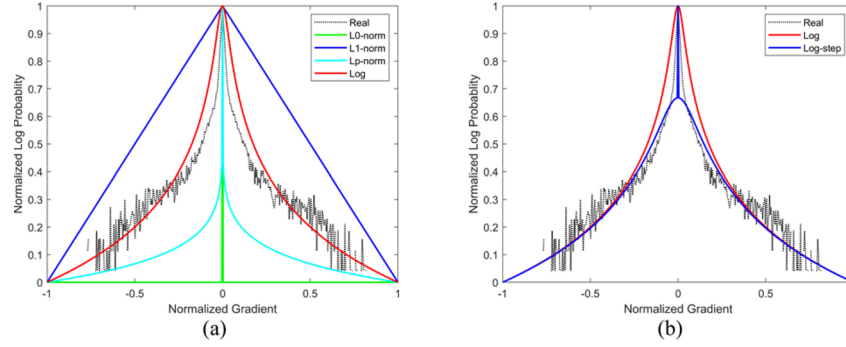$$Step(x) = \begin{cases} 1, & x > 0 \\ 0, & x \le 0 \end{cases}. \tag{4}$$

Fig. 1. Comparison of curves of different sparse priors.
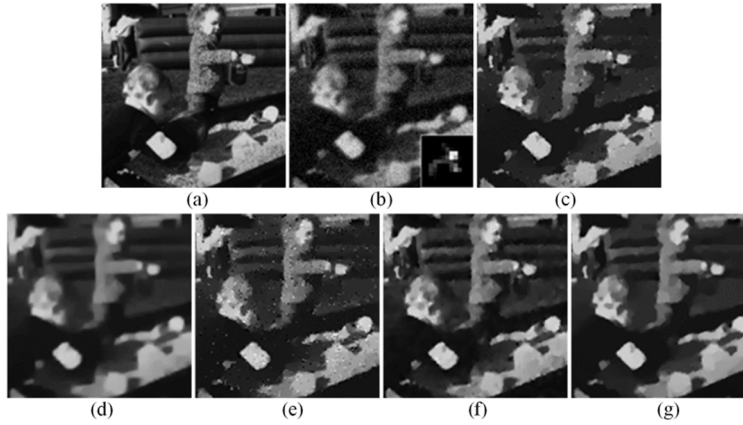

Fig. 2. Performances of different sparse priors. (a) Clear image. (b) Blurred image and blur kernel. (c) Result of L_0-norm gradient prior. (d) Result of L_1-norm gradient prior. (e) Result of L_p-norm gradient prior. (f) Result of log gradient prior. (g) Result of log-step gradient prior.

As pointed out in [52–55], since the image details cannot be accurately estimated in the intermediately restored image in iterative blind image deconvolution methods, they should be suppressed along with the negative artifacts by sparse priors to ensure the quality of estimated blur kernel. In other words, the accuracy of the estimated blur kernel mainly depends on the salient edges in the intermediately restored image. We can see from Fig. 1 (a) that the L_0-norm gradient prior has a non-distinctive heavy penalty on all the gradients except the zero-gradient, thus it tends to suppress image details as much as possible, the edges in the resulted image will be sharpened and some useful edges may also be wrongly removed. Just as shown in Fig. 2 (c), the restored image of L_0-norm gradient prior has sharp edges but some useful contents in the background are removed. In contrast, the L_1-norm gradient prior has a linear penalty on the gradients, it is capable to preserve more image details but tends to result in ambiguous edges,

e.g., the restored image in Fig. 2 (d). The performance of L_p-norm gradient prior (Fig. 2 (e)) is between the above two priors, it is evident that as $p \to 0$, its performance approximates to the L_0-norm gradient prior, while as $p \to 1$, it is like the L_1-norm gradient prior. However, there are two main problems with the L_p-norm gradient prior, the first is that it usually results in a difficult sub-optimization problem if p is not equal to 1/2 and 2/3 [11]. Secondly, we find that both of the L_0-norm and L_p-norm gradient priors are sensitive to heavy noise in the blurred image, e.g., the restored images in Fig. 2 (c) and Fig. 2 (e) contain some amplified noise. If we want to suppress it, we must enhance the penalty, but many useful contents especially some salient edges will be wrongly smoothed out. To overcome these problems, we first propose a sparse prior in Eq. (5) which is formulated by the sum of logarithm of the squared image gradients. For simplicity, we call it log gradient prior and denote it by $\rho(o)$.

$$\rho(\mathbf{o}) = \sum_{i=1}^{N} \log\{1 + \mu[(\mathbf{d}_x\mathbf{o})_i^2 + (\mathbf{d}_y\mathbf{o})_i^2]\}, \qquad (5)$$

where the parameter $\mu > 0$.

The red curve in Fig. 1 (a) shows $\rho(o)$ when $\mu = 1000$. We can easily infer that as $\mu \to +\infty$, it is approximates to the L_0-norm gradient prior while as $\mu \to 0$, it is puts less penalty on all the gradients. We also use it to restore the blurred image in Fig. 2 (b), it can be seen from the result in Fig. 2 (f) that the image edges are as sharp as that of L_0-norm (Fig. 2 (c)) and L_p-norm gradient priors (Fig. 2 (e)), the noise is successfully suppressed and some contents in the background are reserved.

Besides, we also incorporate the step function in Eq. (4) to further improve $\rho(o)$, it is just like incorporating the L_0-norm gradient prior into $\rho(o)$ and the final expression of the proposed prior is show in Eq. (6). For simplicity, we call it log-step gradient prior in the following parts.

$$\psi(\mathbf{o}) = \sum_{i=1}^{N} \log\{1 + (\alpha - 1)Step[(\mathbf{d}_x\mathbf{o})_i^2 + (\mathbf{d}_y\mathbf{o})_i^2] + \mu[(\mathbf{d}_x\mathbf{o})_i^2 + (\mathbf{d}_y\mathbf{o})_i^2]\}, \qquad (6)$$

where the parameter $\alpha \geq 1$.

In contrast to log gradient prior, the log-step gradient prior is more effective in suppressing the wrongly estimated image details and negative artifacts in the restored image. In Fig. 1 (b), we make a comparison between the curves of $\rho(o)$ and $\psi(o)$ where $\mu = 1000$ and $\alpha = 10$. We can see that a larger value of α makes the curve of log-step gradient prior become lower around zero, which means the image details corresponding to small gradients will be further removed than that of $\rho(o)$. We present the restored image of log-step gradient prior in Fig. 2 (g), we can see that the background is of less image details than that in Fig. 2 (f) while the edges are as sharp, which will be beneficial for iterative blur kernel estimation. In addition, the log-step gradient prior will result in a sub-optimization problem with closed-form solution which can be solved efficiently by finding the roots of a cubic equation for any values of α and μ, as will be shown in Section 4.

3.2. Deconvolution problem model

The formation of Poissonian blurred image can be modeled by the following expression:

$$\mathbf{g} = Poisson(\mathbf{ho}), \qquad (7)$$

where $\mathbf{g}$ and $\mathbf{o}$ denote the vector forms of the blurred image and latent image, respectively; $\mathbf{h}$ represents the convolution matrix of the blur kernel. The function *Poisson* represents the Poisson

noise degradation process, i.e.,

$$P(\mathbf{g}|\mathbf{h},\mathbf{o}) = \prod_{i=1}^{N} \frac{\exp[-(\mathbf{ho})_i](\mathbf{ho})_i^{\mathbf{g}_i}}{g_i!}. \tag{8}$$

Blind multi-image deconvolution aims to estimate the latent image and blur kernels from a sequence of blur images. Let $\mathbf{g}_{1\cdots K}$ represents the set of K blurred images $\{\mathbf{g}_1, \mathbf{g}_2, \cdots, \mathbf{g}_K\}$ and $\mathbf{h}_{1\cdots K}$ denotes the set of blur kernels $\{\mathbf{h}_1, \mathbf{h}_2, \cdots, \mathbf{h}_K\}$ for corresponding blurred images. According to the MAP estimation framework, the problem can be formulated by

$$
\begin{aligned}
(\mathbf{h}_{1\cdots K}, \mathbf{o}) &= \arg\min_{(\mathbf{h}_{1\cdots K}, \mathbf{o})} -\log[P(\mathbf{h}_{1\cdots K}, \mathbf{o}|\mathbf{g}_{1\cdots K})] \\
&= \arg\min_{(\mathbf{h}_{1\cdots K}, \mathbf{o})} -\log[P(\mathbf{g}_{1\cdots K}|\mathbf{h}_{1\cdots K}, \mathbf{o})] - \log[P(o)] - \log[P(\mathbf{h}_{1\cdots K})]
\end{aligned}, \tag{9}
$$

where $P(\mathbf{g}_{1\cdots K}|\mathbf{h}_{1\cdots K}, \mathbf{o})$ and $P(\mathbf{h}_{1\cdots K})$ represent the joint distributions of the noise and blur kernels, respectively. $P(\mathbf{o})$ denotes the prior distribution of the latent image.

Considering the topic discussed in this paper, we assume that each pixel in different blurred images is independent and follows the Poisson distribution, then

$$-\log[P(\mathbf{g}_{1\cdots K}|\mathbf{h}_{1\cdots K}, \mathbf{o})] = -\sum_{k=1}^{K} P(\mathbf{g}_k|\mathbf{h}_k, \mathbf{o}) \propto \sum_{k=1}^{K}\sum_{i=1}^{N} \{(\mathbf{h}_k\mathbf{o})_i - (\mathbf{g}_k)_i \ln[(\mathbf{h}_k\mathbf{o})_i]\}, \tag{10}$$

where k is the image and blur kernel index. We also suppose that each $\mathbf{h}_k$ in $\mathbf{h}_{1\cdots K}$ is independent identically distributed and adopt TV for regularization, then we obtain

$$-\log[P(\mathbf{h}_{1\cdots K})] = -\sum_{k=1}^{K} P(\mathbf{h}_k) \propto \sum_{k=1}^{K}\sum_{j=1}^{J} \sqrt{(\mathbf{d}_x\mathbf{h}_k)_j^2 + (\mathbf{d}_y\mathbf{h}_k)_j^2}, \tag{11}$$

where J is the number of pixels in each $\mathbf{h}_k$ and j denotes the pixel index. The TV regularization can help to suppress negative artifacts in the estimated blur kernels.

We adopt the proposed log-step gradient prior to model the term $-\log[P(\mathbf{o})]$ and introduce two regularization parameters λ and ξ, then we combine Eq. (6), (9), (10), and (11) to form the final optimization problem in Eq. (12), i.e.,

$$
\begin{aligned}
(\mathbf{h}_{1\cdots K}, \mathbf{o}) &= \arg\min_{(\mathbf{h}_{1\cdots K}, \mathbf{o})} \lambda \sum_{k=1}^{K}\sum_{i=1}^{N} \{(\mathbf{h}_k\mathbf{o})_i - (\mathbf{g}_k)_i \log[(\mathbf{h}_k\mathbf{o})_i]\} \\
&+ \sum_{i=1}^{N} \log\{1 + (\alpha - 1)Step[(\mathbf{d}_x\mathbf{o})_i^2 + (\mathbf{d}_y\mathbf{o})_i^2] + \mu[(\mathbf{d}_x\mathbf{o})_i^2 + (\mathbf{d}_y\mathbf{o})_i^2]\} + \xi \sum_{k=1}^{K}\sum_{j=1}^{J} \sqrt{(\mathbf{d}_x\mathbf{h}_k)_j^2 + (\mathbf{d}_y\mathbf{h}_k)_j^2}
\end{aligned}. \tag{12}
$$

4. Optimization approach

In this section, we will give a detailed description about how to solve the problem in Eq. (12). To obtain restored image of high quality, just like most blind deconvolution methods [52–55], the proposed algorithm consists of two phases, i.e., multi-blur kernel estimation and non-blind multi-image deconvolution.

4.1. Framework for multi-blur kernel estimation

The problem in Eq. (12) is a difficult non-convex optimization problem, we propose an approach based on the methods of Lagrange multiplier and variable splitting to convert it into a simple convex optimization problem and a blind multi-image deconvolution problem with quadratic fidelity term.

Firstly, we introduce a set of auxiliary variables $\mathbf{u}_{1\cdots K} = \{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_K\}$ to approximate the terms $\{\mathbf{h}_1\mathbf{o}, \mathbf{h}_2\mathbf{o}, \cdots, \mathbf{h}_K\mathbf{o}\}$, then the problem in Eq. (12) is equivalent to the following problem:

$$
(\mathbf{h}_{1\cdots K}, \mathbf{u}_{1\cdots K}, \mathbf{o}) = \arg\min_{(\mathbf{h}_{1\cdots K}, \mathbf{u}_{1\cdots K}, \mathbf{o})} \lambda \sum_{k=1}^{K} \sum_{i=1}^{N} \{(\mathbf{u}_k)_i - (\mathbf{g}_k)_i \log[(\mathbf{u}_k)_i]\}
$$
$$
+ \sum_{i=1}^{N} \log\{1 + (\alpha - 1)Step[(\mathbf{d}_x\mathbf{o})_i^2 + (\mathbf{d}_y\mathbf{o})_i^2] + \mu[(\mathbf{d}_x\mathbf{o})_i^2 + (\mathbf{d}_y\mathbf{o})_i^2]\} + \xi \sum_{k=1}^{K} \sum_{j=1}^{J} \sqrt{(\mathbf{d}_x\mathbf{h}_k)_j^2 + (\mathbf{d}_y\mathbf{h}_k)_j^2} \;,
$$
$$
\text{Subject to } ||\mathbf{h}_k\mathbf{o} - \mathbf{u}_k||_2^2 = 0
$$

(13)

where $k = 1, 2, \cdots K$.

According to the method of Lagrange multiplier, there should exist a set of multipliers $\{\beta_1, \beta_2, \cdots, \beta_K\}$ which can be used to convert the constrained optimization problem in Eq. (13) into a unconstrained optimization problem. To lower the difficulty, we assume that all the multipliers are of equal value, i.e., $\beta_k = \beta$, $k = 1, 2, \cdots K$. Then the problem in Eq. (13) is transformed into the following expression:

$$
(\mathbf{h}_{1\cdots K}, \mathbf{u}_{1\cdots K}, \mathbf{o}) = \arg\min_{(\mathbf{h}_{1\cdots K}, \mathbf{u}_{1\cdots K}, \mathbf{o})} \lambda \sum_{k=1}^{K} \sum_{i=1}^{N} \{(\mathbf{u}_k)_i - (\mathbf{g}_k)_i \log[(\mathbf{u}_k)_i]\} + \sum_{k=1}^{K} \frac{\beta}{2} ||\mathbf{h}_k\mathbf{o} - \mathbf{u}_k||_2^2
$$
$$
+ \sum_{i=1}^{N} \log\{1 + (\alpha - 1)Step[(\mathbf{d}_x\mathbf{o})_i^2 + (\mathbf{d}_y\mathbf{o})_i^2] + \mu[(\mathbf{d}_x\mathbf{o})_i^2 + (\mathbf{d}_y\mathbf{o})_i^2]\} + \xi \sum_{k=1}^{K} \sum_{j=1}^{J} \sqrt{(\mathbf{d}_x\mathbf{h}_k)_j^2 + (\mathbf{d}_y\mathbf{h}_k)_j^2} \;.
$$

(14)

We split Eq. (14) into two optimization problems and alternately solve them to reach the solution, i.e., sub-problem with respect to $\mathbf{u}_{1\cdots K}$ in Eq. (15) and sub-problem with respective to $\mathbf{h}_{1\cdots K}$ and $\mathbf{o}$ in Eq. (16). For simplicity, we call them $\mathbf{u}$-problem and $\mathbf{ho}$-problem, respectively. Just like the methods in [52–55], we only retain the estimated $\mathbf{h}_{1\cdots K}$ for subsequent non-blind multi-image deconvolution.

The $\mathbf{u}$-problem:

$$
\mathbf{u}_{1\cdots K} = \arg\min_{\mathbf{u}_{1\cdots K}} \lambda \sum_{k=1}^{K} \sum_{i=1}^{N} \{(\mathbf{u}_k)_i - (\mathbf{g}_k)_i \log[(\mathbf{u}_k)_i]\} + \sum_{k=1}^{K} \frac{\beta}{2} ||\mathbf{h}_k\mathbf{o} - \mathbf{u}_k||_2^2,
$$

(15)

The $\mathbf{ho}$-problem:

$$
(h_{1\cdots K}, o) = \arg\min_{(h_{1\cdots K}, o)} \sum_{k=1}^{K} \frac{\beta}{2} ||h_k o - u_k||_2^2
$$
$$
+ \sum_{i=1}^{N} \log\{1 + (\alpha - 1)Step[(\mathbf{d}_x\mathbf{o})_i^2 + (\mathbf{d}_y\mathbf{o})_i^2] + \mu[(\mathbf{d}_x\mathbf{o})_i^2 + (\mathbf{d}_y\mathbf{o})_i^2]\} + \xi \sum_{k=1}^{K} \sum_{j=1}^{J} \sqrt{(\mathbf{d}_x\mathbf{h}_k)_j^2 + (\mathbf{d}_y\mathbf{h}_k)_j^2} \;.
$$

(16)

4.2. Algorithm for solving the *u*-problem

The $\mathbf{u}$-problem can be divided into KN independent sub-problems:

$$
(\mathbf{u}_k)_i = \arg\min_{(\mathbf{u}_k)_i} J[(\mathbf{u}_k)_i] = \arg\min_{(\mathbf{u}_k)_i} \lambda\{(\mathbf{u}_k)_i - (\mathbf{g}_k)_i \log[(\mathbf{u}_k)_i]\} + \frac{\beta}{2}[(\mathbf{h}_k\mathbf{o})_i - (\mathbf{u}_k)_i]^2.
$$

(17)

Since $(\mathbf{u}_k)_i > 0$ and $d^2 J[(\mathbf{u}_k)_i]/d(\mathbf{u}_k)_i^2 > 0$, the problem in Eq. (17) is convex and the solution is the positive root of $dJ[(\mathbf{u}_k)_i]/d(\mathbf{u}_k)_i = 0$, i.e.,

$$
\beta(\mathbf{u}_k)_i^2 + [\lambda - \beta(\mathbf{h}_k\mathbf{o})_i](\mathbf{u}_k)_i - \lambda(\mathbf{g}_k)_i = 0.
$$

(18)

Because $-\lambda(\mathbf{g}_k)_i < 0$ and $\beta > 0$, Eq. (18) has only one positive root:

$$
(\mathbf{u}_k)_i = \frac{1}{2\beta} \left\{ \beta(\mathbf{h}_k\mathbf{o})_i - \lambda + \sqrt{[\lambda - \beta(\mathbf{h}_k\mathbf{o})_i]^2 + 4\beta\lambda(\mathbf{g}_k)_i} \right\}.
$$

(19)

Since all $(\mathbf{u}_k)_i$ are independent, they can be solved in parallel to reach the estimation of $\mathbf{u}_k$.

4.3. Algorithm for solving the **ho**-problem

We propose an alternating minimization algorithm to solve the **ho**-problem. For simplicity, we use the **o**-problem and **h**-problem to denote the optimization problem respect to **o** and $\mathbf{h}_{1 \cdots K}$, respectively, i.e.,

The **o**-problem:

$$\mathbf{o} = \arg\min_{\mathbf{o}} \sum_{k=1}^{K} \frac{\beta}{2} ||\mathbf{h}_k \mathbf{o} - \mathbf{u}_k||_2^2 + \sum_{i=1}^{N} \log\{1 + (\alpha - 1)Step[(\mathbf{d}_x \mathbf{o})_i^2 + (\mathbf{d}_y \mathbf{o})_i^2] + \mu[(\mathbf{d}_x \mathbf{o})_i^2 + (\mathbf{d}_y \mathbf{o})_i^2]\}. \tag{20}$$

The **h**-problem:

$$\mathbf{h}_{1 \cdots K} = \arg\min_{\mathbf{h}_{1 \cdots K}} \sum_{k=1}^{K} \frac{\beta}{2} ||\mathbf{h}_k \mathbf{o} - \mathbf{u}_k||_2^2 + \xi \sum_{k=1}^{K} \sum_{j=1}^{J} \sqrt{(\mathbf{d}_x \mathbf{h}_k)_j^2 + (\mathbf{d}_y \mathbf{h}_k)_j^2}. \tag{21}$$

4.3.1. Algorithm for solving the **o**-problem

The algorithm for solving **o**-problem is the core of the paper and we propose to use the quadratic penalty method. We first introduce two auxiliary variables **w** and **v** to approximate the terms $\mathbf{d}_x \mathbf{o}$ and $\mathbf{d}_y \mathbf{o}$, respectively, the **o**-problem in Eq. (20) is converted into the following equation:

$$\begin{aligned} (\mathbf{o}, \mathbf{w}, \mathbf{v}) = \arg\min_{(\mathbf{o}, \mathbf{w}, \mathbf{v})} \sum_{k=1}^{K} \frac{\beta}{2} ||\mathbf{h}_k \mathbf{o} - \mathbf{u}_k||_2^2 + \frac{\eta}{2} ||\mathbf{d}_x \mathbf{o} - \mathbf{w}||_2^2 \\ + \frac{\eta}{2} ||\mathbf{d}_y \mathbf{o} - \mathbf{v}||_2^2 + \sum_{i=1}^{N} \log\{1 + (\alpha - 1)Step[\mathbf{w}_i^2 + \mathbf{v}_i^2] + \mu[\mathbf{w}_i^2 + \mathbf{v}_i^2]\} \end{aligned}, \tag{22}$$

where $\eta > 0$ is the penalty parameter. It is known that as $\eta \to +\infty$, the solution of Eq. (22) converges to that of Eq. (20).

According to the quadratic penalty method, the solution of Eq. (22) can be reached by iteratively implementing the following steps:

Step 1: solving **o** with Eq. (23)

$$\mathbf{o} = \arg\min_{\mathbf{o}} \sum_{k=1}^{K} \frac{\beta}{2} ||\mathbf{h}_k \mathbf{o} - \mathbf{u}_k||_2^2 + \frac{\eta}{2} ||\mathbf{d}_x \mathbf{o} - \mathbf{w}||_2^2 + \frac{\eta}{2} ||\mathbf{d}_y \mathbf{o} - \mathbf{v}||_2^2. \tag{23}$$

Step 2: solving **w** and **v** with Eq. (24)

$$\begin{aligned} (\mathbf{w}, \mathbf{v}) = \arg\min_{(\mathbf{w}, \mathbf{v})} \frac{\eta}{2} ||\mathbf{d}_x \mathbf{o} - \mathbf{w}||_2^2 + \frac{\eta}{2} ||\mathbf{d}_y \mathbf{o} - \mathbf{v}||_2^2 \\ + \sum_{i=1}^{N} \log\{1 + (\alpha - 1)Step[\mathbf{w}_i^2 + \mathbf{v}_i^2] + \mu[\mathbf{w}_i^2 + \mathbf{v}_i^2]\} \end{aligned}. \tag{24}$$

Step 3: updating η by $\eta = \tau \eta$, where $\tau > 1$.

a) Solving the problem in step 1:

For step 1, the solution of Eq. (23) has a closed-form solution in frequency domain,

$$F(\mathbf{o}) = \frac{\beta \sum_{k=1}^{K} F^*(\mathbf{h}_k) \circ F(\mathbf{u}_k) + \eta F^*(\mathbf{d}_x) \circ F(\mathbf{w}) + \eta F^*(\mathbf{d}_y) \circ F(\mathbf{v})}{\beta \sum_{k=1}^{K} F^*(\mathbf{h}_k) \circ F(\mathbf{h}_k) + \eta F^*(\mathbf{d}_x) \circ F(\mathbf{d}_x) + \eta F^*(\mathbf{d}_y) \circ F(\mathbf{d}_y)}, \tag{25}$$

where F denote the Fourier transform, $*$ is the conjugate operator, $\circ$ represents the component-wise multiplication and the division operation is also component-wise. Then **o** can be reached with the inverse Fourier transform.

b) Solving the problem in step 2:

For step 2, since all $(\mathbf{w}_i, \mathbf{v}_i)$ are independent, the problem in Eq. (24) can be divided into N sub-problems, i.e.,

$$
\begin{aligned}
(\mathbf{w}_i, \mathbf{v}_i) &= \arg\min_{(\mathbf{w}_i,\mathbf{v}_i)} J_i(\mathbf{w}_i, \mathbf{v}_i) \\
&= \arg\min_{(\mathbf{w}_i,\mathbf{v}_i)} \tfrac{\eta}{2}[(\mathbf{d}_x\mathbf{o})_i - w_i]^2 + \tfrac{\eta}{2}[(d_y o)_i - v_i]^2 \quad + \log\{1 + (\alpha - 1)Step[\mathbf{w}_i^2 + \mathbf{v}_i^2] + \mu[\mathbf{w}_i^2 + \mathbf{v}_i^2]\}
\end{aligned}
$$
(26)

We can derive the following two propositions. The proofs are given in Supplement 1.
Proposition 1.

1) if $(\mathbf{d}_x\mathbf{o})_i = 0$ and $(\mathbf{d}_y\mathbf{o})_i = 0$, then $(\mathbf{w}_i, \mathbf{v}_i) = (0, 0)$.

2) if $(\mathbf{d}_x\mathbf{o})_i \neq 0$, $(\mathbf{d}_y\mathbf{o})_i \neq 0$ and $\mathbf{w}_i^2 + \mathbf{v}_i^2 \neq 0$, then $\mathbf{w}_i$ and $\mathbf{v}_i$ meet $\mathbf{w}_i/(\mathbf{d}_x\mathbf{o})_i = \mathbf{v}_i/(\mathbf{d}_y\mathbf{o})_i = r_i$, where $0 < r_i < 1$ and r_i is one real root of the cubic equation in Eq. (27).

3) if $(\mathbf{d}_x\mathbf{o})_i \neq 0$, $(\mathbf{d}_y\mathbf{o})_i = 0$ and $\mathbf{w}_i^2 + \mathbf{v}_i^2 \neq 0$, then $\mathbf{w}_i$ meet $\mathbf{w}_i/(\mathbf{d}_x\mathbf{o})_i = r_i$, where $0 < r_i < 1$ and r_i is one real root of the cubic equation in Eq. (27), $v_i = 0$.

4) if $(\mathbf{d}_x\mathbf{o})_i = 0$, $(\mathbf{d}_y\mathbf{o})_i \neq 0$ and $\mathbf{w}_i^2 + \mathbf{v}_i^2 \neq 0$, then $\mathbf{v}_i$ meet $\mathbf{v}_i/(\mathbf{d}_y\mathbf{o})_i = r_i$, where $0 < r_i < 1$ and r_i is one real root of the cubic equation in Eq. (27), $\mathbf{w}_i = 0$.

$$\mu\eta m_i r_i^3 - \mu\eta m_i r_i^2 + (\mu + \alpha\eta)r_i - \alpha\eta = 0, \tag{27}$$

where $m_i = (\mathbf{d}_x\mathbf{o})_i^2 + (\mathbf{d}_y\mathbf{o})_i^2$.

Proposition 2. The cubic equation in Eq. (27) has at least one real root and all the real roots locate in the open interval $(0, 1)$.

The above two propositions present an approach for solving Eq. (26), i.e.,

If $(\mathbf{d}_x\mathbf{o})_i^2 + (d_y o)_i^2 = 0$, according to the conclusion 1) in Proposition 1, $(\mathbf{w}_i, \mathbf{v}_i) = (0, 0)$.

If $(\mathbf{d}_x\mathbf{o})_i^2 + (\mathbf{d}_y\mathbf{o})_i^2 \neq 0$, we should consider two conditions as follows and make several comparisons to reach the optimal solution:

Condition 1: Suppose that $\mathbf{w}_i^2 + \mathbf{v}_i^2 = 0$, we calculate $J_i(0, 0)$.

Condition 2: Suppose that $\mathbf{w}_i^2 + \mathbf{v}_i^2 \neq 0$, we first construct the cubic equation in Eq. (27) according to the conclusions 2)-4) in Proposition 1, then we calculate the three roots with standard formula and eliminate the imaginary roots. Based on Proposition 2, if there is only one real root $r_{i,1} \in (0, 1)$, the possible solution of Eq. (26) is $(\mathbf{w}'_i, \mathbf{v}'_i) = (r_{i,1}\mathbf{w}_{i,1}, r_{i,1}\mathbf{v}_{i,1})$. If there are three real roots $r_{i,s} \in (0, 1)$, $s = 1, 2, 3$, we firstly calculate three two-tuples $(\mathbf{w}_{i,s}, \mathbf{v}_{i,s}) = (r_{i,s}\mathbf{w}_{i,s}, r_{i,s}\mathbf{v}_{i,s})$ and the corresponding values of $J_i(\mathbf{w}_{i,s}, \mathbf{v}_{i,s})$, and afterwards we find the minimum of $J_i(\mathbf{w}_{i,s}, \mathbf{v}_{i,s})$ and the corresponding root $r_{i,m}$, then the possible solution of Eq. (26) is $(\mathbf{w}'_i, \mathbf{v}'_i) = (r_{i,m}\mathbf{w}_{i,m}, r_{i,m}\mathbf{v}_{i,m})$.

Since $J_i(\mathbf{w}_i, \mathbf{v}_i)$ is not continuous at $(\mathbf{w}_i, \mathbf{v}_i) = (0, 0)$, $J_i(\mathbf{w}'_i, \mathbf{v}'_i)$ may be not the minimum of Eq. (26), we should make another comparison between $J_i(0, 0)$ obtained in condition 1 and $J_i(\mathbf{w}'_i, \mathbf{v}'_i)$ to reach the optimal solution $(\mathbf{w}_i, \mathbf{v}_i)$ of Eq. (26), i.e.,

$$
\begin{cases}
(\mathbf{w}_i, \mathbf{v}_i) = (0, 0) & if\ J_i(0, 0) \leq J_i(\mathbf{w}'_i, \mathbf{v}'_i) \\
(\mathbf{w}_i, \mathbf{v}_i) = (\mathbf{w}'_i, \mathbf{v}'_i) & if\ J_i(0, 0) > J_i(\mathbf{w}'_i, \mathbf{v}'_i)
\end{cases}
\tag{28}
$$

Based on the above analysis, we can calculate (w_i, v_i) for all the indices i in parallel and reach the final solution (w, v) of Eq. (24), the algorithm is given in Algorithm 1.

Algorithm 1. Algorithm for solving (w,v)

1.	initialization: α, μ
2.	find the indices where $(\mathbf{d}_x\mathbf{o})_i^2 + (\mathbf{d}_y\mathbf{o})_i^2 = 0$ and denote the set by I_0
3.	find the indices where $(\mathbf{d}_x\mathbf{o})_i^2 + (\mathbf{d}_y\mathbf{o})_i^2 \neq 0$ and denote the set by I_1
4.	for all indices i, do the following in parallel
5.	if $i \in I_0$, $(\mathbf{w}_i, \mathbf{v}_i) = (0,0)$, end
6.	if $i \in I_1$, construct Eq. (27),
7.	if Eq. (27) has one real root $r_{i,1}$,
8.	$(\mathbf{w}_i', \mathbf{v}_i') = (r_{i,1}\mathbf{w}_{i,1}, r_{i,1}\mathbf{v}_{i,1})$
9.	else if Eq. (27) has three real root $r_{i,s}$, $s = 1,2,3$
10.	calculate $(\mathbf{w}_{i,s}, \mathbf{v}_{i,s}) = (r_{i,s}\mathbf{w}_{i,s}, r_{i,s}\mathbf{v}_{i,s})$, find minimum of $J_i(\mathbf{w}_{i,s}, \mathbf{v}_{i,s})$ and the corresponding root $r_{i,m}$, $(\mathbf{w}_i', \mathbf{v}_i') = (r_{i,m}\mathbf{w}_{i,m}, r_{i,m}\mathbf{v}_{i,m})$
11.	end
12.	calculate $J_i(0,0)$ and find $(\mathbf{w}_i, \mathbf{v}_i)$ with Eq. (28)
13.	end
14.	end
15.	output: $(\mathbf{w}, \mathbf{v})$

4.3.2. Algorithm for solving the **h**-problem

We divide the **h**-problem in Eq. (21) into K independent sub-problems to get the solution, i.e.,

$$\mathbf{h}_k = \arg\min_{(\mathbf{h}_k,\mathbf{o})} \frac{\beta}{2}||\mathbf{h}_k\mathbf{o} - \mathbf{u}_k||_2^2 + \xi \sum_{j=1}^{J} \sqrt{(\mathbf{d}_x\mathbf{h}_k)_j^2 + (\mathbf{d}_y\mathbf{h}_k)_j^2}. \tag{29}$$

According to [53], we use the difference operator to modify the fidelity term in Eq. (29) to accelerate the convergence, i.e.,

$$\mathbf{h}_k = \arg\min_{\mathbf{h}_k} \frac{\beta}{2}||\mathbf{dh}_k\mathbf{o} - \mathbf{du}||_2^2 + \xi \sum_{j=1}^{J} \sqrt{(\mathbf{d}_x\mathbf{h}_k)_j^2 + (\mathbf{d}_y\mathbf{h}_k)_j^2}. \tag{30}$$

where $\mathbf{d} = \mathbf{d}_x + \mathbf{d}_y$. Just like in [38], we adopt the iteratively reweighted least squares (IRLS) method to solve for $\mathbf{h}_k$, then the pixel in it should be normalized by $(\mathbf{h}_k)_j = (\mathbf{h}_k)_j / \sum_j (\mathbf{h}_k)_j$. After all $\mathbf{h}_k$ are calculated, we achieve the solution of $\mathbf{h}_{1\cdots K}$.

In order to make an accurate estimation of $\mathbf{h}_{1\cdots K}$, we usually take heavy penalty to suppress the negative artifacts in the intermediately restored image **o**, most details in it are also removed, so we abandon it and only retain the estimated $\mathbf{h}_{1\cdots K}$, the restored image is finally reached with a non-blind multi-image deconvolution method presented in the following section.

4.4. Non-blind multi-image deconvolution

The problem model of non-blind multi-image deconvolution is almost the same as Eq. (12) except that $\mathbf{h}_{1\cdots K}$ is known, i.e.,

$$\begin{aligned}
\mathbf{o} = \arg\min_{\mathbf{o}} \lambda \sum_{k=1}^{K} \sum_{i=1}^{N} \{(\mathbf{h}_k\mathbf{o})_i - (\mathbf{g}_k)_i \log[(\mathbf{h}_k\mathbf{o})_i]\} \\
+ \sum_{i=1}^{N} \log\{1 + (\alpha - 1)Step[(\mathbf{d}_x\mathbf{o})_i^2 + (\mathbf{d}_y\mathbf{o})_i^2] + \mu[(\mathbf{d}_x\mathbf{o})_i^2 + (\mathbf{d}_y\mathbf{o})_i^2]\}
\end{aligned} \tag{31}$$

To reach a final restored image of high quality, we use the quadratic penalty method to solve it. The details are shown in Supplement 1.

4.5. Parameter setting

In the phase of multi-blur kernel estimation, we usually set $1000 \leq \lambda \leq 15000$. The Lagrange multiplier β is empirically set $\lambda/10 < \beta < \lambda/5$ and $\xi = 1000\beta$. η is initialized with 1 and increases with a factor $\tau = 2$, and the total iteration number is five. For the log-step gradient prior, we empirically set $1 < \alpha \leq 10$ and $1000 \leq \mu \leq 10000$. The total iteration number for IRLS is set to five. We find that the parameter setting is very related to the noise level and the size of the blur kernel. Specifically, if the noise is of higher level or the blur kernel is of larger size, the parameters λ and β should be tuned smaller while α and μ should be tuned larger to suppress the negative artifacts. If the noise level and the size of the blur kernel are similar for different blurred images, the setting of the parameters will be of minor change.

In the phase of non-blind multi-image deconvolution, we usually set $1000 \leq \lambda \leq 20000$. The parameters for the log-step gradient prior are set to $1 \leq \alpha \leq 2$ and $100 \leq \mu \leq 2000$, respectively.

5. Experimental results

5.1. Experimental results of synthetic data

We first use synthetic blurred images to evaluate the proposed method. In [56], the authors construct a dataset which contains four clear images and eight blur kernels (Fig. 3), and we use them to generate synthetic multi-blurred images, then we consider two experimental settings: 1) experiment on multi-blurred images with different blur kernels; 2) experiment on multi-blurred images with the same blur kernel.
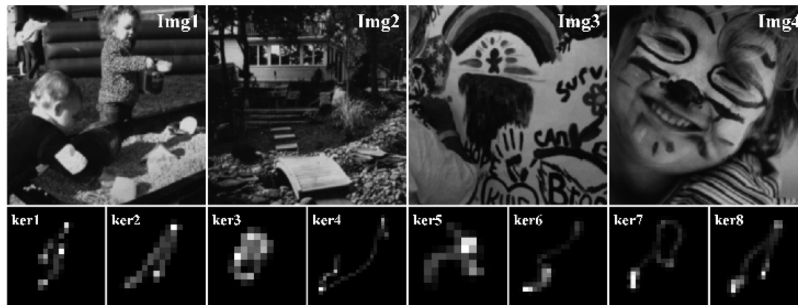

Fig. 3. Clear images and blur kernels in the dataset of [56].

This is because in some conditions, e.g., the random motion caused multi-blurred images, the blur kernels may be different one from the other. While in some other conditions, e.g., continuous exposure in fluorescence microscopy observation, the imaging parameters are fixed and the blur kernels are almost the same. In addition, the noise level may be quite different in various imaging environments which is also a significant factor that affects the quality of the result. Therefore, in each of the above settings, we carry out evaluations in two cases: 1) the multi-blurred images are contaminated by slight Poisson noise; 2) the multi-blurred images are contaminated by heavy Poisson noise.

Based on the above analysis, there will be four experimental configurations, i.e., A: multi-blurred images with different blur kernels and slight Poisson noise; B: multi-blurred images with different blur kernels and heavy Poisson noise; C: multi-blurred images with the same blur kernel and slight Poisson noise; D: multi-blurred images with the same blur kernel and heavy Poisson noise.
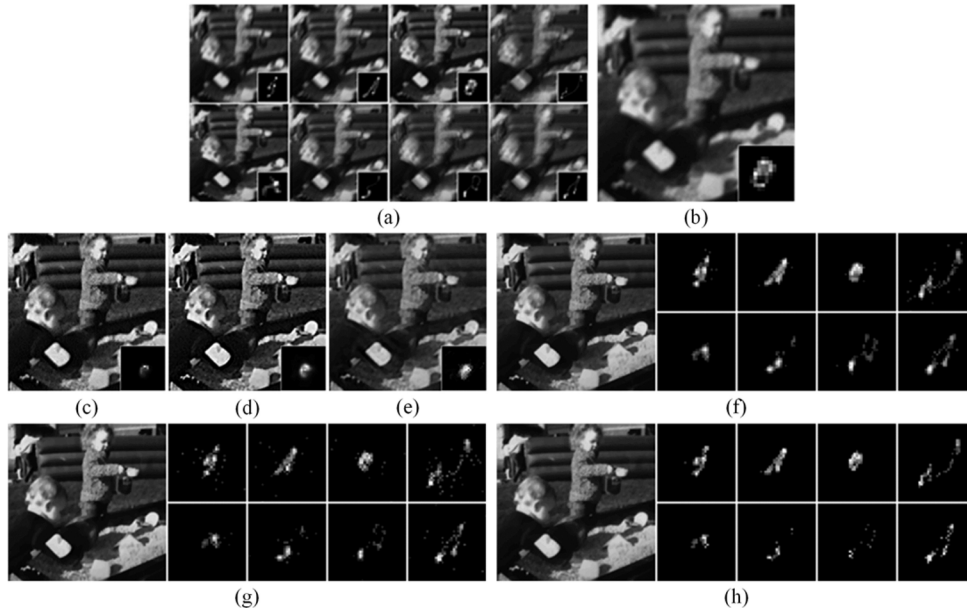
Fig. 4. Experimental results of multi-blurred images with different blur kernels and slight Poisson noise. (a) Eight blurred images and the corresponding blur kernels. (b) The chosen blurred image with the smallest blur kernel. (c) L_1-framelet [36], PSNR: 23.80 dB, SSIM: 0.5588. (d) L_0-framelet [37], PSNR: 23.63 dB, SSIM: 0.7576. (e) L_0-gradient [38], PSNR: 27.63 dB, SSIM: 0.8391. (f) Dual-L_0 prior [57], PSNR: 33.69 dB, SSIM: 0.9462. (g) Surface-aware prior [58], PSNR: 33.12 dB, SSIM: 0.9469. (h) Our log-step gradient prior, PSNR: 34.31 dB, SSIM: 0.9498.

To prove the effectiveness of our approach, we take four experiments for configurations A and B, respectively. We also take 16 experiments for configurations C and D, respectively. We first compare our approach with some state-of-the-art blind deconvolution methods for single Poissonian blurred image, i.e., the methods in [36–38]. For simplicity, we denote them by the methods of L_1-framelet, L_0-framelet, and L_0-gradient in the following subsections. For configurations A and B, we choose to restore the blurred image generated from the smallest blur kernel with these methods, while for configurations C and D, we calculate the average of the multi-blurred images and restore it.

Besides, we also try to use two robust priors which are raised in recent researches to substitute the log-step gradient prior in our problem model in Eq. (12), they are the dual L_0-norm of image intensity and gradient prior [57] and the surface-aware prior [58] (just as mentioned in [58], we only adopt the surface-aware prior to estimate the blur kernels, the latent image is reached with the TV regularized non-blind multi-image deconvolution method). For simplicity, we denote them by the methods of dual-L_0 prior and surface-aware prior in the following subsections, respectively.

We use the popular image quality assessment method, i.e., the peak signal-to-noise ratio (PSNR) and structural similarity index method (SSIM) [59] to evaluate the restored images of different methods. Since there are too many images of all the experiments, due to the length limitation of the paper, we only show the results of one experiments in the following parts. The readers can refer to Supplement 1 for all the other results.

The example belongs to the experimental configuration A. In Fig. 4(a), the input degraded multi-images are generated with the first image and the eight different blur kernels in Fig. 3, they

are contaminated by slight Poisson noise. In Fig. 4(b), we select the blurred image in the first row and third column whose corresponding blur kernel is of the smallest size and use the three single image deconvolution methods to restore it, the results are shown in Fig. 4(c)-(e).

We also use blind multi-image deconvolution methods with the dual-L_0 [57], surface-aware [58], and log-step gradient priors for restoration, the results are shown in Fig. 4(f)-(h). The PSNR and SSIM values for each restored image are given in the figure caption. Comparing the results with the ground truth in Fig. 3, it is evident that our approach achieves the best result.

Furthermore, we calculate the PSNR and SSIM values of all the restored images of different deconvolution methods. Due to the length limitation of the paper, the data are listed in the tables in Supplement 1. For clarity, we plot the PSNR and SSIM curves of different deconvolution methods in Fig. 5 to make a comparison. We can see that the proposed log-step gradient prior performs better than the other methods.
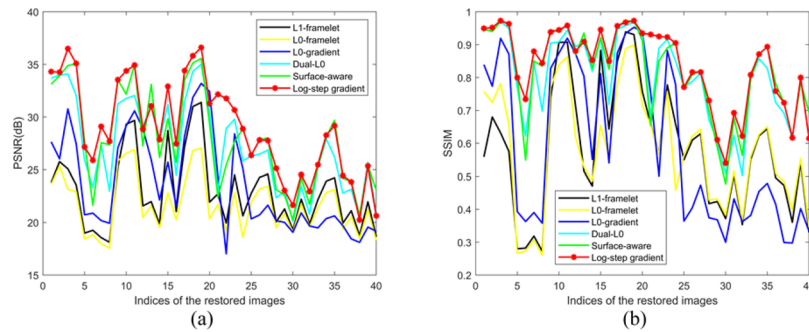

Fig. 5. PSNR and SSIM of different deconvolution methods

5.2. Experimental results of real-world data

We also use real-world multi-blurred images to test our approach. In [60], the authors construct a dataset of fluorescence microscopy images which is used for evaluating denosing algorithms. To acquire the data, they fix the exposure configuration for each field of view and take 50 images continuously, all of which are degraded by heavy Poisson noise, thus the authors take the average of all the 50 images as ground truth. In the dataset, we find that some of the images are blurred and use the single and multi-image deconvolution methods to restore them. Due to the length limitation of the paper, we only show the results of one experiment, in which 50 wide-field microscopy images of the fixed bovine pulmonary artery endothelial (BPAE) cells are taken as input of the algorithms. The blur kernels of them are the point spreading function of the system, which is of low-pass nature. Figure 6(a) shows eight of the 50 blurred images, we can see that they are very blurred and contaminated by heavy Poisson noise, Fig. 6(b) is the average image. Figure 6(c)-(e) show the results of blind single image deconvolution methods for Fig. 6(b), Fig. 6(f)-(h) show the results of blind multi-image deconvolution methods with different priors. We can also see that the methods of L_1-framelet [36] and L_0-framelet [37] fail to remove the blur. While the results of L_0-gradient [38], Dual-L_0 [57] and Surface-aware [58] are degraded by much amplified noise and ringing artifacts. The proposed method achieves result of the best quality, which has rich details and few negative artifacts as shown in Fig. 7 . (Larger and clearer images of Fig. 6 and Fig. 7 as well as the results of all the other experiments are provided in Supplement 1).
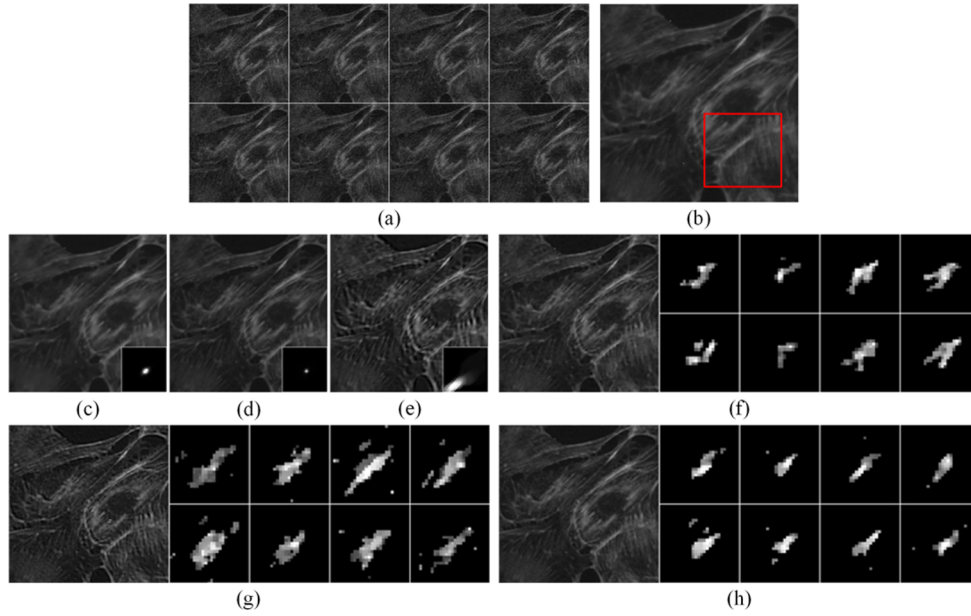
Fig. 6. Experiment on real-world wide-field microscopy images of fixed BPAE cells. (a) Eight of the 50 blurred images. (b) The average of the 50 blurred images. (c) L_1-framelet [36]. (d) L_0-framelet [37]. (e) L_0-gradient [38]. (f) Dual-L_0 prior [57]. (g) Surface-aware prior [58]. (h) Our log-step gradient prior.
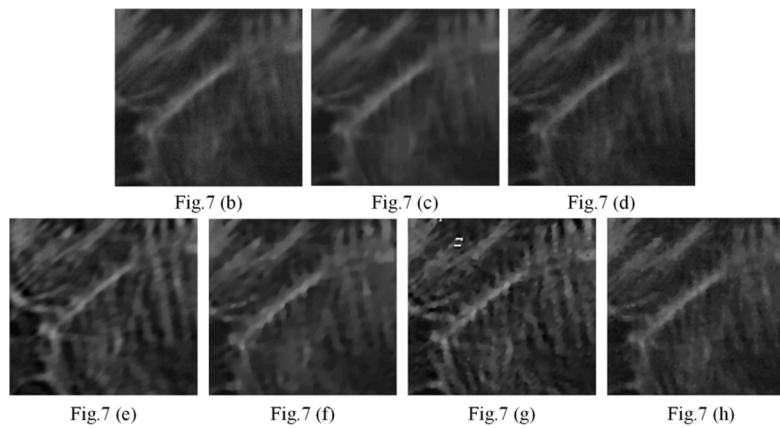

Fig. 7. Zoom of the contents in the red rectangle of corresponding figures in Fig. 6.

6. Discussion

6.1. Comparison of multi-image deconvolution and single image deconvolution

It is true that multiple captures of the same scene or object may lead to some negative issues, e.g., reducing temporal resolution, exacerbated photobleaching, etc. However, the multi-image deconvolution can significantly outperform single image deconvolution in restoration quality even with the same prior, which makes it have great application prospect in certain application scenarios. We take two examples to prove this. In Fig. 8(a), we show the fourth blurred image in the first row in Fig. 4(a). Figure 8(b) is the corresponding clear image and blur kernel. Figure 8(c) is the result of the proposed method but using single blurred image, while Fig. 8(d) is the restoration result of the proposed method using the eight blurred images in Fig. 4(a). We can see that multi-image deconvolution performs much better than single image deconvolution, its restored image is very close to the ground truth. Similarly, Fig. 9(a) shows one of the blurred images in Fig. 6(a). The results of single image deconvolution and multi-image deconvolution are shown in Fig. 9(b) and (c), respectively. We can see that Fig. 9(b) is contaminated by amplified noise due to the ill-posedness, while Fig. 9(c) is of much higher quality.
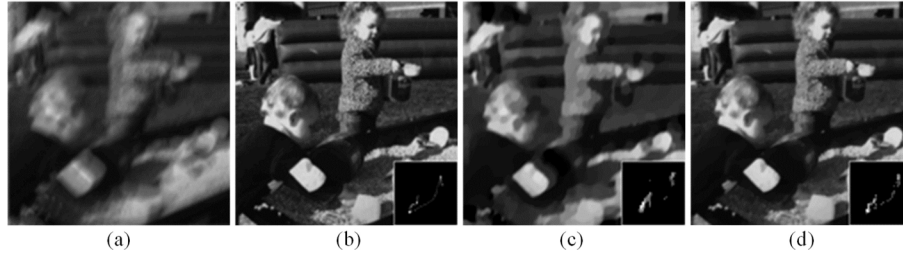

(a)   (b)   (c)   (d)

Fig. 8. Restoration results of multi-image deconvolution and single image deconvolution with synthetic blurred images. (a) One of the original blurred images in Fig. 4(a). (b) The ground truth. (c) Result of single image deconvolution. (d) Result of multi-image deconvolution.
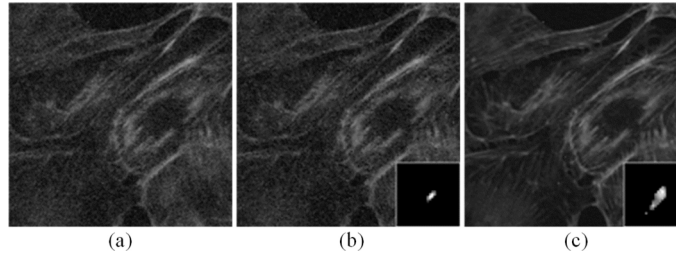

(a)   (b)   (c)

Fig. 9. Restoration results of multi-image deconvolution and single image deconvolution with real-world blurred images. (a) One of the original blurred image in Fig. 6(a). (b) Result of single image deconvolution. (d) Result of multi-image deconvolution.

6.2. Impact of the key hyper-parameters

There are two hyper-parameters μ and α that determine the characteristics of the log-step gradient prior. Besides the analysis in Subsection 3.1, here we take an experiment to further demonstrate the impact of them on the restoration results. Figure 10(a) shows the restored image with $\mu = 100$ and $\alpha = 1.1$, we can see that it has rich details. Then we tune μ to 1000 and keep α unchanged, the restored image is shown in Fig. 10(b), we can see that some tiny details are smoothed. In

Fig. 10(c), the restored image is reached by keeping $\mu = 100$ and tune α to 10, many details are removed and the main edges are sharpened. Thus we can conclude from the experiment that μ mainly influence image details, while α has a strong impact on both image details and edges. This is because μ is the weight of squared image gradient while α is the weight of the step response of squared image gradient, a large α makes the prior perform like the L_0-norm based prior and thus have stronger impact on both image details and edges.
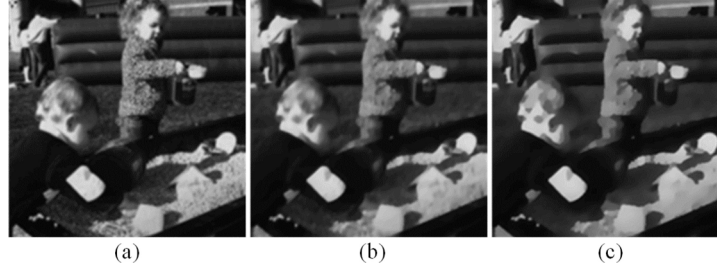

(a)   (b)   (c)

Fig. 10. Restoration results with different μ and α. (a) Result of $\mu=100$, $\alpha=1.1$. (b) Result of $\mu=1000$, $\alpha=1.1$. (d) Result of $\mu=100$, $\alpha=10$.

When the log-gradient prior is used to restore blurred images contaminated by heavy Poisson noise, we can assign large values to both μ and α to suppress the amplified noise due to ill-posedness. A large μ can smooth the noise and prevent the algorithm from wrongly recognizing the amplified noise as image details, then a large α can remove the noise accurately. That is why the log-gradient prior performs better than some classic priors, such as the TV and L_0-norm of image gradients.

6.3.   Algorithm acceleration scheme

For the above algorithm, if the input contains many blurred images with large size, it will be computationally expensive, thus it is very necessary to accelerate the proposed algorithm.

Since most modern computers are equipped with high-performance multi-core CPU and GPU which support parallel computation, we propose a hybrid acceleration scheme based on multi-core CPU and GPU as follow.

(1) For the complicated operations such as forward and inverse Fourier transforms, we load the data onto GPU and carry out the calculation.

(2) For the **u**-problem, we solve for the elements in $u_{1\cdots K}$ in parallel with multi-core CPU.

(3) For the **h**-problem, we also solve for the elements in $\mathbf{h}_{1\cdots K}$ in parallel with multi-core CPU.

(4) For the **o**-problem and non-blind multi-image deconvolution, if the size of the blurred images is large, we first segment all the blurred images into groups of multi-patches and restore them, the results are finally combined to form the whole restored image.

We have provided a detailed demonstration of how to carry out the algorithm acceleration scheme and verify its effectiveness in Supplement 1.

7.   Conclusions

In this paper, we propose a multi-image deconvolution method for restoring blurred images contaminated by Poisson noise. The main contribution is that we design a very robust sparse prior which adopts a mixture of logarithm and step functions to regularize the squared image gradients. In contrast to some other popular sparse priors, the proposed log-step gradient prior can achieve

better results even in the condition of heavy Poisson noise. Furthermore, we design an efficient algorithm to solve the resulted optimization problem. Experimental results on both synthetic and real-world data show that our approach is very effective in processing multi-Poissonian blurred images and the restored image and blur kernels are of high quality.

Disclosures. The authors declare no conflicts of interest.

Data availability. Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

Supplemental document. See Supplement 1 for supporting content.

References

1. M. Bertero and P. Boccacci, *Introduction to inverse problems in imaging* (Institute of Physics Publishing, 1998), Chap. 4.
2. T. F. Chan and W. Chiu-Kwong, "Total variation blind deconvolution," IEEE Trans. on Image Process. **7**(3), 370–375 (1998).
3. Q. Shan, J. Jia, and A. Agarwala, "High-quality motion deblurring from a single image," ACM Trans. Graph. **27**(5), 1–7 (2008).
4. A. Levin, R. Fergus, F. Durand, *et al.*, "Image and depth from a conventional camera with a coded aperture," ACM Trans. Graph. **26**(3), 70–80 (2007).
5. S. K. Nayar and M. Ben-Ezra, "Motion-based motion deblurring," IEEE Trans. Pattern Anal. Machine Intell. **26**(6), 689–698 (2004).
6. J. M. Bioucas-Dias, M. A. T. Figueiredo, and J. P. Oliveira, "Adaptive total variation image deconvolution: A majorization-minimization approach," in *Proceedings of European Conference on Signal Processing* (2016), pp. 1–4.
7. J. Besag, "Spatial Interaction and the Statistical Analysis of Lattice Systems," Journal of the Royal Statistical Society: Series B **36**(2), 192–225 (1974).
8. A. N. Tikhonov, "On the stability of inverse problems," in *Proceedings of Dok. Akad. Nauk Sssr* (1943).
9. J. M. Bardsley and N. Laobeul, "Tikhonov regularized Poisson likelihood estimation: theoretical justification and a computational method," Inverse Probl. Sci. Eng. **16**(2), 199–215 (2008).
10. Y. Wang, J. Yang, W. Yin, *et al.*, "A new alternating minimization algorithm for total variation image reconstruction," SIAM J. Imaging Sci. **1**(3), 248–272 (2008).
11. D. Krishnan and R. Fergus, "Fast image deconvolution using hyper-Laplacian priors," Advances in Neural Information Processing Systems **22**, 1–9 (2009).
12. L. Xu, C. Lu, Y. Xu, *et al.*, "Image smoothing via L_0 gradient minimization," Acm Trans. on Graphics **30**(6), 1–12 (2011).
13. L. Xu, S. Zheng, and J. Jia, "Unnatural L_0 sparse representation for natural image deblurring," in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition* (IEEE, 2013), pp. 1107–1114.
14. S. Roth and M. J. Black, "Fields of experts: a framework for learning image priors," in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition* (IEEE, 2005), pp. 860–867.
15. S. Roth and M. J. Black, "Fields of experts," International Journal of Computer Vision **82**(2), 205–229 (2009).
16. Y. Weiss and W. T. Freeman, "What makes a good model of natural images?" in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition* (IEEE, 2007), pp. 1–8.
17. P. J. Green, "On use of the EM for penalized likelihood estimation," J. Roy. Statistical Soc. Series B. **52**(3), 443–452 (1990).
18. P. J. Green, "Bayesian reconstructions from emission tomography data using a modified EM algorithm," IEEE Trans. Med. Imaging **9**(1), 84–93 (1990).
19. N. Dey, L. Blanc-Feraud, C. Zimmer, *et al.*, "Richardson-Lucy algorithm with total variation regularization for 3D confocal microscope deconvolution," Microsc. Res. Tech. **69**(4), 260–266 (2006).
20. S. Tao, W. Dong, H. Feng, *et al.*, "Non-blind image deconvolution using natural image gradient prior," Optik **124**(24), 6599–6605 (2013).
21. G. Vicidomini, S. W. Hell, and A. Schönle, "Automatic deconvolution of 4Pi-microscopy data with arbitrary phase," Opt. Lett. **34**(22), 3583–3585 (2009).
22. G. Vicidomini, R. Schmidt, A. Egner, *et al.*, "Automatic deconvolution in 4Pi-microscopy with variable phase," Opt. Express **18**(10), 10154–10167 (2010).
23. L. Yuan, J. Sun, L. Quan, *et al.*, "Image deblurring with blurred/noisy image pairs," ACM Trans. Graph. **26**(3), 1 (2007).

24. L. Yuan, J. Sun, L. Quan, *et al.*, "Progressive inter-scale and intra-scale non-blind image deconvolution," ACM Trans. Graph. **27**(3), 1–10 (2008).
25. W. H. Richardson, "Bayesian based iterative method of image restoration," J. Opt. Soc. Am. **62**(1), 55–59 (1972).
26. L. B. Lucy, "An iterative technique for the rectification of observed distributions," Astron. J. **79**(6), 745–754 (1974).
27. M. A. T. Figueiredo and J. M. Bioucas-Dias, "Restoration of Poissonian images using alternating direction optimization," IEEE Trans. on Image Process. **19**(12), 3133–3145 (2010).
28. S. Setzer, G. Steidl, and T. Teuber, "Deblurring Poissonian images by split Bregman techniques," Journal of Visual Communication and Image Representation **21**(3), 193–199 (2010).
29. M. Carlavan and L. Blanc- Féraud, "Sparse Poisson noisy image deblurring," IEEE Trans. on Image Process. **21**(4), 1834–1846 (2012).
30. T. Jeong, H. Woo, and S. Yun, "Frame-based Poisson image restoration using a proximal linearized alternating direction method," Inverse Problems **29**(7), 075007 (2013).
31. J. Liu, T. Z. Huang, X. G. Lv, *et al.*, "High-order total variation based Poissonian image deconvolution with Spatially adapted regularization parameter," Applied Mathematical Modelling **45**, 516–529 (2017).
32. S. Tao, W. Dong, Z. Xu, *et al.*, "Fast total variation deconvolution for blurred image contaminated by Poisson noise," Journal of Visual Communication and Image Representation **38**, 582–594 (2016).
33. D. Chen, "Regularized generalized inverse accelerating linearized alternating minimization algorithm for frame-based Poissonian image deblurring," SIAM Journal on Imaging Sciences **7**(2), 716–739 (2014).
34. Z. Xu and E. Y. Lam, "Maximum a posteriori blind image deconvolution with Huber-Markov random-field regularization," Opt. Lett. **34**(9), 1453–1455 (2009).
35. L Hai, Z. Zhang, S. Liu, *et al.*, "Richardson-Lucy blind deconvolution of spectroscopic data with wavelet regularization," Appl. Opt. **54**(7), 1770–1775 (2015).
36. H. Fang, L. Yan, H. Liu, *et al.*, "Blind Poissonian images deconvolution with framelet regularization," Opt. Lett. **38**(4), 389–391 (2013).
37. X. Gong, B. Lai, and Z. Xiang, "A L_0 sparse analysis prior for blind Poissonian image deconvolution," Opt. Express **22**(4), 3860–3865 (2014).
38. W. Dong, S. Tao, G. Xu, *et al.*, "Blind deconvolution for Poissonian blurred image with total variation and L_0-norm gradient regularizations," IEEE Trans. on Image Process. **30**, 1030–1043 (2021).
39. F. Sroubek and J. Flusser, "Multichannel blind deconvolution of spatially misaligned images," IEEE Trans. on Image Process. **14**(7), 874–883 (2005).
40. F. Sroubek and P. Milanfar, "Robust multichannel blind deconvolution via fast alternating minimization," IEEE Trans. on Image Process. **21**(4), 1687–1700 (2012).
41. X. Zhu, F. Roubek, and P. Milanfar, "Deconvolving PSFs for a better motion deblurring using multiple images," in *Proceedings of European Conference on Computer Vision* (2012), pp. 636–647.
42. E. Faramarzi, D. Rajan, and M. P. Christensen, "Unified blind method for multi-image super-resolution and single/multi-image blur deconvolution," IEEE Trans. on Image Process. **22**(6), 2101–2114 (2013).
43. H. Zhang, D. Wipf, and Y. Zhang, "Multi-image blind deblurring using a coupled adaptive sparse prior," in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition* (IEEE, 2013), pp. 1051–1058.
44. H. Zhang, D. Wipf, and Y. Zhang, "Multi-observation blind deconvolution with an adaptive sparse prior," IEEE Trans. on Pattern Analysis and Machine Intelligence **36**(8), 1628–1643 (2014).
45. M. Sonogashira, T. Funatomi, M. Iiyama, *et al.*, "Variational Bayesian approach to multi-frame image restoration," IEEE Trans. on Image Process. **26**(5), 2163–2178 (2017).
46. T. Lin, L. Hou, H. Liu, *et al.*, "Reconstruction of single image from multiple blurry measured images," IEEE Trans. on Image Process. **27**(6), 2762–2776 (2018).
47. D. Kang and S. I. Yoo, "Multi-image blind deconvolution using low-rank representation," in *Proceedings of IEEE International Conference on Image Processing* (IEEE, 2019), pp. 4669–4673.
48. B. Ma, J. Trisovic, and H. A. Loeliger, "Multi-image blind deblurring using a smoothed NUV prior and iteratively reweighted coordinate descent," in *Proceedings of IEEE International Conference on Image Processing* (IEEE, 2020), pp. 973–977.
49. H. Zhou, Y. Chen, H. Feng, *et al.*, "Rotated rectangular aperture imaging through multi-frame blind deconvolution with hyper-Laplacian priors," Opt. Express **29**(8), 12145–12159 (2021).
50. W. Dong, H. Feng, Z. Xu, *et al.*, "Multi-frame blind deconvolution using sparse priors," Opt. Commun. **285**(9), 2276–2288 (2012).
51. C. Gu, X. Lu, Y. He, *et al.*, "Blur removal via blurred-noisy image pair," IEEE Trans. on Image Process. **30**, 345–359 (2021).
52. L. Xu and J. Y. Jia, "Two-phase kernel estimation for robust motion deblurring," in *Proceedings of European Conference on Computer Vision* (2010), pp. 157–170.
53. S. Cho and S. Lee, "Fast motion deblurring," ACM Trans. Graph. **28**(5), 1–8 (2009).
54. R. Fergus, B. Singh, A. Hertzmann, *et al.*, "Removing camera shake from a single photograph," ACM Trans. Graph. **25**(3), 787–794 (2006).
55. D. Krishnan, T. Tay, and R. Fergus, "Blind deconvolution using a normalized sparsity measure," in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition* (IEEE, 2011), pp. 233–240.

56. A. Levin, Y. Weiss, F. Durand, *et al.*, "Understanding and evaluating blind deconvolution algorithms," in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition* (IEEE, 2009), pp. 1964–1971.

57. J. Pan, Z. Hu, Z. Su, *et al.*, "L_0-regularized intensity and gradient prior for deblurring text images and beyond," IEEE Trans. Pattern Anal. Mach. Intell. **39**(2), 342–355 (2017).

58. J. Liu, M. Yan, and T. Zeng, "Surface-aware blind image deblurring," IEEE Trans. Pattern Anal. Mach. Intell. **43**(3), 1041–1055 (2021).

59. W. Zhou, A. C. Bovik, H. R. Sheikh, *et al.*, "Image quality assessment: from error visibility to structural similarity," IEEE Trans. on Image Process. **13**(4), 600–612 (2004).

60. Y. Zhang, Yinhao Zhu, Evan Nichols, *et al.*, "A Poisson-Gaussian denoising dataset with real fluorescence microscopy images," in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition* (IEEE, 2019), pp. 11702–11710.