Neurocomputing 413 (2020) 305–327

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Gradient-based discriminative modeling for blind image deblurring

Wen-Ze Shao [a,b,*], Yun-Zhi Lin [c], Yuan-Yuan Liu [a], Li-Qian Wang [a], Qi Ge [a], Bing-Kun Bao [a], Hai-Bo Li [a,d]

[a] College of Telecommunications and Information Engineering, Nanjing University of Posts and Telecommunications, China
[b] Key Laboratory of Intelligent Perception and Systems for High-Dimensional Information of Ministry of Education, Nanjing University of Science and Technology, China
[c] School of Electrical and Computer Engineering, Georgia Institute of Technology, United States
[d] School of Computer Science and Communication, KTH Royal Institute of Technology, Sweden

ARTICLE INFO

Article history:
Received 12 June 2019
Revised 20 June 2020
Accepted 22 June 2020
Available online 8 July 2020
Communicated by Yan Bo

Keywords:
Image deblurring
Blind deconvolution
Discriminative prior
Blur kernel estimation
Low-illumination

ABSTRACT

Blind image deconvolution is a fundamental task in image processing, computational imaging, and computer vision. It has earned intensive attention in the past decade since the seminal work of Fergus et al. [1] for camera shake removal. In spite of the recent great progress in this field, this paper aims to formulate the blind problem with a simpler modeling perspective. What is more important, the newly proposed approach is expected to achieve comparable or even better performance towards the real blurred images. Specifically, the core critical idea is the proposal of a pure gradient-based discriminative prior for accurate and robust blur kernel estimation. Numerous experimental results on both the benchmark datasets and real-world blurred images in various imaging scenarios, e.g., natural, manmade, low-illumination, text, or people, demonstrate well the effectiveness and robustness of the proposed approach.

1. Introduction

Blind image deconvolution, a.k.a. blind deblurring, is a fundamental problem in image processing, computational imaging, and computer vision. It has earned intensive attention in the past decade since the seminal work of Fergus *et al.* [1] for camera shake removal.

Among multiple algorithmic ingredients of blind deblurring, it is universally acknowledged that the unnatural image priors [27] are playing the dominant role in the success of reasonable blur kernel estimation. To bring out the proposed approach in a smooth way, representative image priors specifically proposed for blind deblurring are summarized in this section.

According to the literature review, blind deblurring can be divided into three categories of methodologies in general, i.e., Maximum-a-Posterior (MAP), Variational Bayesian (VB), and Representation Learning (RL). In the past several years, the RL particularly the convolutional neural networks (CNN), has been successfully applied to kinds of imaging and vision tasks. It is discovered, however, that the generalization capability of existing CNN-based blind approaches [6–14,68,69,70,73] is far from enough to deal with blurs in different imaging scenarios, e.g., natural, manmade, low-illumination, text, or people. In comparison, the gradient-based algorithms formulated in the MAP framework such as [3,15] are demonstrating greater and greater potentialities in the practical deblurring tasks. Thus, this paper still follows this route and tries to further exploit the potential of gradient-based models. In specific, a simple, robust yet discriminative prior is introduced for blind deblurring. To be noted that, our discussions in this paper are limited to the spatially-invariant image deblurring. One consideration is that the performance of previous advanced non-uniform blind methods such as [16,17] are not competitive at all to those uniform ones, as demonstrated by Lai et al. [2]. Another consideration is that an image prior for uniform blind deconvolution can be naively extended to the spatially-variant problems [3].

1.1. Existing image priors in blind deblurring

For the sake of description clarity, a short review on image priors proposed and used in blind deconvolution is made in this part. About twenty years ago, a total variation (TV)-based blind deblurring approach was presented by Chan and Wong [18], being viewed as the first modern and influential algorithm to the blind task. However, it is interesting to note that little significant progress was made since then till the exciting work of Fergus et al.

* Corresponding author at: College of Telecommunications and Information Engineering, Nanjing University of Posts and Telecommunications, China.
E-mail address: shaowenze@njupt.edu.cn (W.-Z. Shao).

https://doi.org/10.1016/j.neucom.2020.06.093
0925-2312/

[1] about fifteen years ago. It utilizes mixture of Gaussians as the natural image prior for camera shake removal rather than the TV model. The inference is made in the VB framework and afterwards simplified in [19] for better and faster blur removal. As a matter of fact, an empirical observation in [20] indicates that within the VB framework the naive Gaussian prior is even qualified for the blind task [20] to a certain degree. Nonetheless, when exactly the same image priors as the ones in [1,19] are plugged into the MAP framework (e.g., Gaussian, mixture of Gaussians, fields of experts [21], failure results are usually produced [1,20] such as the pair of delta blur kernel and original blurred image. Note that, another naive MAP approach [22] regularized by the framelet-based natural image prior is also reported in the risk of similar failure cases.

An intuitive failure cause of those naive MAP methods is that, the harnessed natural image priors favor a blurred image to its sharp counterpart. To ensure the success of MAP-based blind deblurring, kinds of tricks have to be exploited in practice. For instance, the seminal TV-based method [18] is recently explored once again in [23], which provides an in-depth analysis on its real working principle. After ten years of publication of [18], the first creative trick for boosting it, proposed in [24], is to add shock filtering into TV minimization for better prediction of salient edges as clues to kernel estimation. Inspired by [24], the shock filtering is also applied in [25,26] with higher quality blind deblurring results produced. Actually, as phrased in [27] the natural priors combining with shock filtering essentially plays a role of implicit unnatural sparse representation.

Undergone several years of exploration since the exciting work of Fergus et al. [1] in 2006, unnatural image models have been predominating the blind deblurring literature until now. On this line, the first daring try is harnessing the normalized sparsity measure [4] with the idea that the image prior should favor a sharp image to its blurry one. Nevertheless, the method can not produce state-of-the-art performance on this or that benchmark dataset, let alone blurry images in the wild [2]. The normalized sparsity is mathematically an approximation of the L0-norm in essence, indicating that the salient edges are more important than the faint textures for the success of blind image deconvolution. Inspired by the normalized sparsity measure, a simple yet effective L0-norm-based blind deblurring approach is subsequently presented in [27].

In literature, two unnatural image models [28,29] using the Lp-norm are tried, too, wherein the parameter p is set as 0.3 in [28] and a non-increasing sequence 0.8, 0.8, 0.6, 0.6, 0.6, 0.6, 0.4, ..., 0.4 along iteration in [29]. The TV-based method [18] is recently investigated in [30] again which proposes a nonconvex logarithmic TV prior with improved deblurring performance on some benchmark datasets. Moreover, an unnatural iteration-wise hyper-Laplacian image prior [31] is learned in the MAP framework leading to better results than existing gradient-based approaches. Although the learned prior in [31] is claimed to be discriminative, its generalization performance is questionable to blurry images in the practical diverse imaging scenarios.

In fact, unnatural image priors are not only requested in the MAP framework but also advocated in the VB case in spite of its more robustness in posterior inference. In distinction to [1,19], recent empirical and theoretical findings both prove that the Jeffreys' prior could achieve more accurate blur kernel estimation [32,33], and this non-informative prior is essentially similar to the logarithmic TV prior [30] to a large degree. In addition, the prior is even shown optimal to a certain degree in terms of deblurring quality [33]. Another work by the authors of the present paper has proposed to determine priors for blind image deblurring as a self-learning problem [34]. The learned model resembles the non-informative Jeffreys' prior in a sense, whose negative-logarithm is obviously another approximation to the L0-based model.

Instead of approximating the L0-norm with diverse strategies, three pure L0-based image priors [27,35,36] are proposed in 2013 for blind deblurring. However, they are found not generalized well to large-scale blurs especially in specific imaging scenarios, e.g., face, text, or low-illumination images. In [37], a bi-L0-L2-norm-based regularization term imposed on both sharp image and blur kernel is proposed for higher precision of kernel estimation. In [38], a L0-norm-based joint intensity and gradient prior is presented for the text image deblurring. Furthermore, an exemplar-driven approach with L0-norm-based gradient regularization is proposed in [39] for the facial image deblurring.

An alternative strategy to formulate blind deblurring is to explore the patch-based priors [40,41]. In [40], a novel idea of internal patch recurrence is exploited for the blind problem. While, in [41] the idea of modeling via external patch querying is proposed for edge-based blur kernel estimation. However, it generally takes a higher computational cost to query a large external dataset [3]. Recently, a novel patch-based approach is presented in [42] harnessing the normalized color-line priors, and shows better performance than [40,41]. To be noted that, the text deblurring method [38] has been also extended to a patch-based scheme for handling natural image deblurring [43]. Its core idea is essentially to exploit the structural sparsity prior, implemented via a couple of rank penalty terms on similar patches over both the intensity and gradient domains.

Furthermore, several other methods are proposed to improve the robustness of blind deblurring to noise [44], outlier [45,46], and other possible degradations, e.g., light streak [47]. One may refer to [2,48,63] for a more comprehensive survey on image priors and other technical components in blind deblurring.

According to above discussions, numerous blind deblurring algorithms have been reported in the past decade, achieving better and better performance on one or another synthetic dataset. However, as empirically concluded in [2], the performance of existing methods on the benchmark datasets is generally inferior to that on the real-world blurred images. In other words, existing blind deblurring algorithms are far from being practical in terms of the restoration quality.

Actually, a real breakthrough for blind deblurring is just made recently in [3], which combines the L0-regularized sparsity on both domains of image gradient and dark channel. The experimental results demonstrate its superior performance to all the representative methods in the past decade as studied in [2]. Note that, although the L0-based dark channel prior (DCP) is discriminative as desired, the combinatorial L0-regularized model [3] is not necessarily so. That is to say, the composite prior does not necessarily prefer a sharp image to its blurred counterpart. We note that the work [3] can be thought of as a smart generalization over [38] which is not a pure gradient-based approach, either. More recently, a convolutional neural network (CNN) based classifier is learned in [71] to distinguish whether an input image is sharp or not, which is directly plugged into the MAP framework of [3] as a replacement of the L0-based DCP. In addition, an L0-based bright channel prior (BCP) is proposed in [15]. The BCP is then combined with the L0-regularized sparsity on both image gradient and dark channel of [3], expecting to achievie more robust blur kernel estimation. Note that, the two successors of [3], i.e., [15,71], are natually with higher computational burden than [3]. Besides, the composite priors in [15,71] are not necessarily discriminative as a whole, either.

In this paper, the main motivation is to exploit the full potential of gradient-based methods, attempting to explore a simple, robust yet discriminative image prior for blind deblurring. Specifically, our contributions in this paper are three-fold: Above all, a pure gradient-based heavy-tailed prior is proposed indicating that the success of blind image deblurring requires dual principles of discriminativeness (DPD). On the one hand, since this paper formu-

lates blind deblurring into a model-based minimization problem, the image prior should be able to discriminate between a sharp image and its blurry counterpart. On the other hand, a really good image prior for blind deblurring should be also capable of discriminating between strong edges and faint textures, so as to reduce the possible interfering effects on blur kernel estimation caused by the textures. Such an idea is actually a consensus among existing blind deblurring literature [24–37]. What makes us particularly surprised is that, the above DPD required by blind deblurring can be easily implemented with the well-known gradient-based hyper-Laplacian priors [58]. To the best of our knowledge, it has been the first time that the concept of DPD is clarified and formulated for the blind deblurring problem and to be demonstrated very necessary for a blind deblurring method both understandable and applicable. In specific, inspired by the first daring try for discriminative blind deblurring in [4], our image prior essentially falls into a spatially variant hyper-Laplacian model. Secondly, a plug-and-play algorithm is derived to alternatively estimate the intermediate sharp image and the nonparametric blur kernel. With the numerical scheme, intermediate image estimation is then simplified to a simple image filtering problem. Finally, a great many experiments are performed accompanied with comparisons with state-of-the-art methods on both synthetic benchmark datasets and real blurry images in various scenarios, e.g., natural, manmade, low-illumination, text, or people. Experimental results validate well the effectiveness and robustness of the proposed approach, which is proved to be a promising new candidate solution for blind image deconvolution.

The rest of the paper is organized as follows. In Section 2, a pure gradient-based discriminative image model is provided as a novel candidate solution to blind image deblurring. To test its performance, Section 3 formulates blind deblurring into an unconstrained optimization problem and solves it utilizing a plug-and-play numerical scheme. In Section 4, a series of comprehensive analysis is performed to empirically clarify the motivation of this paper in an intuitive perspective. Section 5 provides numerous experimental results so as to demonstrate the effectiveness and robustness of the proposed solution along with comparisons against the state-of-the-art methods. This paper is finally concluded in Section 6. We should note that, the present journal paper is an extensively extended version of our previous conference paper published in [72]. The difference mainly lies in the following three aspects: (1) A fairly more comprehensive overview is provided on blind image deblurring algorithms, following which the dual principles of discriminativeness are claimed for the first time for blind deblurring; (2) A fairly more detailed analysis is provided on the motivation of proposing dual principles of discriminativeness in this paper, where Section 4 is a completely new part; (3) A fairly more convincing experimental comparison has been made among the proposed approach and existing blind methods, particularly those very recent deep learning-based ones [68,69,70,14,73].

2. Proposed gradient-based discriminative prior

This section explores the potentials of pure gradient-based modeling for blind deblurring, aiming to propose a simpler yet much more robust heavy-tailed discriminative image prior. We are to preliminarily show that the expectation can be realized by integrating those re-refined ideas from existing priors in blind deblurring and image filtering. Hence, another role of this paper is to throw out a brick to attract a jade, just with the hope that more practical yet simpler methods may be produced for the blind deblurring problem.

Our discussion begins with the first daring attempt towards discriminative modeling for blind image deconvolution, i.e., the normalized sparsity [4]. However, as indicated in [2,3], its practical performance is fairly limited according to the experimental results on both synthetic and real-world data, which is the exact reason that the normalized sparsity-based prior receives no further attention after the publication of [4]. In fact, our work in this paper reveals an interesting and instructive fact that, the principle of discriminativeness between a sharp image and its blurry counterpart, being greatly emphasized by the normalized sparsity [4], is really not the solely essential element for nonparametric blind restoration. As noted in Section 1, the dual principles of discriminativeness are the decisive components for achieving successful blind deblurring in practice. Thus, a new candidate prior for blind image deconvolution is advocated in this paper, given as

$$\mathcal{R}(u)=\sum_{p}\varpi_{x,p}(u)\cdot\left|\partial_x u_p\right|^{\alpha}+\sum_{p}\varpi_{y,p}(u)\cdot\left|\partial_y u_p\right|^{\alpha}, \tag{1}$$

where u is a latent sharp image, $p\in\Omega(u)$ a pixel index, α a positive number far less than 1, and ∂ a derivative operator, and $\varpi_{x,p}(u)$ a positive weight related to pixel index and derivative direction. Obviously, model (1) is an adaptive variant of the well-known hyper-Laplacian image priors [58]. Meanwhile, it is not difficult to deduce that the core novelty of prior (1) should be in the definition of $\varpi_{x,p}(u)$ and $\varpi_{y,p}(u)$ which needs incorporate the demanded DPD for plausible intermediate image update.

The first aspect of discriminativeness in (1) is between a latent sharp image and its blurry counterpart, which guarantees that the optimal solution of model-based blind image deblurring should not be a pair of blurry image and delta kernel. While, the second aspect of discriminativeness is between salient edges and faint textures, which guarantees that interfering textures should be removed from the intermediate sharp image for ensuring accurate and robust blur kernel estimation, as validated in existing methods [24,25,26,27]. Guided by the DPD idea, we simply formulate $R(u)$ into a pure gradient-based composite image prior. Specifically, $\varpi_{o,p}(u), o\in\{x,y\}$ is defined as

$$\varpi_{o,p}(u)=\frac{1-t}{(\mathcal{D}_o(u))^{\beta}+\varepsilon}+\frac{t}{(\mathcal{S}_o(p))^{\beta}+\varepsilon} \tag{2}$$

where β is a positive power, t is a number between 0 and 1, and ε is a small positive number to avoid division by zero, and $D_o(u)$ and $S_o(p)$ are expressed respectively as

$$\mathcal{D}_o(u)=\left(\sum_{p\in\Omega(u)}\left|\partial_o u_p\right|^2\right)^{1/2}=\|\,\partial_o u\,\|_2, \tag{3}$$

$$\mathcal{S}_o(p)=\left|\sum_{q\in\Omega(p)}\phi_{p,\,q}\cdot\partial_o u_q\right|, \tag{4}$$

where $\Omega(p)$ in (4) is a rectangular field centered at pixel p, and $\phi_{p,q}$ is defined according to the spatial affinity harnessing the distance function of Gaussianity, i.e.,

$$\phi_{p,\,q}\propto\exp\left(-\frac{(x_p-x_q)^2+(y_p-y_q)^2}{2\sigma^2}\right),$$

where σ is a spatial scale to be specified in implementation.

The adaptive weight defined in (2) is to be demonstrated a very effective and robust choice for blind deblurring. It is not hard to find that, the first aspect of discriminativeness desired by the proposed model (1) is simply achieved by amending the normalized sparsity in [4]. While, the second aspect of discriminativeness for accurate intermediate image update is further ensured by regulating the relative total variation (RTV) in [5]. Actually, the term $S_o(p)$ used in (2) was originally proposed in [5] for image filtering and manipulation, whose value in a window just with textures is found

statistically smaller than that in a window also containing structural edges. To the best of our knowledge, it has been the first time that RTV is successfully borrowed for achieving state-of-the-art blind deblurring performance as a regularization penalty.

Let's dive into (1) and (2) for more details. One finding is that, $\sum_p|\partial_x u_p|^\alpha/(D_x(u))^\beta+\sum_p|\partial_y u_p|^\alpha/(D_y(u))^\beta$ plays a primary role in discriminating sharp images from blurred ones as proper settings are provided to α,β. Apparently, above regularization term naturally degenerates to the normalized sparsity in [4] as α,β are equal to 1. In other words, Eq. (3) exactly originates from [4] in spite that the normalized sparsity measure itself performs inferiorly to state-of-the-art blind deblurring methods. Another finding is that, the performance of above regularization term could be ulteriorly boosted via $\sum_p|\partial_x u_p|^\alpha/(S_x(p))^\beta+\sum_p|\partial_y u_p|^\alpha/(S_y(p))^\beta$. The reason lies in that it is able to remove the interfering textures while making the salient structures stand out more accurately in the intermediate sharp image. Such additional amending is proved critical for higher quality blind deconvolution in spite that the amending strength governed by the parameter t is relatively less. A large amount of experiments demonstrate that t being set as 0.05 satisfactorily serves the deduced plug-and-play algorithm in Section 3. The above statements are to be comprehensively analyzed and systematically validated in Sections 4 and 5, respectively.

3. A plug-and-play numerical scheme to blind deconvolution

As the blur is assumed spatially invariant, the blurred image observation process can be described as

$$g=k*u+n, \tag{5}$$

where u denotes the latent sharp image, g the captured blurry image, k the blur kernel corresponding to trajectory of camera shake or out-of-focus of lens, and n the possible random noise. It is known that blind image deconvolution is mathematically ill-posed because there are infinite solution pairs (u,k) satisfying the formulation (5). Therefore, appropriate regularization should be imposed on both the image u and the kernel k.

Harnessing the proposed model (1), a MAP-based objective function for blind deblurring can be expressed as

$$J(u,k)\triangleq\|g-k*u\|_2^2+\lambda R(u)+\eta\|k\|_2^2, \tag{6}$$

where λ and η are the two positive adjusting parameters. The first quadratic term is for the image fidelity, while the third term is a Tikhonov regularization on the blur kernel k. Note that, the formulation (6) works free of any ad-hoc modeling tricks, e.g., continuation, or additional image processing operations such as bilateral smoothing or shock filtering. Here, it deserves pointing out that the trick of continuation is applied in the recent breakthrough work [3] as well as its three variants [15,71,38]. Actually, continuation is found fairly important for TV workable in blind deblurring, too, being demonstrated in [18] and other state-of-the-art approaches. In consequence, the blind deconvolution performance of the proposed algorithm is overwhelmingly determined by the proposed discriminative image prior (1), considering that the Tikhonov regularization on the blur kernel is a standard configuration in a large majority of existing approaches. In this paper, the tuning parameter η for the kernel regularization is universally set as 2.

Now, the image and the kernel can be obtained by solving the joint minimization problem $\left(\hat{u},\hat{k}\right)=\mathrm{argmin}_{u,k}J(u,k)$ in an alternatingly iterative manner. Provided the $(i-1)$th iterative solution of $k^{(i-1)}$, $u^{(i)}$ and $k^{(i)}$ are then respectively solved by $u^{(i)}=\mathrm{argmin}_u J\left(u,k^{(i-1)}\right)$ and $k^{(i)}=\mathrm{argmin}_k J(u^{(i)},k)$. Particularly, a plug-and-play numerical scheme is introduced in Section 3.1 to estimate the intermediate sharp image $u^{(i)}$ wherein the image filtering step is implemented via reweighted least squares approximation.

3.1. A plug-and-play scheme to intermediate image estimation

The half-quadratic regularization strategy is used to estimate $u^{(i)}$ via decomposing the original minimization problem into two simpler sub-problems. Specifically, an auxiliary variable is introduced corresponding to u, i.e., let $u=z$ a new objective function can be then obtained as

$$J\left(u,z,k^{(i-1)}\right)\triangleq\|g-k^{(i-1)}*u\|_2^2+\lambda R(z)+\rho\|u-z\|_2^2.$$

The minimization solution of $J\left(u,z,k^{(i-1)}\right)$ is expected to approach that of $J\left(u,k^{(i-1)}\right)$ as ρ is close to the infinity. In each alternative minimization over u and z, it is clear that u can be efficiently calculated via use of fast Fourier transform (FFT) as a closed form solution. That is,

$$u=F^{-1}\left(\frac{\bar{F}\left(k^{(i-1)}\right)F(g)+\rho F(z)}{\bar{F}\left(k^{(i-1)}\right)F\left(k^{(i-1)}\right)+\rho}\right), \tag{7}$$

where F and $\bar{F}$ represent the FFT and its complex conjugate, respectively, and F^{-1} denotes the operation of inverse FFT. In addition, z is simply initialized to be a zero image. Given u, z is then numerically computed by minimizing the sub-problem

$$\|u-z\|_2^2+\frac{\lambda}{\rho}R(z). \tag{8}$$

Apparently, solving (8) actually amounts to an amendatory step of image smoothing regularized by (1). In this perspective, the intermediate sharp image estimation falls into a plug-and-play framework seminally proposed in [49,50] and later extended and harnessed in [51,52,53,54,55,56]. It should be noted that, current plug-and-play-based papers mostly make use of state-of-the-art Gaussian denoisers, among which BM3D [57] is the most popular candidate. While, to the best of our knowledge, the present paper has been the first applying the plug-and-play idea for blind image deconvolution via use of a specifically customized discriminative prior. Therefore, the above plug-and-play scheme has a very large extension space as more advanced discriminative image priors are available.

3.2. Image smoothing with the discriminative prior

There are at least two solution strategies for solving (8) in principle. One is to apply the half-quadratic regularization or the ADMM (Alternating Direction Method of Multipliers), e.g., [55,56], for fast alternating computation. While, this paper harnesses another strategy to solve (8) directly for more stable calculation.

First of all, $R(z)$ is rewritten in a slightly different manner as (9), which is the discriminative regularization (1) imposed on z

$$\begin{aligned}R(z)&=\sum_p\varpi_{x,p}(z)\cdot|\partial_x z_p|^\alpha+\sum_p\varpi_{y,p}(z)\cdot|\partial_y z_p|^\alpha\\&\approx\sum_o\sum_p\varpi_{o,p}(z)\cdot\frac{1}{|\partial_o z_p|^{2-\alpha}+\varepsilon}\cdot(\partial_o z_p)^2\\&\triangleq\sum_o\sum_p\varpi_{o,p}(z)\cdot v_{o,p}(z)\cdot(\partial_o z_p)^2,\end{aligned} \tag{9}$$

where the approximation in the second line is due to the introduction of ε for numerical stability, and

$$v_{o,p}(z) = \frac{1}{\left|\partial_o z_p\right|^{2-\alpha} + \varepsilon}. \tag{10}$$

In this case, (9) can be rewritten in a matrix–vector form. Then, (8) is minimized by a following equivalent energy functional

$$\| u - z \|_2^2 + \frac{\lambda}{\rho} R(z), \tag{11}$$

where **u** and **z** are the vectorized representations of u and z, respectively, and $R(z)$ is

$$R(z) = \sum_o z^T D_o^T W_o(z) V_o(z) D_o z, \tag{12}$$

where $W_o(z)$ and $v_o(z)$ are the diagonal weighting matrices constructed by $\varpi_{o,p}(z)$ and $v_{o,p}(z)$ defined at all pixels, and $\mathbf{D}_o$ is a BCCB (block circulant with circulant blocks) matrix corresponding to the derivative operator in the o-direction, and T denotes the transpose operation.

With (12), minimization of (11) boils down to iteratively solving a linear system as

$$(\rho I + \lambda L)\, z = \rho u, \tag{13}$$

where I is a diagonal identity matrix and L is computed on the basis of a previous estimate of **z**, i.e.,

$$L = \sum_o D_o^T W_o(z) V_o(z) D_o. \tag{14}$$

As the forward difference is used to approximate the discrete gradients in $\mathbf{D}_o$, L would turn to a sparse five-point Laplacian matrix. Owing to its nature of symmetric positive definiteness, the incomplete Choleskey decomposition (ICD) can be applied and then the preconditioned conjugate gradient (PCG) method is used to solve (13) efficiently. Approximately, 3 s are expended to process a gray-scale image with 500 thousand pixels on a Laptop computer (Intel i7 2.60 CPU, 12G memory) running MATLAB v7.0. Combining Sections 3.1 and 3.2, the main steps of our intermediate image estimation are summarized in **Algorithm 1**, wherein the discriminative image smoothing step is outlined in **Algorithm 2**.

Algorithm 1 Intermediate image estimation

input : Blurred image g, blur kernel $k^{(i-1)}$, parameter λ.
initialization : $u \leftarrow g$, $z \leftarrow$ zero, $\rho \leftarrow 0.01$.
while $\rho < \rho_{\max}$ **do**
  update u by computing (7) via FFT.
  update z by minimizing (8) via **Algorithm 2**.
  update ρ by $\rho \leftarrow 2\rho$.
end while
output : $u^{(i)}$.

3.3. Blur kernel estimation

With an estimated intermediate image $u^{(i)}$, blur kernel $k^{(i)}$ can be directly produced by solving the Tikhonov-based energy functional $k^{(i)} = \arg\min_k J(u^{(i)}, k)$. In spite of that, a slightly modified functional defined in the gradient domain is used for more accurate estimation, as commonly practiced in existing blind deblurring works [3,19,25,37,38]. That is,

$$k^{(i)} = \arg\min_k \| \nabla g - k * \nabla u^{(i)} \|_2^2 + \eta \| k \|_2^2, \tag{15}$$

wherein $k_{(i)}$ can be solved very efficiently in a closed-form via use of FFT. One more point to be noted is that, blur kernel $k^{(i)}$ should be also projected onto the set $C = \{k \geq 0, \Sigma_i \Sigma_j | k_{i,j}| = 1\}$ considering the physical property of blur kernels. Another consideration is that a multi-scale estimating scheme is applied which is commonly practiced in blind deblurring since Fergus et al. [1] in 2006, aiming to deal with large-scale kernel estimation and reduce the risk of getting stuck in poor local minima. In each scale, the input image is a down-sampled version from the original blurred image g (in the finest scale the input image is g itself). While the initialized blur kernel $k^{(0)}$ is set as an up-sampled version from the estimated result in the coarser scale (in the coarsest scale $k^{(0)}$ is set as a Dirac pulse). For clarity, the steps for blur kernel estimation in a single scale is summarized in **Algorithm 3**.

Algorithm 3 Blur kernel estimation in a single scale

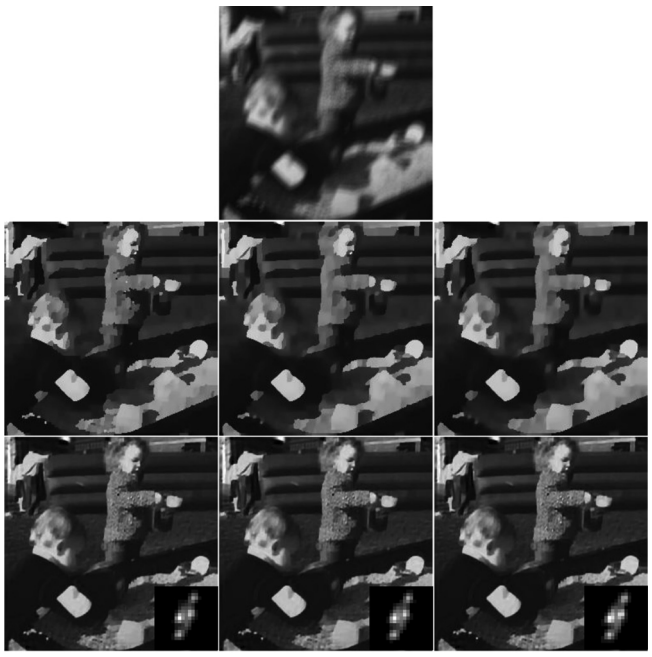
input : downsampled blurry image.
initialization : $k^{(0)} \leftarrow$ upsampled coraser scale kernel.
for $i = 1 : 5$ **do**
  solve $u^{(i)} = \arg\min_u J(u, k^{(i-1)})$ via **Algorithm 1**.
  solve $k^{(i)}$ using (15).
end for
output : intermediate image û and blur kernel k̂.

3.4. Non-blind restoration

Given the intermediate image u and the blur kernel $\hat{k}$, the final deblurred image can be produced by utilizing a non-blind restoration procedure. In literature, the Laplacian prior-based approach [58] is one of the most popular choices for standard non-blind deblurring, which has been used in Lai *et al.* [2] for fair comparisons of different blind approaches. We note that as dealing with real blurred images, the non-blind scheme in [38] is found more appropriate. In addition, as images are partially saturated, the recent non-blind deblurring method by Whyte et al. [59,60] is considered more robust and has been also used in [2]. With the above considerations, all the experimental analysis and comparisons are

Algorithm 2 Discriminative image smoothing

input : vectorized image **u**, parameters $\frac{\lambda}{\rho}$, t, σ, α, β.
initialization : $\mathbf{z} \leftarrow \mathbf{u}$, $W_o \leftarrow$ zero, $V_o \leftarrow$ zero.
for $l = 1 : 4$ **do**
  update $\mathbf{W}_o$, V_o by computing (2) and (10).
  update **z** by solving the linear system in (13).
endfor
output : reshape **z** into a 2D image z.



(a) (b) (c)

Fig. 1. Sensitivity analysis of the proposed model (1) on the spatial scale σ. The first row presents the blurred image image01-kernel01 (225 × 225) from the dataset of Levin et al. [19]. The intermediate image and final deblurred image accompanied with the estimated kernel corresponding to each setting of σ are provided in each of three columns (from the second to the third rows). The PSNR of final deblurred images are: (*a*) 31.93 dB ($\sigma = 1$); (*b*) 32.03 dB ($\sigma = 3$); (*c*) 32.03 dB ($\sigma = 5$).

(a) (b) (c)

Fig. 2. Sensitivity analysis of the proposed model (1) on the spatial scale σ. The first row presents the blurred image image04-kernel05 (225 × 225) from the dataset of Levin et al. [19]. The intermediate image and final deblurred image accompanied by the estimated kernel corresponding to each setting of σ are provided in each of three columns (from the second to the third rows). The PSNR of final deblurred images are: (*a*) 34.92 dB ($\sigma = 1$); (*b*) 34.66 dB ($\sigma = 3$); (*c*) 34.11 dB ($\sigma = 5$).

being made as fair as possible, ensuring that the same non-blind deblurring algorithm is applied in each group of experiments. An exception is the comparison made on the benchmark dataset Köhler et al. [61], where the compared approaches are run by their own authors who are requested to provide their best results as claimed in [61]. As for this dataset, the non-blind scheme in [38] is used to produce our final deblurred images. In spite of that, it is definite that the primary component affecting the final deblurring quality is the precision of kernel estimation.

4. Analysis

The analysis made in this part is to make clear in an intuitive perspective the motivation of the discriminative image prior as formulated in (1). Although there are no any theoretical results or mathematical proofs, a few aspects of empirical analysis are largely getting the point.

4.1. Specification of model parameters

There are four tunable parameters in the proposed model (1), namely, σ,α,β,t We find that, they can be set as 3, 0.1, 0.5, and 0.05, respectively, for acceptable performance in a majority of benchmark and real-world experiments. While, we should note that other candidates for settings on σ, α, β, t are possible to better serve the task. Although automatic estimate or learning of those parameters is appealing, it is beyond the scope of this paper.

According to the specifications, α set as 0.1 implies that the proposed image prior (1) is really heavy-tailed and unnatural, which is similar in form to several previous algorithms [28,29,30]. Yet, the model (1) implements the required DPD due to the newly introduced adaptive weight $\varpi_{o,p}(u)$ defined at each image pixel. In the meanwhile, t set as 0.05 implies that $1/((D_o(u))^{\beta} + \varepsilon)$ plays a different role from $1/((S_o((p))^{\beta} + \varepsilon)$ in terms of the regularization strength. Actually, the former discriminates the blurred image from its sharp one, while the latter is responsible for amending the intermediate image. More concretely, the latter recognizes and smooths out the interfering faint textures or other artifacts, which are known harmful to the accurate kernel estimation. Computations on both synthetic and real-world images show that, $t = 0.05$ and $\beta = 0.5$ are indeed robust choices for making the model (1) discriminative as a whole and also applicable for achieving state-of-the-art blind deblurring performance in various imaging scenarios.

It is additionally found that choice of the scale σ is not much sensitive in terms of deblurring quality as demonstrated in Figs. 1 and 2. As σ is specified to 1, 3, 5, respectively, the corresponding deblurring results including the intermediate sharp image, the estimated kernel and the final deblurred image are perceptually rather similar to each other in both figures. In addition, image PSNR results also demonstrate that σ being set as 3 is an appropriate choice. Hence, the insensitivity of our prior to σ greatly relieves us the annoying parameter tuning, which is much different from what is done in the original RTV [5]. The reason is that, the faint details or other artifacts to be smoothed out by $1/((S_o(p))^{\beta} + \varepsilon)$ are generally not in large scales due to the complementary role played by the term $1/((D_o(u))^{\beta} + \varepsilon)$.

4.2. Discriminativeness and effectiveness

Firstly, two benchmark datasets proposed by Levin et al. [20] and Köhler et al. [61] are used to test the discriminativeness of the prior (1) in discriminating between blurry and sharp images. In [20], each ground truth gray-scale image is convolved by 8 blur kernels of sizes from 13 × 13 to 25 × 25, producing 32 test images

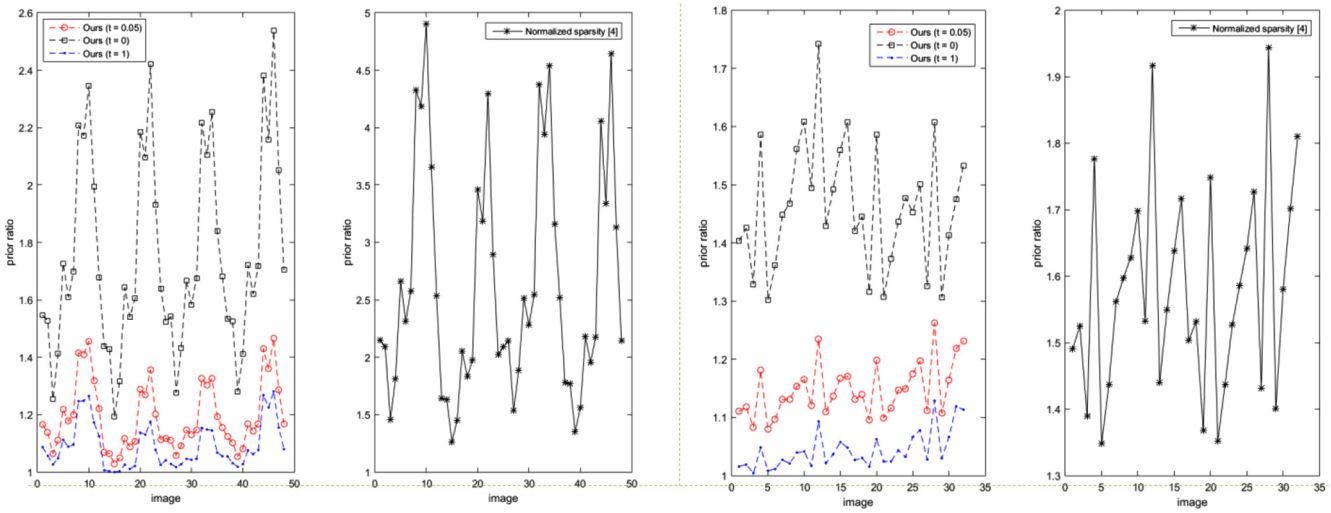


Fig. 3. Ratio curves of our prior ($t = 0.05$, $t = 0$, $t = 1$) and the normalized sparsity prior [4]. (*a*) Levin et al. [20]; (*b*) Köhler et al. [61].

altogether. While in [61], every true color image is convolved by 12 blur kernels created to approximate the real image blurs, with 48 test images generated in total. To be noted, as pointed out by the authors blur kernels numbered 1, 3, 4, 5, 8, 9, 11 tend to be approximately spatially uniform, while blur kernels 2, 6, 7, 10, 12 appear non-uniform. For each image pair (g, u) in the two datasets, we can calucate its prior ratio as $R(g)/R(u)$ with the parameters specified in Section 4.1.

Ideally, prior ratios of all image pairs should be greater than one, reflecting preference of our model over a sharp image to its blurry counterpart. Then, the curve of prior ratios is plotted for each dataset, as shown in Fig. 3. It is observed that both ratio curves are above the level line valued one, therefore demonstrating the discriminativeness of the model (1) on the two image sets. The ratio curve of the normalized sparsity [4], i.e., L1/L2, is plotted for comparison as shown in Fig. 3, too. It is clear that L1/L2 is more dis-

Table 1
Average statistics of PSNR (dB) of the final deblurred images corresponding to each blind deblurring approach on the dataset by Levin et al. [20].

Blurred Image No.	[1]	[25]	[19]	[41]	[23]	[40]	Ours
image01	30.40	29.87	31.59	**31.97**	**31.78**	30.45	**32.20**
image02	29.39	29.55	**30.35**	29.49	**30.02**	29.11	**31.28**
image03	31.61	32.43	**33.18**	32.56	**33.71**	31.42	**32.99**
image04	26.95	**30.60**	29.36	**30.17**	29.44	27.85	**30.99**
Average	29.59	30.61	**31.12**	31.05	**31.24**	29.71	**31.87**

Table 2
Average statistics of PSNR (dB) of the final deblurred images corresponding to each blind deblurring approach on the dataset by Köhler et al. [61].

Blurred Image No.	[1]	[25]	[4]	[5]	[26]	[64]	[62]	[17]	Ours
image01	25.89	30.60	28.20	**31.63**	**30.97**	29.01	28.74	29.83	**31.69**
image02	19.48	**24.87**	22.50	23.77	**25.07**	22.43	23.57	24.05	**25.58**
image03	25.54	30.81	27.35	**30.95**	**31.56**	27.76	28.04	30.55	**31.65**
image04	19.53	**26.78**	23.61	26.49	**26.72**	23.81	24.48	25.50	**27.14**
Average	22.61	**28.29**	25.42	28.21	**28.58**	25.75	26.21	27.48	**29.02**

Table 3
Average statistics of PSNR (dB) of the deblurred images corresponding to each deep model [14,68,73] on Levin et al. [20] and Köhler et al. [61].

Dataset	Levin et al. [20]			Köhler et al. [61]		
Blurred Image No.	[14]	[68]	[73]	[14]	[68]	[73]
image01	25.76	25.51	27.51	28.52	28.60	29.42
image02	24.23	24.57	25.46	22.55	22.46	23.19
image03	25.40	25.51	26.38	28.55	28.66	29.81
image04	26.13	25.88	27.34	23.74	23.68	24.78
Average	25.38	25.37	26.67	25.84	25.85	26.80

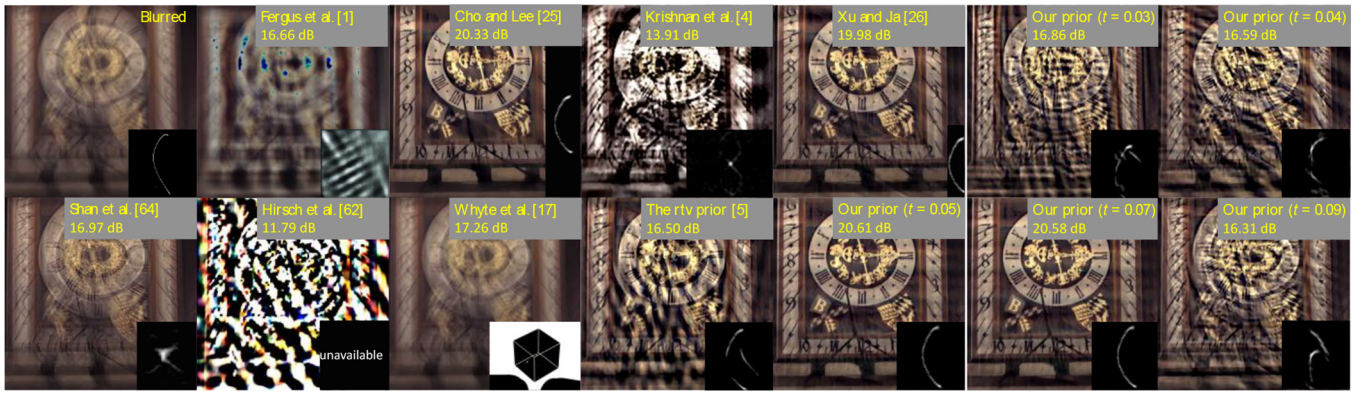


Fig. 4. Restored images with estimated kernels corresponding to each blind method, including the RTV prior and our model (1) with other settings of t on the image image02-kernel08. The results of [1,25,4,26,64,62,17] are the best ones as remarked in [61], whose kernel sizes are different from each other.

criminative than the model (1) on both datasets. In fact, computations on other numerous image pairs have shown that the discriminativeness of L1/L2 is universal to a great degree. However, the success of blind image deblurring is not solely decided by the discriminativeness between blurry and sharp images. On the one hand, the blind deblurring performance of L1/L2 is questionable in both synthetic and real-world experiments as claimed by both [2,3]. On the other hand, existing models for blind image deblurring are mostly non-discriminative including those ones proposed very recently, e.g., [3,15,30,38].

Let's dive into the image prior (1) a little more bit. As the balance coefficient t is set as 0, we see from the prior ratio curve in Fig. 3 that our degenerated prior becomes fairly more discriminative than the case as t is set as 0.05. Comparatively, as t is set as 1 our degenerated prior turns much less discriminative. In consequence, it is concluded that the regularization term $\sum_p|\partial_x u_p|^\alpha/(D_x(u))^\beta+\sum_p|\partial_y u_p|^\alpha/(D_y(u))^\beta$ has rised a primary role in discriminating a sharp image from its blurred counterpart, while the term $\sum_p|\partial_x u_p|^\alpha/(S_x(p))^\beta+\sum_p|\partial_y u_p|^\alpha/(S_y(p))^\beta$ is more capable of driving minimum, i.e., intermediate sharp images such as the ones in Figs. 1 and 2, physically meaningful to reasonable blur kernel estimation. That is because, this term is able to locate and remove the interfering faint textures or other artifacts which are considered harmful to accurate kernel estimation.

On the second, the effectiveness of the proposed prior is to be undergone a preliminary validation. The deblurred images and estimated kernels are produced by the proposed algorithm for the above two benchmark datasets, i.e., Levin et al. [20] and Köhler et al. [61]. For quantitative evaluation, the PSNR values are computed and listed in Tables 1 and 2 corresponding to the blurred images in the datasets [20,61], respectively. In the meanwhile, the PSNR values of several representative model-based algorithms are given for comparison, including both uniform [1,25,19,41,23,40,4,26,64] and non-uniform [17,62] ones.

It is obviously observed that our method has achieved comparable or superior performance to existing state-of-the-art algorithms on both image sets. When running our approach, the regularization parameter λ is set as 0.0005 for the dataset by Levin et al. [20] and 0.0001 for the dataset by Köhler et al. [61]. For visual perception, the image image02-kernel08 in Köhler et al. [61] is taken as an example, whose corresponding deblurred image and estimated kernel by each method are provided in Fig. 4. We see that the kernel by our method resembles those by [25] and [26] very much and naturally the final deblurred images are visually comparable to each other, in spite that the PSNR of our approach in this case is slightly less [25,26].

For the sake of comparison completeness, three recent deep learning-based blind deblurring methods are also tested on the two benchmark datasets. The key differences among those learning-based methods are in the network structure and learning mechanism. Specifically, the multi-scale convolutional neural networks in [14], the generative adversarial networks in [68], and the coupled convolutional-recurrent neural networks in [73] are respectively tested by directly running the corresponding pre-trained models. Correspondingly, Table 3 provides their average PSNR over each image in both datasets for quantitative comparison. It is clearly observed that their performance is really far from enough to deal with blurs in different scales. For example, even the best-performing one with the coupled convolutional-recurrent neural networks [73] among the three deep learning-based methods cannot be comparable to most of model-based blind deblurring methods as shown in Tables 1 and 2.

On the third, to validate the necessity of both terms in model (1), i.e., $\sum_p|\partial_x u_p|^\alpha/(D_x(u))^\beta+\sum_p|\partial_y u_p|^\alpha/(D_y(u))^\beta$ and $\sum_p|\partial_x u_p|^\alpha/(S_x(p))^\beta+\sum_p|\partial_y u_p|^\alpha/(S_y(p))^\beta$, the paper also makes an ablation study with both α,β set as 1 in (1), which allows inspecting the complementary role of its two originating models, i.e., normalized sparsity measure [4] and relative total variation [5]. Specifically, the benchmark dataset of Köhler et al. [61] is harnessed here for testing. According to Table 2, the average PSNR score of normalized sparsity measure [4] across 12 blur kernels corresponding to each image is 28.20 dB (image01), 22.50 dB (image02), 27.35 dB (image03), and 23.61 dB (image04), respectively. Comparatively, the average PSNR score of relative total variation [5] for each image is higher, i.e., 31.63 dB (image01), 23.77 dB (image02), 30.95 dB (image03), and 26.49 dB (image04), respectively. It is observed that, the normalized sparsity measure is indeed far more satisfactory in blind deblurring performance, while RTV performs fairly better than the normalized sparsity measure. A more interesting observation is that, RTV has performed comparably to [25,26] or even superiorly to [1,64,62,17]. While, we note that the composite discriminative prior (1) has outperformed both the normalized sparsity measure and relative total variation. Besides, Fig. 4 also visually presents the influence of other settings of t on the performance of the composite prior (1), revealing the significance of

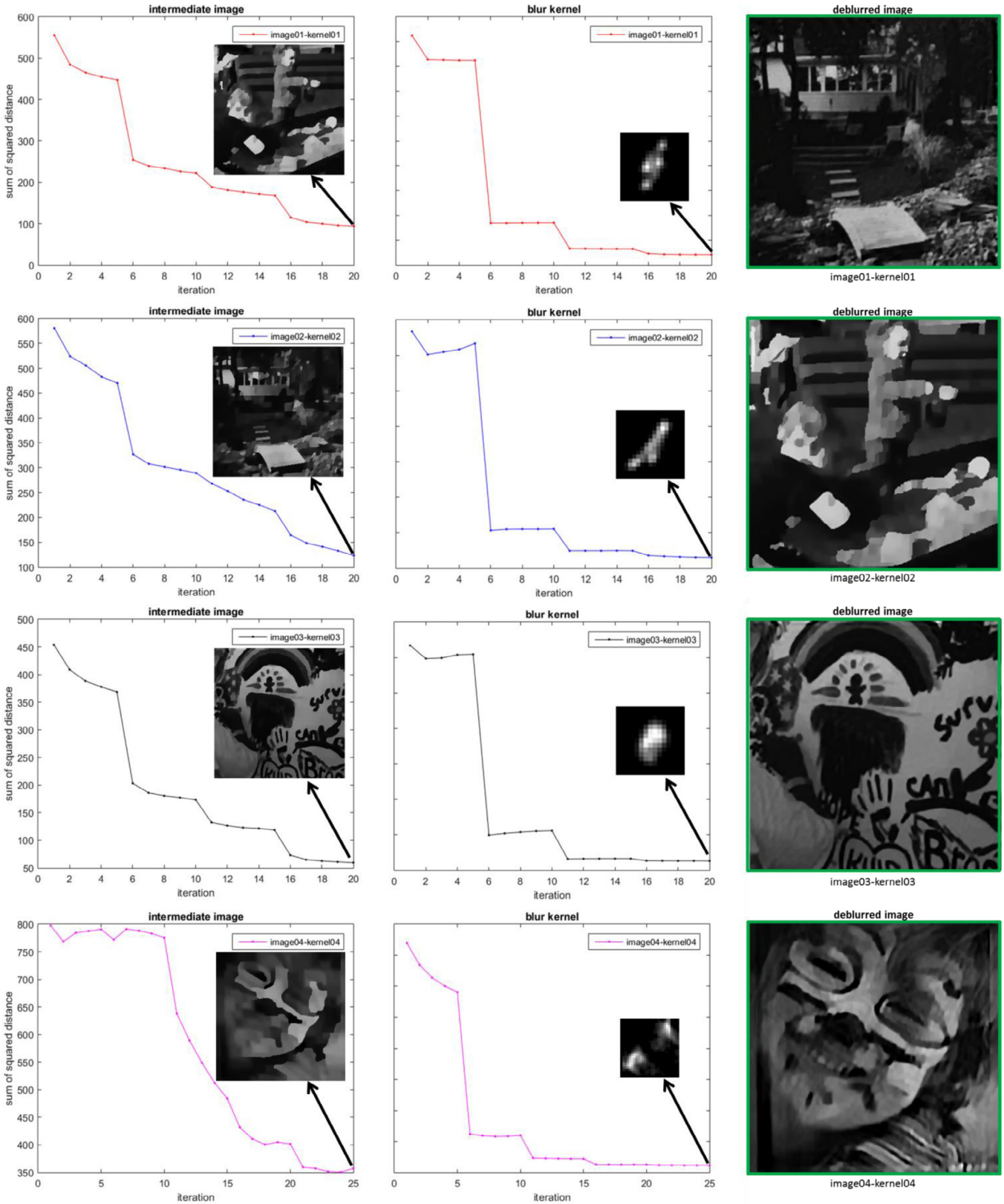


Fig. 5. Analysis on convergence of the proposed method in terms of the sum of squared distance. Left to right: intermediate image, blur kernel, and deblurred image; Top to bottom: image01-kernel01, image02-kernel02, image03-kernel03, and image04-kernel04.

choosing a proper setting for balancing the normalized sparsity measure and relative total variation. Hence, $\sum_p |\partial_x u_p|^\alpha / (S_x(p))^\beta + \sum_p |\partial_y u_p|^\alpha / (S_y(p))^\beta$, $\sum_p |\partial_x u_p|^\alpha / (S_x(p))^\beta + \sum_p |\partial_y u_p|^\alpha / (S_y(p))^\beta$ would be both indispensible to the success of the proposed blind deblurring approach. Experimental results in Section 5 on the more challenging synthetic and real-world datasets

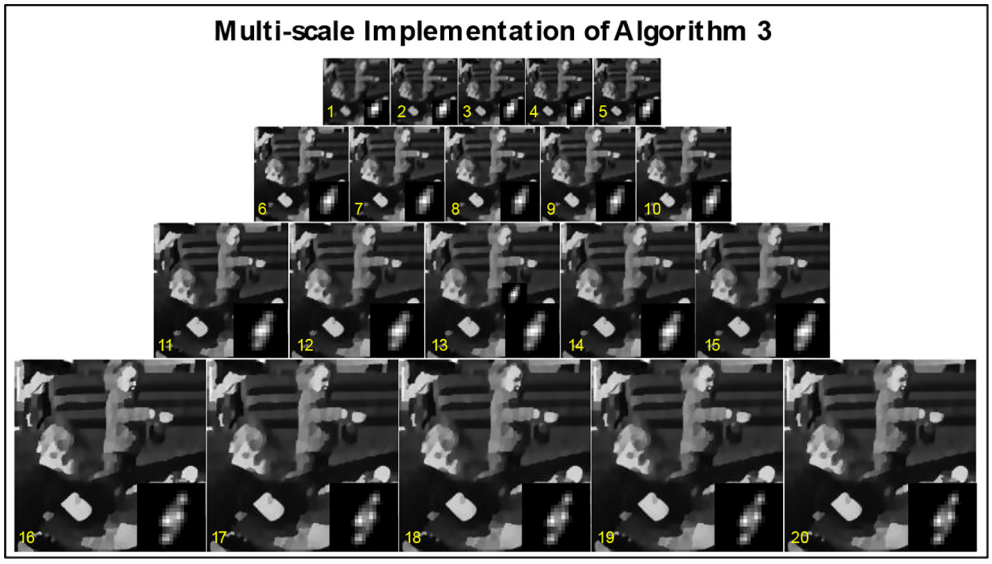


Fig. 6. The multi-scale estimates (5 iterations per scale) of blur kernel and intermediate sharp image for image01-kernel01.

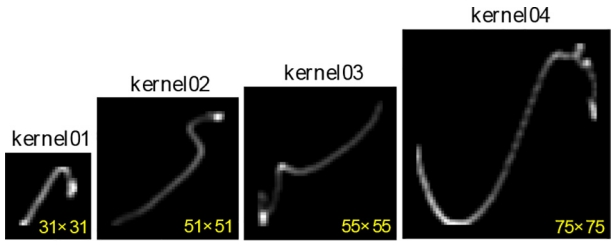


Fig. 7. Blur kernels with different sizes used for generating the 100 synthetic uniform blurry images in Lai et al. [2].

are to further demonstrate the merit of our proposed prior (1) and its superiority to many state-of-the-art models.

4.3. Convergence and computational cost

We take four blurry images image01-kernel01, image02-kernel02, image03-kernel03, and image04-kernel04 in the dataset of Levin et al. [20] for example, demonstrating convergence of the proposed approach in terms of sum of squared distance over both estimated intermediate images and blur kernels. Note that, the sum of squared difference over images makes more sense to monitor the algorithm convergence than that over kernels. The convergence curves for each example are provided in Fig. 5, respectively.

Table 4
Average statistics of PSNR (dB), SSIM [66], and no-reference (no-ref.) metric [67] of the final deblurred images corresponding to each blind deblurring approach on the 25 blurred images generated by **kernel01 (31 × 31)** in the dataset of Lai et al. [2].

Metric	Type	[1]	[25]	[26]	[4]	[19]	[17]	[41]	[27]	[65]	[44]	[40]	[38]	[23]	Ours
PSNR	**ALL**	16.66	19.07	**22.64**	19.78	19.45	18.95	**22.26**	22.11	21.25	21.38	21.44	21.50	21.28	**23.47**
SSIM	**ALL**	0.5429	0.6846	**0.8346**	0.7335	0.7287	0.6576	0.8181	0.8102	0.7826	0.8024	0.7774	0.8272	**0.8191**	**0.8880**
No-ref.	**ALL**	−18.05	−11.94	**−10.43**	−11.15	−11.23	−13.91	−10.53	−10.82	−11.30	−10.86	−10.81	**−10.44**	−10.87	**−10.28**

Table 5
Average statistics of PSNR (dB), SSIM [66], and no-refrence (no-ref.) metric [67] of the final deblurred images corresponding to each blind deblurring approach on the 25 blurred images generated by **kernel02 (51 × 51)** in the dataset of Lai et al. [2].

Metric	Type	[1]	[25]	[26]	[4]	[19]	[17]	[41]	[27]	[65]	[44]	[40]	[38]	[23]	Ours
PSNR	**ALL**	15.67	17.65	**21.65**	18.02	18.77	17.59	**21.05**	20.96	18.58	18.74	19.33	19.85	18.13	**21.92**
SSIM	**ALL**	0.4606	0.5892	**0.8096**	0.6433	0.6812	0.5183	0.7658	**0.7765**	0.6172	0.6774	0.6547	0.7537	0.6460	**0.8322**
No-ref.	**ALL**	−18.02	−13.00	−11.64	**−11.55**	−12.29	−14.75	**−11.38**	−12.10	−12.21	−12.40	−11.73	−13.04	−12.86	**−11.27**

Table 6
Average statistics of PSNR (dB), SSIM [66], and no-refrence (no-ref.) metric [67] of the final deblurred images corresponding to each blind deblurring approach on the 25 blurred images generated by **kernel03 (55 × 55)** in the dataset of Lai et al. [2].

Metric	Type	[1]	[25]	[26]	[4]	[19]	[17]	[41]	[27]	[65]	[44]	[40]	[38]	[23]	Ours
PSNR	**ALL**	15.25	18.06	**20.95**	17.42	18.95	17.39	**21.20**	19.96	19.21	18.84	18.90	19.01	19.15	**21.87**
SSIM	**ALL**	0.4571	0.6251	**0.7739**	0.6123	0.6923	0.5402	**0.7948**	0.7185	0.6802	0.6771	0.6470	0.6939	0.7295	**0.8419**
No-ref.	**ALL**	−17.96	−12.61	**−11.05**	−11.61	−11.73	−14.24	**−11.22**	−12.07	−12.21	−12.19	−11.31	−12.76	−12.30	**−10.91**

Table 7
Average statistics of PSNR (dB), SSIM [66], and no-refrence (no-ref.) metric [67] of the final deblurred images corresponding to each blind deblurring approach on the 25 blurred images generated by **kernel04 (75 × 75)** in the dataset of Lai et al. [2].

Metric	Type	[1]	[25]	[26]	[4]	[19]	[17]	[41]	[27]	[65]	[44]	[40]	[38]	[23]	Ours
PSNR	**ALL**	14.31	15.19	17.71	14.87	16.78	15.56	**17.81**	16.64	16.59	16.14	16.00	16.29	**18.27**	**19.13**
SSIM	**ALL**	0.3346	0.4383	**0.6128**	0.4228	0.5614	0.3988	0.5844	0.5736	0.4994	0.4907	0.4473	0.5304	**0.6474**	**0.6969**
No-ref.	**ALL**	−20.57	−16.32	−16.07	−16.50	−14.23	−14.04	**−13.34**	−18.22	−14.36	−16.66	**−11.70**	−17.69	−14.49	**−13.82**

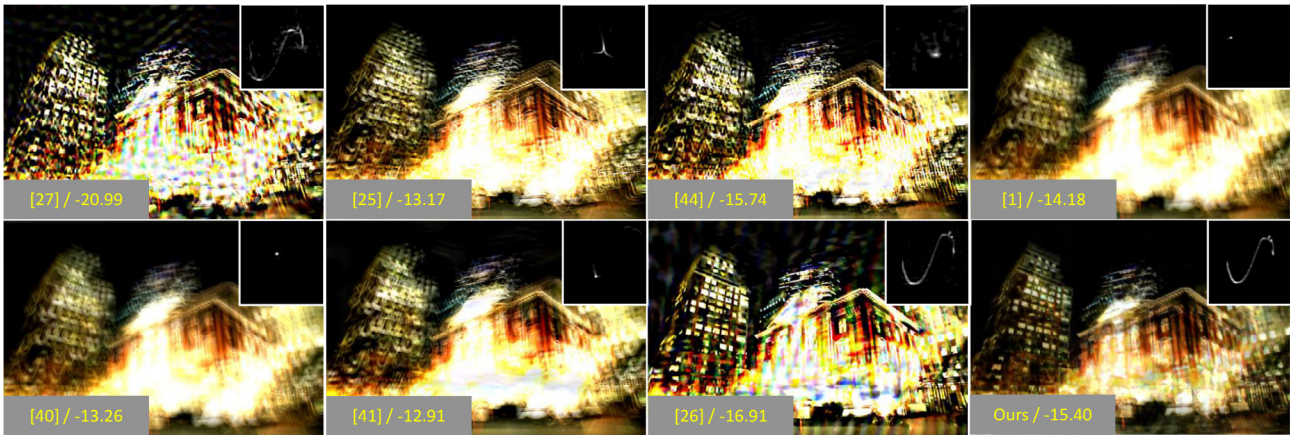


Fig. 8. Deblurring results of the blurred image saturated04-kernel04 in Lai et al. [2] corresponding to the top eight methods in terms of PSNR/SSIM. The value of no-reference metric [67] is shown in each image for visual perception assessment. Their PSNR/SSIM values are [27] (12.81 dB/0.4182), [25] (12.98 dB/0.2957), [44] (13.05 dB/0.4749), [1] (13.14 dB/0.3625), [40] (13.20 dB/0.3563), [41] (**13.53** dB/**0.5203**), [26] (**14.94** dB/**0.4482**), Ours (**16.29** dB/**0.4848**).

Fig. 9. Deblurring results of the blurred image text04-kernel04 in the dataset Lai et al. [2] corresponding to the top eight approaches in terms of PSNR/SSIM. The value of no-reference metric [67] is shown in each image for visual perception assessment. Their PSNR/SSIM values are [65] (13.63 dB/0.5611), [40] (13.72 dB/0.5634), [17] (13.74 dB/0.5605), [19] (13.83 dB/0.5903), [26] (13.91 dB/0.5724), [27] (**14.03** dB/**0.6386**), [23] (**16.30** dB/**0.7166**), Ours (**21.43** dB/**0.8437**).

Note that, as for each blur kernel, the overall alternating iterations are to depend on the size of the kernel. For example, the size of kernel01 is 19 × 19 and then a 4-scale procedure of **Algorithm 3** is to be implemented. Thus, as restoring image01-kernel01, there are 20 alternating iterations in the estimation process for either the kernel or the image. According to Fig. 5, it is observed that most error curves, especially those corresponding to the estimated intermediate images, decrease quickly from one scale to another and become flat after 2 or 3 iterations in each scale, demonstrating the stable performance of the proposed approach. Besides, along with each error curve the finally estimated intermediate image, blur kernel, as well as the deblurred image are provided. And, we also take the blurry image image01-kernel01 for example, presenting iterative estimates of the intermediate image and blur kernel in Fig. 6 for visual perception.

As remarked in Section 3.2, the major computational burden of the proposed approach is on solving Eq. (13) where PCG is harnessed in this paper. We take all the 8 blurred images generated by kernel01 in the dataset of Levin et al. [20] for example. On average, the running cost is about 1 min for each blind deblurring experiment on the computing platform as previously detailed in Section 3.2. In fact, we also implement our approach via a smilar numerical scheme in [58] and the corresponding average cost can be reduced to 15 s or so. In spite of that, all the experiment

Table 8
Average statistics of PSNR (dB) and SSIM [66] of the deblurred images corresponding to each deep model [14,68,73] on the 25 blurred images generated respectively by **kernel01**, **kernel02**, **kernel03**, **kernel04** in the dataset of Lai et al. [2].

Blur kernel		kernel01			kernel02			kernel03			kernel04		
Metric	Type	[14]	[68]	[73]	[14]	[68]	[73]	[14]	[68]	[73]	[14]	[68]	[73]
PSNR	**ALL**	19.09	18.96	19.39	17.43	17.51	17.83	17.35	17.49	17.55	15.80	16.00	15.82
SSIM	**ALL**	0.6209	0.5817	0.6682	0.4742	0.4615	0.5332	0.4890	0.4758	0.5266	0.3898	0.3734	0.4160

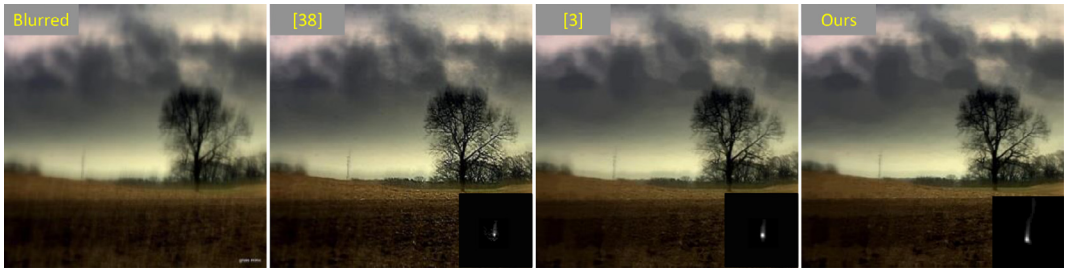


Fig. 10. Results for the natural (**N**) image 'outdoor' corresponding to the blind methods [3,38] and Ours with reasonable kernels produced.

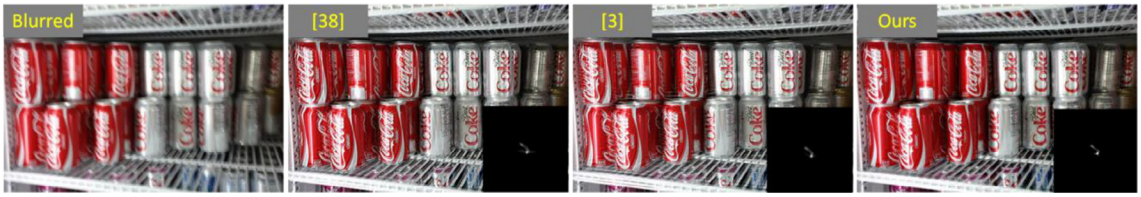


Fig. 11. Results for the manmade (**M**) image 'coke' corresponding to the methods [3,38] and Ours with reasonable kernels produced.

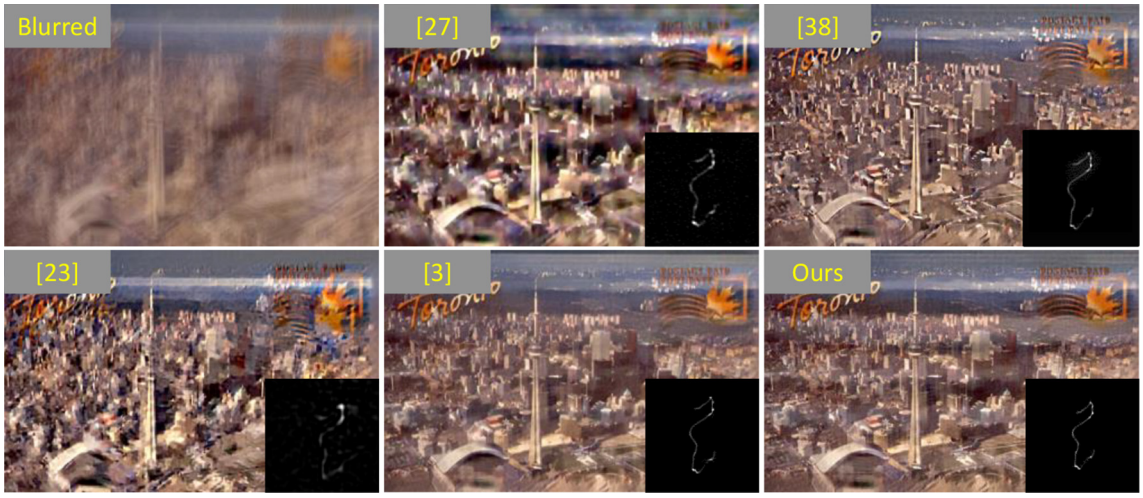


Fig. 12. Results for the manmade (**M**) image 'postcard' corresponding to the methods [3,23,27,38] and Ours with reasonable kernels produced.

results are produced by **Algorithm 3** owing to its more stable and precise estimation of blur kernels.

5. Results

This section validates the proposed approach on the datasets proposed by Lai et al. [2] with comparisons against the current representative blind deblurring algorithms: [1,4,17,19,23,25–27,38,40,41,44,65]. Note that the non-blind deconvolution schemes as detailed in Section 3.4 are used for final restoration, and hence the comparisons are fair and convincing. Besides the PSNR metric, the SSIM in [66] and the no-reference metric in [67] are also harnessed for quantitative assessment of different methods in this part. Note that, [67] is specifically proposed to evaluate the motion deblurring quality which is claimed consistent with human feelings and ratings to a certain degree.

The datasets in Lai et al. [2] include a synthetic one consisting of 100 blurred images generated by 4 blur kernels shown in Figs. 7

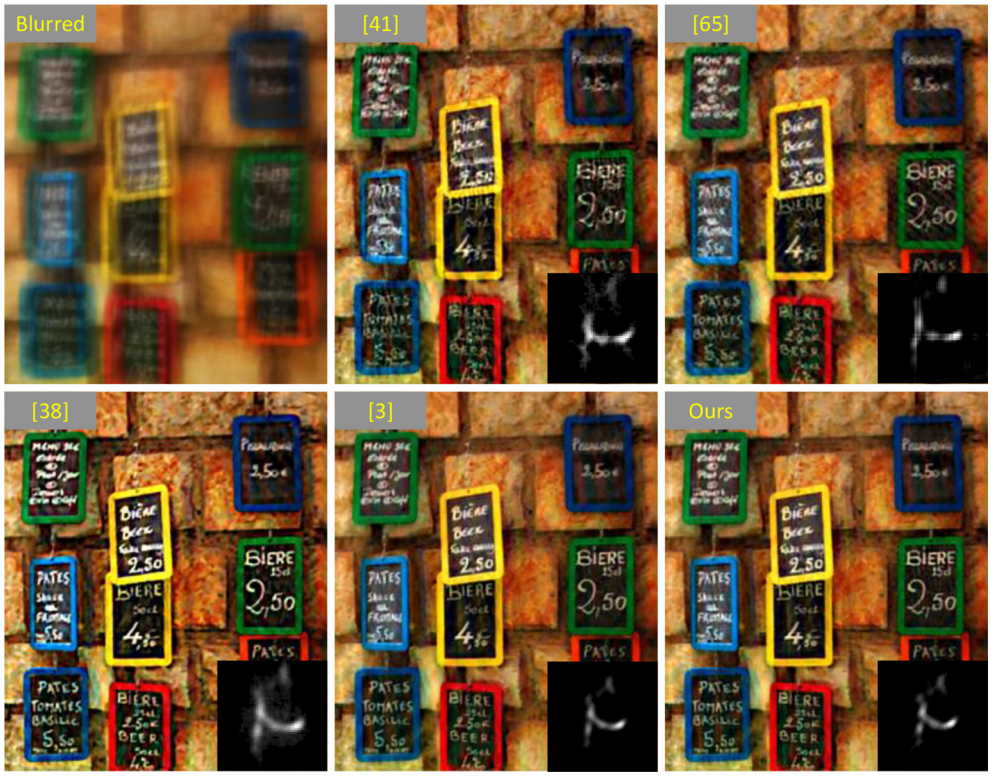


Fig. 13. Results for the manmade (**M**) image 'wall' corresponding to the methods [3,23,38,41,65] and Ours with reasonable kernels produced.

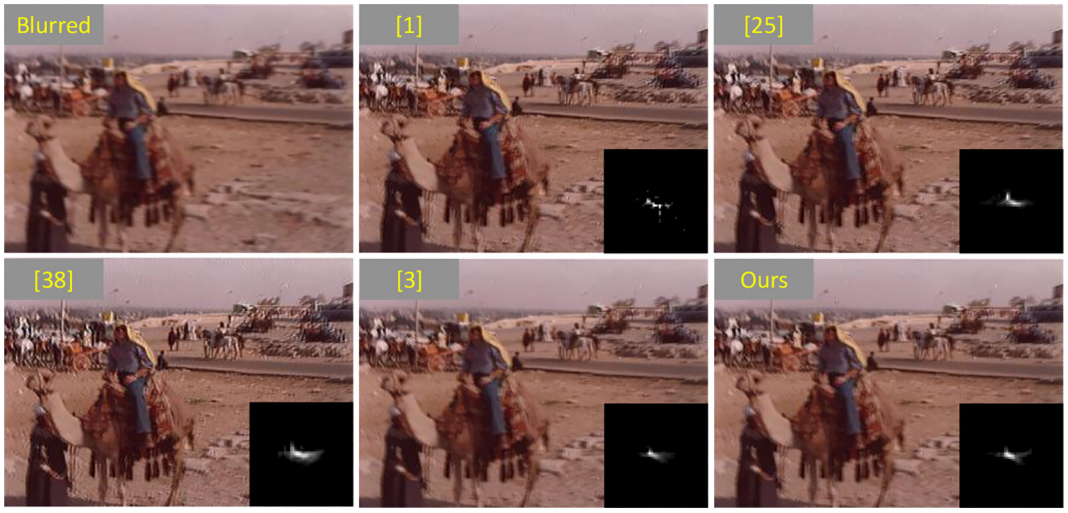


Fig. 14. Results for the people (**P**) image 'cario1979' corresponding to the methods [1,3,25,38] and Ours with reasonable kernels produced.

and 25 true clear images divided into 5 categories, i.e., natural (**N**), manmade (**M**), text (**T**), people (**P**), and saturated (**S**), as well as a real one containing 100 blurred color images collected from either previous deblurring works, or Flicker and Google Search, or those captured by the authors themselves, which also fall into the above five categories.

5.1. Synthetic experimental results

Tables 4–7 show the average statistics of the three metrics of deblurred images corresponding to each of the four blur kernels, respectively, as shown in Fig. 7. In each table, the row named '**ALL**' represent the average evaluation across the five image categories, i.e., **N**, **M**, **T**, **P**, **S**. It is seen that the overall performance of our proposed approach ranks the best in almost all scenarios in terms of either PSNR, or SSIM, or no-reference metric, proving its effective and robust functionality in dealing with different categories of blurred images with various kernel sizes.

We note an exception that in terms of the no-reference metric [67], our approach seems perform slightly inferior to [40,41] as dealing with images convolved by kernel04, as shown in Table 7. However this objective evaluation does not comply with the practical visual perception. In fact, the comparion should be more based on the PSNR and SSIM in this case. That is to say, sometimes [67] will fail to fairly measure the quality of deblurred images, especially for those satuated images shown in the following. A

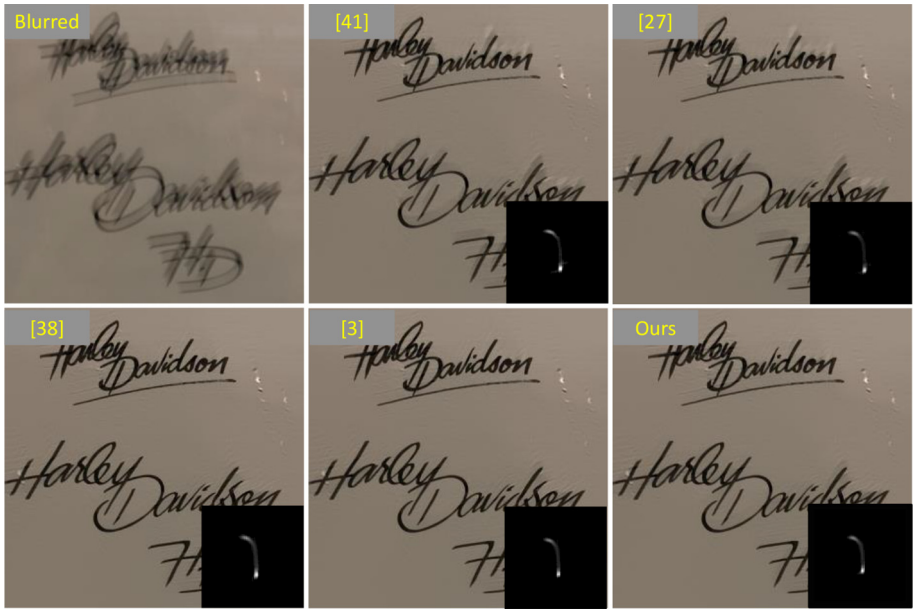


Fig. 15. Results for the text (**T**) image 'text2' corresponding to the methods [3,27,38,41] and Ours with reasonable kernels produced.

Fig. 16. Results for the text (**T**) image 'text3' corresponding to the methods [3,38] and Ours with reasonable kernels produced.

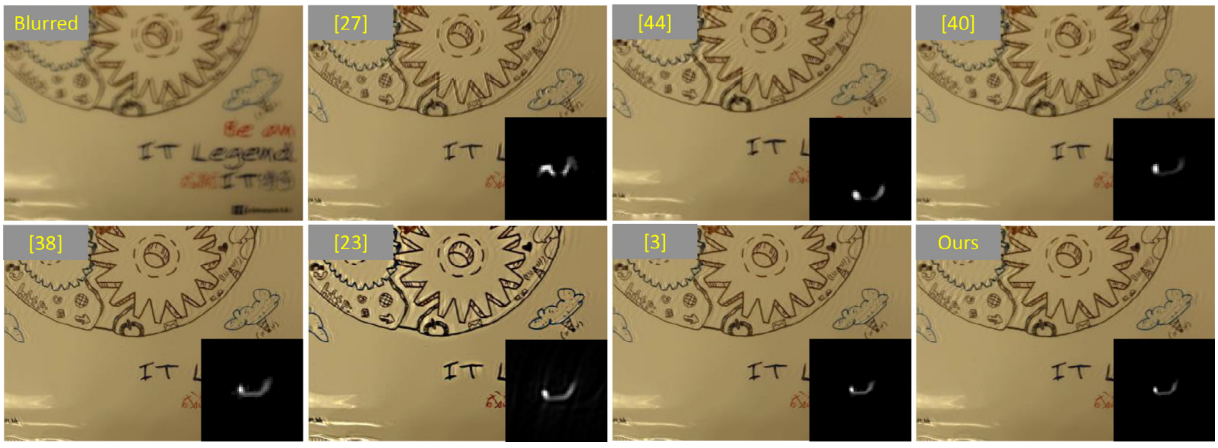


Fig. 17. Results for the text (**T**) image 'text6' corresponding to the methods [3,38] and Ours with reasonable kernels produced.

notable instance can be observed from Fig. 8, where our approach has produced a blur kernel with very high precision and of course a reasonably good deblurred image. It is found that, however, it only ranks the fifth among the eight compared methods in terms of no-reference metric. In fact, all the other approaches completely fail in this example. Thus, visual comparisons demonstrate that [67] is not applicable for fairly measuring the saturated image deblurring quality to some extent. A similar phenomenon is also noted in the text cases, e.g., the example shown in Fig. 9. Therefore, in this case PSNR and SSIM should be more relied on for fair assessement of

Fig. 18. Results for the saturated (**S**) image 'car2' corresponding to the methods [3,19,38,65] and Ours with reasonable kernels produced.

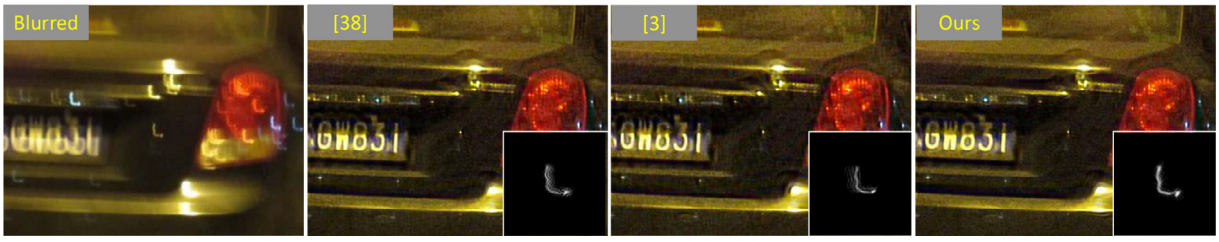


Fig. 19. Results for the saturated (**S**) image 'car5' corresponding to the methods [3,38] and Ours with reasonable kernels produced.

Fig. 20. Results for the saturated (**S**) image 'topresso' corresponding to the methods [3,38] and Ours with reasonable kernels produced.

Fig. 21. Results for the saturated (**S**) image 'family' corresponding to the methods [3,38,65] and Ours with reasonable kernels produced.

Table 9
Running time (in seconds) of [3,38], and the proposed method for each realistic blurry image as shown in Figs. 10–23.

Time	Fig. 10	Fig. 11	Fig. 12	Fig. 13	Fig. 14	Fig. 15	Fig. 16	Fig. 17	Fig. 18	Fig. 19	Fig. 20	Fig. 21	Fig. 22	Fig. 23
[38]	**304**	**741**	**384**	**199**	**613**	**1032**	**1167**	**391**	**948**	**356**	**348**	**471**	**1135**	**751**
[3]	**1383**	**3508**	**1580**	**845**	**2700**	**4508**	**5375**	**1764**	**3904**	**1489**	**1632**	**2012**	**5066**	**3155**
Ours	**974**	**2366**	**1010**	**606**	**1649**	**2879**	**3029**	**1138**	**3199**	**1128**	**1231**	**1547**	**3587**	**2573**

Fig. 22. Visual comparison among [3,38], and Ours on the saturated (**S**) image 'garden', where the proposed approach succeeds in producing a very reasonable kernel which naturally leads to visually acceptable deblurred image while [3] has failed to a great degree.

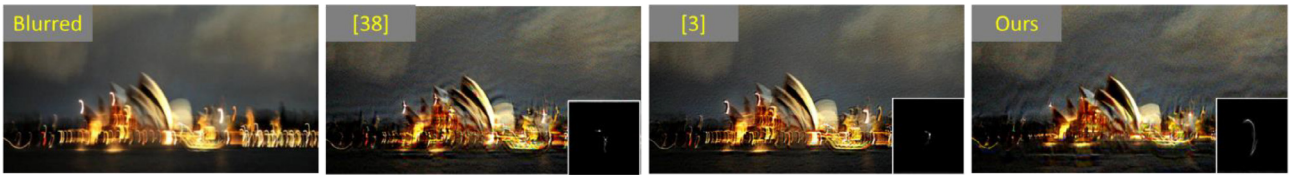


Fig. 23. Visual comparison among [3,38], and Ours on the saturated (**S**) image 'sydney_opera', where both [3,38] have failed to a great degree while Ours succeeds in producing a very reasonable kernel which naturally leads to a visually acceptable deblurred image.

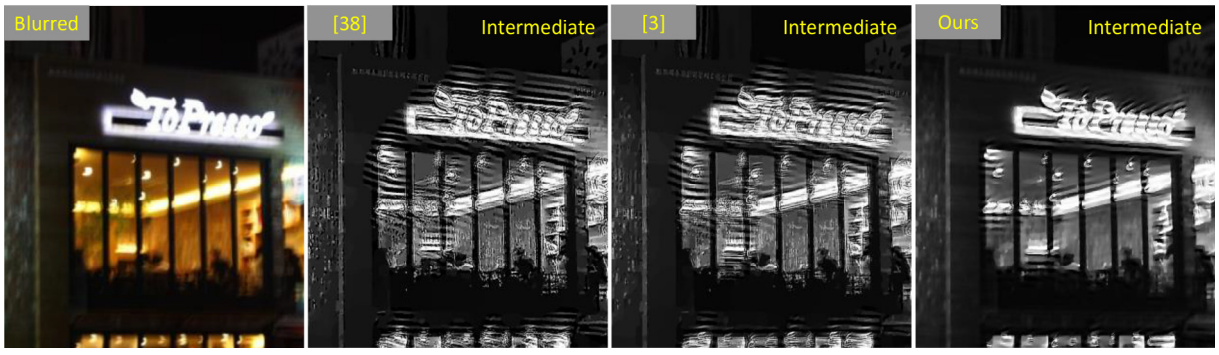


Fig. 24. Intermediate sharp images corresponding to the blur kernels estimated by the three blind approaches [3,38], and Ours shown in Fig. 20.

various algorithms. In brief, the comprehensive evaluation demonstrates that, our method achieves comparabe or superior performance in all the five scenarios of synthetic blur as compared against other approaches in Tables 4–7.

In addition, the aforementioned three deep learning-based approaches [14,68,73] are tested here, too, largely for the sake of comparison completeness. As stated above, the ordinary metrics PSNR and SSIM are found more convincing to evaluate the blind deblurring performance across different categories of blurred images. As such, Table 8 just provides the average statistics of PSNR and SSIM of deblurred images produced by each deep model [14,68,73], corresponding to each group of 25 blurred images generated by kernel01, kernel02, kernel03, and kernel04, respectively, in the dataset of Lai et al. [2]. It is obviously observed that all the three deep models perform either comparably or inferiorly to most of previous algorithms in Tables 4–7, let alone the proposed approach. Therefore, the generalization capability of existing deep learning-based blind deblurring methods [14,68,73] is really problematic as dealing with blurs in different imaging scenarios, e.g., natural, manmade, low-illumination, text, or people.

5.2. Realistic experimental results

Since there is not a reliable quantitative metric of measuring the deblurring quality for the real, particularly the saturated or text images, comparisons are made merely based on our visual perception. In this part, the practical performance of the recent breakthrough work [3] is also tested.

The comprehensive assessment validates that the proposed method is more robust than most of the compared approaches in Section 5.1, which achieves comparable (**N**, **M**, **P**) or better (**T**, **S**) performance on the whole set of 100 real images. In the mean-

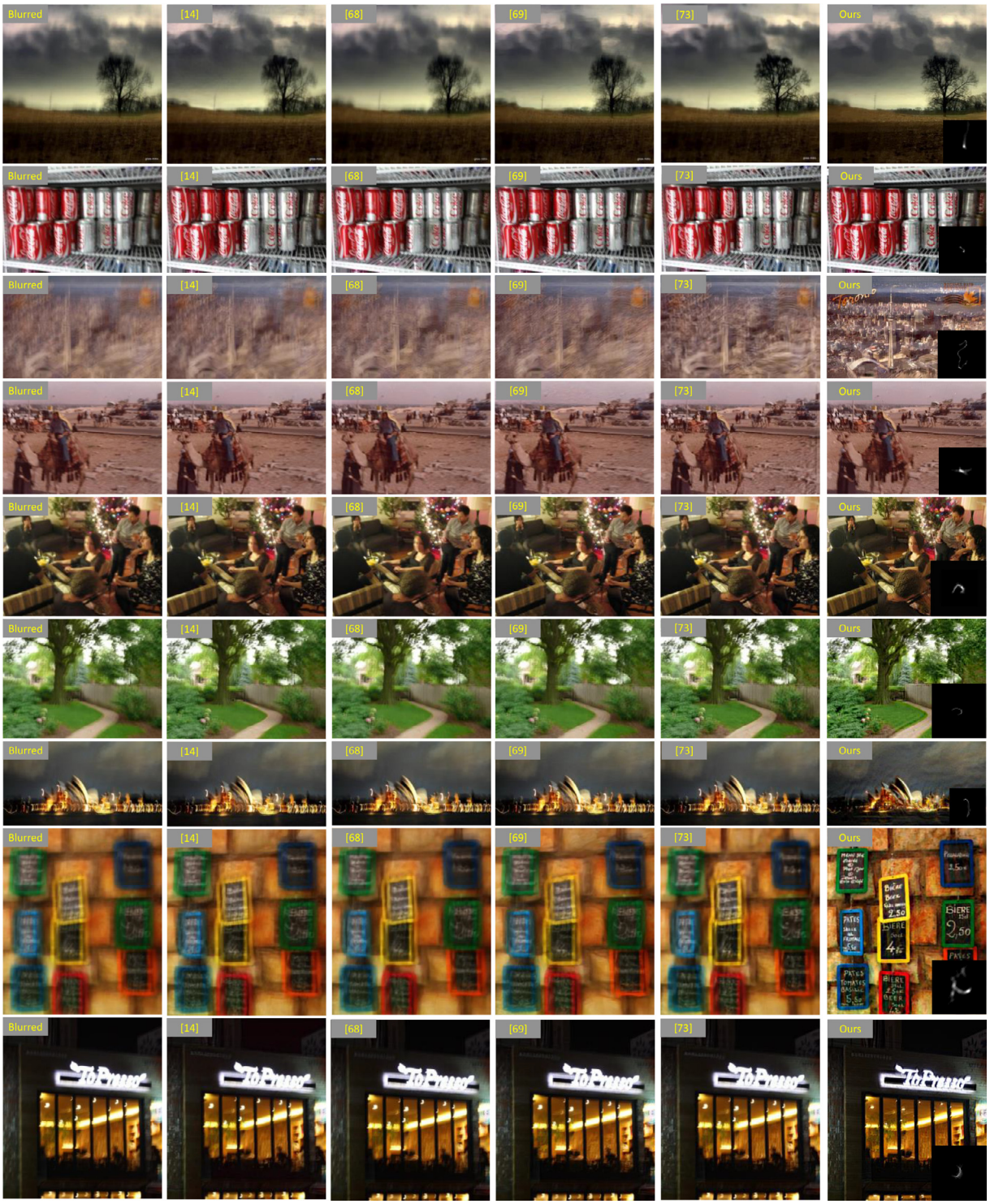


Fig. 25. Deblurring results on the real blurred images corresponding to the four recent deep learning-based methods [14,68,69,73] the proposed approach (Ours).

while, it is found that our approach performs much comparably to both [3,38] on the five categories of blurred images, particularly on those text and saturated ones.

Considering the presentation clarity and restricted paper space, we just take several challenging blurry images in Lai *et al.* [2] for example, and provide deblurred images corresponding to those blind approaches capable of producing reasonable blur kernels. Figs. 10–21 provide the deblurred results corresponding to 12 blurred images for visual perception. It is seen that the proposed method, Pan *et al.* [3], and Pan *et al.* [38] can produce plausible kernels in most scenarios. Most of the kernels are with tiny differences, which naturally lead to visually similar and acceptable deblurred images. Nevertheless, other approaches just get occasional success in those challenging experiments.

In spite of similar deblurring performance among [3,38], and our approach on the above 12 blurred images, we note that as for real blurry images such as 'car4', 'night1', 'night4', 'notredame', 'text1', 'text12', 'garden', and 'sydney_opera' in Lai *et al.* [2], either

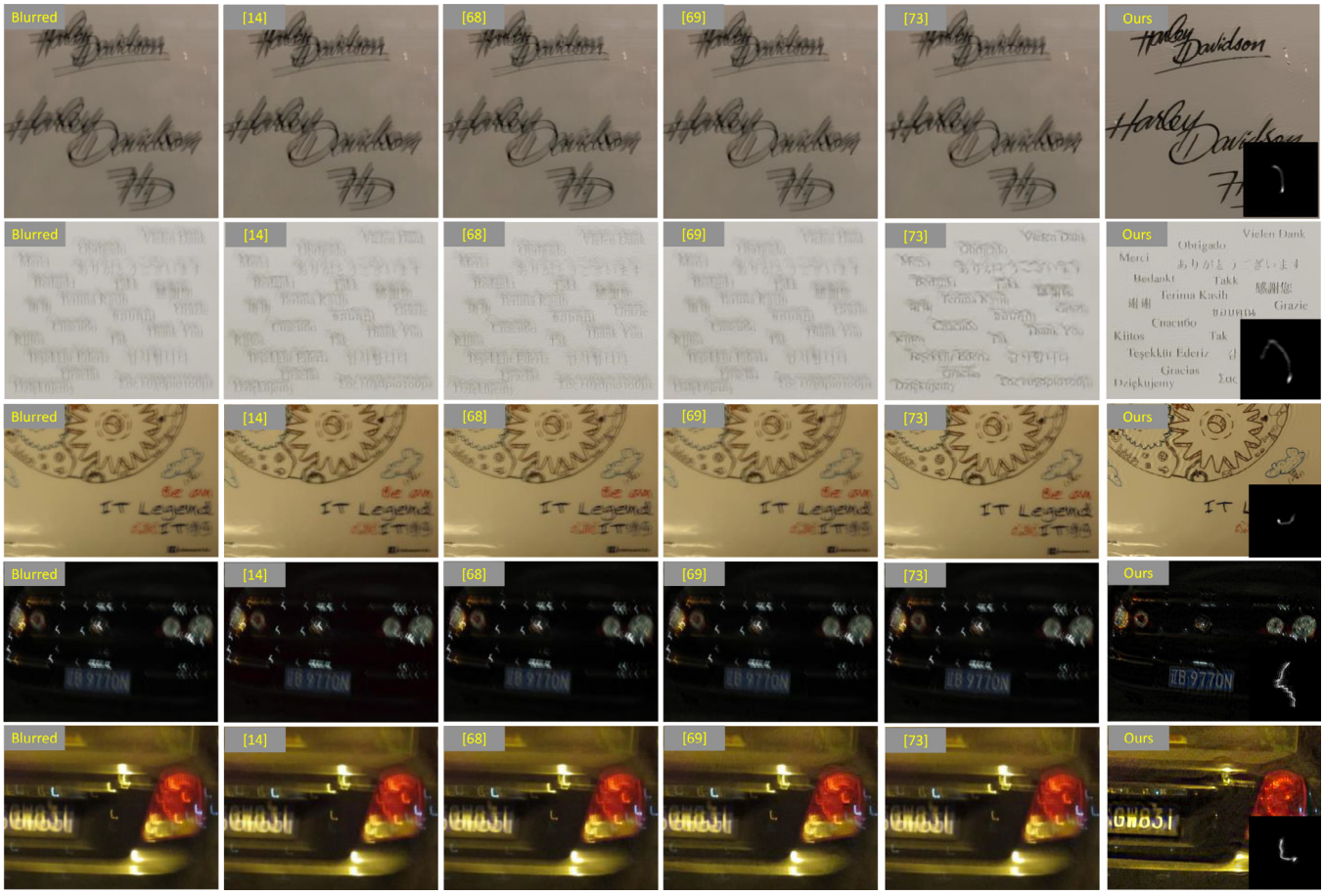


Fig. 26. Deblurring results on the real blurred images (texts and car plates) corresponding to the four recent deep learning-based methods [14,68,69,73] and the proposed approach (Ours).

[38] or [3] or ocassionally both will generate much less accurate kernels than those of our approach. We take the two saturated images 'garden' and 'sydney_opera' for example. It is found that, as for 'sydney_opera' both [3,38] completely fail to a great degree and as for 'garden', [3] fails either while [38] is much less accurate than the proposed method. Figs. 22 and 23 provide the deblurring results for the three methods, from which it is observed our method succeeds in recovering a very plausible blur kernel for each image, validating more robustness of the proposed approach than [3,38].

In addition, two notable advantages deserve noting for the proposed approach. On the one hand, our method works with a discriminative model while both [3,38] rely heavily on the L0-norm-based gradient regularization. However, the hybrid regularizations in [3,38] are apparently not necessarily preferrable to the sharp images, due to which the continuation strategy has been exploited in both [3,38] as well as most of other blind approaches. On the other hand, although Pan *et al.* [3] introduces a new L0-norm-based dark channel prior for blind deblurring, it results in a higher computational cost than the proposed approach. For the sake of clearness, the running time for all the images in Figs. 10–23 is summarized in Table 9 (using the same experimental platform mentioned in Section 3.2), from which the efficiency difference among [3,38], and the proposed approach can be observed. Clearly, both the proposed approach and [3] are computationally more expensive than [38] on all the compared images. Note that, the computational burden of the deduced plug-and-play numerical algorithm is mainly on the discriminative smoothing in Section 3.2, whose fast implementation is another critical topic to be studied in the near future and would naturally speed up the proposed blind deblurring algorithm. In spite of that, the proposed approach in the current is still more efficient than [3]. For instance, the proposed approach runs about 21 min for the saturated image 'topresso' in Fig. 20 which is of size 606×690, while [3] costs about 27 min. And, what is more important, as for those low-illumination blurry images, our approach can generate visually more pleasant intermediate sharp images than [3,38]. We still take 'topresso' for example and Fig. 24 shows the intermediate result corresponding to each method.

As shown in Section 5.1, the generalization capability of existing deep models [14,68,73] is really far from enough to deal with blurs in different imaging scenarios, e.g., natural, manmade, low-illumination, text, or peoples. Prior to closing this paper, this subsection further inspects the practical performance of those deep models on the real blurred images. Besides the aforementioned three deep models [14,68,73], additional two deep models in [69,70] are also tested here. One is our recent improved method upon DeblurGAN [68], dubbed DeblurGAN+ [69]. Another one in [70] is a classical model-inspired deep learning method, also published in the flagship computer vision conference, i.e., CVPR, in the same year as [68,73]. Comparisons are firstly made among [14,68,69,73], and the proposed method on the blurred images in Figs. 10–23. The deblurred images by [68,69] corresponding to the blurry images in Figs. 10–23 are provided in Fig. 25, and those corresponding to the blurry images in Figs. 15–19 are provided in Fig. 26. Since [14,68,69,73] are all fully end-to-end learning-based schemes, there are no estimated blur kernel available. Apparently, observed from Figs. 25 and 26 these deep-learning methods have

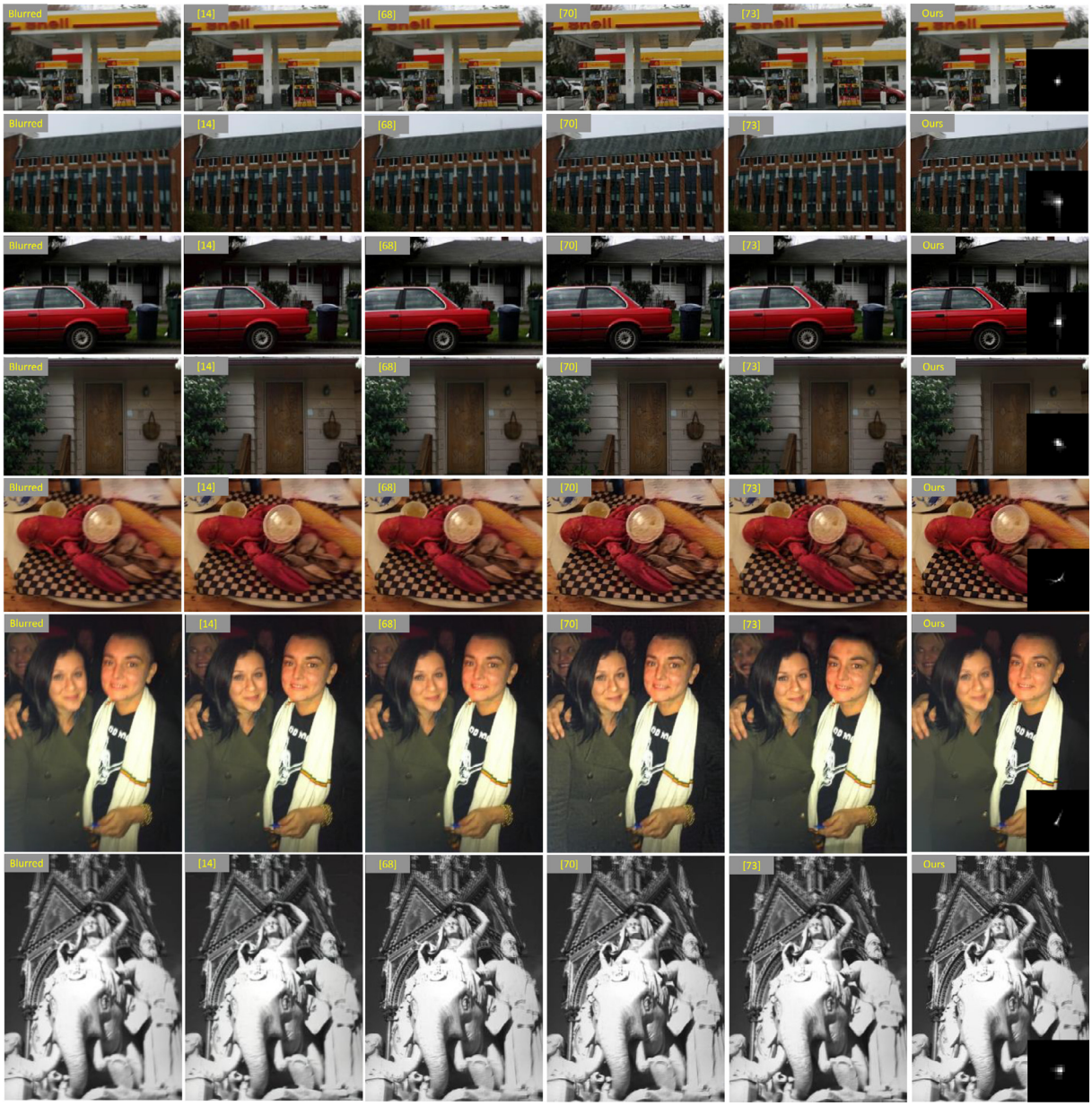


Fig. 27. Deblurring results on the real blurred images (spatially-invariant) corresponding to the four recent deep learning-based methods [14,68,70,73] and the proposed approach (Ours).

almost completely failed in removing the blur except the first and third images (i.e., 'outdoor' and 'cairo1979') in Fig. 25, where [14,68,69,73] generate comparable deblurred results to our proposed method. As for the end-to-end learning-based method in [70], it is not open-sourced while the authors kindly prepare for us the deblurring results corresponding to 15 real blurred images. They are also part of the 100 real images provided in Lai et al. [2]. But, compared with the blurry images in Figs. 10–23 the complexity of blur patterns or trajectories in those 15 images is much less. Actually, most of them are undergone small-scale out-of-focus or linear motion blurs, while most of the blurry images in Figs. 10–23 are undergone complicated middle or large-scale curvilinear motion blurs. For the sake of clearness, the 15 deblurred images estimated by [70] as well as [14,68,73] and our method are shown in Fig. 27 (mainly spatially invariant blur) and Fig. 28 (mainly spatially variant blur). It is clearly seen from both Figs. 27 and 28 that, the proposed approach has achieved fairly comparable performance to [14,68,70,73] on most of those blurred images, except the first spatially variant blurry image in Fig. 28 where [14,68,70,73] could generate clearer results than ours in the upper left corner of the image. However, we should note that the proposed method is not specifically made adaptive to spatially variant blurs. Therefore, according to above comparisons it can be concluded that the proposed approach is not only more effective but also more robust than state-of-the-art deep learning-based methods [14,68,69,70,73], particularly as for images undergone complicated curvilinear motion blurs, no matter they are of natural, manmade, people or low-illumination or text types.

6. Conclusion

Blind image deblurring, as a fundamental low-level vision problem, is far from being solved due to the challenging blur process in practical imaging, e.g., Gaussian-shaped kernels of varying sizes, ellipse-shaped kernels of varying orientations, curvilinear kernels of varying trajectories. It is comforting that in the past decade a

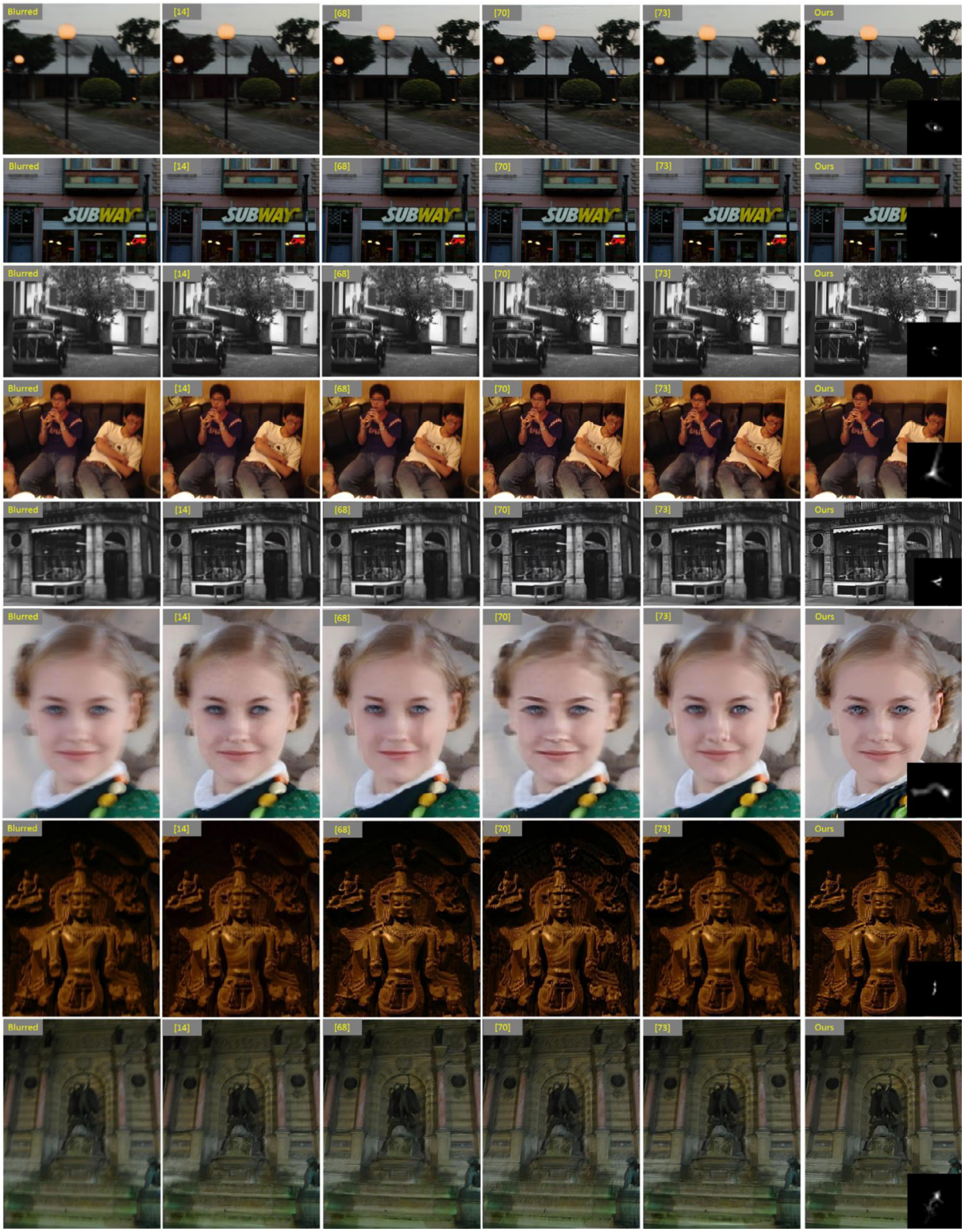


Fig. 28. Deblurring results on the real blurred images (spatially-variant) corresponding to the four recent deep learning-based methods [14,68,70,73] and the proposed approach (Ours).

great progress has been made wherein the most representative is the recent breakthrough work by Pan *et al.* [3] which is a lifting version of their previous work in Pan *et al.* [38]. In distinction to the previous methods, this paper is inspired by a rule of work from Albert Einstein: Out of clutter find simplicity, aiming to exploit the full potential of gradient-based approaches with the new proposal of a simple, robust yet discriminative prior for nonparametric blur kernel estimation.

Our research finds that the success of blind image deblurring requires dual principles of discriminativeness (DPD), which formally leads to a simple gradient-based spatially variant heavy-tailed prior. Numerous experimental results are provided to vali-

date the effectiveness of the new prior along with comparisons against state-of-the-art approaches on both synthetic and real blurred images in different scenarios, e.g., natural, manmade, low-illumination, text, people. We are convinced that the proposed method is indeed a promising new candidate solution for blind image deconvolution in spite of its simplicity in modeling. In particular, the results on the 100 real blurred images demonstrate well that our approach is not only more robust than most of the previous approaches but is also much comparable to Pan *et al.* [3] and Pan *et al.* [38] both of which are not discriminative as a whole and required to apply the additional continuation trick.

To briefly sum up, our new discriminative approach achieves decent performance on both synthetic and realistic blurry images, and could be served as a new start point to develop more reliable, robust, effective and efficient blind deblurring approaches. Considering that the current end-to-end deep learning-based methods cannot produce state-of-the-art deblurring results, we are to newly propose a fast blind deblurring paradigm via embedding discriminative deep edge-aware filters into the proposed regularization framework.

CRediT authorship contribution statement

Wen-Ze Shao: Conceptualization, Methodology, Software, Writing - original draft. **Yun-Zhi Lin:** Methodology, Software, Validation, Investigation. **Yuan-Yuan Liu:** Software, Validation, Visualization. **Li-Qian Wang:** Formal analysis, Writing - review & editing. **Qi Ge:** Formal analysis, Writing - review & editing. **Bing-Kun Bao:** Writing - review & editing. **Hai-Bo Li:** Conceptualization, Supervision, Writing - review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

The work was supported in part by the Natural Science Foundation of China (61771250, 11901299, and 61972213) and in part by the Fundamental Research Funds for the Central Universities (30918014108). Wen-Ze Shao is also grateful to Prof. Zhi-Hui Wei, Prof. Michael Elad, Prof. Yi-Zhong Ma, Dr. Min Wu, and Mr. Ya-Tao Zhang for their kind support in the past years.

References

[1] R. Fergus, B. Singh, A. Hertzmann, S.T. Roweis, W.T. Freeman, Removing camera shake from a single photograph, ACM Trans. Graph. 25 (3) (2006) 787–794.
[2] W.S. Lai, J.B. Huang, Z. Hu, N. Ahuja, M.H. Yang, A comparative study for single image blind deblurring, in: IEEE Conf. Computer Vision and Pattern Recognition (CVPR), 2016, pp.1701–1709
[3] J. Pan, D. Sun, H. Pfister, M.H. Yang, Blind Image Deblurring Using Dark Channel Prior, CVPR, 2016, pp. 1628–1636.
[4] D. Krishnan, T. Tay, R. Fergus, Blind Deconvolution Using a Normalized Sparsity Measure, CVPR, 2011, pp. 233–240.
[5] L.i. Xu, Q. Yan, Y. Xia, J. Jia, Structure extraction from texture via relative total variation, ACM Trans. Graph. 31 (6) (2012) 1–10, https://doi.org/10.1145/2366145.2366158.
[6] L. Xu, J.S. Ren, C. Liu, J. Jia, Deep convolutional neural network for image deconvolution, Adv. Neural Inf. Process. Syst. (NIPS) (2014) 1790–1798.
[7] C. Schuler, M. Hirsch, S. Harmeling, B. Scholkopf, Learning to Deblur, arXiv:1406.7444, 2014.
[8] J. Sun, W. Cao, Z. Xu, J. Ponce, Learning a convolutional neural network for non-uniform motion blur removal, CVPR, 2015, pp. 769–777.
[9] M. Hradis, J. Kotera, P. Zemčík, Convolutional neural networks for direct text deblurring, British Machine Vision Conference, vol. 10, 2015, pp. 6.1–6.13
[10] C. Schuler, M. Hirsch, S. Harmeling, B. Scholkopf, Learning to deblur, IEEE Trans. Pattern Anal. Mach. Intell. 38 (7) (2016) 1439–1451.
[11] A. Chakrabart, A neural approach to blind motion deblurring, in: European Conf. Computer Vision (ECCV), 2016, pp. 221–235
[12] P. Wieschollek, B. Schölkopf, H.P.A. Lensch, M. Hirsch, End-to-end learning for image burst deblurring, Asian Conf. Computer Vision (ACCV), 2016, pp. 35–51
[13] P. Svoboda, M. Hradi, L. Mark, P. Zemck, CNN for license plate motion deblurring, IEEE Int. Conf. Image Processing (ICIP), 2016, pp. 3832–3836
[14] S. Nah, T.H. Kim, K.M. Lee, Deep multi-scale convolutional neural network for dynamic scene deblurring, CVPR, 2017
[15] Y. Yan, W. Ren, Y. Guo, R. Wang, X. Cao, Image deblurring via extreme channels prior, CVPR, 2017.
[16] S. Harmeling, H. Michael, B. Schölkopf, Space variant single-image blind deconvolution for removing camera shake, NIPS, 2010, pp. 829–837.
[17] O. Whyte, J. Sivic, A. Zisserman, J. Ponce, Non-uniform deblurring for shaken images, Int. J. Comput. Vis. 98 (2) (2012) 168–186.
[18] T.F. Chan, C.K. Wong, Total variation blind deconvolution, IEEE Trans. Image Process. 7 (3) (1998) 370–375.
[19] A. Levin, Y. Weiss, F. Durand, W.T. Freeman, Efficient marginal likelihood optimization in blind deconvolution, CVPR, 2011, pp. 2657–2664.
[20] A. Levin, Y. Weiss, F. Durand, W.T. Freeman, Understanding blind deconvolution algorithms, IEEE Trans. Pattern Anal. Mach. Intell. 33 (12) (2011) 2354–2367.
[21] S. Roth, M.J. Black, Fields of experts, Int. J. Comput. Vis. 82 (2) (2009) 205–229.
[22] J.F. Cai, H. Ji, C. Liu, Z. Shen, Framelet-based Blind Motion deblurring from a single Image, IEEE Trans. Image Process. 21 (2) (2012) 562–572.
[23] D. Perrone, P. Favaro, A clearer picture of total variation blind deconvolution, IEEE Trans. Pattern Anal. Mach. Intell. vol. 38, 6 (2016) 1041–1055.
[24] J.H. Money, S.H. Kang, Total variation minimizing blind deconvolution with shock filter reference, Image Vis. Comput. 26 (2) (2008) 302–314.
[25] S. Cho, S. Lee, Fast motion deblurring, ACM Trans. Graph. 28 (5) (2009) 1–8, https://doi.org/10.1145/1618452.1618491.
[26] L. Xu, J. Jia, Two-phase Kernel Estimation for Robust Motion Deblurring, ECCV, Part I, LNCS 6311, 2010, pp. 157–170
[27] L. Xu, S. Zheng, J. Jia, Unnatural L0 sparse representation for natural image deblurring, IEEE Conf. Computer Vision and Pattern Recognition, 2013, pp. 1107–1114
[28] J. Kotera, F. Sroubek, P. Milanfar, Blind Deconvolution using alternating maximum a posteriori estimation with heavy-tailed priors, in: R. Wilson, et al. (Eds.), CAIP, Part II, Lecture Notes in Computer Science, vol. 8048, 2013, pp. 59–66
[29] M. Almeida, L. Almeida, Blind and semi-blind deblurring of natural images, IEEE Trans. Image Process. 19 (1) (2010) 36–52.
[30] D. Perrone, P. Favaro, A logarithmic image prior for blind deconvolution, Int. J. Comput. Vis. 117 (2) (2016) 159–172.
[31] W. Zuo, D. Ren, S. Gu, L. Lin, L. Zhang, Learning iteration-wise generalized shrinkage-thresholding operators for blind deconvolution, IEEE Trans. Image Process. 25 (4) (2016) 1751–1764.
[32] S.D. Babacan, R. Molina, M.N. Do, A.K. Katsaggelos, Bayesian blind deconvolution with general sparse image priors, in: A. Fitzgibbon, et al. (Eds.): ECCV 2012, Part VI, Lecture Notes in Computer Science, vol. 7577, pp. 341–355.
[33] D. Wipf, H. Zhang, Revisiting bayesian blind deconvolution, J. Mach. Learn. Res. 15 (2014) 3595–3634.
[34] W. Shao, H. Deng, Q. Ge, H. Li, Z. Wei, Regularized motion blur-kernel estimation with adaptive sparse image prior learning, Pattern Recogn. 51 (2016) 402–424.
[35] J. Pan, Z. Su, Fast L0-regularized kernel estimation for robust motion deblurring, IEEE Signal Process. Lett. 20 (9) (2013) 1107–1114.
[36] P. Shearer, A.C. Gilbert, A.O. Hero III, Correcting camera shake by incremental sparse approximation, ICIP, 2013, pp. 572–576
[37] W. Shao, H. Li, M. Elad, Bi-L0-L2-norm regularization for blind motion deblurring, J. Visual Commun. Image Represent. 33 (2015) 42–59.
[38] J. Pan, Z. Hu, Z. Su, M.-H. Yang, Deblurring text images via L0-regularized intensity and gradient prior, CVPR, pp. 1628-1636, 2014.
[39] J. Pan, Z. Hu, Z. Su, M.-H. Yang, deblurring face images with exemplars, ECCV, 2014, pp 47–62
[40] T. Michaeli, M. Irani, Blind deblurring using internal patch recurrence, ECCV, 2014, pp. 783–798
[41] L. Sun, S. Cho, J. Wang, J. Hays, Edge-based blur kernel estimation using patch priors, Int. Conf. Computational Photography (ICCP), 2013
[42] W.S. Lai, J.-J. Ding, Y.-Y. Lin, Y.-Y. Chuang, Blur Kernel estimation using normalized color-line prior, CVPR, June 2015.
[43] W. Ren, X. Cao, J. Pan, X. Guo, W. Zuo, M.-H. Yang, Image deblurring via enhanced low-rank prior, IEEE Trans. Image Process. 25 (7) (2016) 3426–3437.
[44] L. Zhong, S. Cho, D. Metaxas, S. Paris, J. Wang, Handling noise in single image deblurring using directional filters, CVPR, 2013, pp. 612–619.
[45] J. Pan, Z. Lin, Z. Su, M.-H. Yang, Robust Kernel Estimation with Outliers Handling for Image Deblurring, CVPR, 2016, pp. 2800–2808.
[46] Z. Hu, S. Cho, J. Wang, M.-H. Yang, Deblurring lowlight images with light streaks, CVPR, 2014, pp. 3382–3389.
[47] J. Kotera, V. Šmídl, F. Šroubek, Blind deconvolution with model discrepancies, IEEE Trans. Image Process. 26 (5) (2017) 2533–2544.
[48] R. Wang, D. Tao, Recent progress in image deblurring, arXiv:1409.6838, 2014.
[49] D. Zoran, Y. Weiss, From learning models of natural image patches to whole image restoration, ICCV, 2011, pp. 479–486

[50] S.V. Venkatakrishnan, C.A. Bouman, B. Wohlberg, plug-and-play priors for model based reconstruction, IEEE Global Conference on Signal and Information Processing, 2013, pp. 945–948
[51] S. Sreehari, S. Venkatakrishnan, B. Wohlberg, L.F. Drummy, J.P. Simmons, C.A. Bouman, Plug-and-play priors for bright field electron tomography and sparse interpolation, arXiv preprint arXiv:1512.07331, 2015.
[52] A. Rond, R. Giryes, M. Elad, Poisson inverse problems by the plug-and-play scheme, J. Visual Commun. Image Represent. 41 (2016) 96–108.
[53] K. Egiazarian, V. Katkovnik, Single Image Superresolution via BM3D Sparse Coding, European Signal Processing Conference, 2015, pp. 2849–2853.
[54] Y. Romano, M. Elad, P. Milanfar, The little engine that cloud: regularization by denoising (RED), arXiv preprint arXiv: 1611.02862, 2016
[55] A. Brifman, Y. Romano, M. Elad, Turning a denoiser into a super-resolver using plug-and-play priors, ICIP, 2016, pp. 1404–1408.
[56] S.H. Chan, X. Wang, O.A. Elgendy, Plug-and-play ADMM for image restoration: fixed-point convergence and applications, IEEE Trans. Comput. Imag. 3 (1) (2017) 84–98.
[57] K. Dabov, A. Foi, V. Katkovnik, K. Egiazarian, Image denoising by sparse 3-D transform-domain collaborative filtering, IEEE Trans. Image Process. 16 (8) (2007) 2080–2095.
[58] D. Krishnan, R. Fergus, Fast image deconvolution using hyper-laplacian priors, NIPS 22 (2009) 1033–1041.
[59] O. Whyte, J. Sivic A. Zisserman, Deblurring Shaken and Partially Saturated Images, IEEE Color and Photometry in Computer Vision Workshop, in conjunction with ICCV, 2011.
[60] O. Whyte, J. Sivic, A. Zisserman, Deblurring shaken and partially saturated images, Int. J. Comput. Vis. 110 (2) (2014) 185–201.
[61] R. Köhler, M. Hirsch, B. Mohler, B. Schölkopf, S. Harmeling, Recording and playback of camera shake: benchmarking blind deconvolution with a real-world database, ICCV, 2012, pp. 27–40.
[62] M. Hirsch, C.J. Schuler, S. Harmeling, B. Schölkopf, Fast removal of non-uniform camera-shake, ICCV, 2011, pp. 463–470.
[63] S. Liu, H. Wang, J. Wang, S. Cho, C. Pan, Automatic Blur-Kernel-size estimation for motion deblurring, Visual Comput. 31 (2015) 733–746.
[64] O. Whyte, J. Sivic A. Zisserman, Deblurring shaken and partially saturated images, IEEE Color and Photometry in Computer Vision Workshop, in Conjunction with ICCV, 2011
[65] H. Zhang, D. Wipf, Y. Zhang, Multi-image blind deblurring using a coupled adaptive sparse prior, IEEE Int. Conf. Computer Vision and Pattern Recognition (CVPR), 2013, pp. 1051–1058
[66] Z. Wang, A.C. Bovik, H.R. Sheikh, E.P. Simoncelli, Image quality assessment: from error visibility to structural similarity, IEEE Trans. Image Process. 13 (4) (2004) 600–612.
[67] Y. Liu, J. Wang, S. Cho, A. Finkelstein, S. Rusinkiewicz, A no-reference metric for evaluating the quality of motion deblurring, ACM Trans. Graph. 32 (6) (2013) 175.
[68] O. Kupyn, V. Budzan, M. Mykhailych, D. Mishkin, J. Matas, DeblurGAN: blind motion deblurring using conditional adversarial networks, IEEE Int. Conf. Computer Vision and Pattern Recognition (CVPR), 2018.
[69] W.-Z. Shao, Y.-Y. Liu, L.-Y. Ye, L.-Q. Wang, Q. Ge, B.-K. Bao, H.-B. Li, DeblurGAN+: revisiting blind motion deblurring using conditional adversarial networks, Signal Process. (2019), https://doi.org/10.1016/j.sigpro.2019.107338.
[70] T.M. Nimisha, A.K. Singh, A.N. Rajagopalan, Blur-invariant deep learning for blind deblurring, IEEE Int. Conf. Comput. Vis. (ICCV) (2018) 4762–4770.
[71] L. Li, J. Pan, W.-S. Lai, C. Gao, N. Sang, M.-H. Yang, Learning a discriminative prior for blind image deblurring, IEEE Int. Conf. Computer Vision and Pattern Recognition (CVPR), 2018
[72] W.-Z. Shao, Y.-Z. Lin, B.-K. Bao, L.-Q. Wang, Q. Ge, H.-B. Li, Blind deblurring using discriminative image smoothing, The Chinese Conference on Pattern Recognition and Computer Vision, 2018
[73] X. Tao, H. Gao, Y. Wang, Scale-recurrent network for deep image deblurring, IEEE Int. Conf. Computer Vision and Pattern Recognition (CVPR), 2018.

Wen-Ze Shao received the B.S. degree in Science of Information and Computation in June 2003, and the Ph. D. degree in Pattern Recognition and Intelligent System in July 2008, both from Nanjing University of Science and Technology, Nanjing, China. From June 2003 to December 2011, he served as an officer in the People's Liberation Army of China. In January 2012, he joined Nanjing University of Posts and Telecommunications as an Assistant Professor. From May 2014 to June 2015, he also worked as a Postdoc Researcher at Department of Computer Science in Technion-Israel Institute of Technology. He is currently an Associate Professor in Nanjing University of Posts and Telecommunications, working in the fields of computational imaging and computer vision harnessing both variational and learning-based methods.

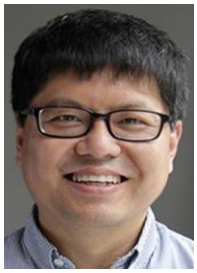
Yun-Zhi Lin received the B.E. degree in Automation in June 2018 from Southeast University, Nanjing, China. From September to December 2017, he was sponsored by National Undegraduate Exchange Scholarship of China working as a reseach assistant at Applied Nonlinear Control Laboratory in University of Alberta, Canada. Since August 2018, he has started his new life as a graduate student in School of Electrical and Computer Engineering at Georgia Institute of Technology, GA, USA. His research interests include computer graphics, computer vision and robotics. Now he mainly focuses on the field of robotics simulation and some other related topics.

Yuan-Yuan Liu received the B.E. degree in Electrical Information Science and Technology in June 2017 from Hebei University, Hebei, China. Since September 2017, she started a Master's program at Nanjing University of Posts and Telecommunications, Nanjing, China. Her current research interests include blind image deblurring, object recognition, deep learning, and generative adversarial modeling.

Li-Qian Wang received the B.S. degree in Science of Information and Computation in June 2006, the M.S. degree in Computer Application Technology in 2008, and the Ph.D. degree in Pattern Recognition and Intelligent System in 2015, all from Nanjing University of Science and Technology, Nanjing, China. She is currently an Assistant Professor at College of Telecommunications and Information Engineering in Nanjing University of Posts and Telecommunications. Her research interests include variational partial differential equations with applications in image processing, image restoration, image enhancement, and pattern recognition.

Qi Ge received the B.S. degree in Science of Information and Computation in 2006 and the M.S. degree in Applied Mathematics in 2009, both from Nanjing University of Information and Engineering, Nanjing, China, and the Ph.D. degree in Pattern Recognition and Intelligent System in 2013 from Nanjing University of Science and Technology, Nanjing, China. Now, she serves as an Assistant Professor at College of Telecommunications and Information Engineering in Nanjing University of Posts and Telecommunications. She is particularly interested in the fields of variational PDE approaches, probabilistic graphical models, sparse representation and their applications to image restoration, segmentation, and so on.

Bing-Kun Bao received the Ph.D. degree in control theory and control application from the University of Science and Technology of China, Hefei, China, in 2009. She is currently a full Professor with College of Telecommunications and Information Engineering, Nanjing University of Posts and Telecommunications, China. Her current research interests include cross-media cross-modal image search, social event detection, image classification and annotation, and sparse/low rank representation. She was the recipient of the 2016 ACM Transactions on Multimedia Computing, Communications and Applications (ACM TOMM) Nicolas D. Georganas Best Paper Award, IEEE Multimedia 2017 Best Paper Award, and the Best Paper Award from ICIMCS'09.



Hai-Bo Li was born in Xuzhou, Jiangsu Province, China. He received the B.E. degree in Wireless Engineering in 1985 and the M.S. degree in Communication and Electronic Systems in 1988, both from NUPT, Nanjing, China, and the Ph.D. degree in Information Theory in 1993 from Lin ping University, Sweden. In 1997, he became docent (UK equivalent senior lecturer, US equivalent associate professor) in image coding, and in 1999 he became the youngest lifetime professor of signal processing in Umea University, Sweden. Now, he is a Professor of innovative media technology in KTH Royal Institute of Technology. His research interest is mainly media signal processing, including image coding, video compression, motion estimation, facial and hand gesture recognition for the next generation of mobile phones, invisible interaction technology, and so on.