

Adapting total generalized variation for blind image restoration

Wen-Ze Shao^{1,2,3} · Feng Wang^{1,4} · Li-Li Huang⁵

Received: 12 September 2016 / Revised: 7 February 2018 / Accepted: 13 May 2018 /
Published online: 18 May 2018
© Springer Science+Business Media, LLC, part of Springer Nature 2018

Abstract In this paper, a fast blind deconvolution approach is proposed for image deblurring by modifying a recent well-known natural image model, i.e., the total generalized variation (TGV). As a generalization of total variation, TGV aims at reconstructing a higher-quality image with high-order smoothness as well as sharp edge structures. However, when it turns to the blind issue, as demonstrated either empirically or theoretically by several previous blind deblurring works, natural image models including TGV actually prefer the blurred images rather than their counterpart sharp ones. Inspired by the discovery, a simple, yet effective modification strategy is applied to the second-order TGV, resulting in a novel L_0 – L_1 -norm-based image regularization adaptable to the blind deblurring problem. Then, a fast numerical scheme is deduced with $O(N \log N)$ complexity for alternately estimating the intermediate sharp images and blur kernels via coupling operator splitting, augmented Lagrangian and also fast Fourier transform. Experiment results on a benchmark dataset and real-world blurred images demonstrate the superiority or comparable performance of the proposed approach to state-of-the-art ones, in terms of both deblurring quality and speed. Another contribution in this paper is the application of the newly proposed image prior to single image nonparametric blind super-resolution, which is a fairly more challenging inverse imaging task than blind deblurring. In spite of that, we have shown that both blind deblurring and blind super-resolution (SR) can be formulated into a common regularization framework. Experimental

✉ Wen-Ze Shao
shaowenze@njupt.edu.cn

¹ College of Telecommunications and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing, China

² National Engineering Research Center of Communications and Networking, Nanjing University of Posts and Telecommunications, Nanjing, China

³ Jiangsu Key Laboratory of Image and Video Understanding for Social Safety, Nanjing University of Science and Technology, Nanjing, China

⁴ Institutes of Advanced Technology, Zhejiang Dahua Technology Co., Ltd., Hangzhou, China

⁵ School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing, China

results demonstrate well the feasibility and effectiveness of the proposed blind SR approach, and also its advantage over the recent method by Michaeli and Irani in terms of estimation accuracy.

Keywords Blind deblurring · Total generalized variation · Camera shake removal · Blind super-resolution · L_0 -norm minimization

1 Introduction

Image deblurring, as a fundamental problem in image and vision computing, has attracted large interest in the past decades (Wang and Tao 2014; Lai et al. 2016). Due to its ill-posed nature, one of the crucial techniques is to develop top-performing natural image models, e.g., fields of experts (FoE) (Roth and Black 2009), block matching and 3D filtering (BM3D) (Dabov et al. 2007), sparse representation (Aharon et al. 2006; Rubinstein et al. 2013), and weighted nuclear norm minimization (Gu et al. 2014). However, as it turns to the blind issue where the blur kernels are unknown, ‘What is a good image prior for reliable blur kernel estimation’ is still an ongoing open problem. Recently, a few blind deblurring works (Benichoux et al. 2013; Fergus et al. 2006; Levin et al. 2011b; Wipf and Zhang 2013) have actually demonstrated that natural image models prefer the blurred images rather than their counterpart sharp ones in the maximum a posterior (MAP) framework. In other words, natural image priors, e.g., Roth and Black (2009), Dabov et al. (2007), Aharon et al. (2006), Rubinstein et al. (2013) and Gu et al. (2014), are no longer applicable to the MAP-based blind deblurring formulation. Naturally, a question arises that if there exists a possible bridge between blind deblurring and its non-blind scenario, inspired by which, a new simple, yet effective modification strategy is applied to Bredies et al.’s recent natural image model (Bredies et al. 2010), i.e., the total generalized variation (TGV), leading to a new L_0 – L_1 -norm-based image regularization specifically adaptable to blind deblurring. To the best of our knowledge, ours is the first paper touching the blind image deblurring task by use of TGV in a critical perspective. With the new regularization, a numerical algorithm with computational complexity $O(N \log N)$ is deduced for estimating alternately the intermediate sharp images and blur kernels by seamlessly coupling operator splitting, augmented Lagrangian and fast Fourier transform (FFT). Experiment results on both Levin et al.’s benchmark dataset (Bredies et al. 2010; Levin et al. 2011a) and real-world blurred images demonstrate the superiority or comparable performance of the present approach to state-of-the-art ones, in terms of both restoration quality and running speed. As a by-product, another contribution is made on nonparametric blind super-resolution via naive application of the newly proposed image regularization, which is a much more challenging computational imaging task than blind deblurring. For readers’ clarity, a detailed review on blind deblurring is provided in Sect. 1.1. We will observe that, just the same as in the present paper, unnatural image priors are advocated in the recent blind algorithms so as to pursue more accurate salient edges as core clues to blur kernel estimation.

1.1 State-of-the-art blind image deblurring

In general, blind image deblurring can be formulated using two inference criteria, namely maximum a posterior and variational Bayes. Due to ease of intuitive understanding, mathematical formulation, and also numerical derivation, the maximum a posterior criterion has attracted fairly more attentions than variational Bayes in the existing literature. A detailed review on variational Bayes-based blind deblurring can be referred to Ruiz et al. (2015). It

is noted that most of variational Bayes-based methods, e.g., Fergus et al. (2006), Levin et al. (2011a, b) and Amizic et al. (2012), exploit natural image models for posterior mean estimates of intermediate sharp images served alternately as clues to kernel estimation. Interestingly, more recent variational Bayes studies (Babacan et al. 2012; Wipf and Zhang 2014; Shao et al. 2016) find that unnatural image models may lead to better blind deblurring performance. For example, Babacan et al. (2012) empirically validate that the non-informative Jeffreys prior achieves more accurate kernel estimation than several other natural options. Furthermore, the theoretical analysis in Wipf and Zhang (2014) introduces several rigorous criteria to the choice of optimal image priors. Interestingly, the Jeffreys model has been proved to be optimal to a certain degree in terms of the deblurring quality. This theoretical result is somewhat consistent to the empirical finding in Babacan et al. (2012). It is worth noting that one of the authors in the present paper treats the choice of image priors for blind deblurring as a self-learning problem (Shao et al. 2016) wherein the clear image is sparsely modeled in a three-layer hierarchical Bayesian form. The experimental results demonstrate that the self-learned image model resembles somewhat the non-informative Jeffreys prior and as a matter of fact its negative-logarithm is relatively more nonconvex indicating that the self-learned image prior is also unnatural.

As being mentioned above, the maximum a posterior principle is exploited more commonly in practice (Cho and Lee 2017), and all the high-performing maximum a posterior approaches, e.g., Almeida and Almeida (2010), Krishnan et al. (2011), Xu et al. (2013), Pan and Su (2013), Shearer et al. (2013) and Shao et al. (2015), advocate non-convex unnatural image models for blind image deblurring. For example, Kotera et al. (2013) formulate blind deblurring utilizing the $L_{0,3}$ -norm-based image prior; while, an L_p -norm-based prior has been harnessed in Almeida and Almeida (2010), where p jumps down from 0.8 until 0.4; a new regularization image prior based on the normalized sparsity measure is described in Krishnan et al. (2011); afterwards, three L_0 -norm-based image priors (Xu et al. 2013; Pan and Su 2013; Shearer et al. 2013) are proposed in the same year for blind deblurring; furthermore, Shao et al. (2015) point out that naive use of L_0 -norm to model images does not necessarily lead to top blind deblurring performance, and propose a bi- L_0 - L_2 -norm-based regularization imposed on both the sharp image and blur kernel; in order to process blurred text images, Pan et al. (2014) propose to use the L_0 -norm of both image gradients and intensities as regularization; more recently, Pan et al. (2016) combine the L_0 -norm of image gradients and the dark channel prior (He et al. 2011), achieving state-of-the-art blind deblurring performance on several types of blurred images. Besides, another technical branch based on shock filters (Osher and Rudin 1990) is also very popular in the blind deblurring literature, whose core idea is to estimate blur kernels from a few predicted step-like salient edges in the original image (Cho and Lee 2009; Xu and Jia 2010). Actually, this idea had previously been exploited in Money and Kang (2008) and afterwards in Ohkoshi et al. (2014) which combines the total variation prior (He et al. 2014; Yan and Lu 2015; Rudin et al. 1992; Chan and Wong 1998) and edge shock filtering for blind image deblurring. Note that since Chan and Wong (1998) proposed in 1998 the total variation blind deconvolution, the work by Money and Kang (2008) was the first influential improvement towards Chan and Wong (1998), implying that harnessing only the total variation is not enough in the blind deblurring problem. Recently, Perrone and Favaro (2016a) further make a daring contribution on the theoretical analysis of Chan and Wong (1998), pointing out that the real success of total variation blind deconvolution actually relies heavily on the details and tricks among algorithmic implementation. With the discovery, more recently they propose a logarithmic total variation-based blind image deblurring approach using a delayed kernel normalization scheme (Perrone and Favaro 2016b). Naturally, direct deployment of the natural image model TGV (Bredies et al.

2010), as an advanced generalization of total variation (He et al. 2014; Yan and Lu 2015; Rudin et al. 1992; Chan and Wong 1998), may offer a great possibility in the blind deblurring performance via harnessing the similar tricks to those in Perrone and Favaro (2016a). In spite of that, we are inspired by Xu et al. (2013), Pan and Su (2013), Shearer et al. (2013), Shao et al. (2015) and Pan et al. (2014, 2016) and propose adapting the TGV for enhanced capability of pursuing higher-quality intermediate images with high-order smoothness and sharp edge structures, which naturally leads to more accurate blur kernel estimation.

2 Motivation and formulation

Prior to formulating the proposed approach, a spatially-invariant convolution model is assumed for the observation process

$$\mathbf{g} = \mathbf{h} * \mathbf{u} + \mathbf{z} \quad (1)$$

where \mathbf{g} is the blurred image, \mathbf{u} is the original clear image, \mathbf{h} is a blur kernel, $*$ denotes a periodic convolution operator, and \mathbf{z} is supposed to be an additive white Gaussian noise. The core task of blind deblurring is to estimate the blur kernel \mathbf{h} , where one of the key techniques is an appropriate model on the unknown image \mathbf{u} . For the convenience of presentation, the observation model (1) in a full-matrix form is reformulated into a matrix–vector version

$$\mathbf{g} = \mathbf{H}\mathbf{u} + \mathbf{z} \quad (2)$$

where \mathbf{g} , \mathbf{u} , \mathbf{z} are the vectorized versions of \mathbf{g} , \mathbf{u} , \mathbf{z} , and \mathbf{H} is the convolutional matrix corresponding to \mathbf{h} .

Since Bredies et al.'s proposal of TGV, a great many applications have benefited from the model such as denoising, super-resolution, non-blind deblurring, diffusion tensor imaging, etc. However, the blind deblurring problem is of another attribute; it is easy to prove just in a similar manner to Benichoux et al. (2013) that the L_1 -norm-based TGV is apt to result in trivial solutions, e.g., a pair of delta kernel and blurry image. Motivated by the recent L_0 -norm-based methods (Xu et al. 2013; Pan and Su 2013; Shearer et al. 2013; Shao et al. 2015; Pan et al. 2014, 2016) which actually originate from total variation blind deconvolution (Money and Kang 2008; Chan and Wong 1998) and also the normalized sparsity-based blind deblurring approach (Krishnan et al. 2011), this paper makes a natural modification towards the second-order TGV, though simple, yet fairly effective for ensuring successful estimation. Specifically, our advocated image prior for blind deblurring is given as

$$\mathcal{R}(\mathbf{u}) = \min_{\mathbf{v}} \left\{ \Phi(\mathbf{u}, \mathbf{v}) + \frac{\alpha_1}{\alpha_0} \Psi(\mathbf{v}) \right\}, \quad (3)$$

where α_0 , α_1 are positive tuning parameters, $\mathbf{v} = (\mathbf{v}_x; \mathbf{v}_y)$ is a cascade of two vectors \mathbf{v}_x and \mathbf{v}_y in column, and $\Phi(\mathbf{u}, \mathbf{v})$ is defined as

$$\Phi(\mathbf{u}, \mathbf{v}) = \|\nabla \mathbf{u} - \mathbf{v}\|_0, \quad (4)$$

where $\nabla \mathbf{u} = (\nabla_x \mathbf{u}; \nabla_y \mathbf{u})$, ∇_x , ∇_y represent the first-order difference operators in the horizontal and vertical directions, respectively, and $\Psi(\mathbf{v})$ is defined as

$$\Psi(\mathbf{v}) = \|(\nabla_x \mathbf{v}_x; \nabla_y \mathbf{v}_y; \nabla_x \mathbf{v}_y + \nabla_y \mathbf{v}_x)\|_1. \quad (5)$$

Apparently, as the L_0 -norm operator in $\Phi(\mathbf{u}, \mathbf{v})$ is replaced by the L_1 -norm, the above L_0 – L_1 -norm-based image prior (3) becomes the original second-order TGV for denoising in Bredies et al. (2010) and for non-blind deconvolution and super-resolution in Zhang et al. (2017)

and Laghrib et al. (2017). According to the analysis made on the original TGV in Bredies et al. (2010), TGV will locally reduce to the total variation (TV) model, i.e., $TGV(\mathbf{u}) \approx \|\nabla \mathbf{u}\|_1$, in regions of strong edge, while to the Hessian-norm-based regularization (HNR), i.e., $TGV(\mathbf{u}) \approx \|(\nabla_{xx} \mathbf{u}; \nabla_{yy} \mathbf{u}; \nabla_{xy} \mathbf{u} + \nabla_{yx} \mathbf{u})\|_1$, in regions of smooth surface and faint texture. Hence, TGV has achieved a decently adaptive combination of TV and HNR which is fairly different from previous approaches, and \mathbf{v} behaves, to a certain degree, as an indicator of the region nature. Returning to the modified TGV in this paper, it is not hard to deduce that $\mathcal{R}(\mathbf{u})$ is expected to behave approximately as the recently advocated blind deblurring component $\|\nabla \mathbf{u}\|_0$ (Xu et al. 2013; Pan and Su 2013; Shearer et al. 2013; Shao et al. 2015; Pan et al. 2014, 2016) in regions with salient edges, while as $\|(\nabla_{xx} \mathbf{u}; \nabla_{yy} \mathbf{u}; \nabla_{xy} \mathbf{u} + \nabla_{yx} \mathbf{u})\|_1$ in regions of flat surface and faint texture, so as to overcome the negative effects of solely utilizing $\|\nabla \mathbf{u}\|_0$, especially the staircase artifacts which are to severely reduce the kernel estimation precision. Hence, harnessing (3) a higher-quality intermediate sharp image with high-order smoothness and sharp edge structures will be produced and thus more accurate blur kernel estimation is to be achieved. Numerous results on Levin et al.’s benchmark dataset and real-world blurred images are to validate the effectiveness of (3). To the best of our knowledge, (3) is the first approach of employing TGV for blind deblurring in the literature, different from existing image models in either the maximum posterior or variational Bayes framework.

To further improve the kernel estimation accuracy, we also exploit an additional composite prior on the blur kernels, exactly the same as the one in Shao et al. (2015). Moreover, the continuation scheme is applied so as to alleviate the tiresome choice of appropriate tuning parameters. Then, given current estimates of the intermediate sharp image and the blur kernel, i.e., $\mathbf{u}_i, \mathbf{h}_i, 0 \leq i \leq I - 1$, the next estimates $\mathbf{u}_{i+1}, \mathbf{h}_{i+1}$ can be solved by minimizing the following cost function

$$\lambda \|\mathbf{g} - \mathbf{H}\mathbf{u}\|_2^2 + c_u^i \alpha_0 \mathcal{R}(\mathbf{u}) + c_h^i \beta_0 \mathcal{Q}(\mathbf{h}), \tag{6}$$

where c_u, c_h are the continuation factors which are fixed as 2/3 and 4/5, respectively, for all the experiments in this paper, and

$$\mathcal{Q}(\mathbf{h}) = \|\mathbf{h}\|_0 + \frac{\beta_1}{\beta_0} \|\mathbf{h}\|_2^2, \tag{7}$$

with β_0, β_1 positive tuning parameters. Notice that, the derivative version of (7) can be served as another possibility. Besides, I is set as 10 in this paper.

3 Numerical solutions

This section turns to the numerical minimization of (6). Using the image and kernel priors, i.e., (3) and (7), blind deblurring can be then addressed into the following alternating minimization problems

$$\mathbf{u}_{i+1}, \mathbf{v}_{i+1} = \arg \min_{\mathbf{u}, \mathbf{v}} \lambda \|\mathbf{H}_i \mathbf{u} - \mathbf{g}\|_2^2 + c_u^i \alpha_0 (\Phi(\mathbf{u}, \mathbf{v}) + \frac{\alpha_1}{\alpha_0} \Psi(\mathbf{v})), \tag{8}$$

and

$$\mathbf{h}_{i+1} = \arg \min_{\mathbf{h}} \lambda \sum_{d \in \Lambda} \|(\mathbf{U}_{i+1})_d \mathbf{h} - \mathbf{g}_d\|_2^2 + c_h^i \beta_0 \mathcal{Q}(\mathbf{h}), \tag{9}$$

where $d \in \Lambda = \{x, y\}$, $\mathbf{g}_d = \nabla_d \mathbf{g}$, $(\mathbf{u}_{i+1})_d = \nabla_d \mathbf{u}_{i+1}$, and $(\mathbf{U}_{i+1})_d$ represents the convolutional matrix corresponding to the image gradient $(\mathbf{u}_{i+1})_d$. Note that estimating the

kernel \mathbf{h}_{i+1} in the image derivative domain is purely an empirical attempt. In the following, both (8) and (9) are to be solved via coupling operator splitting, augmented Lagrangian as well as FFT, finally leading to a numerical scheme with $O(N \log N)$ complexity. Another point to be noted is that the blur kernel should be non-negative and normalized, i.e., the estimated kernel should be projected onto a constraint set $\Omega = \{\mathbf{h} \geq 0, \|\mathbf{h}\|_1 = 1\}$.

3.1 Updating the intermediate sharp image

First of all, apply the operator splitting scheme to (8), obtaining an equivalent minimization problem as follows

$$\begin{aligned} & \mathbf{u}_{i+1}, \mathbf{v}_{i+1}, \mathbf{A}_{i+1}, \mathbf{B}_{i+1}, \mathbf{C}_{i+1}, \mathbf{D}_{i+1}, \mathbf{E}_{i+1} \\ &= \arg \min_{\mathbf{u}, \mathbf{v}, \mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}, \mathbf{E}} \frac{\lambda}{c_u} \|\mathbf{H}_i \mathbf{u} - \mathbf{g}\|_2^2 + \alpha_0 (\|\mathbf{A}\|_0 + \|\mathbf{B}\|_0) \\ &+ \alpha_1 (\|\mathbf{C}\|_1 + \|\mathbf{D}\|_1 + \|\mathbf{E}\|_1), \\ & \text{s.t. } \mathbf{A} = \nabla_x \mathbf{u} - \mathbf{v}_x, \\ & \quad \mathbf{B} = \nabla_y \mathbf{u} - \mathbf{v}_y, \\ & \quad \mathbf{C} = \nabla_x \mathbf{v}_x, \\ & \quad \mathbf{D} = \nabla_y \mathbf{v}_y, \\ & \quad \mathbf{E} = \nabla_x \mathbf{v}_y + \nabla_y \mathbf{v}_x. \end{aligned} \tag{10}$$

Harnessing the augmented Lagrangian scheme, (10) can be solved by transforming it into an unconstrained form, obtaining the iterative estimates of $\mathbf{u}_{i+1}, \mathbf{v}_{i+1}, \mathbf{A}_{i+1}, \mathbf{B}_{i+1}, \mathbf{C}_{i+1}, \mathbf{D}_{i+1}, \mathbf{E}_{i+1}$, i.e.,

$$\begin{aligned} & \mathbf{u}_i^{l+1}, \mathbf{v}_i^{l+1}, \mathbf{A}_i^{l+1}, \mathbf{B}_i^{l+1}, \mathbf{C}_i^{l+1}, \mathbf{D}_i^{l+1}, \mathbf{E}_i^{l+1} \\ &= \arg \min_{\mathbf{u}, \mathbf{v}, \mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}, \mathbf{E}} \frac{\lambda}{c_u} \|\mathbf{H}_i \mathbf{u} - \mathbf{g}\|_2^2 + \alpha_0 (\|\mathbf{A}\|_0 + \|\mathbf{B}\|_0) \\ &+ \alpha_1 (\|\mathbf{C}\|_1 + \|\mathbf{D}\|_1 + \|\mathbf{E}\|_1) \\ &+ \mu_A^{l*} (\nabla_x \mathbf{u} - \mathbf{v}_x - \mathbf{A}) + \frac{\gamma_A}{2} \|\nabla_x \mathbf{u} - \mathbf{v}_x - \mathbf{A}\|_2^2 \\ &+ \mu_B^{l*} (\nabla_y \mathbf{u} - \mathbf{v}_y - \mathbf{B}) + \frac{\gamma_B}{2} \|\nabla_y \mathbf{u} - \mathbf{v}_y - \mathbf{B}\|_2^2 \\ &+ \mu_C^{l*} (\nabla_x \mathbf{v}_x - \mathbf{C}) + \frac{\gamma_C}{2} \|\nabla_x \mathbf{v}_x - \mathbf{C}\|_2^2 \\ &+ \mu_D^{l*} (\nabla_y \mathbf{v}_y - \mathbf{D}) + \frac{\gamma_D}{2} \|\nabla_y \mathbf{v}_y - \mathbf{D}\|_2^2 \\ &+ \mu_E^{l*} (\nabla_x \mathbf{v}_y + \nabla_y \mathbf{v}_x - \mathbf{E}) + \frac{\gamma_E}{2} \|\nabla_x \mathbf{v}_y + \nabla_y \mathbf{v}_x - \mathbf{E}\|_2^2, \end{aligned} \tag{11}$$

where $0 \leq l \leq L - 1$, and $\gamma_A, \gamma_B, \gamma_C, \gamma_D, \gamma_E$ are the augmented Lagrangian penalty parameters. In this paper, these parameters are fixed as the same value, i.e., $\gamma = 100$, for all the experiments. As for the Lagrange multipliers, they are updated via the rules

$$\begin{aligned} \mu_A^{l+1} &= \mu_A^l + \gamma (\nabla_x \mathbf{u}_i^{l+1} - (\mathbf{v}_x)_i^{l+1} - \mathbf{A}_i^{l+1}), \\ \mu_B^{l+1} &= \mu_B^l + \gamma (\nabla_y \mathbf{u}_i^{l+1} - (\mathbf{v}_y)_i^{l+1} - \mathbf{B}_i^{l+1}), \\ \mu_C^{l+1} &= \mu_C^l + \gamma (\nabla_x (\mathbf{v}_x)_i^{l+1} - \mathbf{C}_i^{l+1}), \\ \mu_D^{l+1} &= \mu_D^l + \gamma (\nabla_y (\mathbf{v}_y)_i^{l+1} - \mathbf{D}_i^{l+1}), \\ \mu_E^{l+1} &= \mu_E^l + \gamma (\nabla_x (\mathbf{v}_y)_i^{l+1} + \nabla_y (\mathbf{v}_x)_i^{l+1} - \mathbf{E}_i^{l+1}). \end{aligned} \tag{12}$$

After a series of analytical calculations towards (11), the solutions to $\mathbf{u}_i^{l+1}, \mathbf{v}_i^{l+1}, \mathbf{A}_i^{l+1}, \mathbf{B}_i^{l+1}, \mathbf{C}_i^{l+1}, \mathbf{D}_i^{l+1}, \mathbf{E}_i^{l+1}$ can be derived as follows

$$\begin{aligned}
 \mathbf{u}_i^{l+1} &= \mathcal{F}^{-1} \left\{ \begin{aligned} &\left(\mathcal{F} \left(\frac{\lambda}{c_u^l} \mathbf{H}_i^* \mathbf{H}_i + \gamma \nabla_x^* \nabla_x + \gamma \nabla_y^* \nabla_y \right) \right)^{-1} \\ &\cdot \mathcal{F} \left(\begin{aligned} &\frac{\lambda}{c_u^l} \mathbf{H}_i^* \mathbf{g} + \gamma \nabla_x^* ((\mathbf{v}_x)_i^l + \mathbf{A} - \frac{1}{\gamma} \mu_{\mathbf{A}}^l) \\ &+ \gamma \nabla_y^* ((\mathbf{v}_y)_i^l + \mathbf{B} - \frac{1}{\gamma} \mu_{\mathbf{B}}^l) \end{aligned} \right) \end{aligned} \right\}, \\
 (\mathbf{v}_x)_i^{l+1} &= \mathcal{F}^{-1} \left\{ \begin{aligned} &(\mathcal{F}(\mathcal{U}) \cdot * \mathcal{F}(\mathcal{U}) - \mathcal{F}(\Lambda^*) \cdot * \mathcal{F}(\Lambda))^{-1} \\ &\cdot (\mathcal{F}(\mathcal{U}) \cdot * \mathcal{F}(\Delta_1) - \mathcal{F}(\Lambda) \cdot * \mathcal{F}(\Delta_2)) \end{aligned} \right\}, \\
 (\mathbf{v}_y)_i^{l+1} &= \mathcal{F}^{-1} \left\{ \begin{aligned} &(\mathcal{F}(\mathcal{U}) \cdot * \mathcal{F}(\mathcal{U}) - \mathcal{F}(\Lambda^*) \cdot * \mathcal{F}(\Lambda))^{-1} \\ &\cdot (\mathcal{F}(\mathcal{U}) \cdot * \mathcal{F}(\Delta_2) - \mathcal{F}(\Lambda^*) \cdot * \mathcal{F}(\Delta_1)) \end{aligned} \right\}, \\
 \mathbf{A}_i^{l+1} &= \Theta_{\text{HARD}} \left(\nabla_x \mathbf{u}_i^l + \frac{1}{\gamma} \mu_{\mathbf{A}}^l - (\mathbf{v}_x)_i^l, \left(\frac{2\alpha_0}{\gamma} \right)^{0.5} \right), \\
 \mathbf{B}_i^{l+1} &= \Theta_{\text{HARD}} \left(\nabla_y \mathbf{u}_i^l + \frac{1}{\gamma} \mu_{\mathbf{B}}^l - (\mathbf{v}_y)_i^l, \left(\frac{2\alpha_0}{\gamma} \right)^{0.5} \right), \\
 \mathbf{C}_i^{l+1} &= \Theta_{\text{SOFT}} \left(\nabla_x (\mathbf{v}_x)_i^l + \frac{1}{\gamma} \mu_{\mathbf{C}}^l, \frac{\alpha_1}{\gamma} \right), \\
 \mathbf{D}_i^{l+1} &= \Theta_{\text{SOFT}} \left(\nabla_y (\mathbf{v}_y)_i^l + \frac{1}{\gamma} \mu_{\mathbf{D}}^l, \frac{\alpha_1}{\gamma} \right), \\
 \mathbf{E}_i^{l+1} &= \Theta_{\text{SOFT}} \left(\nabla_x (\mathbf{v}_y)_i^l + \nabla_y (\mathbf{v}_x)_i^l + \frac{1}{\gamma} \mu_{\mathbf{E}}^l, \frac{\alpha_1}{\gamma} \right), \tag{13}
 \end{aligned}$$

where \mathcal{F} and \mathcal{F}^{-1} represent the operations of 2D FFT and its inverse version, $\Lambda = \nabla_y^* \nabla_x$, and $\mathcal{U} = \mathbf{I} + \nabla_x^* \nabla_x + \nabla_y^* \nabla_y$, and as for Δ_1 and Δ_2 , they are calculated according to

$$\begin{aligned}
 \Delta_1 &= \nabla_x^* \left(\mathbf{C}_i^l - \frac{1}{\gamma} \mu_{\mathbf{C}}^l \right) + \nabla_y^* \left(\mathbf{E}_i^l - \frac{1}{\gamma} \mu_{\mathbf{E}}^l \right) + \left(\nabla_x \mathbf{u}_i^l - \mathbf{A}_i^l + \frac{1}{\gamma} \mu_{\mathbf{A}}^l \right), \\
 \Delta_2 &= \nabla_y^* \left(\mathbf{D}_i^l - \frac{1}{\gamma} \mu_{\mathbf{D}}^l \right) + \nabla_x^* \left(\mathbf{E}_i^l - \frac{1}{\gamma} \mu_{\mathbf{E}}^l \right) + \left(\nabla_y \mathbf{u}_i^l - \mathbf{B}_i^l + \frac{1}{\gamma} \mu_{\mathbf{B}}^l \right). \tag{14}
 \end{aligned}$$

Furthermore, Θ_{HARD} and Θ_{SOFT} represent the hard- and soft-thresholding operators, respectively. With above iterative estimates, \mathbf{u}_{i+1} can be finally obtained as $\mathbf{u}_{i+1} = \mathbf{u}_i^l$.

3.2 Updating the blur kernel

Given the updated sharp image \mathbf{u}_{i+1} , a similar numerical scheme can be derived for updating the blur kernel \mathbf{h}_{i+1} . Firstly, apply operator splitting to (9) and obtain an equivalent constrained L_0 - L_2 -minimization problem

$$(\mathbf{w}_{i+1}, \mathbf{h}_{i+1}) = \arg \min_{\mathbf{w}, \mathbf{h}} \mathcal{G}(\mathbf{w}, \mathbf{h}, \mathbf{u}_{i+1}) \text{ s.t. } \mathbf{w} = \mathbf{h}, \tag{15}$$

where $\mathcal{G}(\mathbf{w}, \mathbf{h}, \mathbf{u}_{i+1})$ is defined as

$$\mathcal{G}(\mathbf{w}, \mathbf{h}, \mathbf{u}_{i+1}) = \lambda \sum_{d \in \Lambda} \|(\mathbf{U}_{i+1})_d \mathbf{h} - \mathbf{g}_d\|_2^2 + \beta_0 (\|\mathbf{w}\|_0 + \frac{\beta_1}{\beta_0} \|\mathbf{h}\|_2^2). \tag{16}$$

Secondly, based on the augmented Lagrangian scheme, \mathbf{w}_{i+1} and \mathbf{h}_{i+1} can be alternately estimated by minimization the following unconstrained problem

$$(\mathbf{w}_i^{j+1}, \mathbf{h}_i^{j+1}) = \arg \min_{\mathbf{w}, \mathbf{h}} \tilde{\mathcal{G}}^j(\mathbf{w}, \mathbf{h}, \mathbf{u}_{i+1}), \tag{17}$$

where $0 \leq j \leq J - 1$ and $\bar{\mathcal{G}}^j(\mathbf{w}, \mathbf{h}, \mathbf{u}_{i+1})$ is defined as

$$\bar{\mathcal{G}}^j(\mathbf{w}, \mathbf{h}, \mathbf{u}_{i+1}) = \mathcal{G}(\mathbf{w}, \mathbf{h}, \mathbf{u}_{i+1}) + \boldsymbol{\mu}_h^{j*}(\mathbf{h} - \mathbf{w}) + \frac{\chi}{2} \|\mathbf{h} - \mathbf{w}\|_2^2. \tag{18}$$

In (18), χ is the augmented Lagrangian penalty parameter, $\boldsymbol{\mu}_h^j$ is the Lagrange multiplier of the constraint $\mathbf{w} = \mathbf{h}$ and updated according to the rule

$$\boldsymbol{\mu}_h^{j+1} = \boldsymbol{\mu}_h^j + \chi (\mathbf{h}_i^{j+1} - \mathbf{w}_i^{j+1}). \tag{19}$$

Similar to Sect. 3.1, a fixed, large χ equal to 1×10^6 is selected across all the experiments and found work well. After some straightforward manipulations, $\mathbf{w}_i^{j+1}, \mathbf{h}_i^{j+1}$ can be easily computed from (18) and given as

$$\mathbf{w}_i^{j+1} = \arg \min_{\mathbf{w}} \bar{\mathcal{G}}^j(\mathbf{w}, \mathbf{h}, \mathbf{u}_{i+1}) = \Theta_{\text{HARD}} \left(\mathbf{h}_i^j + \frac{1}{\chi} \boldsymbol{\mu}_h^j, \left(\frac{2\beta_0}{\chi} \right)^{0.5} \right), \tag{20}$$

$$\begin{aligned} \mathbf{h}_i^{j+1} &= \arg \min_{\mathbf{h}} \bar{\mathcal{G}}^j(\mathbf{w}, \mathbf{h}, \mathbf{u}_{i+1}) \\ &= \left(\lambda \sum_{d \in \Lambda} (\mathbf{U}_{i+1})_d^* (\mathbf{U}_{i+1})_d + (\beta_1 + \frac{\chi}{2}) \mathbf{I} \right)^{-1} \left(\lambda \sum_{d \in \Lambda} (\mathbf{U}_{i+1})_d^* \mathbf{g}_d + \frac{\chi}{2} \left(\mathbf{w}_i^j - \frac{1}{\chi} \boldsymbol{\mu}_h^j \right) \right), \end{aligned} \tag{21}$$

where $\mathbf{w}_0^0, \mathbf{h}_0^0, \boldsymbol{\mu}_h^0$ are the provided initializations. Now, the minimizers of (15) can be obtained as $\mathbf{w}_{i+1} = \mathbf{w}_i^J, \mathbf{h}_{i+1} = \mathbf{h}_i^J$. Similar to $\mathbf{A}_i^{l+1}, \mathbf{B}_i^{l+1}, \mathbf{C}_i^{l+1}, \mathbf{D}_i^{l+1}, \mathbf{E}_i^{l+1}$ in (13), it is very simple to compute \mathbf{w}_i^{j+1} because of its pixel-by-pixel processing nature. Besides, the assumption of circular convolution makes it very efficient to compute \mathbf{h}_i^{j+1} using FFT, just similar to the update of \mathbf{u}_i^{l+1} in (13). The pseudo code of updating the blur kernel is provided here for brevity. With the iterative estimates $\{\mathbf{h}_i^{j+1}\}$, \mathbf{h}_{i+1} can be finally obtained as $\mathbf{h}_{i+1} = \mathbf{h}_i^J$, followed with projection onto the constraint set $\Omega = \{ \mathbf{h} \geq 0, \|\mathbf{h}\|_1 = 1 \}$.

Pseudo code of updating the blur kernel:

1: **Input** : sharp image \mathbf{u}_{i+1} , Lagrangian penalty parameter χ .

2: **While** $j < J$ do

– Update \mathbf{w}_i^{j+1} by computing $\mathbf{w}_i^{j+1} = \Theta_{\text{HARD}} \left(\mathbf{h}_i^j + \frac{1}{\chi} \boldsymbol{\mu}_h^j, \left(\frac{2\beta_0}{\chi} \right)^{0.5} \right)$.

– Update \mathbf{h}_i^{j+1} based on 2D FFT by computing

$$\mathbf{h}_i^{j+1} = \mathcal{F}^{-1} \left\{ \left(\mathcal{F} \left(\frac{\lambda}{\chi} \sum_{d \in \Lambda} (\mathbf{U}_{i+1})_d^* (\mathbf{U}_{i+1})_d + (\beta_1 + \frac{\chi}{2}) \mathbf{I} \right) \right)^{-1} \cdot \mathcal{F} \left(\frac{\lambda}{\chi} \sum_{d \in \Lambda} (\mathbf{U}_{i+1})_d^* \mathbf{g}_d + \frac{\chi}{2} \left(\mathbf{w}_i^j - \frac{1}{\chi} \boldsymbol{\mu}_h^j \right) \right) \right\}.$$

– Update $\boldsymbol{\mu}_h^{j+1}$ by computing $\boldsymbol{\mu}_h^{j+1} = \boldsymbol{\mu}_h^j + \chi (\mathbf{h}_i^{j+1} - \mathbf{w}_i^{j+1})$.

3: **End**

From above it is easily seen that the calculations for updating the blur kernel and also the image are either pixel-by-pixel operations or 2D FFTs, and therefore the overall complexity of the proposed method is $O(N \log N)$. In addition, both L and J are set as 10 in this paper.

Upon obtaining the estimated blur kernel \mathbf{h}_{l-1} , we exploit the hyper-Laplacian prior-based non-blind deconvolution method (Krishnan and Fergus 2009) for the final image restoration.

3.3 Multi-scale implementation

In a similar spirit to many previous blind deblurring methods (Bredies et al. 2010; Levin et al. 2011a; Wang and Tao 2014; Almeida and Almeida 2010; Krishnan et al. 2011; Xu et al. 2013; Pan and Su 2013; Shearer et al. 2013; Shao et al. 2015; Babacan et al. 2012; Wipf and Zhang 2014), a multi-scale version of the proposed approach is implemented so as to take into account of large-scale blur kernel estimation. Specifically, we consider 4 scales in total. In each scale s , the blurred image is set as the 2 times down-sampled version of the original blurred image \mathbf{g} ; and during iterative estimation, \mathbf{u}_0^0 is set as a zero image, while \mathbf{h}_0^0 is set as a normalized, up-sampled kernel estimated from the coarser level (and in the coarsest scale \mathbf{h}_0^0 is set as a Dirac pulse); and $\mathbf{v}_0^0, \mathbf{A}_0^0, \mathbf{B}_0^0, \mathbf{C}_0^0, \mathbf{D}_0^0, \mathbf{E}_0^0, \mu_A^0, \mu_B^0, \mu_C^0, \mu_D^0, \mu_E^0, \mathbf{w}_0^0, \mu_h^0$ are also simply set as zeros.

4 Blind deblurring results

We first test the proposed method along with comparisons against several previous algorithms using Levin et al.'s benchmark dataset (Levin et al. 2011a), consisting of 32 motion blurred images generated from 4 natural images (Image01 ~ Image04) and 8 different motion blur kernels (Kernel01 ~ Kernel08). To be noted that, the tuning parameters are fixed for the 32 blurred images, including $\lambda = 0.01, \alpha_0 = 0.25, \alpha_1 = 5, \beta_0 = 0.25, \beta_1 = 5, \gamma = 100, \chi = 1e6, I = J = M = 10$. The SSD (sum of squared difference) ratio between images non-blindly restored using the estimated blur kernel and the ground truth blur kernel is selected as the final evaluation measure, so as to take into account the fact that harder blur kernels lead to larger deblurring errors even if the true kernels are known, since the corresponding non-blind deconvolution problem is also harder. When the SSD ratio is below 3, the blind deblurring is considered to be successful (Levin et al. 2011a; Wang and Tao 2014; Almeida and Almeida 2010; Krishnan et al. 2011; Xu et al. 2013; Pan and Su 2013; Shearer et al. 2013; Shao et al. 2015; Babacan et al. 2012; Wipf and Zhang 2014).

Table 1 lists the SSD ratios of the proposed approach corresponding to the 32 blurred images. It is seen that our approach fails only as restoring Image04-Kernel06 with SSD ratio 3.45, and therefore the successful percentage of our method is 97%. We also compute its average SSD ratio for comparison later on, i.e., 1.48.

A few recent methods are also tested on the benchmark, including four approaches of VB type, i.e., Fergus et al. (2006), Levin et al. (2011a), Babacan et al. (2012), and Shao et al. (2016), as well as seven methods of MAP type including Xu et al. (2013), Dong et al. (2014),¹ Pan and Su (2013), Shearer et al. (2013), Shao et al. (2015), Kim et al. (2015), Kotera et al. (2013), and Perrone and Favaro (2016a). Except Pan and Su (2013) and (Perrone and Favaro 2016a), all the other methods also use (Krishnan and Fergus 2009) for non-blind deblurring. In spite of that, the non-blind method in Pan and Su (2013) and Perrone and Favaro (2016a)

¹ Since 20 March, 2013, the authors of (Xu et al. 2013) have successively released two executable software (implemented in C++) for blind motion deblurring, i.e., Robust Motion Deblurring System. The first version is v3.0.1 which implements the algorithm as detailed in Xu et al. (2013), and the second version is v3.1 which incorporates the algorithms in both Xu et al. (2013) and Dong et al. (2014), i.e. Xu et al. (2013), Dong et al. (2014), for more accurate blur-kernel estimation.

Table 1 SSD ratios of the proposed approach on the benchmark dataset (Levin et al. 2011a)

Image no.	Image01	Image02	Image03	Image04
Kernel01	1.20	1.59	1.17	2.00
Kernel02	0.96	1.42	1.01	1.65
Kernel03	1.35	1.71	1.18	2.39
Kernel04	1.19	1.29	1.42	1.37
Kernel05	1.05	1.58	1.12	2.47
Kernel06	1.48	1.81	1.79	3.45
Kernel07	1.07	1.68	1.15	1.92
Kernel08	0.63	1.45	0.84	0.83

Bold value indicates that the proposed method fails to restore the blurred image Image04-Kernel06 in terms of the SSD ratio 3.45

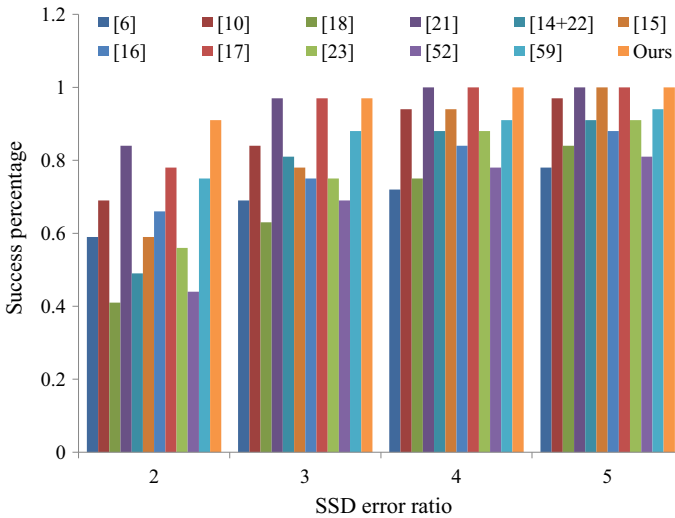


Fig. 1 Cumulative histograms of SSD ratios corresponding to different blind deblurring methods

exploits the same prior as in Krishnan and Fergus (2009), i.e., the anisotropic discrete total variation. And hence, the overall comparison is convincing.

Figure 1 plots the cumulative histogram of SSD ratios according to each method. The r 'th bin in the figure counts the percentage of blurred images in the dataset achieving error ratio below r (Levin et al. 2011b). For each bin, the higher the bar the better the blind deblurring quality. The success percentage, i.e., SSD ratio below 3, of each approach is: 69% (Fergus et al. 2006), 84% (Levin et al. 2011a), 63% (Babacan et al. 2012), 97% (Shao et al. 2016), 81% (Xu et al. 2013; Dong et al. 2014), 78% (Pan and Su 2013), 75% (Shearer et al. 2013), (Shao et al. 2015) 97%, 75% (Kim et al. 2015), 69% (Kotera et al. 2013), 88% (Perrone and Favaro 2016a), 97% (Ours). Apparently, our method achieves the same highest percentage of success with Shao et al. (2015, 2016). Not only that, our method in fact achieves overall better performance. On one hand, compared against (Shao et al. 2015) and (2016), more deblurred images with smaller SSD ratios, e.g., 2, are produced by our method; on the other hand, our approach enjoys smaller average SSD ratio than (Shao et al. 2015) (1.56), and

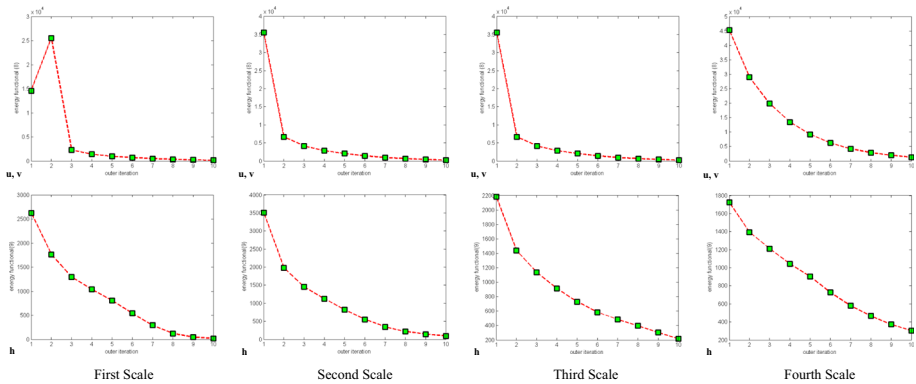


Fig. 2 Objective function curves of 10 outer iterations for each scale of the proposed approach as for Image04-kernel06. Top row: functional (8) for estimating \mathbf{u} and \mathbf{v} ; Bottom row: functional (9) for estimating \mathbf{h}

although (Shao et al. 2016) (1.44) enjoys a slightly smaller average ratio, it is fairly more time-consuming compared to our approach. Using the same lap-top computer configuration (Intel i7-4600 M CPU (2.90 GHz) and 8 GB memory, running Windows 7), our kernel estimation algorithm costs about 14 s for each blurred image, while (Shao et al. 2016) costs about 363 s. Therefore, we conclude that the proposed method achieves a good compromise between deblurring quality and speed.

Figure 2 presents plots of the functionals (8) for updating the variables \mathbf{u} and \mathbf{v} and (9) for updating the blur-kernel \mathbf{h} , so as to demonstrate the convergence tendency of the proposed approach. We take Image04-Kernel06 as an example. The graphs provide the objective function curves of 10 outer iterations for each scale of the running algorithm. From these curves we see that the proposed alternating minimization algorithm is quite effective in pursuing the (possibly local) minimizers of the functionals (8) and (9). Moreover, as suggested by one of the reviewers, we also compare Kotera et al. (2013) with the proposed method due to their similar alternately iterative computing structures, in terms of peak signal-to noise ratio (PSNR), structural similarity image metric (SSIM) (Wang et al. 2004; Sampat et al. 2009), as well as the running time. Table 2 provides comparative results on the eight blurred images corresponding to Image04 in the benchmark dataset (Levin et al. 2011a). It is observed that our proposed method not only achieves higher PSNR or SSIM on most of the eight images, but is also more computationally efficient than Kotera et al. (2013). As kindly suggested by another reviewer, the eight blur kernels estimated by the proposed method are also compared with those by Kotera et al. (2013) in a quantitative way. Specifically, the estimated blur kernels are provided in Fig. 3 accompanied with their errors measured by the sum of squared difference (SSD). According to the calculated figures and estimated trajectories, it is seen that the proposed method achieves more stable performance on average. But we should claim that quality of the final deblurred images tells the real truth about performance of different methods when they exploit the same non-blind deblurring scheme. With this view, the proposed method is proved better than Kotera et al. (2013) according to the PSNR and SSIM values in Table 2. In Fig. 4, we take Image04-Kernel08 for example, presenting the estimated kernels by the two algorithms as well as their corresponding final deblurred images. Apparently, (Kotera et al. 2013) has failed to produce a reasonable blur kernel in this example while ours has generated a final deblurred image (28.17 dB) even comparable to the non-blind version (27.36 dB).

Table 2 Comparisons on PSNR, SSIM, and running time between the proposed method and Kotera et al. (2013) on the eight blurred images corresponding to Image04 in the benchmark dataset (Levin et al. 2011a)

Image no.	PSNR (dB)		SSIM		Time (s)	
	Kotera et al. (2013)	Ours	Kotera et al. (2013)	Ours	Kotera et al. (2013)	Ours
Image04-Kernel01	29.78	32.31	0.8822	0.9133	29	14
Image04-Kernel02	26.55	29.83	0.7847	0.8831	30	14
Image04-Kernel03	32.80	32.52	0.9313	0.9265	29	14
Image04-Kernel04	22.04	26.47	0.6746	0.7837	29	14
Image04-Kernel05	32.30	32.97	0.9308	0.9350	34	14
Image04-Kernel06	30.85	32.21	0.9095	0.9131	31	15
Image04-Kernel07	22.56	29.49	0.6827	0.8909	31	15
Image04-Kernel08	22.43	28.17	0.6859	0.8331	30	14

Bold values indicate that the deblurring performance of the proposed method is better than Kotera et al. (2013) in terms of PSNR, SSIM, or running time

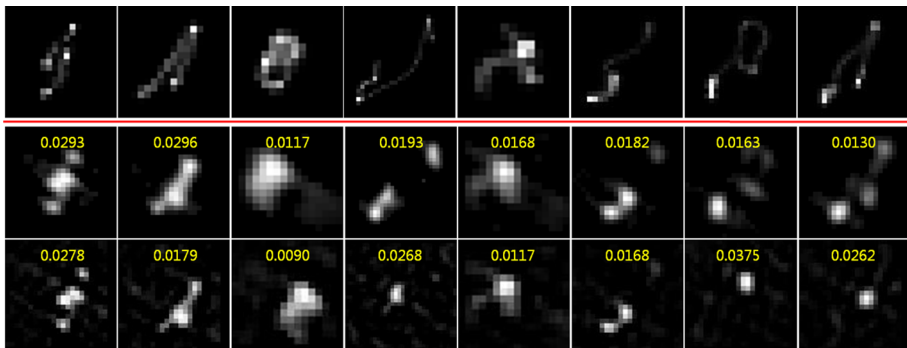


Fig. 3 The estimated blur kernels by the proposed method and Kotera et al. (2013) for Image04. The figure embedded into each blur kernel is the error measured by the sum of squared difference (SSD). Top: ground truth; middle: ours; bottom: Kotera et al. (2013)

However, it should be noted that the selected parameters in the above may not be applicable to all types of images. For example, Fig. 5 presents a text image example revealing the failure of the proposed approach. According to several recent works about text image blind deblurring, e.g., Pan et al. (2014), blind deblurring approaches for natural images are generally not applicable to the text images. As a matter of fact, Pan et al. (2014) have designed an L_0 -norm-based image prior in both gradient and intensity domains for removing kinds of motion blur in text images, which is quite different from the modeling principle of existing natural image blind deblurring approaches. Hence, the proposed method fails in Fig. 5 majorly because of the inherent problem in our modified total generalized variation, thus tuning parameters would not work in this case. In spite of that, it is really the case that there are probably not universal parameters for all the natural images. It is much interesting that Zhu and Milanfar (2010) have made a daring try very recently, proposing an automatic parameter selection scheme for image denoising algorithms. They experiment on two denoising methods including the steering kernel regression (SKR) (Takeda et al. 2007) and the block-matching and 3D filtering (BM3D) (Dabov et al. 2007). It is noticed that in either Takeda et al. (2007) or

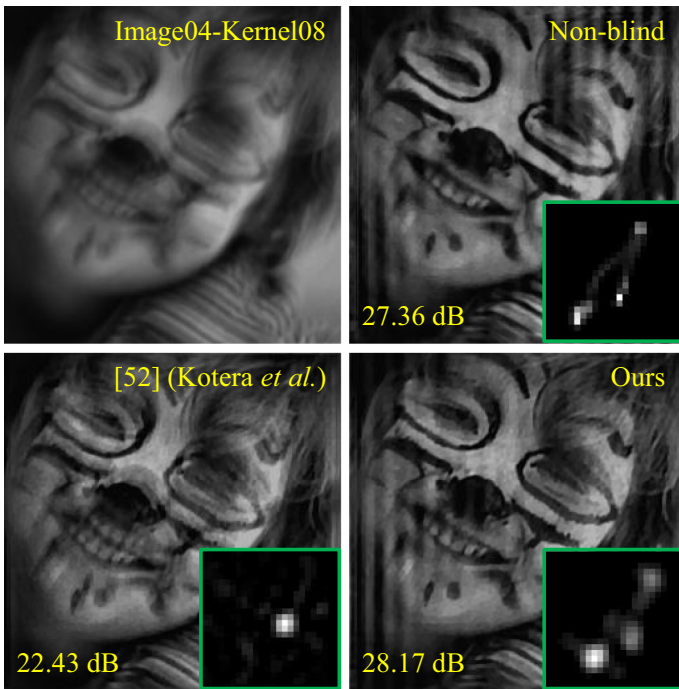


Fig. 4 The estimated blur kernels by Kotera et al. (2013) and the proposed method for Image04-Kernel08 and their corresponding final deblurred images



Fig. 5 One failure case of the proposed approach on the text image

Dabov et al. (2007), there are only one parameter to be automatically estimated along with the iteration process. One of the reviewers kindly suggests that the proposed blind deblurring algorithm should benefit from the excellent ideas in Zhu and Milanfar (2010). Though this is really a well-deserving studied topic, we currently do not have clear thoughts on this due to the involved too many parameters in the proposed method which is in fact a disadvantage of this paper. Furthermore, we show in Fig. 6 that the proposed approach may achieve better blind deblurring visual quality if the two parameters α_0, β_1 are adjusted respectively to $\alpha_0 = 0.15, \beta_1 = 0.5$. As a matter of fact, it is empirically found that one or two of the involved parameters need tuning as the fixed set of parameter values does not produce acceptable deblurring performance.

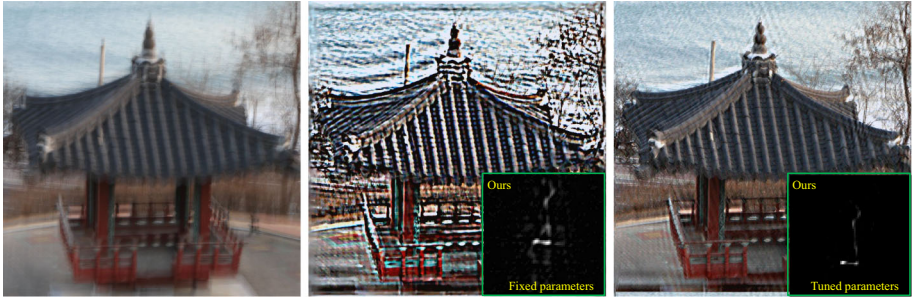


Fig. 6 The proposed approach may achieve better visual quality as using tuned parameters for many natural images. Left: blurred image; middle: blind deblurred results with the fixed parameters specified for Levin et al.'s benchmark dataset (Wipf and Zhang 2014); right: blind deblurred results as α_0 is changed as 0.15, and β_1 is changed as 0.5. In both cases, the blur kernel size is set as 71×71

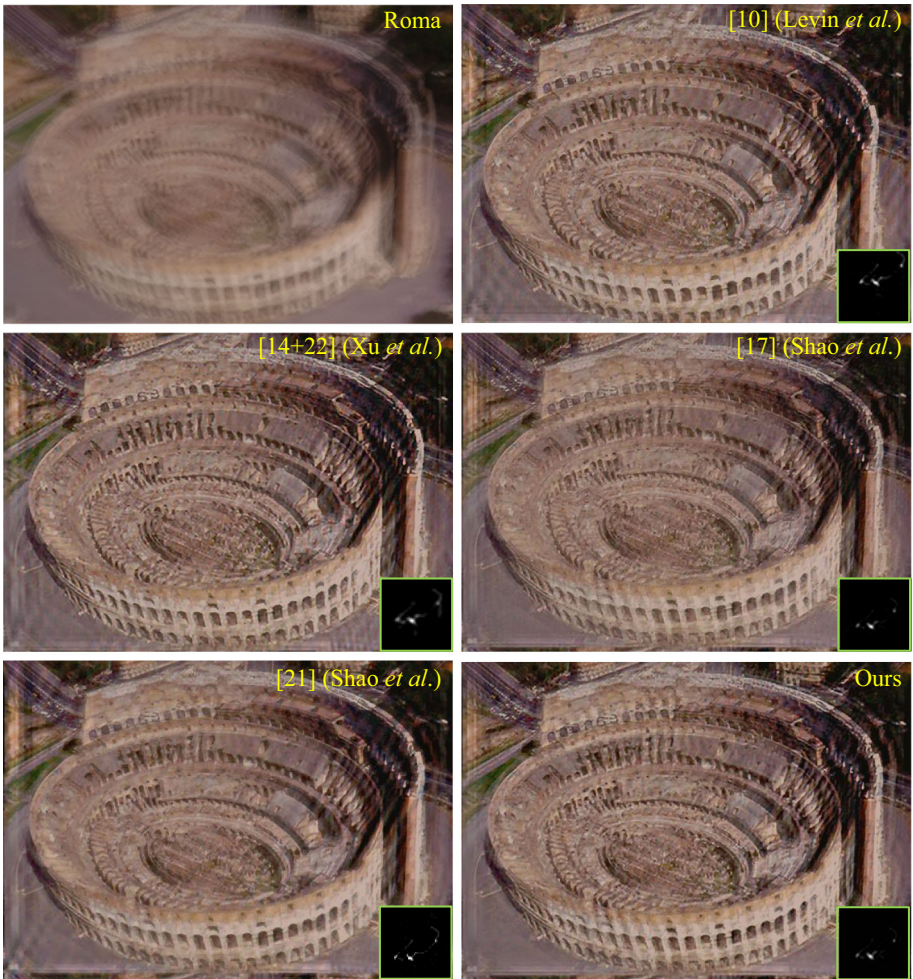


Fig. 7 Blind deblurring results with the real-world blurred image Roma, and all the methods assume the kernel size 51×51 . Note that in our method, α_0 is set as 0.15



Fig. 8 Blind deblurring results by our method and (Xu et al. 2013; Dong et al. 2014) assuming two kernel sizes. Note that in our method, α_0 is set as 0.15

In the following, we test the proposed method using real-world blurry color images while just tuning the parameter α_0 which is set universally as 0.15. The first group of experiments compares the proposed method to Levin et al. (2011a), Shao et al. (2016), Xu et al. (2013, 2014) and Shao et al. (2015). The deblurred images and the estimated kernels are shown in Fig. 7. We see that all the methods are capable of removing the large motion blur in Roma to a large degree, and the estimated blur kernels are of similar trajectories, demonstrating the practicality of the proposed approach to real-world blind deblurring problems. It is observed that Xu et al. (2013) and Dong et al. (2014) achieves the best compromise between the deblurring quality and speed in this example. However, Fig. 8 shows that the proposed approach is more robust than Xu et al. (2013) and Dong et al. (2014) in terms of the blur kernel size. That is, the deblurred images by our method are of similar visual perception with the two kernel sizes, but it is not the case for Xu et al. (2013) and Dong et al. (2014). To show the acceptable performance achieved by the proposed approach more convincingly, several more real-world blind deblurring experiments are conducted with the estimated images and blur kernels provided in Fig. 9. Note that in all the experiments above we just change α_0 from 0.25 to 0.15, and hence it is very possible that the proposed approach may achieve better results as jointly tuning other parameters as well. It is believed, however, that the automatic estimation of those parameter values is a more exciting and meaningful topic to be studied in the future.

5 Applications to single image nonparametric blind super-resolution

Single image super-resolution (SISR), aiming to generate a high-resolution (HR) image from a blurred and low-resolution (LR) one, has undergone a rapid development particularly in



Fig. 9 Left: blurry image; right: finally estimated image and kernel of the proposed approach with α_0 set as 0.15

the recent decade. However, a non-ignorable problem inherently in SISR is that most state-of-the-art methods (Dong et al. 2014; Kim et al. 2015; Yang et al. 2010, 2012, 2013, 2014; Glasner et al. 2009; Zeyde et al. 2012; Timofte et al. 2013; Peleg and Elad 2014; Timofte et al. 2014; Chang et al. 2004; Fattal 2007) including the latest deep learning-based approaches (Dong et al. 2014; Kim et al. 2015; Yang et al. 2010) suppose a known blur kernel in certain parametric forms. Actually, the influence of an accurate blur kernel, concluded by a recent analysis (Efrat et al. 2013) in both practical and theoretical aspects, is significantly larger than that of an advanced image prior. Therefore, blind SISR is more critical compared against existing non-blind ones. While there are a few parametric methods addressing the blind SISR (Begin and Ferrie 2007; Wang et al. 2005; He et al. 2009), they are based on a Gaussian kernel assumption. The low-quality SR results are naturally obtained as actual blur models are far away from the Gaussian hypothesis; one common instance is the combination of out-of-focus and camera shake blur.

In this section, nonparametric blind SISR is of our particular interest, and current literature seldom touches on this topic. In the past 15 years, Michaeli and Irani’s recent approach (Michaeli and Irani 2013) has been the most representative one. An inherent recurrence property of small image patches across distinct scales is explored and a $\text{MAP}_{\mathbf{h}}$ -based estimation scheme (Levin et al. 2011b) finally produces the blur kernel. While Joshi et al. (2008) aims to present a nonparametric kernel estimation method for both blind SISR and blind deblurring in a unified fashion, it merely restricts its treatment to single-mode blur kernels (Michaeli and Irani 2013; Marquina and Osher 2008), and heuristically builds on the detection and prediction of step edges as core clues to kernel estimation, i.e., unlike Michaeli and Irani (2013), not originating from a strict optimization principle.

In distinction to the nonparametric blind SISR methods above, our proposed method is also inspired by Bredies et al.’s TGV model (Bredies et al. 2010) and motivated by an idea that a successful blind SISR method builds heavily on jaggging artifact-free salient edges as well as staircase artifact-free flat regions pursued from the LR image. Specifically, our proposed approach is formulated by embedding a so-called convolutional consistency constraint and imposing the modified L_0 – L_1 -norm-based second-order TGV model (3), with the objective of getting more accurate estimation of the blur kernel from the single LR image.

Before formulating the proposed method, we first of all set up the blind SISR task. Let \mathbf{o} be the LR image of size $N_1 \times N_2$, and \mathbf{u} be the corresponding HR image of size $rN_1 \times rN_2$, with $r > 1$ an up-sampling factor. The relation between \mathbf{o} and \mathbf{u} can be expressed in the following way:

$$\mathbf{o} = \mathbf{S}\mathbf{H}\mathbf{u} + \mathbf{n} \quad \text{or} \quad \mathbf{o} = \mathbf{S}\mathbf{U}\mathbf{h} + \mathbf{n} \tag{22}$$

where \mathbf{U} and \mathbf{H} are the BCCB (block-circulant with circulant blocks) convolution matrices corresponding to the vectorized representations of the HR image \mathbf{u} and the blur kernel \mathbf{h} , and \mathbf{S} denotes the down-sampling matrix. In implementation, image boundaries are smoothed in order to prevent border artifacts. The task turns to estimate \mathbf{u} and \mathbf{h} given only the LR image \mathbf{g} and the up-sampling factor r .

According to our experience in blind deblurring, a successful blind SISR method dwells heavily on the jaggging artifact-free salient edges as well as staircase artifact-free smooth areas pursued from the LR image. To achieve this goal, we introduce a new blind SISR formulation inspired by the modified TGV model described above, given as

$$\min_{\mathbf{u}, \mathbf{h}} \lambda \|\mathbf{S}\mathbf{H}\mathbf{u} - \mathbf{o}\|_2^2 + c_{\mathbf{u}}^i \alpha_0 \mathcal{R}(\mathbf{u}) + c_{\mathbf{h}}^i \beta_0 \mathcal{Q}(\mathbf{h}) + \eta \|\mathbf{H}\mathbf{u} - \bar{\mathbf{u}}\|_2^2, \tag{23}$$

where η is also a positive tuning parameter. The same as in blind deblurring, $\mathcal{R}(\mathbf{u})$, $\mathcal{Q}(\mathbf{h})$ are chosen as (3) and (7), respectively. The first term in (23) models the image fidelity, formulated based on the assumption that \mathbf{n} in (22) is a white Gaussian noise, and the last term is the convolutional consistency constraint, playing a similar role as the first term, whose rationale is when a bicubic low-pass filter kernel is assumed for advanced non-blind SISR methods, the super-resolved image $\tilde{\mathbf{u}}$ is actually a blurred version of the true solution \mathbf{u} and we approximately get the relationship $\mathbf{H}\mathbf{u} \approx \tilde{\mathbf{u}}$. In this paper, the learning-based non-blind SISR method (Peleg and Elad 2014), i.e., Anchored Neighborhood Regression (ANR), is selected for blur-kernel estimation in all the blind SISR experiments. In fact, deep learning-based methods (Dong et al. 2014; Kim et al. 2015; Yang et al. 2010) may serve as better candidates, but this is not the emphasis in this paper.

With $\mathcal{R}(\mathbf{u})$ in (3), it is expected to serve approximately as an L_0 -norm-based Potts image model in the edge region, while as a higher-order total variation (TV), i.e., HNR, in the smooth area. And hence, by adapting the TGV as above as well as using the convolutional consistency constraint, we are allowed to get a more accurate salient edge image not suffering from the staircase artifacts in flat regions, which can be used as a more reliable clue to nonparametric blind SISR. Now, as provided the blur kernel \mathbf{h}_i , the super-resolved sharp image \mathbf{u}_{i+1} can be then estimated via

$$\mathbf{u}_{i+1}, \mathbf{v}_{i+1} = \arg \min_{\mathbf{u}, \mathbf{v}} c_{\mathbf{u}}^i \alpha_0 \Phi(\mathbf{u}, \mathbf{v}) + c_{\mathbf{u}}^i \alpha_1 \Psi(\mathbf{v}) + \lambda \|\mathbf{S}\mathbf{H}_i \mathbf{u} - \mathbf{g}\|_2^2 + \eta \|\mathbf{H}_i \mathbf{u} - \tilde{\mathbf{u}}\|_2^2. \quad (24)$$

When it turns to estimating the blur kernel, we follow the same empirical choice as Sect. 3.2, i.e., the blur kernel \mathbf{h}_{i+1} is estimated in the image derivative domain just due to the achieved better performance. Thus, \mathbf{h}_{i+1} is estimated via

$$\mathbf{h}_{i+1} = \arg \min_{\mathbf{h}} c_{\mathbf{h}}^i \beta_0 \|\mathbf{h}\|_0 + c_{\mathbf{h}}^i \beta_1 \|\mathbf{h}\|_2^2 + \sum_{d \in \{x, y\}} \{ \lambda \|\mathbf{S}(\mathbf{U}_{i+1})_d \mathbf{h} - \mathbf{o}_d\|_2^2 + \eta \|(\mathbf{U}_{i+1})_d \mathbf{h} - \tilde{\mathbf{u}}_d\|_2^2 \}, \quad (25)$$

subject to the constraint set $\Omega = \{ \mathbf{h} \geq 0, \|\mathbf{h}\|_1 = 1 \}$. In (25), $(\mathbf{U}_{i+1})_d$ denotes the convolutional matrix corresponding to the image gradient $(\mathbf{u}_{i+1})_d = \nabla_d \mathbf{u}_{i+1}$, and $\mathbf{g}_d = \nabla_d \mathbf{g}$, $\tilde{\mathbf{u}}_d = \nabla_d \tilde{\mathbf{u}}$. Obviously, both (24) and (25) can be solved in a similar spirit to the counterpart blind deblurring problem. But due to the sampling operator \mathbf{S} , \mathbf{u}_{i+1} , \mathbf{h}_{i+1} cannot be calculated via the FFT, and the conjugate gradient (CG) method is utilized instead which naturally makes our blind SISR algorithm computationally expensive. Besides, to take into account large-scale kernel estimation, the proposed approach is implemented in a multi-scale fashion, too. With the final output blur kernel, the super-resolved HR image can be generated by the simple non-blind TV-based SR approach (Marquina and Osher 2008). This is because Efrat et al. (Begin and Ferrie 2007) has concluded that an accurate reconstruction constraint (i.e., with a precise kernel) combined with a simple gradient regularization achieves SR results almost as good as those of state-of-the-art algorithms with sophisticated image priors.

We conduct a group of synthetic experiments using ten test images from the Berkeley Segmentation Dataset, as shown in Fig. 10. Each image is blurred by a 19×19 Gaussian kernel with standard deviation 2.5, 3 times down-sampled, and degraded by a white Gaussian noise with noise level equal to 1. We make comparisons between our approach and the recent method by Michaeli and Irani (2013). For fair comparison, the same non-blind SR method (Marquina and Osher 2008) is utilized to produce the final super-resolved image for the two compared methods. In the meanwhile, as estimating blur kernels the same kernel size 19×19 is set for the two methods. It should be also acknowledged that the authors of Michaeli and Irani (2013) have kindly prepared all the estimated kernels for us and therefore the comparison fairness can be guaranteed.



Fig. 10 Test images from the Berkeley Segmentation Dataset for quantitative evaluation of nonparametric blind SISR methods

Table 3 Image PSNR, SSIM, and kernel SSD corresponding to Michaeli and Irani (Michaeli and Irani 2013) and the proposed approach

Image No.	Image PSNR (dB)		Image SSIM		Kernel SSD	
	Michaeli and Irani (2013)	Ours	Michaeli and Irani (2013)	Ours	Michaeli and Irani (2013)	Ours
(a)	22.7	25.3	0.54	0.63	0.0022	0.0007
(b)	30.4	31.8	0.81	0.82	0.0015	0.0009
(c)	23.5	24.4	0.65	0.68	0.0038	0.0009
(d)	25.5	31.8	0.77	0.88	0.0026	0.0002
(e)	21.7	24.5	0.64	0.72	0.0063	0.0002
(f)	26.3	28.0	0.72	0.76	0.0030	0.0006
(g)	22.1	24.0	0.57	0.66	0.0083	0.0002
(h)	24.4	25.6	0.61	0.66	0.0035	0.0002
(i)	24.0	26.4	0.61	0.69	0.0055	0.0002
(j)	30.4	30.8	0.74	0.75	0.0009	0.0004

The PSNR (peak signal to noise ratio) metric is used for quantitative comparison. Table 3 provides the PSNR of super-resolved images by different approaches. Observed from the experimental results, it is clear that our method performs much better than Michaeli and Irani (2013) in all the ten images, showing the advantage and robustness of the newly advocated image prior for blind SISR. Moreover, we include the image SSIM and kernel SSD corresponding to the two methods in Table 3. Obviously, the proposed approach is demonstrated outperform Michaeli and Irani (2013), too, in terms of both image SSIM and kernel SSD.

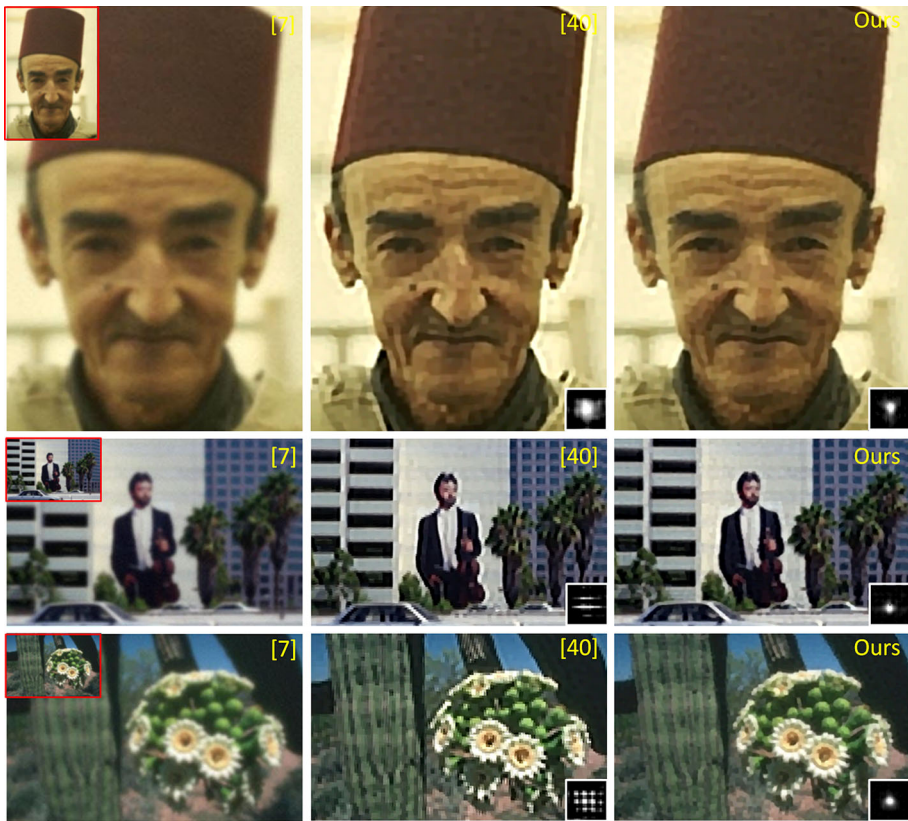


Fig. 11 Three times super-resolved images for Fig. 6b, g, i. Left to right: non-blind ANR (Levin et al. 2011b) with the low-res images, Michaeli and Irani (2013), and our proposed method

Note that the involved parameters are also fixed across all the experiments. Since the two blind restoration problems in this paper are formulated and calculated very similarly, we may also provide the parameter values for blind SISR without much confusion. Specifically, they are set as $\eta = 100$, $\lambda = 0.01$, $\alpha_0 = 0.25$, $\alpha_1 = 15$, $\beta_0 = 0.25$, $\beta_1 = 1.5$, $\gamma = 100$, $\chi = 1e6$, $I = J = M = 10$, $c_u = 2/3$, $c_h = 4/5$. Besides, in CG the error tolerance and the maximum number of iterations are set respectively as $1e^{-5}$ and 15 for calculation of both image and kernel. For visual comparison of Michaeli and Irani (2013) and the proposed method, we take the images Fig. 10b, g, i for example and show the super-resolved images in Fig. 11 accompanied by the blur kernels estimated by the two methods.

In Fig. 12, we conduct another synthetic blind SISR experiment in the scenario of motion blur. The original high-resolution image and the ground truth blur kernel are downloaded from the link: http://cg.postech.ac.kr/research/fast_motion_deblurring. Then, a low-resolution image is generated by applying the motion blur kernel and two-times down-sampling. It is clearly observed that (Michaeli and Irani 2013) fails to a great degree in this example, i.e., the super-resolved image is too blurry to be visually acceptable. Comparatively, the proposed approach performs well in recovering the ground truth kernel, which naturally leads to a super-resolved image of higher quality.



Fig. 12 Super-resolution with a synthetic LR image (Hollywood, the up-sampling factor is 2). Left to right, top to bottom: HR image (ground truth blur kernel), non-blind ANR (Levin et al. 2011b), Michaeli and Irani (2013), and our method

Moreover, we present three real-world nonparametric blind SISR experimental results in Figs. 13, 14 and 15, where the LR images are downloaded from the Internet. The images in Figs. 13 and 14 are of mixture of motion and Gaussian blur. It is seen that both our method and (Michaeli and Irani 2013) produce visually reasonable SR results in the two examples. While as for Fig. 15, our method produces a visually more pleasant SR image because one may obviously observe the ringing artifacts along the salient edges in the SR image corresponding to Michaeli and Irani (2013).

6 Conclusions

This paper concentrates on the topic of single image blind restoration including both the intensively studied blind deconvolution and the largely ignored blind super-resolution. It is empirically demonstrated that the two imaging tasks can be formulated in a common regularization perspective while differing in how we could deal well with the problem of jagg-ing artifacts along the salient edges, about which our introduced convolutional consistency constraint is proved a preliminary candidate. With this finding, single image blind restoration amounts to a proper image prior for pursuing the salient edges as core clues to kernel estimation. Inspired by the recent several efforts on unnatural image modeling for blind



Fig. 13 Super-resolution with a realistic LR image (Crowd, the up-sampling factor is 2). Left to right, top to bottom: LR image, non-blind ANR (Levin et al. 2011b), Michaeli and Irani (2013), and our method



Fig. 14 Super-resolution with a realistic LR image (Flower, the up-sampling factor is 3). Left to right: LR image, non-blind ANR (Levin et al. 2011b), Michaeli and Irani (2013), and our proposed method

image deblurring, this paper proposes a simple, yet effective modification strategy to Bredies et al.'s total generalized variation model (Bredies et al. 2010), which has been successfully

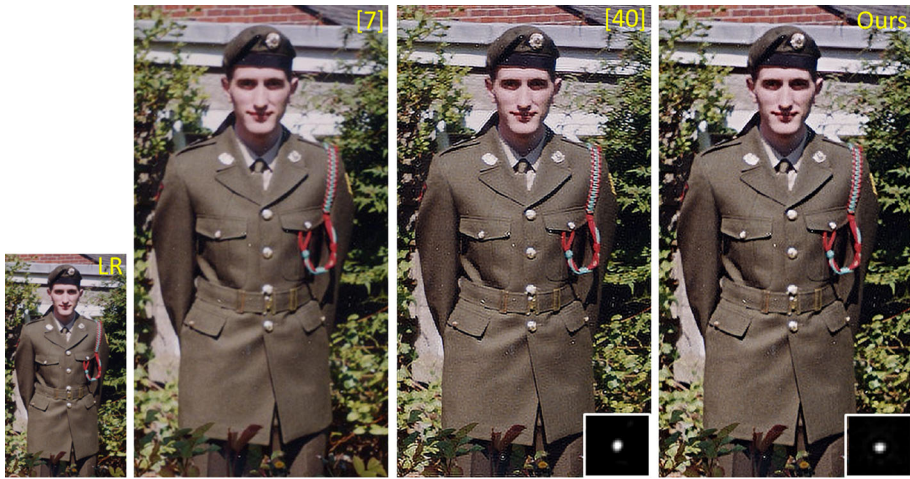


Fig. 15 Super-resolution with a realistic LR image (Soldier, the up-sampling factor is 2). Left to right: LR image, non-blind ANR (Levin et al. 2011b), Michaeli and Irani (2013), and our proposed method

harnessed for denoising and non-blind deconvolution and super-resolution. Specifically, a new L_0 – L_1 -norm-based image regularization is introduced as an unnatural prior for blur kernel estimation. To the best of our knowledge, ours has been the first paper touching the blind deblurring task by use of total generalized variation in a critical perspective. With the new image model, a maximum a posterior formulation is established for blind deconvolution as well as blind super-resolution. The estimated kernels and final images in both tasks demonstrate well the feasibility and effectiveness of our proposed approach, and also its advantage over state-of-the-art methods in terms of estimation accuracy. Lastly, motivated by this paper, we wish to establish a better bridge between blind deblurring and blind super-resolution in the near future.

Acknowledgements Many thanks are given to the anonymous reviewers for their pertinent and helpful comments which have improved the paper a lot. The first author is very grateful to Prof. Zhi-Hui Wei, Prof. Yi-Zhong Ma, Dr. Min Wu, and Mr. Ya-Tao Zhang for their kind supports in the past years. The study was supported in part by the Natural Science Foundation (NSF) of China (61771250, 61402239, 61602257, 11671004), NSF of Jiangsu Province (BK20130868, BK20160904) and Guangxi Provinces (2014GXNSFAA118360), NSF for the Jiangsu Institutions (16KJB520035), and also the Open Project Fund of both Jiangsu Key Laboratory of Image and Video Understanding for Social Safety (Nanjing University of Science and Technology, 30920140122007) and National Engineering Research Center of Communications and Networking (Nanjing University of Posts and Telecommunications, TXKY17008).

References

- Aharon, M., Elad, M., & Bruckstein, A. M. (2006). The K-SVD: An algorithm for designing of overcomplete dictionaries for sparse representation. *IEEE Transaction On Signal Processing*, 54(11), 4311–4322.
- Almeida, M., & Almeida, L. (2010). Blind and semi-blind deblurring of natural images. *IEEE Transactions on Image Processing*, 19(1), 36–52.
- Amizic, B., Molina, R., & Katsaggelos, A. K. (2012). Sparse Bayesian blind image deconvolution with parameter estimation. *EURASIP Journal on Image and Video Processing*, 20(2012), 1–15.

- Babacan, S. D., Molina, R., Do, M. N., & Katsaggelos, A. K. (2012). Bayesian blind deconvolution with general sparse image priors. In A. Fitzgibbon et al. (Ed.), *ECCV 2012, Part VI, LNCS* (Vol. 75771, pp. 341–355).
- Begin, I., & Ferrie, F. R. (2007). PSF recovery from examples for blind super-resolution. In *Proceedings of IEEE conference on image processing* (pp. 421–424).
- Benichoux, A., Vincent, E., & Gribonval, R. (2013). A fundamental pitfall in blind deconvolution with sparse and shift-invariant priors. In *Proceedings of international conference on acoustics, speech and signal processing* (pp. 6108–6112).
- Bredies, K., Kunisch, K., & Pock, T. (2010). Total generalized variation. *SIAM Journal on Imaging Sciences*, 3(3), 492–526.
- Chan, T. F., & Wong, C. K. (1998). Total variation blind deconvolution. *IEEE Transactions on Image Processing*, 7(3), 370–375.
- Chang, H., Yeung, D.-Y., & Xiong, Y. (2004). Super-resolution through neighbor embedding. In *Proceedings of IEEE international conference on computer vision* (pp. 275–282).
- Cho, S., & Lee, S. (2009). Fast motion deblurring. *ACM Transactions on Graphics*, 28(5), 145.
- Cho, S., & Lee, S. (2017). *Convergence analysis of MAP based blur kernel estimation*. Unpublished. [arXiv: 1611.07752v1](https://arxiv.org/abs/1611.07752v1).
- Dabov, K., Foi, A., Katkovnik, V., & Egiazarian, K. (2007). Image denoising by sparse 3D transform-domain collaborative filtering. *IEEE Transactions on Image Processing*, 16(8), 2080–2095.
- Dong, C., Loy, C. C., He, K., & Tang, X. (2014). Learning a deep convolutional network for image super-resolution. In *Proceedings of european conference on computer vision* (pp. 194–199).
- Efrat, N., Glasner, D., Apartsin, A., Nadler, B., & Levin, A. (2013). Accurate blur models versus image priors in single image super-resolution. In *Proceedings of IEEE conference on computer vision* (pp. 2832–2839).
- Fattal, R. (2007). Image upsampling via imposed edge statistics. *ACM Transactions on Graphics*, 26, 95.
- Fergus, R., Singh, B., Hertzmann, A., Roweis, S. T., & Freeman, W. T. (2006). Removing camera shake from a single photograph. *ACM Transactions on Graphics*, 25(3), 787–794.
- Glasner, D., Bagon, S., & Irani, M. (2009). Super-resolution from a single image. In *Proceedings of IEEE international conference on computer vision* (pp. 349–356).
- Gu, S., Zhang, L., Zuo, W., & Feng, X. (2014). Weighted nuclear norm minimization with application to image denoising. In *IEEE international conference on computer vision and pattern recognition (CVPR)* (pp. 2862–2869).
- He, C., Hu, C., Zhang, W., & Shi, B. (2014). A fast adaptive parameter estimation for total variation image restoration. *IEEE Transactions on Image Processing*, 23(12), 4954–4967.
- He, K., Sun, J., & Tang, X. (2011). Single image haze removal using dark channel prior. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(12), 2341–2353.
- He, Y., Yap, K. H., Chen, L., & Chau, L. P. (2009). A soft MAP framework for blind super-resolution image reconstruction. *Image and Vision Computing*, 27, 364–373.
- Joshi, N., Szeliski, R., & Kriegman, D. J. (2008). PSF estimation using sharp edge prediction. In *Proceedings of IEEE conference on CVPR* (pp. 1–8).
- Kim, J., Lee, J. K., & Lee, K. M. (2015). *Deeply-recursive convolutional network for image super-resolution*. [arXiv:1511.04491](https://arxiv.org/abs/1511.04491).
- Kotera, J., Sroubek, F., & Milanfar, P. (2013). Blind deconvolution using alternating maximum a posteriori estimation with heavy-tailed priors. In R. Wilson et al. (Eds.), *CAIP, Part II, LNCS* (Vol. 8048, pp. 59–66).
- Krishnan, D., & Fergus, R. (2009). Fast image deconvolution using hyper-laplacian priors. In *Proceedings of international conference on neural information and processing systems* (pp. 1033–1041).
- Krishnan, D., Tay, T., & Fergus, R. (2011). Blind deconvolution using a normalized sparsity measure. In *Proceedings of international conference on computer vision and pattern recognition* (pp. 233–240).
- Laghrib, A., Ezzaki, M., Rhabi, M., Hakim, A., Monasse, P., & Raghay, S. (2018). Simultaneous deconvolution and denoising using a second order variational approach applied to image super resolution. *Computer Vision and Image Understanding*, 168, 50–63.
- Lai, W.-S., Huang, J.-B., Hu, Z., Ahuja, N., & Yang, M.-H. (2016). A comparative study for single image blind deblurring. In *IEEE international conference on computer vision and pattern recognition (CVPR)* (pp. 1701–1709).
- Levin, A., Weiss, Y., Durand, F., & Freeman, W. T. (2011a). Efficient marginal likelihood optimization in blind deconvolution. In *Proceedings of international conference on computer vision and pattern recognition* (pp. 2657–2664).
- Levin, A., Weiss, Y., Durand, F., & Freeman, W. T. (2011b). Understanding blind deconvolution algorithms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(12), 2354–2367.
- Marquina, A., & Osher, S. J. (2008). Image super-resolution by TV-regularization and Bregman iteration. *Journal of Scientific Computing*, 37, 367–382.

- Michaeli, T. & Irani, M. (2013). Nonparametric blind super-resolution. In *Proceedings of IEEE conference on computer vision* (pp. 945–952).
- Money, J. H., & Kang, S. H. (2008). Total variation minimizing blind deconvolution with shock filter reference. *Image and Vision Computing*, 26(2), 302–314.
- Ohkoshi, K., Sawada, M., Goto, T., Hirano, S., & Sakurai, M. (2014). Blind image restoration based on total variation regularization and shock filter for blurred images. In *2014 IEEE international conference on consumer electronics (ICCE)*, Las Vegas, NV (pp. 217–218).
- Osher, S., & Rudin, L. I. (1990). Feature-oriented image enhancement using shock filters. *SIAM Journal on Numerical Analysis*, 27(4), 919–940.
- Pan, J., Hu, Z., Su, Z., & Yang, M.-H. (2014). Deblurring text images via L_0 -regularized intensity and gradient prior. In *IEEE international conference on computer vision and pattern recognition (CVPR)* (pp. 1628–1636).
- Pan, J., & Su, Z. (2013). Fast L_0 -regularized kernel estimation for robust motion deblurring. *IEEE Signal Processing Letters*, 20(9), 1107–1114.
- Pan, J., Sun, D., Pfister, H., & Yang, M.-H. (2016). Blind image deblurring using dark channel prior. In *IEEE international conference on computer vision and pattern recognition (CVPR)* (pp. 2901–2908).
- Peleg, T., & Elad, M. (2014). A statistical prediction model based on sparse representations for single image super-resolution. *IEEE Transactions on Image Processing*, 23, 2569–2582.
- Perrone, D., & Favaro, P. (2016a). A clearer picture of total variation blind deconvolution. *IEEE Transactions on Pattern Analysis Machine Intelligence*, 38(6), 1041–1055.
- Perrone, D., & Favaro, P. (2016b). A logarithmic image prior for blind deconvolution. *International Journal of Computer Vision*, 117(2), 159–172.
- Roth, S., & Black, M. J. (2009). Fields of experts. *International Journal of Computer Vision*, 82(2), 205–229.
- Rubinstein, R., Peleg, T., & Elad, M. (2013). Analysis K-SVD: A dictionary-learning algorithm for the analysis sparse model. *IEEE Transaction on Signal Processing*, 61(3), 661–677.
- Rudin, L. I., Osher, S., & Fatemi, E. (1992). Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena*, 60, 259–268.
- Ruiz, P., Zhou, X., Mateos, J., Molina, R., & Katsaggelos, A. K. (2015). Variational Bayesian blind image deconvolution: A review. *Digital Signal Processing*, 52, 116–127.
- Sampat, M., Wang, Z., Gupta, S., Bovik, A., & Markey, M. (2009). Complex Wavelet structural similarity: A new image similarity index. *IEEE Transactions on Image Processing*, 18(11), 2385–2401.
- Shao, W.-Z., Deng, H.-S., Ge, Q., Li, H.-B., & Wei, Z.-H. (2016). Regularized motion blur-kernel estimation with adaptive sparse image prior learning. *Pattern Recognition*, 51, 402–424.
- Shao, W.-Z., Li, H.-B., & Elad, M. (2015). Bi- l_0 - l_2 -norm regularization for blind motion deblurring. *Journal of Visual Communication and Image Representation*, 33, 42–59.
- Shearer, P., Gilbert, A. C., & Hero A. O., III. (2013). Correcting camera shake by incremental sparse approximation. In *Proceedings of IEEE international conference on image processing* (pp. 572–576).
- Takeda, H., Farsiu, S., & Milanfar, P. (2007). Kernel regression for image processing and reconstruction. *IEEE Transactions on Image Processing*, 16(2), 349–366.
- Timofte, R., Smet, V. D., & Gool, L. V. (2013). Anchored neighborhood regression for fast example-based super-resolution. In *Proceedings of IEEE international conference on computer vision* (pp. 1920–1927).
- Timofte, R., Smet, V. D., & Gool, L. V. (2014). A+: Adjusted anchored neighborhood regression for fast super-resolution. In *Proceedings of Asian conference on computer vision* (pp. 111–126).
- Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P. (2004). Image quality assessment: From error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4), 600–612.
- Wang, Q., Tang, X., & Shum, H. (2005). Patch based blind image super resolution. In *Proceedings of IEEE conference on computer vision* (pp. 709–716).
- Wang, R. & Tao, D. (2014). *Recent progress in image deblurring*. Unpublished. [arXiv:1409.6838](https://arxiv.org/abs/1409.6838)
- Wipf, D., & Zhang, H. (2013). Analysis of Bayesian blind deconvolution. In *Proceedings of international conference on energy minimization methods in computer vision and pattern recognition* (pp. 40–53).
- Wipf, D., & Zhang, H. (2014). Revisiting Bayesian blind deconvolution. *Journal of Machine Learning Research*, 15, 3595–3634.
- Xu, L., & Jia, J. (2010). Two-phase kernel estimation for robust motion deblurring. In *ECCV, Part I, LNCS* (Vol. 6311, pp. 157–170).
- Xu, L., Zheng, S., & Jia, J. (2013). Unnatural L_0 sparse representation for natural image deblurring. In *Proceedings of international conference on computer vision and pattern recognition* (pp. 1107–1114).
- Yan, J., & Lu, W.-S. (2015). Image denoising by generalized total variation regularization and least squares fidelity. *Multidimensional Systems and Signal Processing*, 26(1), 243–266.
- Yang, J., Lin, Z., & Cohen, S. (2013). Fast image super-resolution based on in-place example regression. In *Proceedings of IEEE conference on CVPR* (pp. 1059–1066).

- Yang, C.-Y., Ma, C., & Yang, M.-H. (2014). Single image super resolution: A benchmark. In *Proceedings of european conference on computer vision* (pp. 372–386).
- Yang, J., Wang, Z., Lin, Z., Cohen, S., & Huang, T. (2012). Coupled dictionary training for image super-resolution. *IEEE Transactions on Image Processing*, 21, 3467–3478.
- Yang, J., Wright, J., Huang, T., & Ma, Y. (2010). Image super-resolution via sparse representation. *IEEE Transactions on Image Processing*, 19, 2861–2873.
- Zeyde, R., Elad, M., & Protter, M. (2012). On single image scale-up using sparse-representations. In *Proceedings of curves and surfaces* (pp. 711–730).
- Zhang, H., Tang, L., Fang, Z., Xiang, C., & Li, C. (2017). Nonconvex and nonsmooth total generalized variation model for image restoration. *Signal Processing*, 143, 69–85.
- Zhu, X., & Milanfar, P. (2010). Automatic parameter selection for denoising algorithms using a no-reference measure of image content. *IEEE Transactions on Image Processing*, 19(12), 3116–3132.



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 Systems 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.



Feng Wang received the B.S. degree in Communication Engineering from China Jiliang University in June 2014, Hangzhou, China, and the M.S. degree in Signal and Information Processing from Nanjing University of Posts and Telecommunications in May 2017, Nanjing, China. He is currently an Intelligent Algorithm Engineer in Institutes of Advanced Technology, Zhejiang Dahua Technology Co., Ltd, working in the fields of computer vision, natural language processing, and deep learning.



Li-Li Huang received her Ph.D. in Pattern Recognition and Intelligent Systems from Nanjing University of Science and Technology, Nanjing, China, in 2013. She is currently an Associate Professor in Jinling Institute of Technology, Nanjing, China. Her research interests cover image processing, image modeling, and superresolution reconstruction.