

# Blind Motion Deblurring Based on Fused $\ell_0$ - $\ell_1$ Regularization

Kai Wang<sup>1(✉)</sup>, Yu Shen<sup>1</sup>, Liang Xiao<sup>1,2</sup>, Zhihui Wei<sup>1</sup>, and Linxue Sheng<sup>3</sup>

<sup>1</sup> School of Computer Science and Engineering, Nanjing University of Science and Technology, Xiaolingwei 200, Nanjing 210094, Jiangsu, China  
mfranciswong@163.com, shenyu0305@126.com, gswei@mail.njust.edu.cn

<sup>2</sup> Key Lab of Intelligent Perception and Systems for High-Dimensional Information of Ministry of Education, Xiaolingwei 200, Jiangsu 210094, Nanjing, China  
xiaoliang@mail.njust.edu.cn

<sup>3</sup> School of Science, Nanjing University of Science and Technology, Xiaolingwei 200, Jiangsu 210094, Nanjing, China  
njustslx@outlook.com

**Abstract.** In blind motion deblurring, various regularization models using different priors of either the image or the blur kernel are proposed to estimate the blur kernel, with tendency towards naive  $\ell_0$  norm or its approximations. In this paper, we propose a novel fused  $\ell_0$ - $\ell_1$  regularization approach to estimate the motion blur kernel by imposing sparsity and continuation properties on natural images. A fast numerical scheme is then deduced by coupling operator splitting and the Augmented Lagrangian method to solve the proposed model efficiently. Experimental results on both synthetic and real data demonstrate the effectiveness of the proposed method and the superiority over the state-of-the-art methods.

**Keywords:** Blind motion deblurring · Blur kernel estimation · Split augmented lagrangian

## 1 Introduction

Blind motion deblurring, also known as blind deconvolution, has been extensively studied since the pioneering work of [1]. Although it achieves significant success with a few encouraging solutions, the results are still far from perfect. In the current literature, most blind deconvolution approaches share a common Bayesian framework, which fall into two camps, Max a Posterior (MAP) [2–8] and Variational Bayesian (VB) [1, 9–12]. An invigorating work [9] primarily analyzes the failure of natural image statistics that would lead to no-blur solutions. Recently, an insightful theoretical analysis [12] further demonstrates that the VB methodology can be recast as an unconventional MAP problem with a particular prior. Although VB methods have theoretical and empirical superiority to conventional MAP algorithms, the posterior inference of VB methods remains less tractable [9, 11, 12]. Both of these methods mentioned above will be discussed as follows, particularly in aspect of prior choice.

## 1.1 Related Works

Prior blind deconvolution approaches are basically with regard to exploiting techniques to utilize salient edges, either explicitly or implicitly. For instance, Cho *et al.* [2], and Xu *et al.* [3] have utilized bilateral filtering or Gaussian filtering together with shock filtering to predict sharp edges, which involve explicit filtering and edge selection. Shan *et al.* [4] have adopted a sparse image prior and used a weighting scheme to remove harmful subtle image structures. Krishnan *et al.* [5] have imposed a scale-invariant sparsity measure to better discriminate sharp images from blurry ones. It is obviously to see that the new sparsity prior is a good approximation of  $\ell_0$  norm, the natural interpretation of sparsity. Another alternative approximation is the unnatural  $\ell_0$  regularizer which is proposed by Xu *et al.* [6]. Coinstantaneously, Pan *et al.* [7] have directly imposed  $\ell_0$  norm on image gradients. On the VB camp, Babacan *et al.* [11] have developed a general sparsity-inspired prior on the image which has been practically demonstrated more powerful than others [1, 9]. It is highly consistent with the theoretical presentation in Wipf *et al.*'s work [12]. More recently, Krishnan *et al.* [8] shows that all these successful approaches rely on a key principle, that is, sparsity promotion in the image gradient domain.

Summarizing above discussions, the top-performing methods have placed a great deal of effort on the choice of proper priors both on the image and the blur kernel. Essentially, the most effective unnatural image priors may not reflect the true statistics of natural images in blind deconvolution. There is a common feature that all these sparsity-inducing priors can be regarded as some approximations of the  $\ell_0$  norm via different techniques or naive  $\ell_0$  norm prior.

## 1.2 Our Contributions

In this paper, we propose a novel sparsity-inducing model fused with  $\ell_0$  norm and a difference penalty term, which is found to be akin to the fused LASSO [13]. It is worth noting that we use  $\ell_0$  norm and not  $\ell_1$  norm based on the theory [9] that an  $\ell_1$  norm based image prior naturally leads to a trivial no-blur solution. By introducing this prior, namely the fused  $\ell_0$ - $\ell_1$  prior, the proposed method is fairly different from previous works in the following aspects:

- (1) The fused  $\ell_0$ - $\ell_1$  regularization, as a new sparsity-inducing model on the sharp image, is specifically proposed for the kernel estimation. Benefit from enforcing continuation in the new model, it helps to produce higher quality sharp image which has continuous salient edges. Since salient edges serves as an important clue for blur kernel estimation, the model further generate accurate blur kernel with nearly true support of a desired blur kernel.
- (2) The new model for motion blur kernel estimation is then practically formulated as a compound regularized problem with respect to alternatively estimating sharp image and blur kernel, and can be transformed into two subproblems using variable splitting technique. Furthermore, by coupling operator splitting and the Augmented Lagrangian method as well as exploiting fast Fourier transform (FFT), we deduce a fast numerical algorithm to efficiently solve the bilinear optimization problem.

By comparing our results with some of the leading motion deblurring methods, both in synthetic experiments and real-world motion blurred images, we demonstrate that our approach outperforms most of the current methods.

## 2 Fused $\ell_0$ - $\ell_1$ -Regularized Kernel Estimation

### 2.1 The Proposed Method

Generally, the formation process of motion blur can be modeled as

$$\mathbf{B} = \mathbf{I} \otimes \mathbf{k} + \mathbf{n}, \quad (1)$$

where  $\mathbf{B}$ ,  $\mathbf{I}$ ,  $\mathbf{k}$  and  $\mathbf{n}$  represent the blurred image, latent image, blur kernel and the additive Gaussian noise, respectively.  $\otimes$  denotes the convolution operator. Currently, most of the previous methods [1, 5, 7–9] have demonstrated that the sharp edges are the vital clue of the kernel estimation. Therefore, instead of recovering latent images, we estimate the blur kernel in the derivative space. Specifically, the kernel estimation model is written as

$$\min_{\mathbf{x}, \mathbf{k}} \|\mathbf{x} \otimes \mathbf{k} - \mathbf{y}\|_2^2 + \lambda \Phi(\mathbf{x}) + \eta \Psi(\mathbf{k}), \quad (2)$$

where  $\mathbf{x}$  denotes  $\nabla \mathbf{I} = [\nabla_h \mathbf{I}, \nabla_v \mathbf{I}]$ ,  $\mathbf{y}$  denotes  $\nabla \mathbf{B} = [\nabla_h \mathbf{B}, \nabla_v \mathbf{B}]$ , and  $\lambda, \eta$  are regularization weights.

Model (2) consists of three terms. The first term is the data fitting term, the second term  $\Phi(\mathbf{x})$  is the regularization term on  $\mathbf{x}$ , and the third term  $\Psi(\mathbf{k})$  is the regularization term on blur kernel  $\mathbf{k}$ .

Since the unnatural  $\ell_0$  norm or its approximating priors, no matter explicitly or implicitly [5–8, 12], have shown their effectiveness. It roughly reveals that the desired sharp image compromising of salient edges performs better than original image in motion blur kernel estimation. A good candidate of sharp image should have two essential factors: (1) preserving as many continuous salient edges as possible; (2) reducing the staircase and ringing artifacts along the salient edges, while eliminating ringing artifacts inside the homogeneous regions as much as possible. In order to impose the continuation on latent image, as well as enforce unnatural sparsity, we propose a new prior to produce a better intermediate image results as far as possible to fulfill aforementioned factors.

Similar to Eq. (2), a cost function with the new sparsity-inducing prior is given as follows

$$\min_{\mathbf{x}, \mathbf{k}} \|\mathbf{x} \otimes \mathbf{k} - \mathbf{y}\|_2^2 + \lambda_1 \|\mathbf{x}\|_0 + \lambda_2 \|\mathbf{D}\mathbf{x}\|_1 + \eta \|\mathbf{k}\|_2^2, \quad (3)$$

where  $\lambda_1$ ,  $\lambda_2$  and  $\eta$  are regularization weights. The new image prior combines  $\ell_0$  norm and a continuation constraint to suppress isolated points along the generated salient edges, and to prevent the unpleasant staircase and ringing artifacts at the same time. The so called fused  $\ell_0$ - $\ell_1$  regularization term inherit the effectiveness of recovering salient edges from  $\ell_0$  norm and the ability of preserving continuation from the difference penalty term.

## 2.2 Numerical Algorithm

By now, we come to the issue of alternatively estimating motion blur kernel and sharp edges. Therefore, the proposed method addresses the motion kernel estimation by solving the following two splitting optimization problems with respect to  $\mathbf{x}$  and  $\mathbf{k}$ .

**Sharp Edges Restoration.** In the first step, i.e., the sharp edges restoration, the problem can be formulated as

$$\min_{\mathbf{x}} \|\mathbf{x} \otimes \mathbf{k} - \mathbf{y}\|_2^2 + \lambda_1 \|\mathbf{x}\|_0 + \lambda_2 \|D\mathbf{x}\|_1. \quad (4)$$

Here, we formulate the difference penalty term by introducing the difference operator  $D = (D_1, D_2)$  for computation simplicity, where  $D_1 = [1, -1]$ ,  $D_2 = [1, -1]^T$ . It is difficult to directly solve the discrete optimization problem (4) because of its non-convexity. It should be noted that a similar numerical scheme is considered in [14, 15] by coupling variable splitting and augmented Lagrangian. Borrowing the idea of the so called Split Augmented Lagrangian (SAL) method, we first convert model (4) to the following equivalent constrained problem,

$$\begin{aligned} \min_{\mathbf{u}, \mathbf{v}, \mathbf{x}} \quad & \|\mathbf{x} \otimes \mathbf{k} - \mathbf{y}\|_2^2 + \lambda_1 \|\mathbf{u}\|_0 + \lambda_2 \|\mathbf{v}\|_1, \\ \text{s.t.} \quad & \mathbf{u} = \mathbf{x}, \mathbf{v} = D\mathbf{x}. \end{aligned} \quad (5)$$

Then, we have the following associated Lagrangian function,

$$\begin{aligned} L(\mathbf{u}, \mathbf{v}, \mathbf{x}, J_1, J_2) = & \|\mathbf{x} \otimes \mathbf{k} - \mathbf{y}\|_2^2 + \lambda_1 \|\mathbf{u}\|_0 + \lambda_2 \|\mathbf{v}\|_1 + \langle J_1, \mathbf{x} - \mathbf{u} \rangle \\ & + \langle J_2, D\mathbf{x} - \mathbf{v} \rangle + \frac{\xi_1}{2} \|\mathbf{x} - \mathbf{u}\|_2^2 + \frac{\xi_2}{2} \|D\mathbf{x} - \mathbf{v}\|_2^2, \end{aligned} \quad (6)$$

where  $J_1, J_2$  are the vectors of Lagrange multipliers, and  $\xi_1, \xi_2$  are positive constants. Hence, the corresponding iterative scheme of the SAL method is given as follows,

$$\begin{cases} \mathbf{u}^{t+1} = \arg \min_{\mathbf{u}} L(\mathbf{u}, \mathbf{v}^t, \mathbf{x}^t, J_1^t, J_2^t), \\ \mathbf{v}^{t+1} = \arg \min_{\mathbf{v}} L(\mathbf{u}^{t+1}, \mathbf{v}, \mathbf{x}^t, J_1^t, J_2^t), \\ \mathbf{x}^{t+1} = \arg \min_{\mathbf{x}} L(\mathbf{u}^{t+1}, \mathbf{v}^{t+1}, \mathbf{x}, J_1^t, J_2^t), \\ J_1^{t+1} = J_1^t + \xi_1(\mathbf{x}^{t+1} - \mathbf{u}^{t+1}), \\ J_2^{t+1} = J_2^t + \xi_2(D\mathbf{x}^{t+1} - \mathbf{v}^{t+1}), \end{cases} \quad (7)$$

where  $0 \leq t \leq T - 1$  denotes the  $t$ -th iteration and  $T$  represents the maximum iterations.

For the two sub-problems of updating auxiliary variables  $\mathbf{u}$  and  $\mathbf{v}$ , the solutions  $\mathbf{u}^{t+1}$  and  $\mathbf{v}^{t+1}$  can be easily computed through some straightforward manipulations and given as

$$\begin{aligned} \mathbf{u}^{t+1} &= \arg \min_{\mathbf{u}} \frac{\xi_1}{2} \|\mathbf{x}^t + \frac{1}{\xi_1} J_1^t - \mathbf{u}\|_2^2 + \lambda_1 \|\mathbf{u}\|_0 \\ &= \Theta_{\text{Hard}}(\mathbf{x}^t + \frac{1}{\xi_1} J_1^t, (\frac{2\lambda_1}{\xi_1})^{\frac{1}{2}}), \end{aligned} \quad (8)$$

$$\begin{aligned} \mathbf{v}^{t+1} &= \arg \min_{\mathbf{v}} \frac{\xi_2}{2} \|D\mathbf{x}^t + \frac{1}{\xi_2} J_2^t - \mathbf{v}\|_2^2 + \lambda_2 \|\mathbf{v}\|_1 \\ &= \Theta_{\text{Soft}}(D\mathbf{x}^t + \frac{1}{\xi_2} J_2^t, \frac{\lambda_2}{\xi_2}), \end{aligned} \quad (9)$$

where  $\Theta_{\text{Hard}}(\cdot)$  and  $\Theta_{\text{Soft}}(\cdot)$  are the hard-thresholding operator and soft-thresholding operator respectively.

Since the  $\mathbf{x}$  sub-problem is a least square problem, the closed form solution can be given by

$$\begin{aligned} \mathbf{x}^{t+1} &= \arg \min_{\mathbf{x}} \|\mathbf{x} \otimes \mathbf{k} - \mathbf{y}\|_2^2 + \frac{\xi_1}{2} \|\mathbf{x} + \frac{1}{\xi_1} J_1^t - \mathbf{u}^{t+1}\|_2^2 + \frac{\xi_2}{2} \|D\mathbf{x} + \frac{1}{\xi_2} J_2^t - \mathbf{v}^{t+1}\|_2^2 \\ &= \left( \mathbf{K}^T \mathbf{K} + \frac{\xi_1}{2} + \frac{\xi_2}{2} D^T D \right)^{-1} \cdot \\ &\quad \left( \mathbf{K}^T \mathbf{y} + \frac{\xi_1}{2} \mathbf{u}^{t+1} - \frac{1}{2} J_1^t + \frac{\xi_2}{2} D^T (\mathbf{v}^{t+1} - \frac{1}{\xi_2} J_2^t) \right). \end{aligned} \quad (10)$$

In fact, a circular convolution is assumed for the observation model (1), and hence (10) can be computed efficiently using the FFT instead of the gradient descent based methods.

To summarize, the numerical algorithm of the sharp edges restoration is presented in Algorithm 1. Note that the Augmented Lagrangian method has been proved to be equivalent to the Split Bregman method [16], thus the convergence property can be guaranteed in theory.

---

**Algorithm 1.** SAL method for Sharp Edge Restoration

---

**Input:** blurred image  $\mathbf{B}$ ; blur kernel  $\mathbf{k}$ ; Lagrangian multipliers  $J_1^0 = J_2^0 = 0$ ; regularization parameters  $\lambda_1, \lambda_2$ ; Augmented Lagrangian penalty parameters  $\xi_1, \xi_2$ ; index of iterations  $t = 0$ .

**while**  $\xi_1 \leq 2^8$  **and**  $\xi_2 \leq 2^8$  **do**

**Step 1:** Update  $\mathbf{u}^{t+1}$  by computing (8);

**Step 2:** Update  $\mathbf{v}^{t+1}$  by computing (9);

**Step 3:** Update  $\mathbf{x}^{t+1}$  by computing (10) based on FFT;

**Step 4:** Update  $J_1^{t+1} = J_1^t + \xi_1(\mathbf{x}^{t+1} - \mathbf{u}^{t+1})$  and  $J_2^{t+1} = J_2^t + \xi_2(D\mathbf{x}^{t+1} - \mathbf{v}^{t+1})$ ;

$t \leftarrow t + 1$ ;

**end while**

**Output:** sharp edges  $\mathbf{x}^t$ .

---

**Motion Kernel Estimation.** In the kernel estimation step, the latent blur kernel is obtained by minimizing the following optimization functional

$$\min_{\mathbf{k}} \|\mathbf{x} \otimes \mathbf{k} - \mathbf{y}\|_2^2 + \eta \|\mathbf{k}\|_2^2. \quad (11)$$

Evidently, it is a least square problem whose solution can be also computed by using the FFT, i.e.,

$$\mathbf{k} = \mathcal{F}^{-1} \left( \frac{\overline{\mathcal{F}(\mathbf{x})} \mathcal{F}(\mathbf{y})}{\overline{\mathcal{F}(\mathbf{x})} \mathcal{F}(\mathbf{x}) + \eta} \right), \quad (12)$$

where  $\mathcal{F}(\cdot)$  and  $\mathcal{F}^{-1}(\cdot)$  denote the FFT and inverse FFT, respectively, and  $\overline{\mathcal{F}(\cdot)}$  is the complex conjugate operator.

Finally, we summarize the proposed method in Algorithm 2.

---

**Algorithm 2.** Algorithm for fused  $\ell_0$ - $\ell_1$  regularized kernel estimation

---

**Input:** blurred image  $\mathbf{B}$ ; kernel size  $S$ ; regularization parameters  $\lambda_1, \lambda_2, \eta$ ; parameter decreasing factor  $\rho_1, \rho_2$ ; maximum iteration number  $L$ .

**for**  $l = 1 \rightarrow L$  **do**

**Step 1:** Update sharp edges  $\mathbf{x}_l$  by Algorithm 1;

**Step 2:** Update blur kernel  $\mathbf{k}_l$  by computing (12);

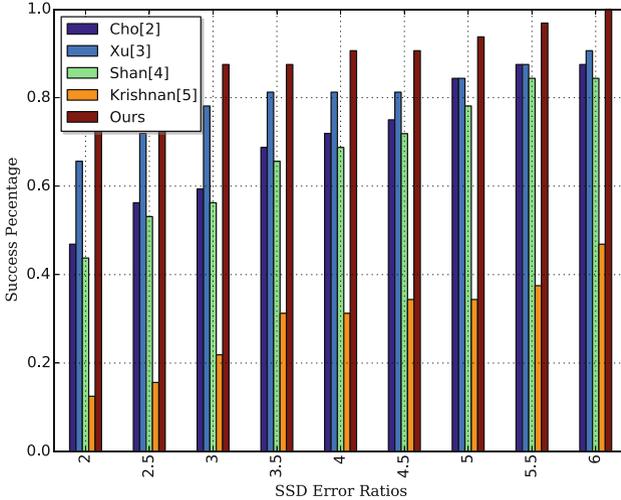
**Step 3:** Compute  $\lambda_1 \leftarrow \rho_1 \lambda_1, \lambda_2 \leftarrow \rho_2 \lambda_2$ ;

**end for**

**Output:** Ultimate blur kernel  $\mathbf{k}$ .

---

In order to further reduce the risk of trapping into local extrema, a widely used technique is the multi-scale ( $S$  scales) mechanism which is similar to all top-performing [1–8] methods. This efficient mechanism progressively estimates the motion blur kernel from coarse to fine. At the coarsest scale  $s = 1$ , we use the down-sampled version of the blurred image as an initialization of latent image, and delta function as the initialization of the blur kernel. At each scale  $s$ , the input of the motion kernel estimation can be up-sampled by bilinear interpolation from the final restored sharp edges obtained at the coarser scale  $s - 1$ . After all, once the kernel  $\mathbf{k}$  has been estimated, a non-blind deconvolution approach [17] is exploited to restore the final sharp image  $\mathbf{I}$ .



**Fig. 1.** Cumulative histograms of the SSD error ratio across test examples from benchmark [9].



**Fig. 2.** Blind motion deblurring with the real-world example (image size:  $640 \times 480$ , kernel size:  $21 \times 21$ ). From left to right, top to bottom: blurred image, deblurred images with estimated blur kernel by Cho *et al.* [2], Xu and Jia [3], Krishnan *et al.* [5], Xu *et al.* [6] and the proposed approach.

### 3 Experimental Results

In this section, we present some experimental results of our approach which requires a few parameters given below. As the kernel size is an important requirement for most algorithms, it also should be specified by users in our approach and does not have much influence on kernel estimation if its support area can accommodate the blur kernel. As for the initial parameters, the maximum iteration number is set to 10, and the regularization parameters  $\lambda_1, \lambda_2, \eta, \xi_1, \xi_2$  are uniformly provided as  $\lambda_1 = 0.03, \lambda_2 = 0.01, \eta = 0.001, \xi_1 = 0.03, \xi_2 = 0.01$  across most of the experiments. In order to achieve faster convergence, the decreasing factor  $\rho_1, \rho_2$  are fixed as  $\rho_1 = 0.55, \rho_2 = 0.8$  to progressively reduce the image regularization weights, leading to accommodating more edges into  $\mathbf{x}$ .

#### 3.1 Synthetic Data

In this subsection, we first test the proposed method on the benchmark image dataset of [9], which consists of 32 examples generated with 4 natural images and 8 different motion blur kernels. To quantitatively measure the quality of restored images, we use the Sum of Square Difference (SSD) ratio defined in [9] which quantifies the error between the estimated and the original images. The formula is given as

$$\text{SSD ratio} = \frac{\|\mathbf{x}_{est} - \mathbf{x}_{sharp}\|_2^2}{\|\mathbf{x}_{gt} - \mathbf{x}_{sharp}\|_2^2}, \quad (13)$$

where  $\mathbf{x}_{est}$  and  $\mathbf{x}_{gt}$  are respectively the restored image of estimated and ground-truth blur kernel adopting the same settings.



**Fig. 3.** Blind motion deblurring with the real-world example (image size:  $800 \times 800$ , kernel size:  $51 \times 51$ ). From left to right, top to bottom: blurred image, deblurred images with estimated blur kernel by Cho *et al.* [2], Xu and Jia [3], Krishnan *et al.* [5], Xu *et al.* [6] and the proposed approach.

For fair comparison, we use the non-blind deconvolution approach [17] to perform all the methods [2–5] to be compared. The cumulative histograms of the SSD error ratio are shown for different methods in Fig. 1. The  $r$ -th bin in the figure counts the percentage of the motion blurred images in the dataset achieving error ratio below  $r$ . For instance, the bar in Fig. 1 corresponding to bin 5 indicates the percentage of test images with SSD error ratios below 4. For each bin, the higher the bar the better the deblurring performance. As pointed out by Levin *et al.* [9], deblurred images are visually plausible in general as their SSD error ratios are below 3, and in this case the blind motion deblurring is considered to be successful. As can be seen from Fig. 1, the performance of the proposed method is better than the other methods quantitatively.

### 3.2 Real Data

To further demonstrate the effectiveness, we apply the proposed method to the real-world motion blur images and compare it with four state-of-the-art MAP methods [2, 3, 5, 6]. As reviewed in the introduction section, two of them [2, 3] are based on explicit filtering and edge selection, the rest two methods [5, 6] involve implicit regularization. The point to be noted is that our approach estimate the blur kernel using the gray version of the color image.

As shown in Figs. 2 and 3, two real-world examples with different image sizes and blurring levels are utilized to compare the performance of these methods. It is observed that our approach achieves visually more accurate motion blur kernels with stronger continuity and less singular points, as well as produces visually plausible restored images. Specifically, in Fig. 2, our approach generates reasonable blur kernels and visually clear deblurred image but the other four methods fail in kernel estimation in some sense and the restored images are accounted to be unnatural. In Fig. 3, our approach also perform well on the moderate blurring level. The corresponding estimated blur kernels of [3, 6] seem still less accurate. By comparison, our approach can successfully recover the true support of the blur kernel in the case of medium size blur kernel. And the visual quality is also comparable with the top-ranking method [6] which has even incorporated with some post-processing procedures.

## 4 Conclusion

We have presented an effective and efficient method for blind motion deblurring, providing a novel fused  $\ell_0$ - $\ell_1$  prior on sharp images. A fast numerical scheme based on split Augmented Lagrangian has also been developed to solve the deduced compound regularized problem. Extensive experiments both on a synthetic benchmark image dataset and two real-world motion blurred images have intensively validated the performance of the proposed method and testified the superiority over the state-of-the-art methods.

**Acknowledgements.** This work was supported in part by the Fundamental Research Funds for the Central Universities under Grant No. 30915012204, the Natural Science Foundation of China under Grant No. 61171165 and 11431015 and also sponsored by The Six Top Talents of Jiangsu Province under Grant No. 2012DZXX-036.

## References

1. Fergus, R., Singh, B., Hertzmann, A., Roweis, S.T., Freeman, W.T.: Removing camera shake from a single photograph. *ACM Trans. Graph.* **25**(3), 787–794 (2006)
2. Cho, S., Lee, S.: Fast motion deblurring. *ACM Trans. Graph.* **145**(8), 145:1–145:8 (2009)
3. Xu, L., Jia, J.: Two-phase kernel estimation for robust motion deblurring. In: Daniilidis, K., Maragos, P., Paragios, N. (eds.) *ECCV 2010, Part I*. LNCS, vol. 6311, pp. 157–170. Springer, Heidelberg (2010)
4. Shan, Q., Jia, J., Agarwala, A.: High-quality motion deblurring from a single image. *ACM Trans. Graph.* **27**(3), 73:1–73:10 (2008)
5. Krishnan, D., Tay, T., Fergus, R.: Blind deconvolution using a normalized sparsity measure. In: *CVPR*, pp. 233–240 (2011)
6. Xu, L., Zheng, S., Jia, J.: Unnatural L0 sparse representation for natural image deblurring. In: *CVPR*, pp. 1107–1114 (2013)
7. Pan, J., Su, Z.: Fast L0-regularized kernel estimation for robust motion deblurring. *IEEE Sig. Process. Lett.* **20**(9), 841–844 (2013)
8. Krishnan, D., Bruna, J., Fergus, R.: Blind deconvolution with re-weighted sparsity promotion (2013). [arXiv:1311.4029](https://arxiv.org/abs/1311.4029)
9. Levin, A., Weiss, Y., Durand, F., Freeman, W. T.: Efficient marginal likelihood optimization in blind deconvolution. In: *CVPR*, pp. 2657–2664 (2011)
10. Levin, A., Weiss, Y., Durand, F., Freeman, W.T.: Understanding blind deconvolution algorithms. *IEEE Trans. Pattern Anal. Mach. Intell.* **33**(12), 2354–2367 (2011)
11. Babacan, S.D., Molina, R., Do, M.N., Katsaggelos, A.K.: Bayesian blind deconvolution with general sparse image priors. In: Fitzgibbon, A., Lazebnik, S., Perona, P., Sato, Y., Schmid, C. (eds.) *ECCV 2012, Part VI*. LNCS, vol. 7577, pp. 341–355. Springer, Heidelberg (2012)
12. Wipf, D., Zhang, H.: Revisiting Bayesian blind deconvolution (2013). [arXiv:1305.2362](https://arxiv.org/abs/1305.2362)
13. Tibshirani, R., Saunders, M., Rosset, S., Zhu, J., Knight, K.: Sparsity and smoothness via the fused lasso. *J. Royal Stat. Soc. Ser. B* **67**(1), 91–108 (2005)
14. Wenze, S., Haisong, D., Zhihui, W.: The magic of split augmented Lagrangians applied to K-frame-based  $\ell_0$ - $\ell_2$  minimization image restoration. *Sig. Image Video Process.* **8**, 975–983 (2014)
15. Afonso, M., Bioucas-Dias, J., Figueiredo, M.: Fast image recovery using variable splitting and constrained optimization. *IEEE Trans. Image Process.* **19**(19), 2345–2356 (2010)
16. Yin, W., Osher, S., Goldfarb, D., Darbon, J.: Bregman iterative algorithms for L1 minimization with applications to compressed sensing. *SIAM J. Imag. Sci.* **1**, 143–168 (2008)
17. Krishnan, D., Fergus, R.: Fast image deconvolution using Hyper-Laplacian priors. In: *NIPS*, pp. 1033–1041 (2009)