This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPAMI.2016.2551244, IEEE Transactions on Pattern Analysis and Machine Intelligence

L_0-Regularized Intensity and Gradient Prior for Deblurring Text Images and Beyond

Jinshan Pan, Zhe Hu, Zhixun Su, and Ming-Hsuan Yang

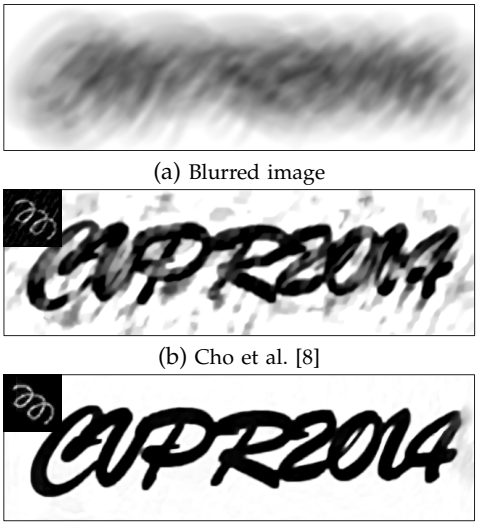
Abstract—We propose a simple yet effective L_0-regularized prior based on intensity and gradient for text image deblurring. The proposed image prior is based on distinctive properties of text images, with which we develop an efficient optimization algorithm to generate reliable intermediate results for kernel estimation. The proposed algorithm does not require any heuristic edge selection methods which are critical to the state-of-the-art edge-based deblurring methods. We discuss the relationship with other edge-based deblurring methods and present how to select salient edges more principally. For the final latent image restoration step, we present an effective method to remove artifacts for better deblurred results. We show the proposed algorithm can be extended to deblur natural images with complex scenes and low illumination, as well as non-uniform deblurring. Experimental results demonstrate that the proposed algorithm performs favorably against the state-of-the-art image deblurring methods.

Index Terms—Image deblurring, L_0-regularized prior, text images, low-illumination images, natural images

1 Introduction

The recent years have witnessed significant advances in single image deblurring. Much success of the state-of-the-art algorithms [1], [2], [3], [4], [5], [6], [7] can be attributed to the use of statistical priors on natural images and selection of salient edges for kernel estimation. Although numerous methods [1], [2], [5], [6], [7] have been proposed for deblurring generic images, these priors are less effective for cases with rich text that do not follow the heavy-tailed gradient statistics of natural images and can be better modeled by two-tone distributions.

Text image deblurring has attracted considerable attention due to its wide range of applications. In [9], Chen et al. propose a content-aware prior based on an intensity density function of documents and foreground segmentation rather than the heavy-tailed gradient prior of natural images. However, this method is developed specifically for document images (i.e., binary text images) and is unlikely to perform well for complex and cluttered images containing text. A direct method that exploits sparse characteristics of natural scenes is proposed for deblurring natural and document images [10]. Nevertheless, the blur kernel is not explicitly estimated from an input image and the computational load for learning an over-complete dictionary is significant. Li and Lii [11] propose an optimization method to estimate a blur kernel until the latent image is two-tone. However, this method is only applied to two-tone images and is less effective for text images with complex backgrounds. In [8] Cho et al. develop a method to incorporate text-specific properties (i.e., sharp contrast between text and background, uniform gradient within text, and background gradients based on natural image statistics) for deblurring. While this algorithm achieves the state-of-the-art deblurring results, the kernel estimation process is complicated and the performance depends on whether the stroke width transform (SWT) [12] separates an image into text and non-text regions well or not. However, the SWT is unlikely to perform well when the characters in a text image are small or clustered. Figure 1 shows one example where blurred characters are clustered due to large camera motion and the deblurred result from the algorithm [8].



(a) Blurred image

(b) Cho et al. [8]

(c) Ours

Fig. 1. A challenging blurred text image.

In this paper, we propose an effective algorithm to

- *J. Pan, Z. Hu, and M.-H. Yang are with School of Engineering, University of California, Merced, CA, 95344. e-mail: {jpan24, zhu, mhyang}@ucmerced.edu.*
- *Z. Su is with School of Mathematical Sciences, Dalian University of Technology, Dalian, 116024, P. R. China. e-mail: zxsu@dlut.edu.cn.*

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPAMI.2016.2551244, IEEE Transactions on Pattern Analysis and Machine Intelligence

deblur text images. The contributions of this paper are summarized as follows:

1) We propose an L_0-regularized intensity and gradient prior based on distinctive properties of text images for text image deblurring.
2) We present an efficient optimization algorithm based on the half-quadratic splitting technique. This approach guarantees that each sub-problem has a closed-form solution and ensures fast convergence.
3) We discuss the relationship with other methods in terms of salient edge selection, and show that the proposed algorithm generates reliable intermediate results for kernel estimation without ad-hoc selection processes. Compared with the state-of-the-art methods [8], [9], the proposed algorithm is efficient and effective as it requires no additional operations (e.g., adaptive segmentation [9], smoothing intermediate latent images [3], or SWT [8]).
4) For the latent image restoration step, we present an effective method to deal with artifacts and evaluate it against other alternatives.
5) We show that the proposed algorithm can effectively process natural blurry images including low-illumination inputs which are not well handled by most state-of-the-art deblurring methods. In addition, the proposed algorithm can be effectively applied to non-uniform image deblurring.

2 Related Work

Image deblurring has been studied extensively and numerous algorithms have been proposed, which can be categorized in three main approaches based on variational Bayesian inference, Maximum a Posteriori (MAP) estimation, and edge prediction.

In [1] Fergus et al. present an algorithm using a mixture of Gaussians to learn an image gradient prior via variational Bayesian inference. Levin et al. [13] analyze the method based on variational Bayesian inference [1] and show that it is able to avoid trivial solutions while naive MAP based methods may not. Since the optimization process of variational Bayesian inference is computationally expensive, methods based on MAP formulations have been developed with different likelihood functions and image priors [2], [6], [7], [10], [14], [15], [16]. In addition, methods that explicitly select sharp edges for kernel estimation been proposed [3], [4], [17] with demonstrated success on benchmark datasets [18]. However, the edge selection step is often based on heuristics and the assumption that there exist strong edges in the latent images may not always hold. To better reconstruct sharp edges for kernel estimation, exemplar based methods [19], [20], [21] are recently presented to exploit information contained both in an blurred input and example images of an external dataset.

Blurred images acquired from moving cameras (e.g., rotational and translational movements) can be better modeled by non-uniform blur models. Shan et al. [22] solve the in-plane rotation deblurring problem based on transparency maps. Tai et al. [23] propose a general projective motion model for non-uniform image deblurring where a blurred image is considered as an integration of a latent image under a sequence of projective transformations that describe camera path. Whyte et al. [24] simplify this model and propose a variational Bayesian approach in a way similar to [1] for non-uniform image deblurring. In [25], a similar model is proposed in which motion density functions are used to represent camera motion trajectories. To make the non-uniform methods computationally efficient, methods based on a locally uniform blur models [26], [27] are developed where the deconvolution step can be computed by Fast Fourier Transforms (FFTs). In addition to camera motion, blurred images caused by different object motions are analyzed [28], [29], [30]. The methods based on the depth variation of scenes are also proposed by [31], [32]. We note most of the aforementioned deblurring methods are developed for generic scenes and few of them exploit properties of text images.

Since the properties of text images are different from natural images, Chen et al. [9] propose a prior based on the image intensity rather than the heavy-tailed gradient prior of natural scenes to characterize text images. Cho et al. [8] consider image properties specific to text images in a way similar to [9] and present a deblurring method on detected regions. However, the above method has limited application domains as it entails text detection [12] and heuristic filtering [3]. In contrast, the proposed algorithm does not require additional filtering or text segmentation in the deblurring process. Although the proposed prior is based on text images, we show that it is able to describe the convolution as well as blur process in generic scenes and applicable to deblur natural and low-illumination images.

3 L_0-Regularized Image Prior

In this section, we present an L_0-regularized prior of intensity and gradient for text image deblurring.

3.1 L_0-Regularized Intensity and Gradient Prior

The proposed intensity and gradient prior is based on the observation that text and background regions usually have nearly uniform intensity values in clear images without blurs. Figure 2(b) illustrates that the pixel intensity distribution of a clear text image (Figure 2(a)) is peaked at two values (near 0 and 255). For a blurred text image, the pixel intensity distribution is significantly different from that of a clear image. Figure 2(e) shows the histogram of pixel intensities (from a blurred image in (d)) with fewer pixels of value 0 and 255. The reason is that each pixel in a

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPAMI.2016.2551244, IEEE Transactions on Pattern Analysis and Machine Intelligence

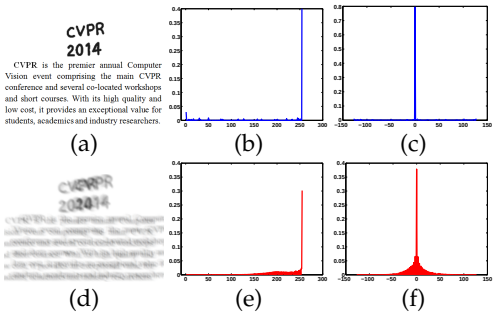


Fig. 2. Intensity and gradient properties of text images. (a) A clear text image. (b) Pixel intensity distribution from (a). (c) Distribution of horizontal gradient from (a). (d) A blurred image. (e) Pixel intensity distribution from (d). (f) Distribution of horizontal gradient from (d).

blurred image can be viewed as the weighted sum of a few neighbors of a clear image. Thus, the intensity distribution is squeezed from both ends of the intensity range. As a result, there are fewer pure black pixels (intensity value 0) in a blurred text image than a clear one. This intensity property does hold for generic images, and appears more obviously in text images (e.g., document images). For an image x, we describe the property with a regularization term in the proposed model,

$$P_t(x) = \|x\|_0, \tag{1}$$

where $\|x\|_0$ counts the number of nonzero-intensity pixels in x. With this intensity property, clear and blurred images can be differentiated. We note that, for an image with more white pixels, we can also reverse the pixel intensity by $1 - x \to x$ (for both latent and blurred images) and use the same blur model.

Gradient priors are widely used for image deblurring as they have been shown to be effective in suppressing artifacts [2], [3]. As the intensity values of a clear text image are close to two-tone, the pixel gradients are likely to have a few nonzero values. Figure 2(c) and (f) show the horizontal gradient histograms of a clear text image and the corresponding blurred one. It is clear that the nonzero values of blurred image gradients are denser than those of the clear one. Thus we use a similar L_0-regularized prior, $P_t(\nabla x)$, to model image gradients.

With the aforementioned regularized priors on intensity and gradient, the prior for text image deblurring is defined by

$$P(x) = \sigma P_t(x) + P_t(\nabla x), \tag{2}$$

where σ is a weight to balance two priors. Although $P(x)$ is developed based on the assumption that background regions of a text image are uniform, we show this prior can also be applied to deblur complex scenes effectively.

3.2 Text Image Deblurring via Proposed Prior

A blurred image y can be formulated as the result of a convolution process with a spatially invariant kernel or point spread function,

$$y = x * k + e, \tag{3}$$

where x and e denote the latent image and noise; k is a blur kernel; and $*$ is the convolution operator. Given a blurred image y, we estimate the latent image x and blur kernel k with a regularized formulation based on the proposed prior $P(x)$,

$$\min_{x,k} \|x * k - y\|_2^2 + \gamma\|k\|_2^2 + \lambda P(x), \tag{4}$$

where the first term is concerned with image data, and the remaining two terms are constraints for the blur kernel and the latent image, with respective weights, γ as well as λ. We note that we introduce the prior for uniform deblurring first, and extend that to non-uniform deblurring in Section 7.

4 Deblurring Text Images

The deblurring process is modeled as the optimization problem by alternatively solving the latent image x

$$\min_x \|x * k - y\|_2^2 + \lambda P(x), \tag{5}$$

and the blur kernel k,

$$\min_k \|x * k - y\|_2^2 + \gamma\|k\|_2^2. \tag{6}$$

The details of the two sub-problems are described in the following sections.

4.1 Estimating Latent Image x

Due to the L_0 regularization term in (5), minimizing (5) is computationally intractable. Based on the half-quadratic splitting L_0 minimization approach [33], we propose an efficient alternating minimization method to solve this problem. We introduce the auxiliary variables u with respect to x and $g = (g_h, g_v)$ corresponding to image gradients in horizontal and vertical directions. The objective function can be rewritten as

$$\begin{aligned}\min_{x,u,g} \|x * k - y\|_2^2 + \beta\|x - u\|_2^2 \\ + \mu\|\nabla x - g\|_2^2 + \lambda(\sigma\|u\|_0 + \|g\|_0),\end{aligned} \tag{7}$$

where σ is the weight defined in (2), β and μ are penalty parameters. When β and μ are close to infinity, the solution of (7) approaches that of (5) [34]. With this formulation, (7) can be efficiently solved through alternatively minimizing x, u, and g independently by fixing the other variables.

The values of u and g are initialized to be zeros. In each iteration, the solution of x is obtained by solving

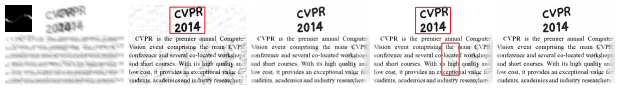
$$\min_x \|x * k - y\|_2^2 + \beta\|x - u\|_2^2 + \mu\|\nabla x - g\|_2^2, \tag{8}$$

and the closed-form solution for this least squares minimization problem is

$$x = \mathcal{F}^{-1}\left(\frac{\overline{\mathcal{F}(k)}\mathcal{F}(y) + \beta\mathcal{F}(u) + \mu F_G}{\overline{\mathcal{F}(k)}\mathcal{F}(k) + \beta + \mu(\sum_{i\in\{h,v\}}\overline{\mathcal{F}(\nabla_i)}\mathcal{F}(\nabla_i))}\right), \tag{9}$$

where $\mathcal{F}(\cdot)$ and $\mathcal{F}^{-1}(\cdot)$ denote the Fourier transform and its inverse transform, respectively; the $\overline{\mathcal{F}(\cdot)}$ is the

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPAMI.2016.2551244, IEEE Transactions on Pattern Analysis and Machine Intelligence



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

Fig. 3. Effectiveness of Algorithm 1. (a) Blurred image and kernel. (b)-(c) Results by only $P_t(x)$ and $P_t(\nabla x)$. (d)-(e) Results by posing the estimation of g in the outer and inner loop. The smooth regions enclosed in the red boxes in (b) and (d) are not preserved well, while some characters in (c) are over smoothed.

complex conjugate operator; and $F_G = \overline{\mathcal{F}(\nabla_h)}\mathcal{F}(g_h) + \overline{\mathcal{F}(\nabla_v)}\mathcal{F}(g_v)$ where ∇_h and ∇_v denote the horizontal and vertical differential operators.

Given x, we compute u and g separately by

$$\min_u \beta\|x-u\|_2^2 + \lambda\sigma\|u\|_0, \tag{10a}$$

$$\min_g \mu\|\nabla x - g\|_2^2 + \lambda\|g\|_0. \tag{10b}$$

We note that (10) is a pixel-wise minimization problem. Thus, the solutions of u and g are obtained based on [33],

$$u = \begin{cases} x, & |x|^2 \geqslant \frac{\lambda\sigma}{\beta}, \\ 0, & \text{otherwise}, \end{cases} \tag{11}$$

and

$$g = \begin{cases} \nabla x, & |\nabla x|^2 \geqslant \frac{\lambda}{\mu}, \\ 0, & \text{otherwise}. \end{cases} \tag{12}$$

The main steps for solving (7) are summarized in Algorithm 1. We pose the subproblem u in the outer loop for the following reasons. If we pose the subproblem for g (10b) in the outer loop, the algorithm in the inner loop is equivalent to solve the minimization problem

$$\min_x \|x*k-y\|_2^2 + \lambda\sigma\|x\|_0 + \mu\|\nabla x - g\|_2^2, \tag{13}$$

by introducing auxiliary variable u. Similarly, posing the sub-problem u (10a) in the outer loop indicates solving

$$\min_x \|x*k-y\|_2^2 + \beta\|x-u\|_2^2 + \lambda\|\nabla x\|_0, \tag{14}$$

by introducing auxiliary variable g in the inner loop. As the intensity prior is based on independent pixels instead of disparities of neighboring pixels (i.e., gradients), it introduces significant noise and artifacts in image restoration (See Figure 3(b)). In contrast, the gradient prior is based on disparities of neighboring pixels, which enforces smooth results with fewer artifacts in the recovered image (Figure 3(c)). As the inner loop involves image restoration, we pose the sub-problem u in the outer loop to reduce artifacts generated by the intensity prior. Figure 3(e) demonstrates the advantages of posing the sub-problem u in the outer loop.

4.2 Estimating Blur Kernel k

Given x, (6) is a least squares minimization problem in which a closed-form solution can be computed by FFTs. As the estimation based on gradients has been shown to be more accurate [3], [5], [7], we estimate the blur kernel k by

Algorithm 1 Solving (7)

Input: Blurred image y and blur kernel k.
$x \leftarrow y$, $\beta \leftarrow 2\lambda\sigma$.
repeat
  solve for u using (11).
  $\mu \leftarrow 2\lambda$.
  repeat
    solve for g using (12).
    solve for x using (9).
    $\mu \leftarrow 2\mu$.
  until $\mu > \mu_{\max}$
  $\beta \leftarrow 2\beta$.
until $\beta > \beta_{\max}$
Output: Intermediate latent image x.

Algorithm 2 Blur kernel estimation algorithm

Input: Blurred image y.
initialize k with the results from the coarser level.
for $i = 1 \rightarrow 5$ **do**
  solve for x using Algorithm 1.
  solve for k using (15).
  $\lambda \leftarrow \max\{\lambda/1.1, 1e^{-4}\}$.
end for
Output: Blur kernel k and intermediate latent image x.

$$\min_k \|\nabla x * k - \nabla y\|_2^2 + \gamma\|k\|_2^2, \tag{15}$$

and the solution can be efficiently computed by FFTs [3]. After obtaining k, we set the negative elements to 0, and normalize it so that the sum of its elements is 1. Similar to the state-of-the-art methods, the proposed kernel estimation process is carried out in a coarse-to-fine manner using an image pyramid [3]. Algorithm 2 shows the main steps for kernel estimation algorithm on one pyramid level. The step in (12) performs similar to the edge selection methods [3], [4]. As suggested by [3], [4], we decrease λ gradually to include more informative gradients for kernel estimation.

4.3 Removing Artifacts

Although the latent images can be estimated from (5) as shown in Figure 4(c), this formulation is less effective for scenes with complex backgrounds or fine texture details. We note that the non-blind deblurring method with a hyper-Laplacian prior [35] has been shown to preserve fine details. However, significant ringing artifacts caused by deconvolution are likely to appear using this prior as shown in Figure 4(b). In contrast, the method with the proposed L_0-regularized prior generates fewer fine details and ringing artifacts as shown in Figure 4(c).

Based on the properties of the results generated by these two aforementioned priors, we propose a non-blind deconvolution method that preserves fine details. This method uses the residue of two recovered images similar to the ringing suppression method [2]. First, we estimate the latent image I_l (See Figure 4(b))

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPAMI.2016.2551244, IEEE Transactions on Pattern Analysis and Machine Intelligence

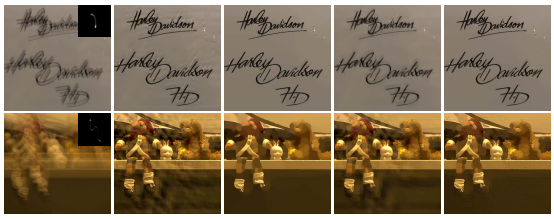


Fig. 4. Non-blind deconvolution examples. (a) Blurred images and the estimated kernels. (b) Results by [35] with Laplacian prior. (c) Results by setting $\sigma = 0$ in (5). (d) Ringing suppression results by [2]. (e) Our results.

using the method with Laplacian prior [35]. Second, we estimate the latent image I_0 (See Figure 4(c)) using the proposed algorithm via (5) with only the gradient regularization $P_t(\nabla x)$ at this stage (i.e., setting $\sigma = 0$). Similar to [2], we compute a difference map between these two estimated images and apply bilateral filtering to it. The filtered residue can be viewed as the artifacts that non-blind deconvolution methods do not handle well. We then subtract the filtered difference map from I_l to remove these artifacts. The results in Figure 4(e) show that this approach is effective for text and natural images, and performs favorably against the ringing suppression method [2].

5 Analysis of Algorithm

In this section, we analyze how the proposed algorithm performs on text image deblurring. We also demonstrate the importance of intensity prior for text image deblurring and discuss its relationship with other methods in terms of the edge selection. Furthermore, we show that the proposed algorithm can be applied to deblur natural images.

5.1 Effectiveness of the L_0-Regularized Prior

The solution for u using (11) leads to results containing segments with intensity value of 0. The appearance of those segments enhances the contrast of u, and therefore drives the solution of x in (9) to have salient edges around the segment boundaries (See Figure 5(f)).

The text image deblurring method [8] involves a hard-thresholding step ((7) in [8]) with values determined by SWT [12] to segment texts. In addition, this text image deblurring algorithm uses the sparse gradient minimization method [33] to remove ringing artifacts from the intermediate latent images, which increases the computational load significantly. Figure 5 shows one example where the method [8] does not perform well. The reason is that SWT is not effective in detecting text when the blurred characters are cluttered as illustrated in the intermediate results of Figure 5(e).

The success of recent deblurring methods hinges on latent image estimation explicitly [3], [4] or implicitly [1], [2], [6]. The proposed method is distinguished from existing methods as it does not involve ad-hoc

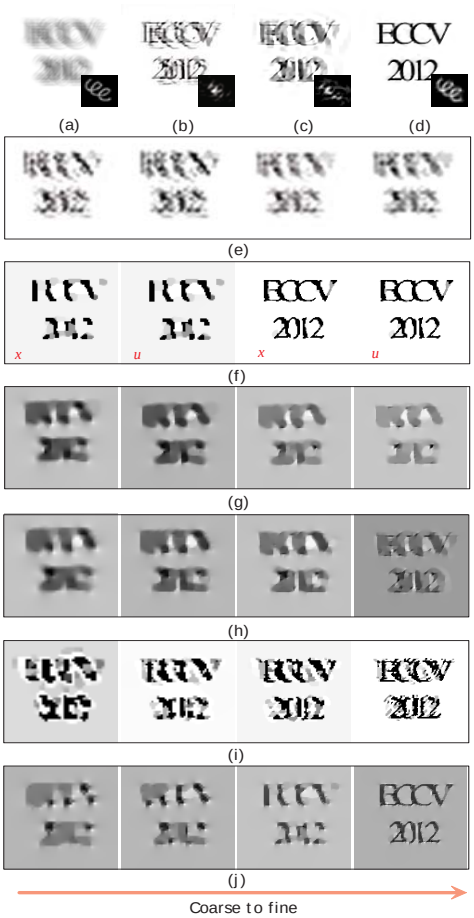


Fig. 5. An example presented in [8]. (a) Blurred image and kernel; (b) Results of [8]; (c) Our results without using $P_t(x)$ in the kernel estimation; (d) Our final results; (e) Intermediate results of [8]; (f) Our intermediate results (including x and u); (g) Intermediate salient edges of [7]; (h) Intermediate salient edges using only $P_t(\nabla x)$; (i) Intermediate results using only $P_t(x)$; (j) Our intermediate salient edges, i.e., g in (12).

edge selection (e.g., spatial filtering [8], [3], [4], or edge re-weighting [2], [6]) for kernel estimation. Instead of finding one good threshold to remove subtle image structures such as filter-based edge selection methods [3], [4], the proposed algorithm computes intermediate latent images iteratively by solving the optimization problems in a way similar to [7]. By using (11) and (12) of the proposed algorithm, pixels with small intensity values or tiny structures can be removed while salient edges are retained. Furthermore, our method exploits the gradient prior with $P_t(\nabla x)$. If σ of (7) is set to 0, then the proposed algorithm is reduced to the recent methods based on L_0 gradient priors [7], [36] which achieve the state-of-the-art results for deblurring natural images. On the other hand, these L_0-based methods [7], [36] (L0Deblur for short) do not perform well for text images. Figure 5(g) shows intermediate salient edges extracted by [7]. As no sharp edges are extracted, the blur kernel is not estimated well by this method.

We note that image deblurring using only intensity

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPAMI.2016.2551244, IEEE Transactions on Pattern Analysis and Machine Intelligence

prior is less effective (See Figure 5(i)) as $P_t(x)$ is based on independent pixels without considering image gradients. On the other hand, image deblurring with only gradient prior $P_t(\nabla x)$ is not effective (See Figure 5(h)) as no salient edges are extracted.

5.2 Convergence of the Proposed Algorithm

The proposed kernel estimation algorithm is mainly based on the alternating minimization method which ensures that each sub-problem has a closed-form solution. Thus, the proposed algorithm has fast convergence property.

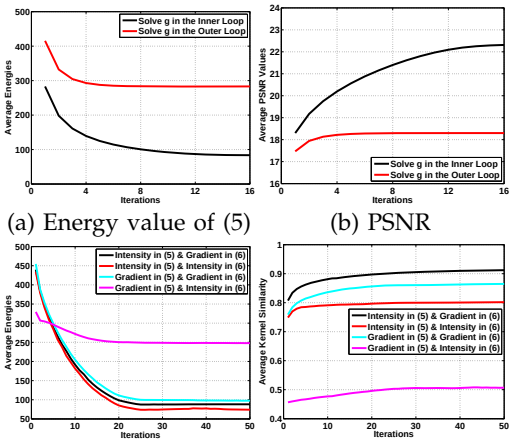
We evaluate the convergence rate of the proposed method using 96 blurred text images from the proposed dataset. The average energy of the objective function (5) decreases with respect to the number of iterations as shown in Figure 6(a). The results demonstrate that the proposed algorithm (Algorithm 1) exhibits good convergence. We also measure the quality of the recovered images in terms of PSNR. Figure 6(b) demonstrates fast convergence of Algorithm 1 in terms of PSNR. Furthermore, we analyze how the proposed method converges by posing g estimation as the outer loop and the inner loop in Algorithm 1. The plots in Figure 6(a) and (b) show that the approach with estimating g in the inner loop converges to better solutions with lower energy values and higher PSNR values of the recovered images. On the other hand, the approach with estimating g in the outer loop does not perform well. The difference between these two approaches is in line with our analysis in Section 4.1.

We evaluate the convergence and performance of Algorithm 2 with respect to the domain of the data term. The proposed kernel estimation is carried out by alternatively solving (5) and (6), where the first term can be based on either image intensity or gradient. The plots in Figure 6(c)[1] show that similar convergence is achieved for both approaches. However, the combination of using image intensity for the first term of (5) and image gradient for that of (6) performs better than the others as shown in Figure 6(d).

5.3 Deblurring Saturated Images

Estimating blur kernels from blurred images with saturated regions is a known difficult problem. Although a few non-blind deblurring methods [37], [38], [39] have been proposed to deal with such images, it remains challenging to develop effective blind deblurring algorithms. Saturated regions usually appear sparsely in clear images and these areas are much larger (e.g., blobs or streaks) after blurring. Figure 7(b) shows two examples of saturated blurred images from Figure 7(a). We use the binary images to display the saturated regions in Figure 7(c) and (d), where there are more non-zero elements in the binary images of the blurred images than those of the clear binary images. As the L_0-norm in $P_t(x)$ minimizes the number of non-zero coefficients, the proposed deblurring algorithm favors solutions with few blobs or streaks in the clear images, which leads to non-trivial solutions. We present results from challenging images in Section 8.2.

1. We use the image intensity to compute the energy of (4), if the data terms of (5) and (6) are in different domains.



(a) Energy value of (5) (b) PSNR

(c) Energy value of (4) (d) Kernel similarity

Fig. 6. Convergence properties of Algorithms 1 and 2. (a)-(b) show the convergence of Algorithm 1. (c)-(d) show the convergence of Algorithm 2 and the corresponding energies and kernel similarity values are computed from the finest level of Algorithm 2. The legend Intensity in (5) & Gradient in (6) in (c) and (d) indicates that the first terms in (5) and (6) use image intensity and image gradient, respectively.

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

Fig. 7. Saturated images. (a) Clear images with saturated regions and ground-truth kernel. (b) Blurred images with saturated regions. (c) Binary images of (a). (d) Binary images of (b). (c) and (d) are obtained from (a) and (b) with the same threshold value.

6 Deblurring Natural Images

We analyze how the proposed algorithm performs on text images in complex backgrounds and present a method to deal with generic blurred images.

The success of the state-of-the-art generic image deblurring methods stems mainly from restoration of intermediate images with salient edges for kernel estimation. The restored intermediate images do not necessarily resemble the sharp natural images in terms of rich textures and contain structures of large gradients [3], [4], [7] or matte maps [40]. In [40], [41], Jia shows that the matte map of an image helps estimate the blur kernel for natural image deblurring. In image matting, the matte map segments an image into background, foreground, and ambiguous regions. The values are zeros for background pixels, ones for

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPAMI.2016.2551244, IEEE Transactions on Pattern Analysis and Machine Intelligence

foreground pixels, and float numbers between zero and one for ambiguous pixels [42]. In the deblurring process, the intermediate result u can be viewed as the matte map of multiple layers (where an object is represented by one layer). The modeled blur process generates ambiguous regions around the boundaries between different objects due to the mixture of pixels belonging to different objects. If the intermediate result contains sharp edges or strong contrast around the object boundaries, it helps the kernel estimation as the intermediate result approximates the latent image in terms of clear boundaries. In the proposed method, the solution of u in (11) is determined by thresholding x with the value of $\frac{\lambda\sigma}{\beta}$. This process results in many zero-intensity pixels in u and enforces higher contrast in the generated result. Similarly, the sub-problem of the gradient prior in (10) maintains only large gradient values, but removes small gradients. Thus, the effect of the L_0-regularized gradient prior is able to retain salient edges in the optimization process. With two components in (8), the intermediate result x is likely to maintain the properties of u and g, which leads to sharper object boundaries and improves kernel estimation. In contrast, the intermediate latent image estimation step from [7] relies only on g, which does not help preserve sharp object boundaries in text image deblurring (See Figure 5(g)). This is the main reason that our method performs better than the method that uses only the image gradient prior.

As discussed above, the sub-problem on the L_0-regularized intensity prior enforces the intermediate image x to contain many pixels of zero values. This step is based on the varying parameter $\frac{\lambda\sigma}{\beta}$ and its initial value is critical for the success of the method. To handle text or natural images with numerous zero-intensity values, the initial value can be easily determined as there exist pixels of low intensity values that are close to zero in the blurred image as well as in the clear image. We note that the document image is nearly two-tone and the value 0.5 usually separates the intensity histogram into two components (text and background) [43]. Since the initial value of β is set to be $2\lambda\sigma$ in Algorithm 1, the initial value of $\frac{\lambda\sigma}{\beta}$ is 0.5. However this approach does not work well for natural images as they contain more complex intensity histograms than those of text images. To address this issue, we propose to find a threshold to segment the intensity histogram of the natural images into two parts (i.e., foreground and background). We determine the initial value of $\frac{\lambda\sigma}{\beta}$ according to the intensity histogram of a blurred image with the threshold selection method by Otsu [44] where it determines the optimal value for two modes of image pixels for separation. We denote the value determined by Otsu algorithm as o, and then set the initial value of β to be $\frac{\lambda\sigma}{o^2}$ in Algorithm 1.

Figure 8 shows an example where the proposed algorithm with adaptive initial threshold value is able to generate reliable intermediate results for kernel estimation, while the algorithm [43] with a pre-defined value fails in this natural image (See Figure 8(c) and (f)). This is mainly because that the initial value of $\frac{\lambda\sigma}{\beta}$ in [43] is 0.5, which keeps the pixels whose intensity values are larger than 0.5 in the computation of u (See (11)). According to the aforementioned analysis, this will improperly segment x (See the white areas in Figure 8(f)), which accordingly alters the structures of intermediate results. Compared to the method [7], our intermediate latent image restoration step with the intensity prior maintains salient structure of the object boundaries. The proposed prior with adaptive threshold value helps enhance the contrast and preserve more salient edges in the intermediate latent image (See the comparisons in Figure 8(e) and (g)).

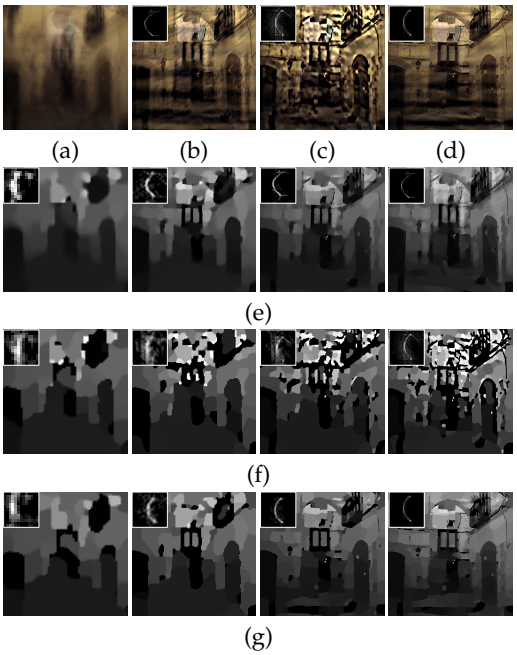


Fig. 8. Intermediate results with different priors in natural image deblurring. (a) Blurred image. (b) Intermediate result with $P_t(\nabla x)$. (c) Deblurred result using [43] with fixed initial threshold 0.5 for $\frac{\lambda\sigma}{\beta}$. (d) Deblurred result using the proposed algorithm with the adaptive initial threshold for $\frac{\lambda\sigma}{\beta}$, estimated by Otsu algorithm [44]. (e) Intermediate results with $P_t(\nabla x)$. (f) Intermediate results [43]. (g) Our intermediate results.

We further consider a modification of the proposed algorithm as there may exist multiple peaks in the intensity histogram of a natural image. As the height of the highest peak in an intensity histogram of a natural or text image is reduced after the blur process, we modify the intensity prior to reflect this property,

$$P_c(x) = \sigma\|x - c\|_0 + \|\nabla x\|_0, \tag{16}$$

where c is the intensity value corresponding to the highest peak of an intensity histogram. The modification, $x - c$, assumes that the solution x contains more pixels whose intensity values are c in the sharp image rather than the blurred one due to the shift of c.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPAMI.2016.2551244, IEEE Transactions on Pattern Analysis and Machine Intelligence

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

Fig. 9. Effect of different priors used in intermediate latent image restoration. (a) Ground-truth clear image and the blur kernel. (b) Blurred image. (c)-(d) Non-blind deconvolution results with prior $P_c(x)$ and $P(x)$. The red boxes in (c) enclose some artificial structures.

Intuitively, the solution using this prior $P_c(x)$ favors a clear image. However, this process does not necessarily facilitate the kernel estimation process. The term $\|x - c\|_0$ used in the intermediate latent image estimation enforces the pixel values approximate to c and thus reduces the contrast in the restored result. Furthermore, it is likely to result in segmentation effects, i.e., generating segments with constant values (See Figure 9(c)). In comparison, the term $\|x\|_0$ used in the proposed prior $P(x)$ of (2)) and its solution enforces the pixels of small intensities to be dark pixels of intensity value 0. Although the use of the prior $P(x)$ may generate unnatural results, it is able to enhance the contrast and preserve sharp edges in the restored result, which is critical in image deblurring. This is the main reason that we use (2) rather than (16) in kernel estimation.

7 Non-Uniform Deblurring

Camera shake during the exposure time often leads to spatially varying blurring effect on the image. Based on the geometric model of camera motion [23], [24], we represent a blurred image as the weighted sum of a clear image under transformations,

$$\mathbf{y} = \sum_{l=1}^{t} w_l \mathbf{K}(\theta_l)\mathbf{x} + \mathbf{e}, \tag{17}$$

where a blurred image $\mathbf{y}$, a latent image $\mathbf{x}$, noise $\mathbf{e}$ are in vector forms, w_l is the weight corresponding to the camera pose θ_l that satisfies $w_l \geq 0$, $\sum_l w_l = 1$, and t is the number of sampled camera poses. In (17), $\mathbf{K}(\theta_l)$ is a matrix derived from the homography matrix $\mathbf{H}(\theta_l)$ that warps the latent image $\mathbf{x}$. As in [24], the homography matrix $\mathbf{H}(\theta_l)$ is defined by

$$\mathbf{H}(\theta_l) = \mathbf{T}\mathbf{R}(\theta_l)\mathbf{T}^{-1}, \tag{18}$$

where $\mathbf{T}$ is the intrinsic matrix of the camera and $\mathbf{R}(\theta_l)$ is the rotation matrix describing the motion of the camera.

Similar to [24], we use the bilinear interpolation when applying $\mathbf{K}(\theta_l)$ to a latent image $\mathbf{x}$. Thus, (17) can be rewritten as

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{e} = \mathbf{B}\mathbf{w} + \mathbf{e}, \tag{19}$$

where $\mathbf{A} = \sum_l w_l\ \mathbf{K}(\theta_l)$, $\mathbf{B} = [\mathbf{K}(\theta_1)\mathbf{x},\ \mathbf{K}(\theta_2)\mathbf{x},\ \cdots,\ \mathbf{K}(\theta_t)\mathbf{x}]$, and $\mathbf{w} = [w_1, w_2, \cdots, w_t]^\top$. Based on (19), the non-uniform deblurring process is carried out by alternatively minimizing

$$\min_{\mathbf{x}} \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2^2 + \lambda P(\mathbf{x}), \tag{20}$$

and

$$\min_{\mathbf{w}} \|\mathbf{B}\mathbf{w} - \mathbf{y}\|_2^2 + \gamma\|\mathbf{w}\|_2^2. \tag{21}$$

Similar to the optimization process of (5), we introduce the same auxiliary variables and rewrite the objective function as

$$\begin{aligned}\min_{\mathbf{x},\mathbf{u},\mathbf{g}} &\|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2^2 + \beta\|\mathbf{x} - \mathbf{u}\|_2^2 \\ &+ \mu\|\nabla\mathbf{x} - \mathbf{g}\|_2^2 + \lambda(\sigma\|\mathbf{u}\|_0 + \|\mathbf{g}\|_0),\end{aligned} \tag{22}$$

where $\mathbf{u}$ and $\mathbf{g}$ are vector forms of u and g defined in (5), β and μ are the same to those in (5). We note that $\|\cdot\|_0$ is defined on each pixel. Thus, the solution of minimization problems with respect to $\mathbf{u}$ and $\mathbf{g}$ can be still obtained by (11) and (12).

The minimization problem,

$$\min_{\mathbf{x}} \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2^2 + \beta\|\mathbf{x} - \mathbf{u}\|_2^2 + \mu\|\nabla\mathbf{x} - \mathbf{g}\|_2^2, \tag{23}$$

cannot be solved directly using FFTs. Since the blur kernels at a small region are similar, we use the locally uniform blur model to approximate the non-uniform blur as the fast approximation method [26]. It divides the image into Q patches and the matrix $\mathbf{A}$ can be represented by

$$\mathbf{A} \approx \sum_{r=1}^{Q} \mathrm{diag}(\mathcal{M}_r)\mathbf{A}_r, \tag{24}$$

where $\mathrm{diag}(v)$ is a diagonal matrix with the element of vector v on the main diagonal, $\mathcal{M}_r$ is a window function that are zero-padded near the border which helps blend overlaid patches, and $\mathbf{A_r}$ is the matrix corresponding to the blur kernel a_r for the r-th patch. To compute $\mathbf{A}$ using FFTs, $\mathbf{A}$ can be expressed as

$$\mathbf{A} \approx \mathbf{Z}_\mathbf{y}^{-1} \sum_{r=1}^{Q} \mathbf{C}_r^{-1}(\mathcal{F}^{-1}(\mathrm{diag}(\mathcal{F}(\mathbf{Z}_a a_r)))\mathcal{F}(\mathbf{C}_r \mathrm{diag}(\mathcal{M}_r))), \tag{25}$$

where $\mathbf{Z_y}$ and $\mathbf{Z}_a$ are the zero-padding matrix that prepends zeros to a vector such that its size matches the size of the vector resulting from the summation, $\mathbf{C}_r(\cdot)$ is a matrix that chops r-th patch from a vector, and $\mathbf{C}_r^{-1}(\cdot)$ is a matrix that paste the r-th patch to the original vector.

By using the approximation (25) of $\mathbf{A}$, the solution of (23) can be obtained by

$$\mathbf{x} = \mathbf{W} \sum_{r=1}^{Q} \mathbf{C}_r^{-1} \left(\mathcal{F}^{-1} \left(\frac{\Delta_n}{\Delta_d} \right) \right), \tag{26}$$

where $\mathbf{W}$ is a weight to suppress visual artifacts caused by the window functions [26], $\Delta_d = \overline{\mathcal{F}(\mathbf{Z}_a a_r)}\mathcal{F}(\mathbf{Z}_a a_r) + \beta + \mu\overline{\mathcal{F}(\mathbf{C}_r(\nabla))}\mathcal{F}(\mathbf{C}_r(\nabla))$, and $\Delta_n = \overline{\mathcal{F}(\mathbf{Z}_a a_r)}\mathcal{F}(\mathbf{C}_r(\mathrm{diag}(\mathcal{M}_r)\mathbf{y})) + \beta\mathcal{F}(\mathbf{C}_r(\mathbf{u})) + \mu(\overline{\mathcal{F}(\mathbf{C}_r(\nabla_v))}\mathcal{F}(\mathbf{C}_r(g_v)) + \overline{\mathcal{F}(\mathbf{C}_r(\nabla_h))}\mathcal{F}(\mathbf{C}_r(g_h)))$.

The algorithm for solving (22) is the same in Algorithm 1, where we only need to replace (9) with (26). For the optimization of the kernel estimation model (21), we use the same optimization

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPAMI.2016.2551244, IEEE Transactions on Pattern Analysis and Machine Intelligence

process proposed by [7]. The proposed non-uniform deblurring method is achieved by alternatively minimizing (22) and (21). We use the same settings as the uniform deblurring presented in Algorithm 2.

8 Experimental results

We present experimental evaluations of the proposed algorithm against the state-of-the-art deblurring methods for text and generic images. All the experiments are carried out on a desktop computer with an Intel Core i7-4790 processor and 24 GB RAM. The execution time for a 255×255 image is 15 seconds on MATLAB without code optimization. In all the experiments, we set $\lambda = 0.004$, $\gamma = 2$, and $\sigma = 1$, respectively. We empirically set $\beta_{\max} = 8$ and $\mu_{\max} = 10^5$ in Algorithm 1. In the final stage, the method presented in Section 4.3 is employed to estimate the latent image, and the non-blind deconvolution method [38] is used to recover a saturated image. More experimental results can be found at http://faculty.ucmerced.edu/mhyang/project/text-deblur/, and the MATLAB code as well as datasets will be released.

8.1 Text Images

Synthetic text images: We first use the example from [8] (See Figure 5) for comparisons. Table 1 shows the structural similarity (SSIM) [45] and kernel similarity [46] values of the recovered images and estimated kernels by the state-of-the-art deblurring methods for text images [8] and generic images [3], [4], [5], [7], [47]. We also compare the proposed deblurring method for text images in Section 4 (referred as L0RIG) and the deblurring method for natural images in Section 6 (referred as IL0RIG). Overall, the proposed algorithm performs well in terms of both metrics as well as visual quality.

In addition, we construct a dataset containing 15 ground truth document images and 8 kernels from [13] (See the project web page for images). For each sharp image, we compute the average PSNR on the blurred images generated by different blur kernel estimation methods [2], [3], [4], [6], [5], [7], [47] as shown in Figure 10. Although the IL0RIG method does not perform well for some examples in this text image dataset (e.g., im11), it performs better than existing methods for text or natural images as shown in the last column of Figure 10. We also note that the IL0RIG method usually generates sharper text images than the L0RIG approach despite the fact that they perform similarly in terms of PSNR values (more results can be found on the project web page).

Real text images: We evaluate the proposed algorithm and other methods using real text images. For fair comparisons with [8], we use an example from [8] and show the deblurred results in Figure 11. The natural image deblurring methods do not perform well on text images. The deblurred result of [9] contains ringing artifacts and some strokes are not recovered well. Although the state-of-the-art method by Cho et al. [8] performs well, the motion blur is not fully removed as shown in the red box in Figure 11(h). In addition, the deblurred result contains unnatural colors due to the SWT process. Compared with [8], the proposed algorithm generates sharper and visually pleasing deblurred results. We note that the L0Deblur method [7] does not estimate the blur kernel or deblur the image well which also demonstrates the importance of $P_t(x)$ in the proposed prior $P(x)$. Both L0RIG and IL0RIG methods perform well on this text image.

Images containing text and complex background: We present an example in Figure 12 where the image contains rich text and cluttered background regions. The state-of-the-art natural image deblurring methods [3], [4], [6], [7] are not effective in handling this text image with complex background. Although the text image deblurring method [8] handles this example well (as shown in Figure 12(f)), the estimated kernel retains certain amount of noise and the deblurred result contains some unnatural colors as a result of the SWT process. In contrast, the proposed algorithm generates the deblurred image (clear text, sharp edges, and natural color) and blur kernel well. Figure 12(g) and (h) show the results using the proposed algorithm with only $P_t(\nabla x)$ or $P_t(x)$. The results in (g) and (h) show that sharp images cannot be obtained by using only the gradient or intensity prior, which indicates that the proposed prior $P(x)$ plays a critical role in text image deblurring.

8.2 Low-Illumination Images

Most state-of-the-art deblurring methods are less effective in processing blurred images with saturated regions [37] which often appear in low-illumination scenes. As discussed in Section 5.3, the proposed algorithm can be used to deblur such images.

Figure 13 shows a real captured image which contains several saturated regions (red boxes in (a)). We compare the proposed algorithm with the state-of-the-art methods [3], [4], [6], [5], [7], [47]. Since the priors of the state-of-the-art methods are developed to exploit salient edges for motion deblurring, these algorithms do not perform well for images containing numerous saturated regions. As the recent method [47] is developed to handle large Gaussian noise, it is less effective for saturated images. Although the saturated areas (e.g., highlighted blobs, streaks, and the characters in Figure 13(a)) are large due to motion blur, the L_0-regularized prior $P(x)$ favors a clear image with few blobs and streaks as a solution of the optimization of (4) in Section 3.2. The recovered image shown in Figure 13(h) is sharper and clearer. We note that while our method is able to estimate the blur kernel well, there still exist some ringing artifacts due to the limitation of the final latent image estimation process as shown in Figure 13(h). To generate better deblurred results, we use the non-blind deconvolution method [38]. The deblurred results shown in Figure 13(i) contains clearer text information and finer

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPAMI.2016.2551244, IEEE Transactions on Pattern Analysis and Machine Intelligence

TABLE 1
Quantitative comparison using the example shown in Figure 5(a).

	[3]	[4]	[5]	[7]	[47]	[8]	Using $P_t(\nabla x)$	Using $P_t(x)$	L0RIG	IL0RIG
SSIM of images	0.6457	0.6269	0.5611	0.4867	0.6190	0.5526	0.5963	0.6718	0.8916	0.8819
Kernel similarity	0.5200	0.5200	0.4170	0.6407	0.4938	0.6456	0.5849	0.6285	0.8699	0.9298

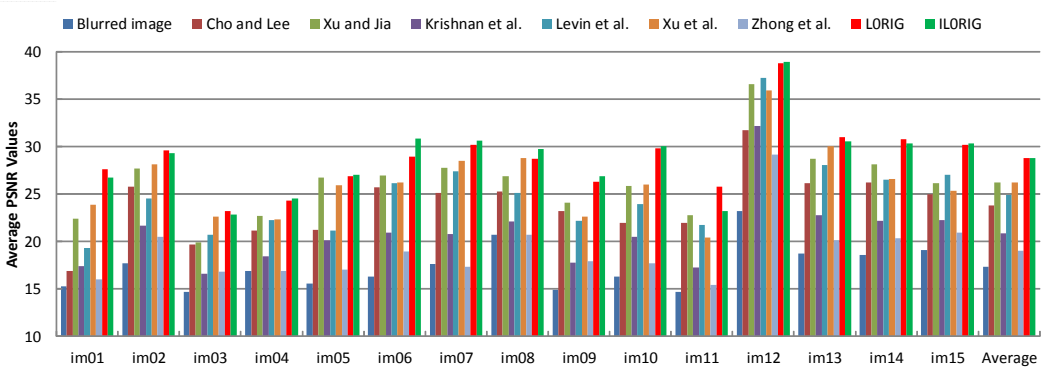
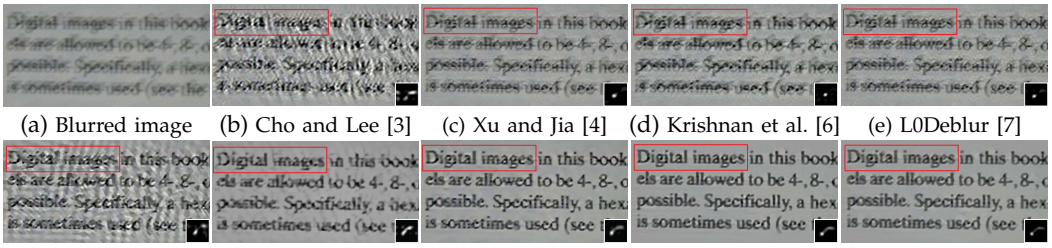


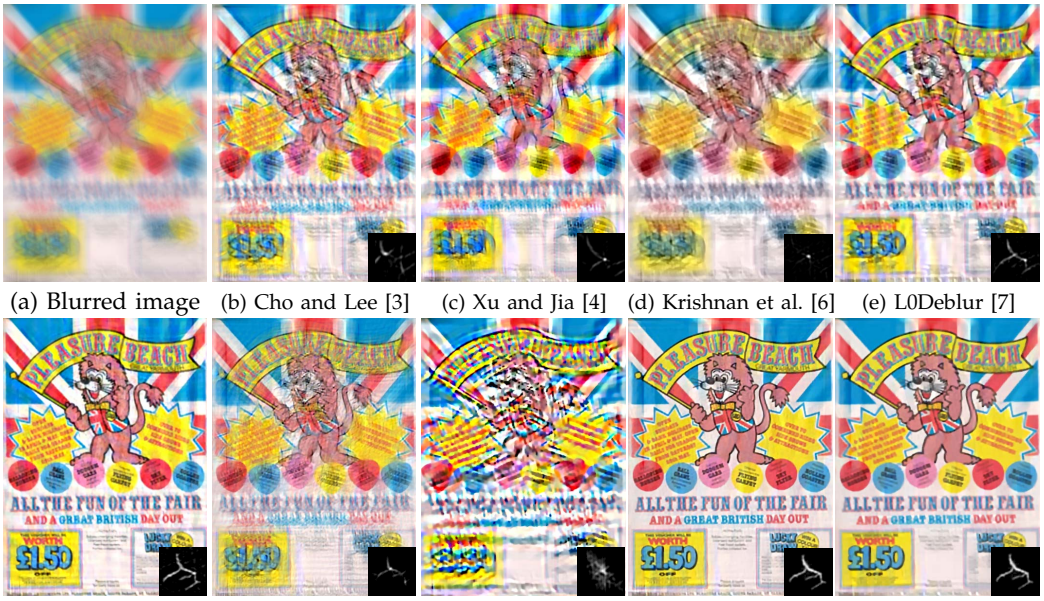
Fig. 10. Quantitative comparison on the proposed text image dataset. The x-axis denotes the image index and the average PSNR values of all the images are shown on the rightmost column.



(a) Blurred image (b) Cho and Lee [3] (c) Xu and Jia [4] (d) Krishnan et al. [6] (e) L0Deblur [7]

(f) Zhong et al. [47] (g) Chen et al. [9] (h) Cho et al. [8] (i) L0RIG (j) IL0RIG

Fig. 11. A real blurred image from [8]. The part in the red box in (h) contains blurry and unnatural results.



(a) Blurred image (b) Cho and Lee [3] (c) Xu and Jia [4] (d) Krishnan et al. [6] (e) L0Deblur [7]

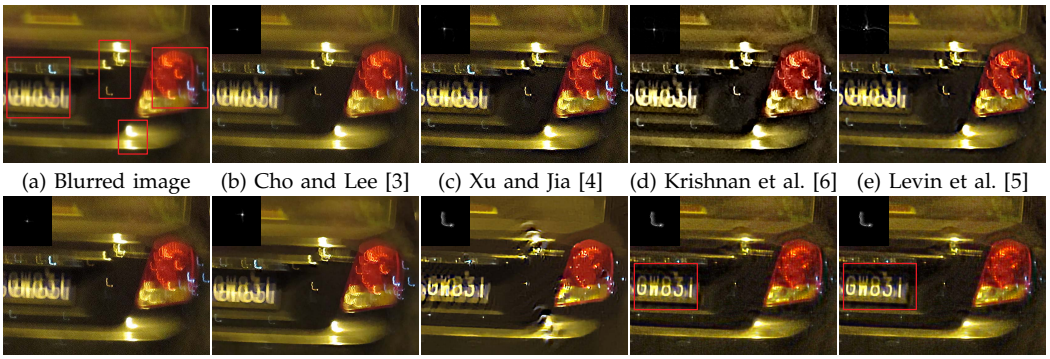
(f) Cho et al. [8] (g) Using only $P_t(\nabla x)$ (h) Using only $P_t(x)$ (i) L0RIG (j) IL0RIG

Fig. 12. A blurred image with text and complex background. Our method performs well in the kernel estimation and deblurred image.

textures which demonstrates the effectiveness of the proposed algorithm for kernel estimation. Both the L0RIG and IL0RIG methods are able to deblur this low-illumination image well.

We note that one recent work [48] is developed for detecting light streaks in low-illumination images for kernel estimation and deblurring. Different from [48], the proposed method does not require any

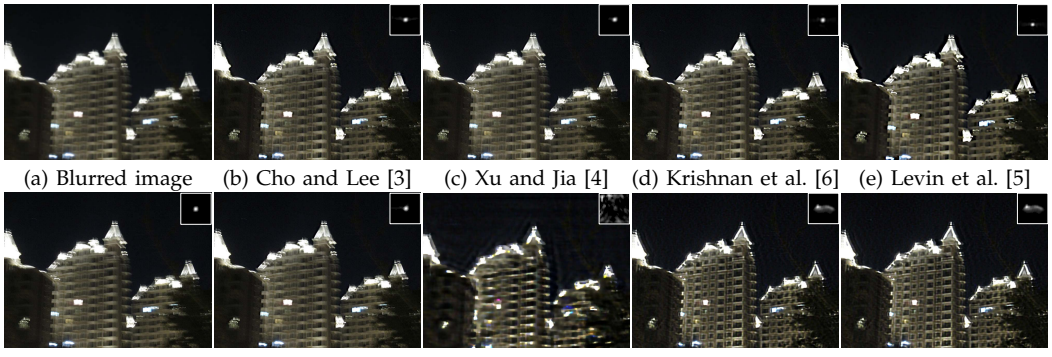
This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPAMI.2016.2551244, IEEE Transactions on Pattern Analysis and Machine Intelligence



(a) Blurred image (b) Cho and Lee [3] (c) Xu and Jia [4] (d) Krishnan et al. [6] (e) Levin et al. [5]

(f) L0Deblur [7] (g) Zhong et al. [47] (h) L0RIG results (i) L0RIG with [38] (j) IL0RIG with [38]

Fig. 13. Real captured blurred image with numerous saturated regions. The red boxes in (a) enclose some saturated pixels (e.g., the highlighted blobs, streaks, and the characters).



(a) Blurred image (b) Cho and Lee [3] (c) Xu and Jia [4] (d) Krishnan et al. [6] (e) Levin et al. [5]

(f) L0Deblur [7] (g) Zhong et al. [47] (h) Hu et al. [48] (i) L0RIG (j) IL0RIG

Fig. 14. Real captured blurred image with a lot of saturated areas.

pre-processing and performs well when light streaks cannot be detected in low-illumination images. Figure 14 shows one example with numerous saturated regions and deblurred results from the state-of-the-art methods [3], [4], [6], [5], [7], [47], [48]. As the saturated areas are large and light streaks cannot be detected in this image, the method [48] does not perform well in this case. For fair comparisons, we use the non-blind deconvolution method [38] to generate the final latent images in all the evaluated algorithms. The results show that the state-of-the-art deblurring methods are less effective for estimating blur kernels from saturated images, whereas the proposed algorithm is able to estimate motion blur kernels well for reconstructing deblurred results.

To further analyze the deblurring performance for low-illumination images, we create a dataset containing 6 ground truth low-illumination images and 8 kernels from [13]. Similar to [48], [37], we stretch the intensity histogram range of each image into [0, 2.2] and then apply 8 different blur kernels to generate blurred images where the pixel intensities are clipped into the range of [0, 1]. For fair comparisons, we use the same non-blind deconvolution method [48] to generate the final results and use PSNR to evaluate the quality of the restored images. Figure 15(a) shows that the proposed algorithm performs favorably against the state-of-the-art deblurring methods and generates clear images (See the project web page for images).

8.3 Natural Images

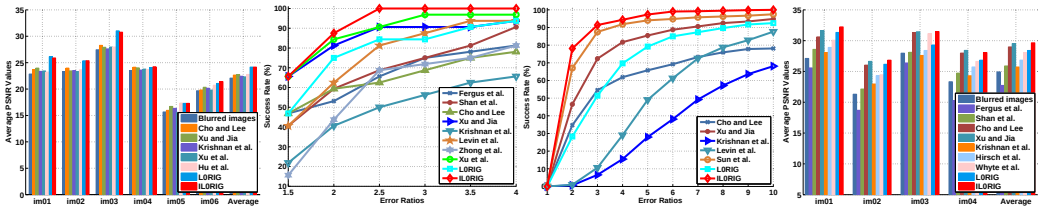
We show that the proposed algorithm can be applied to deblur natural images. We first evaluate our method on the benchmark image dataset [13] on uniform image deblurring, and compare with the state-of-the-art methods [1], [2], [3], [4], [5], [6], [7], [43], [47]. For fair comparisons, all the final deblurred results are generated by the same non-blind deconvolution method [5], and the error ratio metric [13] is used for evaluation.

Overall, the L0RIG algorithm performs well on natural scenes although it is designed to deblur text images. However, the L0RIG method alters the structures of intermediate results due to the complex intensity histograms of natural images according to the analysis in Section 6. Thus, the results are not comparable to those by the state-of-the-art approaches designed for generic scenes [4], [7]. As the IL0RIG algorithm extends the L0RIG method by exploiting the properties of input images to recover reliable salient edges for kernel estimation, the results in Figure 15(b) show that it is able to deblur natural images well.

We evaluate the proposed algorithm on the benchmark dataset and protocol [19] for fair comparisons. As shown in Figure 15(c), the proposed algorithm performs favorably against the state-of-the-art methods, and the curves for the L0RIG and IL0RIG methods demonstrate the effectiveness of the proposed prior.

We also evaluate our method on the benchmark

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPAMI.2016.2551244, IEEE Transactions on Pattern Analysis and Machine Intelligence



(a) Results on LI dataset (b) Results on [13] (c) Results on [19] (d) Results on [18]

Fig. 15. Quantitative evaluation on the proposed low-illumination image dataset and the benchmark datasets [13], [19], [18]. The x-axis in (b) and (c) denotes the error ratio values.

(a) Blurred image (b) Cho and Lee [3] (c) L0Deblur [7] (d) L0RIG (e) IL0RIG

Fig. 16. Deblurring real natural images. The proposed methods generate visually comparable results compared to the state-of-the-art methods.

(a) Blurred image (b) Gupta et al. [25] (c) Hirsh et al. [26] (d) Hu and Yang [49]

(e) Whyte et al. [24] (f) L0Deblur [7] (g) L0RIG (h) IL0RIG

Fig. 17. Non-uniform image deblurring [25]. The images shown in (b)-(f) are obtained from the reported results.

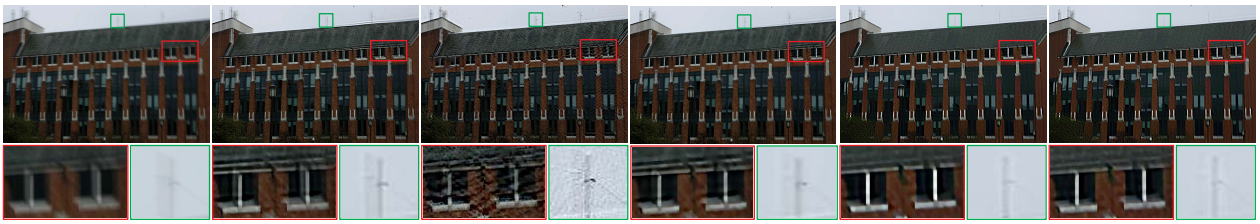
dataset for blind deconvolution [18], which contains 4 images with 12 blur kernels including several challenging cases. The average PSNR values are shown in Figure 15(d). Overall, the proposed L0RIG and IL0RIG methods perform well on this dataset against the state-of-the-art algorithms.

In Figure 16, we show deblurring results using real natural images [7]. As the L0RIG and IL0RIG methods with the proposed prior are able to preserve the salient edges, the deblurred results are sharper with fewer ringing artifacts. More results and quantitative evaluations are presented on the project web page.

8.4 Non-Uniform Image Deblurring

In this section, we compare the proposed algorithms with the state-of-the-art non-uniform deblurring methods [7], [24], [25], [26], [49]. Figure 17 shows an example from [25] which has been used for evaluation on non-uniform image deblurring. Figure 17 shows the recovered letters of (b)-(f) containing some ringing artifacts. Compared to the reported non-uniform deblurring results, the proposed L0RIG and IL0RIG methods generate comparable results with clear textures. Figure 18 shows another example of non-uniform image deblurring. Compared with other methods, the proposed L0RIG and IL0RIG methods generate results with few ringing artifacts and fine textures. More experimental results can be found on the project web page.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPAMI.2016.2551244, IEEE Transactions on Pattern Analysis and Machine Intelligence



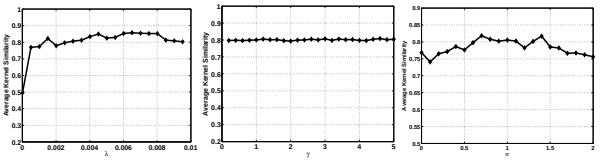
(a) Blurred image (b) Gupta et al. [25] (c) Whyte et al. [24] (d) L0Deblur [7] (e) L0RIG (f) IL0RIG

Fig. 18. Non-uniform image deblurring. The images shown in (b)-(d) are obtained directly from the reported results or generated by the original codes.

8.5 Sensitivity Analysis

The proposed model involves three main parameters, λ, σ, and γ. In this section, we show how they affect the image deblurring performance.

To analyze the effects of these parameters on image deblurring method for the IL0RIG method, we collect 16 blurred images for tests. For each parameter, we carry out experiments with different parameter settings by varying one and fixing the others with the kernel similarity metric to measure the accuracy of estimated kernels. For parameter λ, we set its values from 10^{-5} to 0.01 with the step size of 0.001. Figure 19(a) demonstrates that blur kernels can be well estimated by a wide range of λ, i.e., within $[0.001, 0.01]$. Similarly, we set the values of γ from 0.02 to 5 with the increment of 0.2, and the values of σ from 0 to 2 with the increase of 0.1. Figure 19(b) and (c) show that the proposed IL0RIG algorithm performs well with a wise range of parameter settings.



(a) λ (b) γ (c) σ

Fig. 19. Sensitivity analysis of λ, γ, and σ for the IL0RIG algorithm.

(a) (b) (c)

Fig. 20. A failure example. (a) Blurred image. (b) Intermediate result. (c) Deblurring result and kernel.

9 Concluding Remarks

In this paper, we propose a simple yet effective prior for text image deblurring. We discuss how the proposed prior facilitates preserving salient edges in image deblurring, and extend it to deal with natural images. With this prior, we present an effective optimization algorithm based on the half-quadratic splitting approach, which ensures that each sub-problem has a closed-form solution. Experimental results show that the proposed algorithms perform favorably against the state-of-the-art methods for deblurring text images without additional pre-processing steps (e.g., filtering, adaptive segmentation and SWT). In addition, we develop a latent image restoration method which helps reduce artifacts effectively. The proposed algorithm is also extended to deblur natural images and low-illumination scenes, as well as non-uniform cases.

Limitations: We note that the prior $\|x\|_0$ counts the number of nonzero-intensity pixels of an image x. If x does not contain zero-intensity pixels, we have $\|x\|_0 = \mathcal{C}$, where $\mathcal{C}$ is the total number of pixels in x and it is a constant for an image x. Mathematically, the minimization problem (4) would reduce to the model used in [7]. Furthermore, we would have $\|x\|_0 = \|y\|_0$ as the blurred image y would also not contain zero-intensity pixels according to the properties of convolution. This further demonstrates that the prior $\|x\|_0$ does not help the kernel estimation.

The proposed methods are likely to fail when a blurred image contains a large amount of noise (e.g., Gaussian and non-Gaussian noise) as the data term used in the proposed model (4) is based on L_2 norm which is less robust to noise. In addition, the proposed intensity prior counts the number of pixels with nonzero-intensity values. As the pixels are treated independently, the proposed algorithm is sensitive to large image noise and exacerbates its effect in the intermediate results (See Figure 20(b)). Our future work will focus on simultaneously denoising and deblurring for text as well natural images.

References

[1] R. Fergus, B. Singh, A. Hertzmann, S. T. Roweis, and W. T. Freeman, "Removing camera shake from a single photograph," *ACM Transactions on Graphics*, vol. 25, no. 3, pp. 787–794, 2006.

[2] Q. Shan, J. Jia, and A. Agarwala, "High-quality motion deblurring from a single image," *ACM Transactions on Graphics*, vol. 27, no. 3, p. 73, 2008.

[3] S. Cho and S. Lee, "Fast motion deblurring," *ACM Transactions on Graphics*, vol. 28, no. 5, p. 145, 2009.

[4] L. Xu and J. Jia, "Two-phase kernel estimation for robust motion deblurring," in *ECCV*, 2010, pp. 157–170.

[5] A. Levin, Y. Weiss, F. Durand, and W. T. Freeman, "Efficient marginal likelihood optimization in blind deconvolution," in *CVPR*, 2011, pp. 2657–2664.

[6] D. Krishnan, T. Tay, and R. Fergus, "Blind deconvolution using a normalized sparsity measure," in *CVPR*, 2011, pp. 2657–2664.

[7] L. Xu, S. Zheng, and J. Jia, "Unnatural L_0 sparse representation for natural image deblurring," in *CVPR*, 2013, pp. 1107–1114.

[8] H. Cho, J. Wang, and S. Lee, "Text image deblurring using text-specific properties," in *ECCV*, 2012, pp. 524–537.

[9] X. Chen, X. He, J. Yang, and Q. Wu, "An effective document image deblurring algorithm," in *CVPR*, 2011, pp. 369–376.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPAMI.2016.2551244, IEEE Transactions on Pattern Analysis and Machine Intelligence

[10] Y. Lou, A. L. Bertozzi, and S. Soatto, "Direct sparse deblurring," *Journal of Mathematical Imaging and Vision*, vol. 39, no. 1, pp. 1–12, 2011.

[11] T.-H. Li and K.-S. Lii, "A joint estimation approach for two-tone image deblurring by blind deconvolution," *IEEE Transactions on Image Processing*, vol. 11, no. 8, pp. 847–858, 2002.

[12] B. Epshtein, E. Ofek, and Y. Wexler, "Detecting text in natural scenes with stroke width transform," in *CVPR*, 2010, pp. 2963–2970.

[13] A. Levin, Y. Weiss, F. Durand, and W. T. Freeman, "Understanding and evaluating blind deconvolution algorithms," in *CVPR*, 2009, pp. 1964–1971.

[14] H. Zhang, J. Yang, Y. Zhang, and T. S. Huang, "Sparse representation based blind image deblurring," in *ICME*, 2011, pp. 1–6.

[15] J.-F. Cai, H. Ji, C. Liu, and Z. Shen, "Framelet based blind motion deblurring from a single image," *IEEE Transactions on Image Processing*, vol. 21, no. 2, pp. 562–572, 2012.

[16] H. Takeda, S. Farsiu, and P. Milanfar, "Deblurring using regularized locally adaptive kernel regression," *IEEE Transactions on Image Processing*, vol. 17, no. 4, pp. 550–563, 2008.

[17] N. Joshi, R. Szeliski, and D. J. Kriegman, "PSF estimation using sharp edge prediction," in *CVPR*, 2008.

[18] R. Köhler, M. Hirsch, B. J. Mohler, B. Schölkopf, and S. Harmeling, "Recording and playback of camera shake: Benchmarking blind deconvolution with a real-world database," in *ECCV*, 2012, pp. 27–40.

[19] L. Sun, S. Cho, J. Wang, and J. Hays, "Edge-based blur kernel estimation using patch priors," in *ICCP*, 2013.

[20] Y. HaCohen, E. Shechtman, and D. Lischinski, "Deblurring by example using dense correspondence," in *ICCV*, 2013, pp. 2384–2391.

[21] J. Pan, Z. Hu, Z. Su, and M.-H. Yang, "Deblurring face images with exemplars," in *ECCV*, 2014, pp. 47–62.

[22] Q. Shan, W. Xiong, and J. Jia, "Rotational motion deblurring of a rigid object from a single image," in *ICCV*, 2007, pp. 1–8.

[23] Y.-W. Tai, P. Tan, and M. S. Brown, "Richardson-lucy deblurring for scenes under a projective motion path," *IEEE Transactions on Pattern Analysis Machine Intelligence*, vol. 33, no. 8, pp. 1603–1618, 2011.

[24] O. Whyte, J. Sivic, A. Zisserman, and J. Ponce, "Non-uniform deblurring for shaken images," *International Jouunal on Computer Vision*, vol. 98, no. 2, pp. 168–186, 2012.

[25] A. Gupta, N. Joshi, C. L. Zitnick, M. F. Cohen, and B. Curless, "Single image deblurring using motion density functions." in *ECCV*, 2010, pp. 171–184.

[26] M. Hirsch, C. J. Schuler, S. Harmeling, and B. Schölkopf, "Fast removal of non-uniform camera shake," in *ICCV*, 2011, pp. 463–470.

[27] S. Harmeling, M. Hirsch, and B. Schölkopf, "Space-variant single-image blind deconvolution for removing camera shake," in *NIPS*, 2010, pp. 829–837.

[28] T. H. Kim, B. Ahn, and K. M. Lee, "Dynamic scene deblurring," in *ICCV*, 2013, pp. 3160–3167.

[29] T. H. Kim and K. M. Lee, "Segmentation-free dynamic scene deblurring," in *CVPR*, 2014, pp. 2766–2773.

[30] A. Chakrabarti, T. Zickler, and W. T. Freeman, "Analyzing spatially-varying blur," in *CVPR*, 2010, pp. 2512–2519.

[31] L. Xu and J. Jia, "Depth-aware motion deblurring," in *ICCP*, 2012, pp. 1–8.

[32] Z. Hu, L. Xu, and M. Yang, "Joint depth estimation and camera shake removal from single blurry image," in *CVPR*, 2014, pp. 2893–2900.

[33] L. Xu, C. Lu, Y. Xu, and J. Jia, "Image smoothing via L_0 gradient minimization," *ACM Transactions on Graphics*, vol. 30, no. 6, p. 174, 2011.

[34] Y. Wang, J. Yang, W. Yin, and Y. Zhang, "A new alternating minimization algorithm for total variation image reconstruction," *SIAM Journal on Imaging Sciences*, vol. 1, no. 3, pp. 248–272, 2008.

[35] D. Krishnan and R. Fergus, "Fast image deconvolution using Hyper-Laplacian priors," in *NIPS*, 2009, pp. 1033–1041.

[36] J. Pan and Z. Su, "Fast ℓ^0-regularized kernel estimation for robust motion deblurring," *IEEE Signal Processing Letters*, vol. 20, no. 9, pp. 841–844, 2013.

[37] S. Cho, J. Wang, and S. Lee, "Handling outliers in non-blind image deconvolution," in *ICCV*, 2011, pp. 495–502.

[38] O. Whyte, J. Sivic, and A. Zisserman, "Deblurring shaken and partially saturated images," in *ICCV Workshops*, 2011, pp. 745–752.

[39] L. Xu, J. S. Ren, C. Liu, and J. Jia, "Deep convolutional neural network for image deconvolution," in *NIPS*, 2014, pp. 1790–1798.

[40] J. Jia, "Single image motion deblurring using transparency," in *CVPR*, 2007.

[41] ——, *Mathematical models and practical solvers for uniform motion deblurring*. Cambridge University Press, 2014.

[42] A. Levin, D. Lischinski, and Y. Weiss, "A closed form solution to natural image matting," in *CVPR*, 2006, pp. 61–68.

[43] J. Pan, Z. Hu, Z. Su, and M.-H. Yang, "Deblurring text images via L_0-regularized intensity and gradient prior," in *CVPR*, 2014, pp. 2901–2908.

[44] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, no. 9, pp. 62–66, 1979.

[45] Z. Wang, A. C. Bovik, H. R. Sheikha, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.

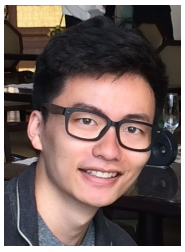
[46] Z. Hu and M.-H. Yang, "Good regions to deblur," in *ECCV*, 2012, pp. 59–72.

[47] L. Zhong, S. Cho, D. Metaxas, S. Paris, and J. Wang, "Handling noise in single image deblurring using directional filters," in *CVPR*, 2013, pp. 612–619.

[48] Z. Hu, S. Cho, J. Wang, and M.-H. Yang, "Deblurring low-light images with light streaks," in *CVPR*, 2014, pp. 3382–3389.

[49] Z. Hu and M.-H. Yang, "Fast non-uniform deblurring using constrained camera pose subspace," in *BMVC*, 2012, pp. 1–11.

Jinshan Pan is currently a joint-training Ph.D. student in School of Mathematical Sciences at Dalian University of Technology, China, and Electrical Engineering and Computer Science at University of California, Merced, CA, USA. His research interest includes image deblurring, image/video analysis and enhancement, and related vision problems.



Zhe Hu is a research scientist at Light Co.. He received the Ph.D. Degree in Computer Science at University of California, Merced, USA, in 2015. His research interests include computer vision, computational photography and image processing.

Zhixun Su is a professor at Dalian University of Technology, China. He received Ph.D. in Computational Mathematics at Dalian University of Technology in 1993. His current research interests include computer graphics, computer vision, computational geometry, and machine learning. He is now a member of the Standing Committee of China Society of Computational Mathematics.

Ming-Hsuan Yang is an associate professor of Electrical Engineering and Computer Science with the University of California, Merced, CA, USA. He received the Ph.D. degree in computer science from the University of Illinois at Urbana-Champaign, USA, in 2000. He served as an Associate Editor of the IEEE Transactions on Pattern Analysis and Machine Intelligence from 2007 to 2011, and is an Associate Editor of the International Journal of Computer Vision, Image and Vision Computing, and Journal of Artificial Intelligence Research. He received the NSF CAREER Award in 2012, and the Google Faculty Award in 2009. He is a senior member of the IEEE and ACM.