The Visual Computer
https://doi.org/10.1007/s00371-018-1562-2

ORIGINAL ARTICLE

Rolling bilateral filter-based text image deblurring

Hang Yang[1] · Zhongbo Zhang[2] · Yujing Guan[2]

Abstract
Although many competitive approaches have been developed for image deblurring, the priors which rely on natural image are less effective for text images which have special properties, lack of heavy-tailed gradient, and clean background. Considering the specific structure of text images, in this work, we present an effective yet simple deblurring method based on rolling bilateral filtering, an improved rolling guidance filtering specifically. The rolling bilateral filter can remove texture and preserve image structures. Thus, it is appropriate for processing the document image whose background regions are uniform. According to this property, we propose an efficient iterative algorithm to estimate the image and point spread function. Experimental results demonstrate that our method outperforms the traditional deblurring algorithms designed for typical natural images and text images.

Keywords Rolling bilateral filter · Text image deblurring · Texture removal · Edge preserving

1 Introduction

During text image acquisition, it can be degraded by many factors such as geometrical warping [1], noise distortions [2], and image blurring [3,4]. Here, we mainly concern the problem of text image deblurring. In recent years, document image deblurring has received extensive concerns because of its wide applications [5,6].

In single image deblurring, the model of blur processing is often written as the convolution of the latent image L with a blur kernel k:

$$B = k * L + \gamma \tag{1}$$

where B is an observed blurry image, γ denotes additive noise, and $*$ is the convolution operator. The inversion of the blurring is a well-known ill-condition problem, and the goal of deblurring is to estimate both L and k from a blurry image B.

This work is supported by the National Science Foundation of China under Grant 61401425.

✉ Hang Yang
yanghang09@mails.jlu.edu.cn

[1] Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Science, Changchun 130033, China

[2] Departments of Mathematics, Jilin University, Changchun 130012, China

In recent years, researchers have made great progress in the single blind image deblurring technology [7–17]. Many state-of-the-art methods present various priors for optimization, for instance, Gaussian prior [18], sparsity prior [19,20], and image statistics prior [4,21,22]. Despite numerous methods are suitable for natural images deblurring, they are often ineffective for document images whose contents are mainly binary—white and black—and the text images do not satisfy the heavy-tailed gradient prior of natural images. By applying a natural image restoration algorithm directly to the text image, it will cause an erroneous result. Previous text images deblurring algorithms [23–25] employ two-tone or bi-level priors, ink-bleeding artifacts, and small blurs can be eliminated by the thresholding approach. Since these methods [23–25] do not consider the motion blur, they tend to fail when dealing with blurred document images caused by large motion blur kernels.

In many deblurring methods [18,20,26–28], sharp edges are predicted and used to estimate the blur kernel. In this way, the natural image can be restored well. However, it cannot directly be implemented to document images. Text characters in image are often small and connected, while their edges are treated as small outliers to be removed in kernel estimation.

Lately, Chen et al. [4] propose a new document image deblurring algorithm based on the image intensity. However, this approach depends on accurate document segmentation,

Published online: 25 May 2018

and it does not work effectively for general blurred document images. Lou et al. [29] present a direct deblurring method for text images that exploits sparsity prior of natural images. However, the estimated blur kernel is not clear and it spends much time on training an over-complete dictionary. In [23], a joint estimation method is proposed to obtain the blur kernel from binary images. Cho et al. [30] analyze useful properties of document images and present an approach to enforce the domain-specific properties with the stroke width transform (SWT) [31]. Although this method yields the competitive results, the process of kernel estimation is complex and the performance depends largely on whether SWT divides the image into nontext and text parts well or not. Pan et al. [5,32] develop a document image restoration approach based on l_0-regularized prior of image gradient and image intensity without additional techniques(e.g., adaptive segmentation [4] or SWT [30]). But the l_0-based deblurring method would be less effective when text image does not contain black pixels. A multiscale dictionaries-based approach for text image deblurring is presented in [33], in which a series of text-specific multiscale dictionaries (TSMD) and a natural scene dictionary are learned.

In this work, we develop a rolling bilateral filter-based algorithm for document image deblurring and propose an effective optimization method based on splitting approach. The rolling bilateral filter can remove almost all texture patterns and preserve the edge of images, so it is very suitable for the special properties of text images. The text-specific properties are also applied to the final deconvolution that recovers the latent image with the estimated kernel. Compared to state-of-the-art algorithms [4,30,33], our approach is much simpler and easier to implement. We summarize the contributions of this paper as follows:

1. We propose a rolling bilateral filter 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 subproblem has a closed-form solution.
3. For the latent image deconvolution step, we present an effective method to deal with artifacts and evaluate it against other alternatives.

This paper is organized as follows. A brief introduction for rolling bilateral filter is provided in Sect. 2. The proposed text image deblurring formulation is described in Sect. 3. We analyze how the proposed method performs on text image deblurring in Sect. 4. In Sect. 5, the efficacy of the proposed method is demonstrated by experiments. Concluding remarks are provided in Sect. 6.

2 Rolling bilateral filter

Edge-aware image processing technique is broadly studied for smoothing images without destroying different levels of structures. It is widely applied in the computer graphics community. Edge-preserving filters can be roughly divided into two broad categories: average-based methods and optimization-based methods.

Many structure-preserving smoothing operators can be achieved by local filtering [34–36]. These filters are widely used for its simplicity; however, they often fail to extract structures from the texture. Optimization-based methods [37,38] are also widely used. In terms of texture removal, some optimization-based methods outperform local filtering. However, these methods are difficult to implement and often guarantee high computational complexity [39].

Recently, Zhang et al. [39] propose a rolling guidance filter (RGF) derived from an effective modification to bilateral filter [35] method. It is essentially a joint bilateral filtering [40], which can be effectively remove texture and preserve the structure, while the standard bilateral filter is hard to achieve this. The rolling guidance filter is a simple extension to the joint bilateral filter, and it has advantages such as simple, fast, easy to implement, scalable, and adaptable. The joint bilateral filter (JBF) is defined as follows:

$$\mathrm{JBF}(p) = \frac{\sum_{q\in\mathcal{N}(p)} W_{p,q} I(q)}{\sum_{q\in\mathcal{N}(p)} W_{p,q}} \tag{2}$$

where the bilateral kernel is:

$$W_{p,q} = e^{-\frac{(p-q)^2}{2\sigma_s^2}} e^{-\frac{(\mathrm{RI}(p)-\mathrm{RI}(q))^2}{2\sigma_r^2}} \tag{3}$$

I is the input image, RI is the reference image, σ_s and σ_r are the parameters to control weights, and $\mathcal{N}(p)$ is a window centered at the pixel p. If $\mathrm{RI} = I$, the JBF is the bilateral filter (BF).

The rolling guidance filter comprises two steps, i.e., the removal of small structures and the edge restoration. The input image is first filtered by applying a Gaussian filter whose variance is σ_s^2, and a filtered image L_{σ_s} is obtained. If the image structure size is less than σ_s, it can be completely eliminated in the output image, and other image structures are also reduced. Then, the other important step in RGF is the iterative edge restoration. A reference image J^p is iteratively updated in this step. For initialization, J^0 is defined as L_{σ_s}. J^{p+1} is the result of iteration $(p+1)$-th, and it can be generated by a joint bilateral filtering, whose input image is original image I and the reference image is previous iteration J^p.

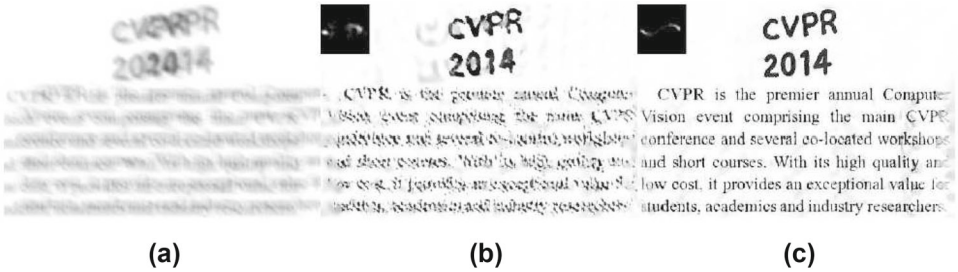


Fig. 1 A text image deblurring example. **a** Blurred image, **b** results obtained by RGF, **c** results obtained by RBF

Unlike previous algorithms which involve JBF, this approach extracts the main structure of I. It employs J^p to acquire the similarity between pixels, and the output image structures are similar to J^p.

In the process of deblurring processing, to estimate the blur kernel and avoid local minima, the coarse-to-fine approach is performed. In practice, we find that the rolling guidance filter can blur narrow-space texts and cannot remove the strong ringing artifacts, so it will produce a blur kernel with large estimation error. The reason is that the Gaussian filter used in the first step removes small-scale structures and keeps strong noise in the deblurring process.

At the coarser level, the decimation image has low resolution, the structure scale is small, and the Gaussian filter cannot guarantee to retain that useful scale while removing noise effectively. In this work, we use the bilateral filter instead of Gaussian filter in the first step of the rolling guidance filter. The bilateral filter with large variance is not only able to save small useful structure scale but also can remove the strong noise. We call this new filter as rolling bilateral filter (RBF). Figure 1 shows the blur kernel estimations and deblurring results for the RBF-based method compared to the RGF-based approach, and we can clearly find out that there are a lot of errors in the blur kernel estimation based on RGF. The proposed rolling bilateral filter is depicted in Algorithm 1.

Algorithm 1: Rolling bilateral filter (RBF)

1. Input: $n_{\text{itr}}, \sigma_s, \sigma_r$, $J^0 = \text{BF}(I, \sigma_s, \sigma_r)$
2. for $p = 0 : n_{\text{itr}} - 1$
   $J^{p+1} = \text{JBF}(I, J^p, \sigma_s, \sigma_r)$
   end
3. Filtered result: $J^{n_{\text{itr}}}$.

The only difference between RGF and RBF is the first step in Algorithm 1; we used a bilateral filter instead of Gaussian filter in this step, so the computational complexity of RGF is $O(Nr) + n_{\text{itr}}O(Nr^2)$, and RBF's computational complexity is $(n_{\text{itr}} + 1)O(Nr^2)$, where N is the number of pixels in the image and the kernel size of $\mathcal{N}(p)$ is $r \times r$.

The rolling bilateral filter inherits the main advantage of bilateral filtering and has the similar property with RGF. It can not only eliminate small-scale structures, but also preserve other contents. Algorithm 1 is simple to implement and achieves promising performance. In this paper, we introduce RBF into text images deblurring problem, which leads to an effective method that produces promising results.

3 Text image deblurring via rolling bilateral filter

We present a new rolling bilateral filter-based method for text image deblurring in this section.

It is observed that the magnitude values of text characters and background regions are nearly uniform in a clean image. The main motivation of our proposed method is based on this fact. That is to say, the document image is piecewise smooth and contains little texture information. Since the pixel values of the clean document image are close to two values, it is likely to be a cartoon part of natural image, and the rolling guidance filter has shown to be effective in suppressing artifacts for textureless image. Thus, we implement the proposed rolling bilateral filter to text images.

We use the regularization term $\varphi(\cdot)$ for document image restoration, which is defined as:

$$\varphi(L) = \|L - \boldsymbol{RBF}(L)\|_2^2 \tag{4}$$

To incorporate the regularization term into the text image deblurring process, we minimize the following cost function to estimate the clean image and the blur kernel:

$$\arg\min_{L,k} \| B - k * L \|_2^2 + \lambda\varphi(L) + \beta\psi(k), \tag{5}$$

where the first term is imposed to the data fitting, $\varphi(L)$ is used to retain main structure and removes artifacts, and $\psi(k) = \| \nabla k \|_2^2$ is the regularization term for the blur kernel; λ and β are the regularization parameters.

The solution of Eq. (5) can be obtained by alternatively solving

$$\arg\min_{L} \| B - k * L \|_2^2 + \lambda\varphi(L) \tag{6}$$

$$\arg\min_{k} \| B - k * L \|_2^2 + \beta\psi(k) \tag{7}$$

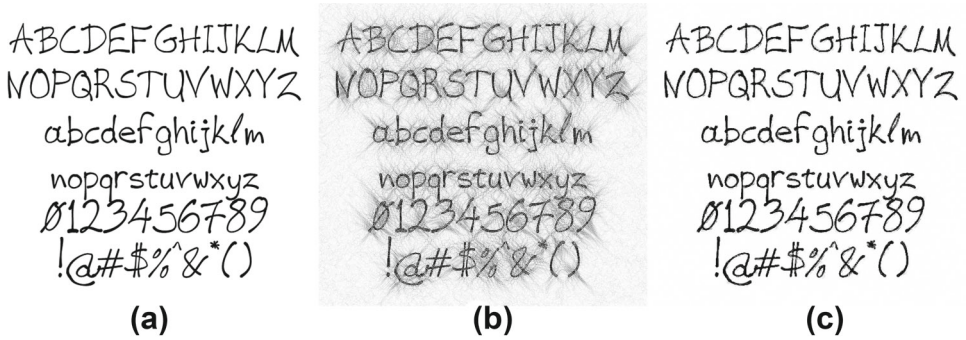


Fig. 2 An intermediate process of non-blind text image deconvolution. **a** Clear image, **b** intermediate results obtained by Eq. (10), **c** denoised result of (**b**) using Eq. (9)

However, directly solving Eq. (6) is difficult because **RBF** is highly nonlinear. So, we present a decouple approach to resolve Eq. (6), and it can achieve a desirable result in practice.

3.1 Non-blind text image deconvolution

Minimizing Eq. (6) is commonly regarded as computationally intractable because of the rolling bilateral filter regularization term. An auxiliary variable v is introduced based on the half-quadratic splitting method [41], and it is useful to process Eq. (6). v balances intensity values and domain-specific properties (sharp contrast between text and background) of the latent image L.

Split variables approaches motivate us to propose this iterative method. In practice, a good result can be obtained by iterating the following two steps:

$$v^{p+1} = \arg\min_{v} \| B - k * v \|_2^2 + \lambda \| v - L^p \|_2^2 \tag{8}$$

$$L^{p+1} = \mathbf{RBF}(v^{p+1}) \tag{9}$$

With this formulation and fixing the other variables, it can efficiently obtain the solution through alternatively minimizing v and L. For initialization, L^0 is set to be zero.

The solution of Eq. (8) can be computed by using the convolution theorem for Fourier transform:

$$\mathcal{F}(v^{p+1}) = \frac{\mathcal{F}^*(k) \cdot \mathcal{F}(B) + \lambda \mathcal{F}(L^p)}{| \mathcal{F}(k) |^2 + \lambda} \tag{10}$$

where $\mathcal{F}$ and $\mathcal{F}^*(\cdot)$ are the fast Fourier transform (FFT) operator and the complex conjugate operator, respectively.

The resulting image v^p obtained by Eq. (10) usually contains useful high-frequency structures and a special form of distortions. Since rolling bilateral filter has good structure-aware smoothing property near the edges, it is used to suppress the amplified noise and artifacts introduced by Eq. (10) in this work.

The main reason that the traditional deblurring algorithms perform poorly for text image is that the natural image priors cannot well characterize the text properties [33]. The rolling bilateral filter can remove nearly all artifacts and textures while keeping the image edges. This property is very suitable for document image whose background regions are uniform.

Therefore, in our work, the rolling bilateral filter is integrated into the non-blind deconvolution model and a reliable sharp text image is obtained.

The main steps of intermediate latent image estimation are summarized in Algorithm 2.

Algorithm 2: Estimation of intermediate latent image

1. Input: Blurred and noisy image B, blur kernel k, initialize λ, σ_s and σ_r.
2. Repeat $p = 1 : \text{Max}_{\text{iter}}$:
   solve for v^p using Eq. (10);
   solve for L^p using Eq. (9);
   update parameters: $\lambda \leftarrow 2\lambda$, $\sigma_s \leftarrow 1.25\sigma_s$ and $\sigma_r \leftarrow 1.25\sigma_r$.
3. Output: Intermediate latent image L.

Parameters λ, σ_s, and σ_r are automatically adapted in iterations starting from small values, and they are multiplied by fixed values each time. This scheme is inspired by Wang et al. [41], who show that this scheme is effective to speed up convergence.

In Fig. 2, we show the effect of RBF in terms of some intermediate results. In the non-blind image deconvolution step [Eqs. (10) and (9)], when the blur kernel is not accurate or the blurred image contains noise, the result obtained from Eq. (10) will contain ring artifacts; we can see that there are many small-scale features such as texture, small objects, and noise in Fig. 2b. Almost all of the text image information is large-scale features, which generally encode boundaries, slow spatial color transition, and smooth regions. In Fig. 2c, we show that the rolling bilateral filter can remove small-scale features and retain large-scale information.

3.2 Blur kernel estimation

With the estimated image L, Eq. (7) is a least-squared minimization problem. However, in [5,18], the authors show that the solution directly from Eq. (7) is not precise. In this work,

the blur kernel is estimated by minimizing the following energy function:

$$R(k) = \| k * \nabla L - \nabla B \|_2^2 + \beta \| \nabla k \|_2^2 \tag{11}$$

The elements $k_{i,j}$ in blur kernel k are subject to the constraints that $k_{i,j} \geq 0$ and $\sum_{i,j} k_{i,j} = 1$, and FFT can be used to compute the solution efficiently.

This problem (11) has widely been discussed in blind deconvolution. To obtain the solution of Eq. (11), we use a gradient descent scheme and the constraints are enforced on the result of each iteration as follows:

$$k^{p+1} = k^p - \eta \frac{\mathrm{d}R}{\mathrm{d}k} \tag{12}$$

where

$$\frac{\mathrm{d}R}{\mathrm{d}k} = \mathcal{F}^{-1}\Big((|\mathcal{F}(\nabla L)|^2 + \beta|\mathcal{F}(\nabla)|^2)\mathcal{F}(k) - \mathcal{F}^*(\nabla L)\mathcal{F}(B)\Big) \tag{13}$$

and $\mathcal{F}(\nabla)$ denotes the Fourier transform of differential operator ∇.

After optimizing Eq. (12), we set elements smaller than 0.05 of the biggest one to zero. Then, the remaining nonzero values are normalized so that their sum becomes one. This gradient descent method repeats NIter times.

In Algorithm 3, we show the proposed kernel estimation algorithm. Similar to the competitive algorithms, our kernel estimation method works in a coarse-to-fine manner following an image pyramid.

Algorithm 3: Estimation of the Blur Kernel

1. Input: Blurred and noisy image B, k is initialized as the result of the coarser level. Set parameters β, η, and NIter.
2. for $j = 1$: Levels
   solve for L using **Algorithm 2**;
   Update k by using Eqs. (12)–(13).
   end
3. Output: Blur kernel k.

3.3 Final image deconvolution

Once the point spread function k is determined, we use it to obtain the final deblurring result L from the input blurred image B. This is not a trivial task when B contains severe noise [42]. Although latent images can be obtained from Eqs. (8)–(9), these formulations are not optimal for images with rich texture details.

We note that another edge-preserving filter-guided filter (GF) has been shown to be able to preserve fine details [43]. The guided filter output is a locally linear transformation of the guidance image. This filter gets the edge-preserving smoothing property like the bilateral filter, but does not suffer from the gradient reversal artifacts. And spatial-based deconvolution algorithms can obtain high-quality denoising image with fewer artifacts. However, guided filter cannot remove the strong texture and artifacts, so it might fail to obtain a good kernel estimation. In contrast, our algorithm with rolling bilateral filter loses some fine details and contains ringing artifacts. According to these properties, we use the rolling bilateral filter to estimate the point spread function, and the guided filter is exploited to obtain the final output with finer details and fewer artifacts.

In our method, we estimate the latent image L using the following method:

$$L_{\text{final}} = \arg\min_L \| B - k * L \|_2^2 + \lambda \| L - \boldsymbol{GF}(L) \|_2^2 \tag{14}$$

One can find that the solution of Eq. (14) can be obtained in a similar way as Eqs. (8)–(9), and the difference is that we use the guided filter (GF) instead of the rolling bilateral filter (RBF) in Eq. (9):

$$v^{s+1} = \arg\min_v \| B - k * v \|_2^2 + \lambda \| v - L^s \|_2^2 \tag{15}$$

$$L^{s+1} = \mathbf{GF}(v^{s+1}) \tag{16}$$

Figure 3d shows our final result. One can see that the proposed approach can restore text images well. The result of [44] is shown in Fig. 3b, from which the result is noisy and contains ringing, and Fig. 3c shows the estimated result of Eqs. (8)–(9). In our method, we processed the RGB color channels separately.

4 Algorithm analysis

In this part, we discuss the performance of our algorithm in document image restoration and investigate the reasons behind that. We show the importance of RBF-based image deconvolution for text deblurring and analyze the convergence of our method.

4.1 Importance of RBF

In Eq. (10), the improper regularization prior causes a noisy estimate v with ringing artifacts. Since the edges and textures caused by these artifacts can be effectively removed by rolling bilateral filter, leaving only the intact edges of the large amplitude, it is of great help to the blur kernel estimation, as mentioned in [20]. Our approach estimates intermediate results by iterative approach, and the optimization subproblems are solved by similar way as [5,44]. In our

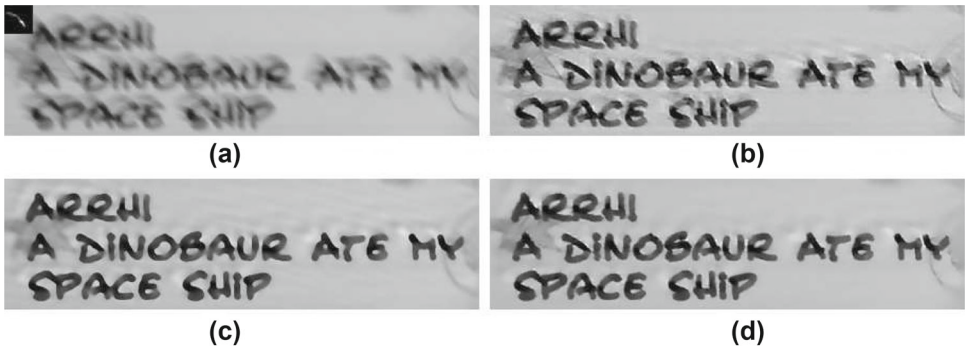


Fig. 3 Image deconvolution example with known blur kernel. **a** Blurred image and blur kernel, **b** result by [44], **c** result by Eqs. (8)–(9), **d** our final result

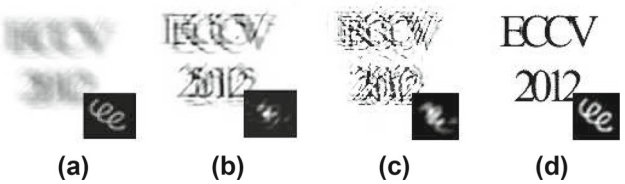


Fig. 4 A text image deblurring experiment provided in [30]. **a** Blurred image and blur kernel, **b** results by [30], **c** the results without utilizing RBF, **d** results by Eqs. (8)–(9)

method, we use Eq. (9) to remove tiny pixels and structures with small intensity values while retaining salient edges.

Figure 4 shows an example to test the effectiveness of the RBF. The blurred image is presented in Fig. 4a. Figure 4b shows results of Cho et al.'s [30] algorithm, which fails to estimate the blur kernel. As the noise and ringing artifacts in latent image greatly disturb the kernel estimation, we note that image deblurring without rolling bilateral filter is less effective (Fig. 4c). In contrast, applying rolling bilateral filter ensures the successful kernel estimation, which is shown in Fig. 4d.

4.2 Convergence of the proposed algorithm

The alternating least squares iteration is the main approach in the proposed image deblurring method. In practice, for a variety of inputs, we find the proposed approach converges well. We show the kernel similarity [45] with respect to iterations in Fig. 5c. It is observed that the similarity between estimated and real kernel becomes higher with more iterations.

The kernel similarity between K and $\hat{K}$ can be computed as:

$$S(K, \hat{K}) = \max_t \rho(K, \hat{K}, t); \tag{17}$$

where ρ means the normalized cross-correlation function and t is the offset between K and $\hat{K}$ [45]. The element coordinates are represented by τ, and $\rho()$ is given by

$$\rho(K, \hat{K}, t) = \frac{\sum_\tau K(\tau)\hat{K}(\tau + t)}{\| K \|_2 \| \hat{K} \|_2} \tag{18}$$

where $\|\|_2$ is l_2-norm.

In our method, the sensitivity of the method is mainly depending on parameters λ in Algorithm 1 and β, η in Algorithm 3. λ is a regularization parameter, and its value is adapted from their initial values over iterations of the optimization. A larger λ will produce a over smooth result, and a smaller one will obtain a noisy result. β is also a regularization parameter, a larger value will lose some detail of blur kernel, and a smaller one will introduce some noise. η is a step size, a smaller one will lead to slow convergence and even miss the optimum solution, and the algorithm may not be converged when η is too large.

4.3 Natural images deblurring

Although the rolling bilateral filter is developed based on the assumptions that the document image has uniform backgrounds, the proposed approach can be utilized to the natural image deblurring as shown in Fig. 6.

In Algorithm 2, the filtered image L (Eq. (9)) contains large-scale structures, that is to say, the intermediate result v (Eq. (10)) inherits the main structures of the image L. Unlike the natural image deblurring algorithm [42], rolling bilateral filter is introduced in the intermediate latent image estimation step. Rolling bilateral filter can preserve important structures of the intermediate image. An intermediate image estimated from a blurred image is shown in Fig. 6b. Fig. 6c shows the result, from which our method removes tiny details while retaining salient edges in natural images; thereby, it is helpful for estimating the blur kernel.

5 Experiments

We compare our algorithm with some competitive methods for document image deblurring in this section. We imple-

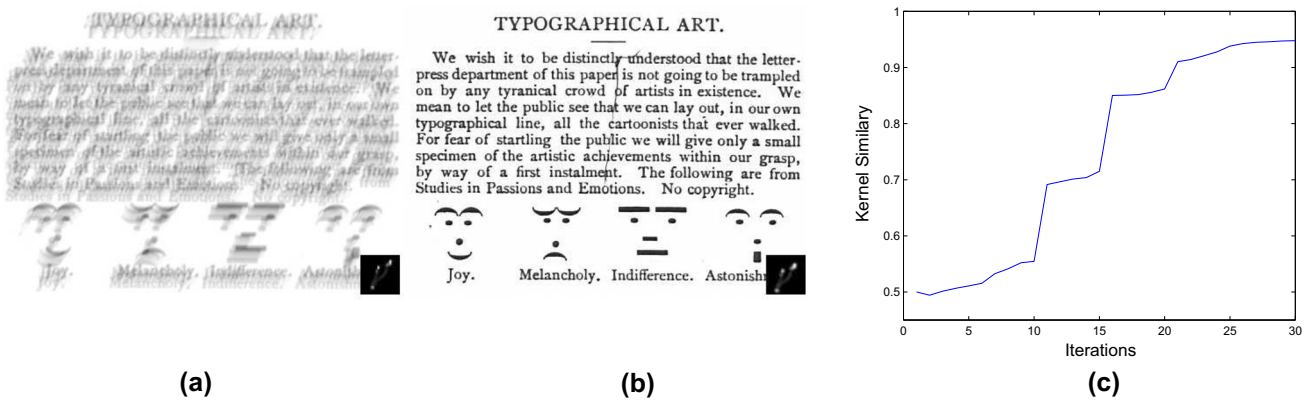


Fig. 5 Convergence of our approach. **a** Blurred image and blur kernel, **b** the results of the proposed method, **c** similarity between estimated kernel and real kernel

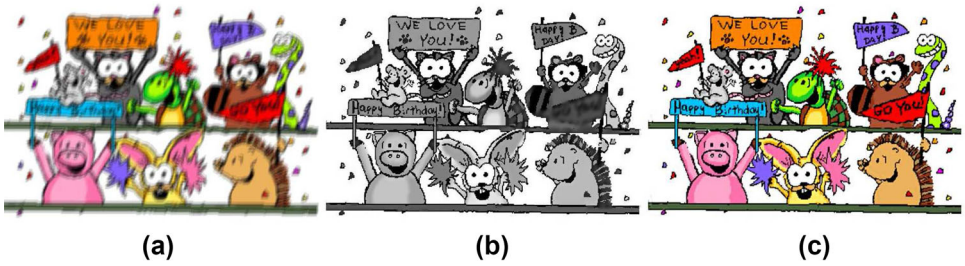


Fig. 6 An example with non-document image. **a** Observation image, **b** intermediate latent image, **c** the result of final deconvolution

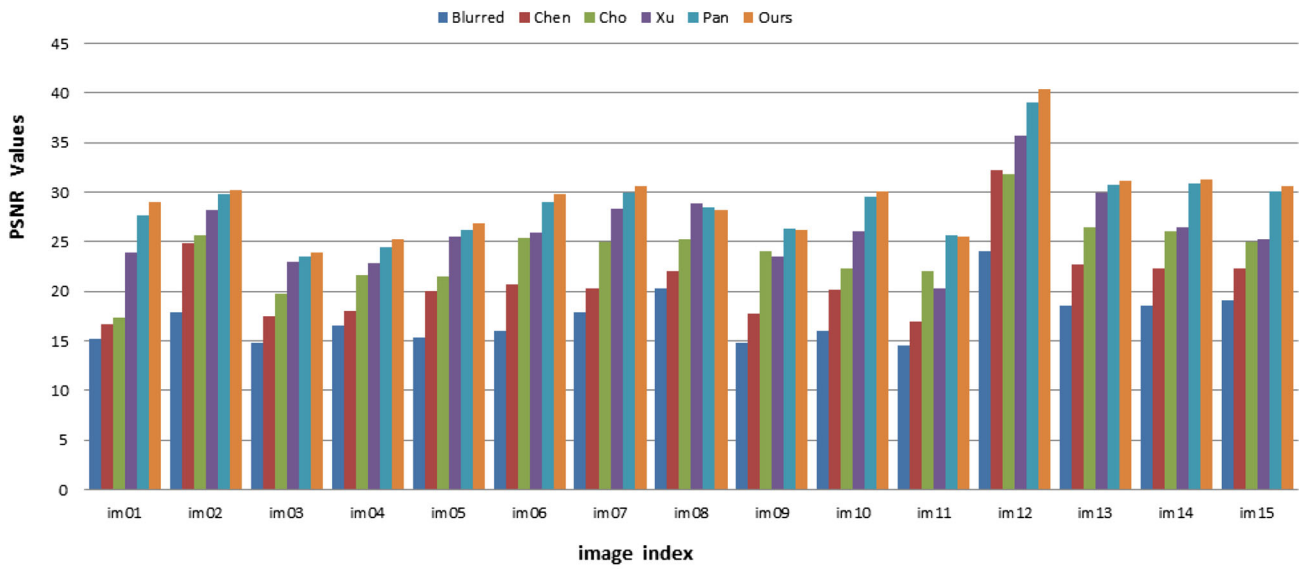


Fig. 7 Quantitative comparison on the dataset in terms of PSNR. The numbers below the horizontal axis denote the image index, and the average PSNR values of all the images are shown in the rightmost column. Our method performs the best in the most cases

mented the proposed approach with MATLAB 2010b in a computer, which runs Windows 7 64 bit version with Intel Core i5 CPU and 8 GB RAM. The proposed method takes less than 35 s for 255×255 image. In all the experiments, we empirically initialize $n_{iter} = 4$ in Algorithm 1, $\lambda = 2e^{-4}$, $\sigma_s = 2$, $\sigma_r = 0.1$ in Algorithm 2, and set $\beta = 2$, $\eta = 0.2$, NIter $= 15$, $\text{Max}_{\text{iter}} = 5$, Levels $= \lfloor -2\log_2(5/\chi) \rfloor$ in Algorithm 3, respectively, where χ is maximum blur size. It performs well for most cases.

5.1 Document images

We compare our method with four competitive algorithms: Chen et al. [4], Cho et al. [30], Xu et al. [44], and Pan et al. [32].

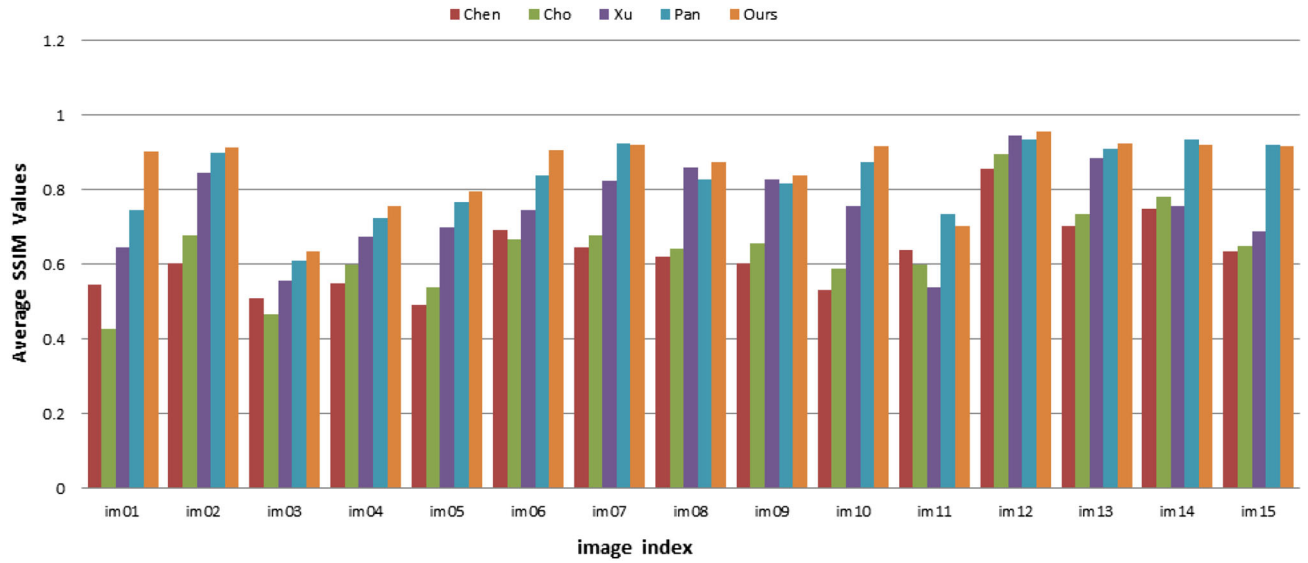


Fig. 8 Quantitative comparison on the dataset in terms of SSIM. The numbers below the horizontal axis denote the image index, and the average SSIM values of all the images are shown in the rightmost column

Fig. 9 Comparisons on the synthesized im01 image. **a** Blurred image and ground-truth PSF, **b** Chen et al. [4], **c** Cho et al. [30], **d** Xu et al. [44], **e** Pan et al. [32], **f** our results

5.1.1 Synthetic images

We use a dataset from [32], which contains 15 ground-truth text images and eight blur kernels extracted from [46]. Thus, we can generate 120 different blurred images. In addition, all images in this dataset are added 0.5% additive Gaussian noises to the blurry images to model sensor noises for better tests on deblurring effects.

For each clear image, the average PSNR and SSIM are computed for the blurred text images generated by different PSFs and compared among different deblurring algorithms in Figs. 7 and 8, respectively. From Figs. 9, 10, 11, and 12, we

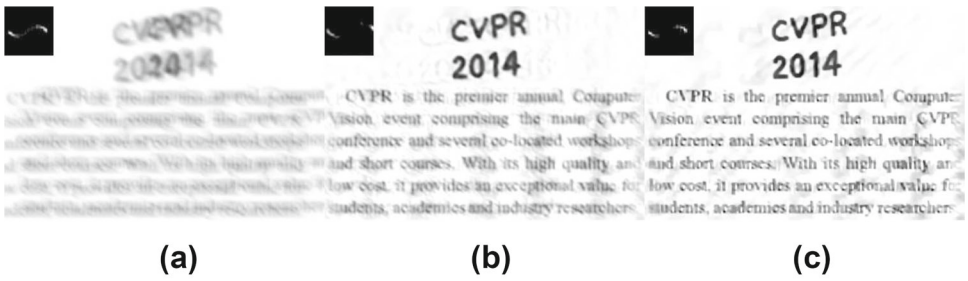


Fig. 10 Deblurred images with different algorithms. **a** Blurred image and ground-truth PSF, **b** Chen et al. [4], **c** Cho et al. [30], **d** Xu et al. [44], **e** Pan et al. [32], **f** our results

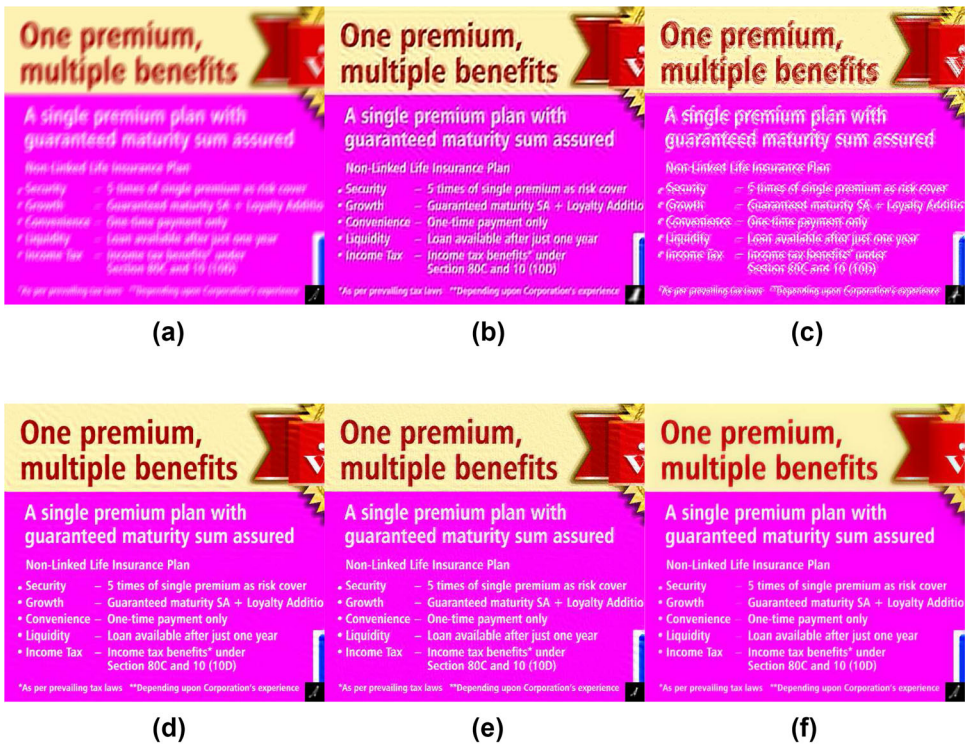


Fig. 11 Comparisons on the synthesized im07 image blurred by 02 kernel. **a** Blurred image and ground-truth PSF, **b** Chen et al. [4], **c** Cho et al. [30], **d** Xu et al. [44], **e** Pan et al. [32], **f** our results

show four deblurring results of synthesized blurry images, where the ground-truth blur kernel is known.

In order to show a general view of the robustness of the method compared to the state-of-the-art methods, we present the average PSNR of all the images in Table 1.

Table 2 compares the processing times of our method and the state-of-the-art methods.

In Fig. 9, image deblurring methods [30,32,44] generate the noisy results. In contrast, method [4] does not remove the blur well, although it is less noisy. Since more prior knowledge of the document image is harnessed in our estimation model, we generate a more accurate point spread function and sharper latent image. It is noticeable that in all of the following figures, notorious rings caused by deconvolution are successfully suppressed by our approach.

In Fig. 10, the natural image deblurring method [44] obtains a noisy result. Chen et al.'s method [4] obtains a blurry result. [30] and [32] work well, but the results still

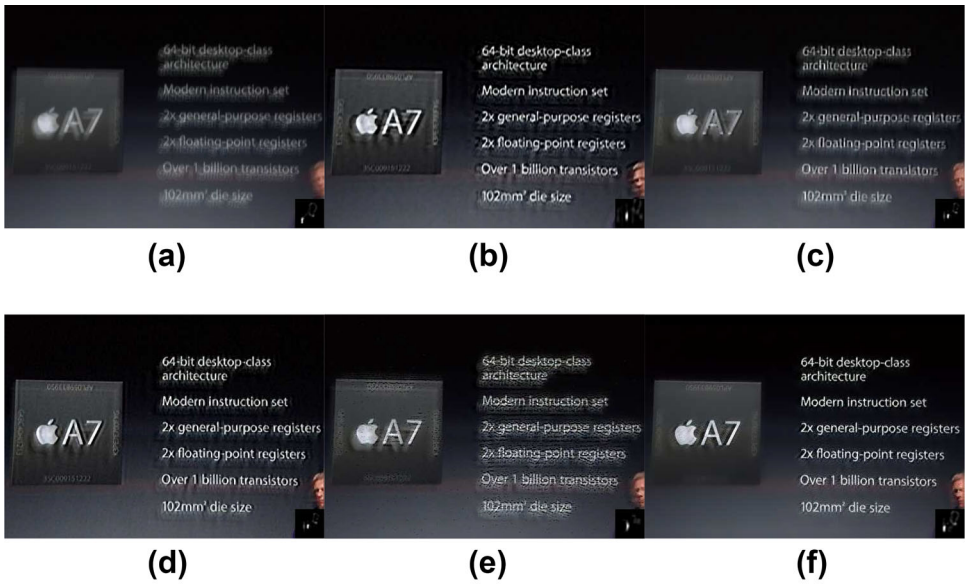


Fig. 12 Comparisons on the synthesized im08 image blurred by 07 kernel. **a** Blurred image and ground-truth PSF, **b** Chen et al. [4], **c** Cho et al. [30], **d** Xu et al. [44], **e** Pan et al. [32], **f** our results

Table 1 Accuracy of deblurred text images in dataset [32]

	Blurry images	[4]	[30]	[44]	[32]	Ours
Average PSNR	17.35	23.97	24.13	26.21	28.80	29.12

Table 2 Processing time comparison

Size image	Processing time (s)				
	[4] MATLAB	[30] C++	[44] MATLAB	[32] MATLAB	Ours MATLAB
255 × 255	12.72	39.76	1.44	37.40	34.62
512 × 512	62.04	186.91	3.68	169.58	159.52
1024 × 1024	276.71	781.24	12.27	726.79	713.80

contain noise and are not clear. Our algorithm also produces a clean and sharp result.

In Figs. 11 and 12, we show deblurring results using more text images. As the RBF-based methods are able to preserve the salient edges, the deblurred results are sharper with fewer ringing artifacts.

In some cases, such as Fig. 10, our estimated kernels look similar to those obtained by Pan's algorithm [32]. But our deblurred images look better. We provide a quantitative analysis of kernel estimates for both methods using the kernel similarity metric [Eq. (17)], which is shown in Table 3.

For the results shown in Table 3, our method performs well in terms of the kernel similarity metric. The values of [32] are smaller than ours. In Fig. 13, we show the average PSNR values obtained by the same non-blind deblurring algorithm, which can also be used for quantitative analysis of kernel estimates. Overall, our algorithm performs well in the most cases.

Table 3 Kernel similarity comparison using the example shown in Figs. 9 and 10

	[4]	[30]	[44]	[32]	Ours
Figure 9	0.6459	0.6983	0.7969	0.7625	0.9173
Figure 10	0.7120	0.7537	0.8644	0.9161	0.9215

5.1.2 Real images

The proposed method is also compared with recently developed image deblurring approaches for real images. An experiment in [30] is used, and the results are shown in Fig. 14. Method [4] generates a blurry and noisy image and misses some details. Although the competitive algorithm [30] works well, from Fig. 14c, one can find that the blur is not removed completely. Method [44] does not obtain a satisfactory result. Compared with [32], our results are sharper and have few artifacts.

Figure 15 shows the experimental results of deblurring optical motion-blurred images which contain complex texts.

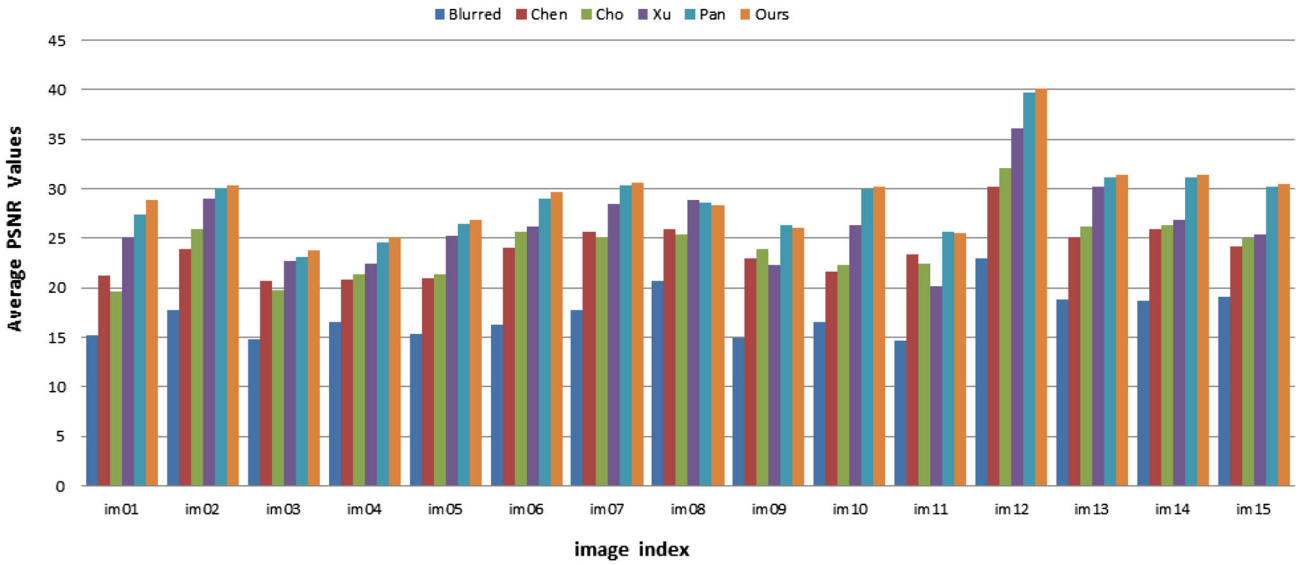


Fig. 13 Quantitative comparison on the dataset using the same non-blind deblurring algorithm. The numbers below the horizontal axis denote the image index, and the average PSNR values of all the images are shown in the rightmost column

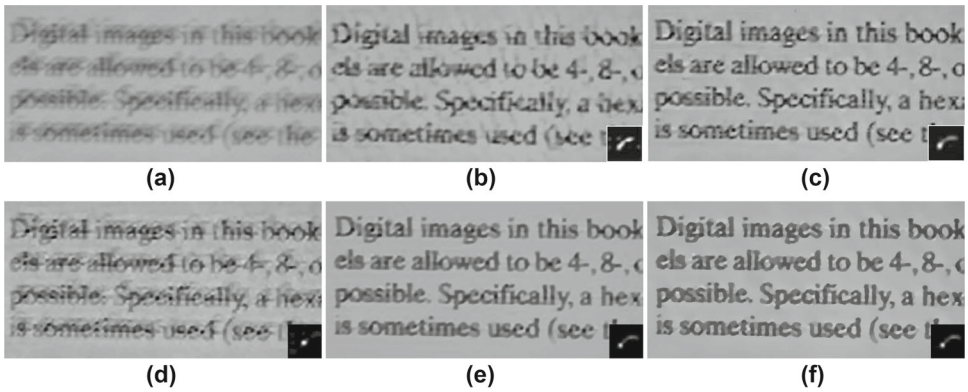


Fig. 14 Deblurred images with different algorithms. **a** Blurred image, **b** Chen et al. [4], **c** Cho et al. [30], **d** Xu et al. [44], **e** Pan et al. [32] result, **f** our results

Again, our estimation results are clearer. White background regions in our estimated images are smoother, and the rings are unnoticeable. There are textures and color information as well as documents in the image, and the proposed method restores them well. In Fig. 15b and d, we can see that the results obtained by [4] and [44] are still blurry and the deblurred image of [30] shown in Fig. 15c contains some ringing artifacts. Our result shows sharper texts compared to them. Pan et al.'s approach [32] obtains similar results as ours. In Fig. 15, we can see that the clear image is smooth, and the edge is not strong. In addition, the distribution of text regions is uniform. For this case, the RBF and L_0 gradient minimization, which are used in Pan et al.'s approach [32], produce similar smoothing result.

5.2 Natural images

In this subsection, an example with the complex image containing rich details is shown. We compare with two competitive natural image deblurring algorithms: Xu et al. [44] and Krishnan et al. [47].

The example is presented in Fig. 16 where the picture contains rich details. Figure 16b shows the result of [47], which contains noise and ringing artifacts, and the estimation of the point spread function is not good. In Fig. 16c, although method [44] handles this image well, the estimated latent image is too smooth, losing some details, such as the "stars" in the image. In contrast, in Fig. 16d, our approach obtains the blur kernel and the deblurred image (clear text, sharp edges) well.

5.3 Compared with more methods

There are quite a lot of methods proposed recently about deblurring. In this subsection, we compare the performance between the proposed method and two latest methods presented in 2017: Gong et al. [15] and Yan et al. [16].

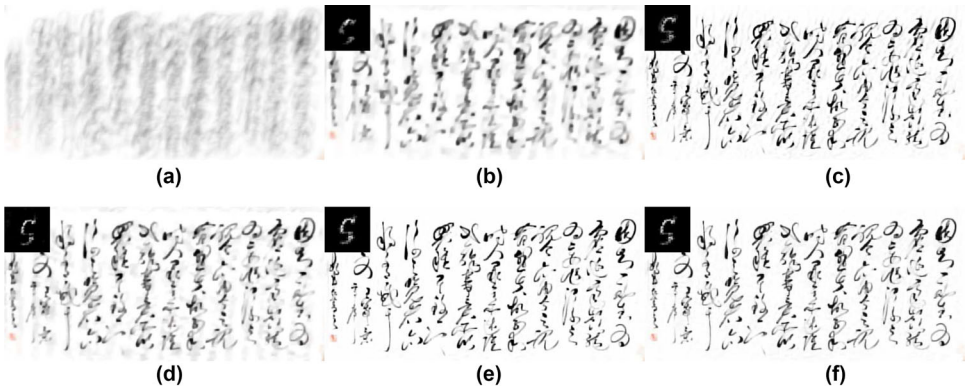


Fig. 15 Deblurred images with different algorithms. **a** Blurred image, **b** Chen et al. [4], **c** Cho et al. [30], **d** Xu et al. [44], **e** Pan et al. [32] result, **f** our results

Fig. 16 Visual quality comparison of image deblurring. **a** Blurred image, **b** Krishnan et al. [47], **c** Xu et al. [44], **d** our results

Table 4 Comparison in average PSNRs on parts of dataset as proposed in [32]

	Blurry images	[15]	[16]	Ours
Average PSNR	17.96	30.16	30.48	30.71

Similar to Table 1, we conduct experiments to compare with the two competitive methods on parts of dataset from [32], which contains seven ground-truth text images (im09–im15) and eight blur kernels. Thus, we can generate 56 different blurred images. The average PSNRs are listed in Table 4. We show the results obtained by the three methods in Fig. 17, running on the blurry image im15 with kernel 4 as proposed in [32].

Figure17a is the blurry input image and the truth kernel. The deblurring result using [15] is shown in Fig. 17b. The result using the extreme channels prior-based sparse regularization [16] is shown in Fig. 17c. As it is shown in Fig. 17b–c, the results deblurred by [15] and [16] are similar and restore the point spread function well. Our result which is shown in Fig. 17d contains less artifacts compared with [16] and is more clear than [15].

6 Conclusion

A simple but effective approach to document image restoration is proposed in this work. Due to the specific properties of document images, rolling bilateral filter is proposed and applied to text image deblurring. Then, a powerful deblur-

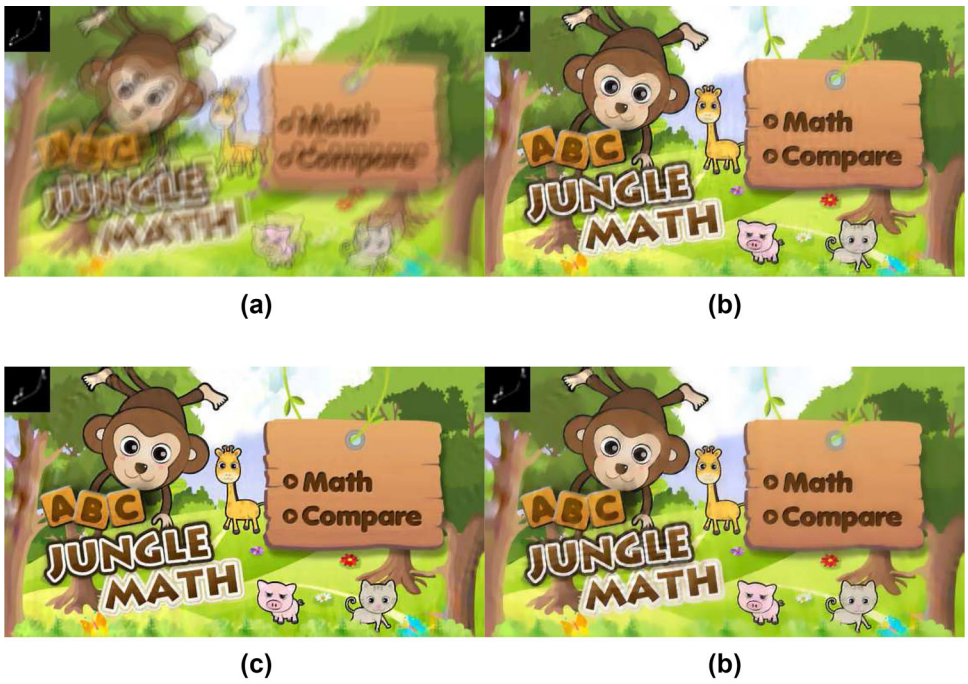


Fig. 17 Visual quality comparison of image deblurring. **a** Blurred image and truth kernel, **b** Gong et al. [15], **c** Yan et al. [16], **d** our results

ring algorithm based on rolling bilateral filter is presented for estimating the latent image and blur kernel. Any trivial preprocessing approaches are not required, such as image segmentation. Finally, we propose an effective image non-blind deconvolution algorithm, which helps to preserve details. In the future, we will extend our approach to non-uniform text image restoration.

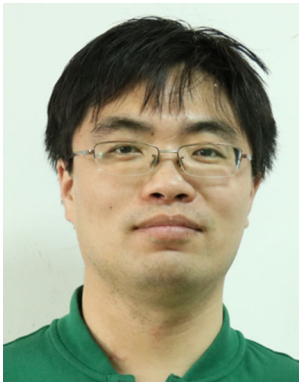
6.1 Limitations

As a limitation, our proposed method is not well suited for removing large noise because the mixed geometric feature and noise cannot be distinguished by RBF. Another limitation is that our deblurring results might produce ringing artifacts if the image contains a lot of texture.

References

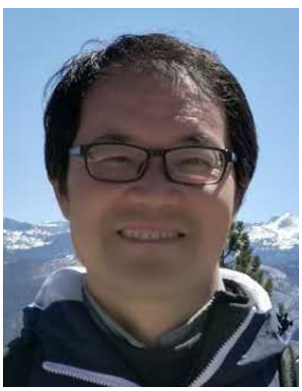
1. Brown, M.S., Seales, W.B.: Image restoration of arbitrarily warped documents. IEEE Trans. PAMI **26**(10), 1295–1306 (2004)
2. Likforman-Sulem, L., Darbon, J., Smith, E.B.: Pre-processing of degraded printed documents by non-local means and total variation. In: ICDAR (2009)
3. Kenneth, L.: Text image deblurring by high-probability word selection. US Patent 6282324, Publication Date (2001)
4. Chen, X., He, X., Yang, J.: An effective document image deblurring algorithm. In: Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on, pp. 369–376. IEEE (2011)
5. Pan, J., Hu, Z., Su, Z.: Deblurring text images via L_0-regularized intensity and gradient prior. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2901–2908 (2014)
6. Jiang, X., Yao, H., Zhao, S.: Text image deblurring via two-tone prior. Neurocomputing **242**, 1–14 (2017)
7. Fergus, R., Singh, B., Hertzmann, A.: Removing camera shake from a single photograph. In: ACM Transactions on Graphics (TOG), vol. **25**. No. 3, pp. 787–794. ACM (2006)
8. Xu, Y., Hu, X., Peng, S.: Sharp image estimation from a depth-involved motion-blurred image. Neurocomputing **171**, 1185–1192 (2016)
9. Song, C., Deng, H., Gao, H., Zhang, H., Zuo, W.: Bayesian non-parametric gradient histogram estimation for texture-enhanced image deblurring. Neurocomputing **197**, 95–112 (2016)
10. Ren, W., Cao, X., Pan, J.: Image deblurring via enhanced low rank prior. IEEE Trans. Image Process. **25**, 3426–3437 (2016)
11. Pan, J., Sun, D., Pfister, H.: Blind image deblurring using dark channel prior. In: CVPR 2016, pp. 1–8 (2016)
12. Dong, J., Pan, J., Su, Z., Yang, M.-H.: Blind image deblurring with outlier handling. In: IEEE International Conference on Computer Vision (ICCV), pp. 2478–2486 (2017)
13. Wieschollek, P., Hirsch, M., Scholkopf, B., Lensch, H.P.A.: Learning blind motion deblurring. IEEE International Conference on Computer Vision (ICCV), pp. 231–240 (2017)
14. He, K., Zhang, X., Ren, S., et al.: Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 770–778 (2016)
15. Gong, D., Tan, M., Zhang, Y., et al.: Self-paced kernel estimation for robust blind image deblurring. In: IEEE International Conference on Computer Vision(ICCV), pp. 1661–1670 (2017)
16. Yan, Y., Ren, W., Guo, Y., et al.: Image deblurring via extreme channels prior. In: Proceedings of the IEEE Conference Computer Vision and Pattern Recognition (CVPR), pp. 4003–4011 (2017)
17. Nah, S., Kim, T.H., Lee, K.M.: Deep multi-scale convolutional neural network for dynamic scene deblurring. In: Proceedings of the IEEE Conference Computer Vision and Pattern Recognition (CVPR) (2017)
18. Cho, S., Lee, S.: Fast motion deblurring. ACM Trans. Graph. **28**, 145:1–145:8 (2009)

19. Levin, A., Weiss, Y., Freeman, B., Durand, F.: Efficient approximations to the marginal likelihood in blind deconvolution. In: Proceedings of the CVPR 2011, pp. 1–8 (2011)
20. Xu, L., Jia, J.: Two-phase kernel estimation for robust motion deblurring. In: ECCV 2010, pp. 157–170 (2010)
21. Shan, Q., Jia, J., Agarwala, A.: High-quality motion deblurring from a single image. ACM Trans. Graph. **27**, 73:1-73:10 (2008)
22. Liu, S., Wang, H., Wang, J.: Automatic blur-kernel-size estimation for motion deblurring. Vis. Comput. **31**(5), 733–746 (2015)
23. Li, T.H., Lii, K.S.: A joint estimation approach for two-tone image deblurring by blind deconvolution. IEEE Trans. Image Process. **11**, 847–858 (2002)
24. Donaldson, K., Myers, G.: Bayesian super-resolution of text in video with a text-specific bimodal prior. In: Proceedings of the CVPR, pp. 1188–1195 (2005)
25. Banerjee, J., Namboodiri, A., Jawahar, C.: Contextual restoration of severely degraded document images. In: Proceedings of the CVPR, pp 517–524 (2009)
26. Fang, H., Shi, Y., Pan, D.: Iteratively reweighted blind deconvolution for passive millimeter-wave images. Signal Process. **138**, 182–194 (2017)
27. Joshi, N., Szeliski, R., Kreigman, D.: PSF estimation using sharp edge prediction. In: Proceedings of the CVPR, pp. 1–8 (2008)
28. Li, J., Lu, W.: Blind image motion deblurring with L_0-regularized priors. J. Vis. Commun. Image Represent. **40**, 14–23 (2016)
29. Lou, Y., Bertozzi, L.A., Soatto, S.: Direct sparse deblurring. J. Math. Imaging Vis. **39**(1), 1–12 (2011)
30. Cho, H., Wang, J., Lee, S.: Text image deblurring using text-specific properties. In: ECCV, pp. 524–537 (2012)
31. Epshtein, B., Ofek, E., Wexler, Y.: Detecting text in natural scenes with stroke width transform. In: CVPR, pp. 2963–2970 (2010)
32. Pan, J., Hu, Z., Su, Z.: L_0-Regularized intensity and gradient prior for deblurring text images and beyond. IEEE Trans. Pattern Anal. Mach. Intell. **39**, 342–355 (2016)
33. Cao, X., Ren, W., Zuo, W., Guo, X., Foroosh, H.: Scene text deblurring using text-specific multiscale dictionaries. IEEE Trans. Image Process. **24**(4), 1302–1314 (2015)
34. Paris, S., Hasinoff, S.W., Kautz, J.: Local Laplacian filters: edge-aware image processing with a Laplacian pyramid. ACM Trans. Graph. **30**, 68 (2011)
35. Tomasi, C., Manduchi, R.: Bilateral filtering for gray and color images. In: IEEE International Conference on Computer Vision, pp. 839–846 (1998)
36. He, K., Sun, J., Tang, X.: Guided image filtering. IEEE Trans. Pattern Anal. Mach. Intell. **35**(6), 1397–1409 (2013)
37. Buades, A., Le, T.M., Morel, J.M., Vese, L.A.: Fast cartoon+texture image filters. IEEE Trans. Image Process. **19**(8), 1978–1986 (2010)
38. Xu, L., Lu, C., Xu, Y., Jia, J.: Image smoothing via L_0 gradient minimization. ACM Trans. Graph. **30**, 174 (2011)
39. Zhang, Q., Shen, X.Y., Xu, L. Jia, J.Y.: Rolling guidance filter. In: European Conference on Computer Vision, pp. 815–830. Springer International Publishing (2014)
40. Petschnigg, G.: Digital photography with flash and no-flash image pairs. In: SIGGRAPH, pp. 664–672 (2004)
41. Wang, Y., Yang, J., Yin, W., Zhang, Y.: A new alternating minimization algorithm for total variation image reconstruction. SIAM J. Imaging Sci. **1**(3), 248–272 (2008)
42. Zhong, L., Cho, S., Metaxas, D., Paris, S., Wang, J.: Handling noise in single image deblurring using directional filters. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 612–619 (2013)
43. Yang, H., Zhang, Z., Guan, Y.: An adaptive parameter estimation for guided filter based image deconvolution. Signal Process. **138**, 16–26 (2017)
44. Xu, L., Zheng, S., Jia, J.: Unnatural L_0 sparse representation for natural image deblurring. In: CVPR, pp. 1107–1114 (2013)
45. Hu, Z., Yang, M.H.: Good regions to deblur. In: European Conference on Computer Vision, pp. 59-72. Springer, Berlin (2012)
46. Levin, A., Weiss, Y., Durand, F., Freeman, W.T.: Understanding and evaluating blind deconvolution algorithms. In: CVPR (2009)
47. Krishnan, D., Tay, T., Fergus, R.: Blind deconvolution using a normalized sparsity measure. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2657–2664 (2011)



Hang Yang received his B.S. and Ph.D. degrees in mathematics from the Jilin University in 2007 and 2012, respectively. He is currently an Associate Research Fellow at the Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Science. His current research interests include image deblurring and visual tracking.

Zhongbo Zhang received his M.S. and Ph.D. degrees in mathematics from the Jilin University in 2000 and 2003, respectively. He is currently an Associate Professor at the School of Mathematics, Jilin University, China. His current research interests include image denoising, image deconvolution, and pattern recognition.



Yujing Guan received his B.S. and Ph.D. degrees in mathematics from the Jilin University in 1992 and 2003, respectively. He is currently a Professor at the School of Mathematics, Jilin University, China. His current research interests include image processing and convolution neural network.