Exploring Aligned Complementary Image Pair for Blind Motion Deblurring

Wen Li[1,2], Jun Zhang[1,2], and Qionghai Dai[3]

School of Electronics and Information Engineering, Beihang University[1],
National Key Laboratory of CNS/ATM[2],
Department of Automation, Tsinghua University[3]

`liwen.buaa@gmail.com`, `buaazhangjun@vip.sina.com`, `qhdai@tsinghua.edu.cn`

Abstract

Camera shake during long exposure is ineluctable in light-limited situations, and results in a blurry observation. Recovering the blur kernel and the latent image from the blurred image is an inherently ill-posed problem. In this paper, we analyze the image acquisition model to capture two blurred images simultaneously with different blur kernels. The image pair is well-aligned and the kernels have a certain relationship. Such strategy overcomes the challenge of blurry image alignment and reduces the ambiguity of blind deblurring. Thanks to the aided hardware, the algorithm based on such image pair can give high-quality kernel estimation and image restoration. The experiments on both synthetic and real images demonstrate the effectiveness of our image capture strategy, and show that the kernel estimation is accurate enough to restore superior latent image, which contains more details and fewer ringing artifacts.

1. Introduction

Camera motion blur is one of the prime causes for poor image quality. In many imaging situations, a long shutter speed is required for adequate exposure. However, the relative motion between the camera and the scene during exposure is ineluctable and results in a blurry image. Using a faster exposure with large aperture and high ISO may be able to bring some improvements, while leading to reduced depth-of-field and more sensor noise.

As for capturing the static scenes, the image blur caused by camera shake can be well described by a single, spatial-invariant blur kernel. The blurring process is commonly modeled as a convolution of the latent image $\mathbf{I}$ with the blur kernel $\mathbf{K}$, plus noise $\mathbf{n}$:

$$\mathbf{B} = \mathbf{I} \otimes \mathbf{K} + \mathbf{n}, \tag{1}$$

where $\otimes$ is the convolution operator.

In general, recovering the latent image from the blurry observation is a deconvolution problem, which is further
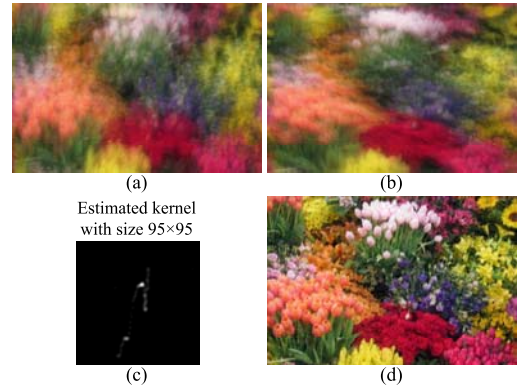

Figure 1: (a)-(b) Two blurred images. (c)-(d) The estimated kernel and the deblurred image obtained by our method.

separated into the non-blind and blind cases. If the blur kernel $\mathbf{K}$ is known as a prior or computed in other ways, recovering the image $\mathbf{I}$ from the blurred image $\mathbf{B}$ is the so-called non-blind deconvolution problem. In the case of blind deconvolution problem, both the blur kernel and the latent image are unknown, which is a more challenging task.

1.1. Previous work

For non-blind deconvolution, some traditional algorithms, such as Wiener filter [22] and Richardson-Lucy (RL) deconvolution [14], are widely used and known to be simple and efficient. However, the estimation suffers from unpleasant noise and ringing artifacts. The sparse natural image prior is used in [12] to produce sharper edges and reduce the undesirable ringing artifacts. In [26], the progressive deconvolution is performed with a coarse-to-fine scheme. The sharp edges restored in the coarser level are used as a guide in the finer level, so that more and more image details and sharp edges can be restored.

In recent years, many researchers have been working on the general blind motion deblurring problem using single blurred image. Since that it is a highly ill-posed problem, all the algorithms make strong prior assumptions about the blur kernel and the image to be recovered. The blind deconvolution is formulated as a joint minimization problem with

regularizations:

$$\mathbf{I}^*, \mathbf{K}^* = \min_{\mathbf{I},\mathbf{K}} \{\|\mathbf{I} \otimes \mathbf{K} - \mathbf{B}\|_2 + \lambda_K E_K(\mathbf{K}) + \lambda_I E_I(\mathbf{I})\}. \quad (2)$$

Here $\|\mathbf{I} \otimes \mathbf{K} - \mathbf{B}\|_2$ is the data fitting term, the ℓ_2 norm of the image intensity is the common choice, while the image derivatives are also used in [7, 17]. The image deconvolution in [23, 24] exploits ℓ_1 norm data fitting term to make it robust to the impulsive noise. $E_K(\mathbf{K})$ and $E_I(\mathbf{I})$ are the regularization terms on the blur kernel and the latent image. For kernel estimation, sparsity and continuity priors are commonly used to remove the ambiguity. The approaches in [3, 5] maximize the sparsity and spatial smooth of the blur kernel under curvelet system. In [7, 11, 23], the edges of large-scale objects are selected to perform high-quality kernel estimation. The Iterative Support Detection (ISD) algorithm is employed in [23] to adaptively impose the sparsity constraint on the blur kernel. For image deconvolution, different algorithms have various options, such as ℓ_1 norm, ℓ_2 norm, Total Variation (TV) norm and its variations. In [8], the zero-mean mixture-of-Gaussians model is used to approximate the heavy-tailed distribution of image gradients. Shan *et al.* [17] proposed a unified probabilistic model to suppress visual artifacts. The representation of the image under framelet system is introduced in [3, 4] to constraint the sparsity of the solution.

To overcome the ambiguity of blind motion deblurring, many algorithms using multiple images have been proposed. In [4, 5, 6, 16], the images are obtained by consecutive shooting with the same exposure time, which provide more information of the scene. Yuan *et al.* [25] introduced a blurred/noisy image pair for deblurring. The sharp edges in the noisy image are used to estimate the blur kernel. Since the camera pose is varying during the imaging process, the pre-processing of accurate image alignment is very important. In previous works, blind motion deblurring using multiple images assumes that the blurry images are well-aligned. However, accurate automatic alignment of images blurred by different kernels is difficult to achieve by existing methods.

Another group of methods employ specific hardware to control the capturing process. Ben-Ezra and Nayar [2] combined a high-resolution still camera and a low-resolution video camera to form a hybrid camera system. The motion blur point spread function (PSF) can be estimated from the image sequence by optical flow computation. A similar system is used in [18], and the low-resolution image sequence also provides super-resolution-based constraints for image deconvolution. In [10], the inertial measurement sensors are used to capture the camera motion during exposure.

In summary, natural images not always meet the strong priors made for the single-image deblurring and the requirement on accurate alignment of blurry images limits the performance of multiple-image deblurring algorithm. Using some aided hardware can capture more information during exposure, which makes significant benefits to solve the highly ill-posed problem.

1.2. Our approach

The goal of this paper is to analyze and propose the image capture strategy, then using a consumer camera with aided hardware can capture two well-aligned blurry images simultaneously of the same scene. Our approach overcomes the challenge of blurry image alignment. Moreover, the two blur kernels are different and have a certain relationship, so that more information of the scene is captured and the inherent ambiguity of blind deblurring is reduced. With such image pair, our algorithm can handle large-scale and complex motion blur and give more accurate kernel estimation. The ringing artifacts are suppressed obviously and more sharp edges are restored in the recovered image.

2. Problem formulation and analysis

In this section, we analyze the image acquisition model and propose some aided hardware to capture two aligned blurry images simultaneously with different motion blurs. Since the aided hardware is specially designed, the two blur kernels have a certain relationship. Such relationship can be used as a constraint in the kernel estimation and also bring benefits to the image deconvolution, which will be discussed in Section 2.2. A instantiation of the image capture system is described in Section 2.3.

2.1. Camera motion blur

The image acquisition model designed in our work is using some aided hardware to establish a mapping $f : \mathbf{R}^n \mapsto \mathbf{R}^n$ in the image space. Then the captured image pair has the relationship of $\mathbf{I}_2 = f(\mathbf{I}_1)$. When the hand-held camera is shaken, the image pair will suffer from the same motion blur. Therefore, we have the observation model:

$$\mathbf{B}_1 = \mathbf{I} \otimes \mathbf{K} + \mathbf{n}_1, \quad \mathbf{B}_2 = f(\mathbf{I}) \otimes \mathbf{K} + \mathbf{n}_2. \quad (3)$$

Here the mapping f is required to be a bijection. For any $\mathbf{a}, \mathbf{b} \in \mathbf{R}^n$, the inverse mapping f^{-1} should satisfy that

$$f^{-1}(\mathbf{a} \otimes \mathbf{b}) = f^{-1}(\mathbf{a}) \otimes f^{-1}(\mathbf{b}). \quad (4)$$

Then we have

$$f^{-1}(\mathbf{B}_2) = f^{-1}(f(\mathbf{I}) \otimes \mathbf{K} + \mathbf{n}_2) = \mathbf{I} \otimes f^{-1}(\mathbf{K}) + \mathbf{n}_2. \quad (5)$$

Moreover, f should also be an isometric isomorphism or ε-isometry, such that for any $\mathbf{a}, \mathbf{b} \in \mathbf{R}^n$, the following holds

$$|d(f(\mathbf{a}), f(\mathbf{b})) - d(\mathbf{a}, \mathbf{b})| < \varepsilon, \quad (6)$$

where $d(\cdot)$ is the distance metric and ε is a given positive real number. So the noise will not be amplified by the inverse mapping.

After transformation, the problem is equivalent to

$$\begin{cases} \mathbf{B}_1 = \mathbf{I} \otimes \mathbf{K}_1 + \mathbf{n}_1, \\ f^{-1}(\mathbf{B}_2) = \mathbf{I} \otimes f^{-1}(\mathbf{K}_1) + \mathbf{n}_2 = \mathbf{I} \otimes \mathbf{K}_2 + \mathbf{n}_2. \end{cases} \quad (7)$$

The observed image pair $\mathbf{B}_1$ and $f^{-1}(\mathbf{B}_2)$ can be considered as the convolutions of the same image with different blur kernels, just like the input in many dual motion deblurring algorithms. While the unknown blur kernels $\mathbf{K}_1$ and $\mathbf{K}_2$ are not arbitrary, but have the relationship

$$\mathbf{K}_1 = f(\mathbf{K}_2). \quad (8)$$

Then blind motion deblurring is performed based on the blurry image pair. Similar to the optimization problem of single-image deblurring in Eq. (2), the latent image and two blur kernels can be estimated by minimizing the following energy function:

$$\begin{aligned} E(\mathbf{I}, \mathbf{K}_1, \mathbf{K}_2) = & \Sigma_{i=1}^2 \| \mathbf{I} \otimes \mathbf{K}_i - \mathbf{B}_i \|_2 \\ & + \lambda_K \Sigma_{i=1}^2 E_K(\mathbf{K}_i) + \lambda_I E_I(\mathbf{I}). \end{aligned} \quad (9)$$

The problem can be divided into two parts: blur kernel estimation and image deconvolution. The $\mathbf{I}$ and $\mathbf{K}_i$ are alternatively optimized by solving the minimizations with the physical constraints:

$$\widehat{\mathbf{K}}_i = \arg \min_{\mathbf{K}_i} \{ \| \mathbf{I} \otimes \mathbf{K}_i - \mathbf{B}_i \|_2 + \lambda_K E_K \mathbf{K}_i \} \quad (10)$$

$$s.t. \quad \widehat{\mathbf{K}}_i(j) \geqslant 0, \quad \sum_j \mathbf{K}(j) = 1;$$

$$\widehat{\mathbf{I}} = \arg \min_I \{ \Sigma_{i=1}^2 \| \mathbf{I} \otimes \mathbf{K}_i - \mathbf{B}_i \|_2 + \lambda_I E_I(\mathbf{I}) \} \quad (11)$$

$$s.t. \quad \widehat{\mathbf{I}}(j) \geqslant 0.$$

Here the relationship between the two kernels can be used as a constraint. For kernel estimation, the relationship given by Eq. (3) is substituted. We rewrite Eq. (10) as

$$\begin{aligned} \widehat{\mathbf{K}} = \arg \min_K \{ & \| \mathbf{I} \otimes \mathbf{K} - \mathbf{B}_1 \|_2 \\ & + \| f(\mathbf{I}) \otimes \mathbf{K} - \mathbf{B}_2 \|_2 + \lambda_K E_K(\mathbf{K}) \}. \end{aligned} \quad (12)$$

Similarly, substituting Eq. (7) and (8), we get

$$\begin{aligned} \widehat{\mathbf{I}} = \arg \min_I \{ & \| \mathbf{I} \otimes f(\mathbf{K}) - \mathbf{B}_1 \|_2 \\ & + \| \mathbf{I} \otimes \mathbf{K} - f^{-1}(\mathbf{B}_2) \|_2 + \lambda_I E_I(\mathbf{I}) \}. \end{aligned} \quad (13)$$

2.2. Benefits analysis

The benefits of using multiple images have been discussed in [1, 4, 16]. The proposed method used in this article has all the advantages of these works. Firstly, using two images can eliminate the inherent ambiguity between the

kernel and the image estimation. Since the blind motion deblurring is highly ill-posed, Eq. (1) has infinitely many solutions. If the blur kernel can be decomposed into a convolution of a factor and a low-pass filter, such as $\mathbf{K} = k_0 \otimes \mathbf{K}'$, then $(\mathbf{I} \otimes k_0, \mathbf{K}')$ is also a solution of motion deblurring. When the two blur kernels have the relationship as Eq. (8), the possibility that $\mathbf{K}_1$ and $\mathbf{K}_2$ have such a common factor is greatly reduced.

Secondly, two images provide more information of the scene and fill the nulls of blur kernels in the spectrum. The blur kernel is considered as a low-pass filter, so the frequency transform of it usually contains some nulls at high frequencies. The information of the scene at these frequencies is lost in the image acquisition process. In addition, the nulls in the spectrum make the motion blur non-invertible. Eq. (1) in the frequency domain is given by:

$$\mathcal{F}(\mathbf{B}) = \mathcal{F}(\mathbf{I}) \cdot \mathcal{F}(\mathbf{K}) + \mathcal{F}(\mathbf{n}), \quad (14)$$

where $\mathcal{F}(\cdot)$ denotes the Fourier transform. The image deconvolution can be simply calculated as

$$\mathcal{F}(\widehat{\mathbf{I}}) = \frac{\mathcal{F}(\mathbf{B})}{\mathcal{F}(\mathbf{K})} = \mathcal{F}(\mathbf{I}) + \frac{\mathcal{F}(\mathbf{n})}{\mathcal{F}(\mathbf{K})}. \quad (15)$$

Here $\frac{\mathcal{F}(\mathbf{n})}{\mathcal{F}(\mathbf{K})}$ is the noise in the image estimation. If $|\mathcal{F}(\mathbf{K})|$ has small values, the noise will be severely enlarged. These can be improved by introducing two images for motion deblurring. Then the deblurred image is obtained by

$$\mathcal{F}(\widehat{\mathbf{I}}) = \mathcal{F}(\mathbf{I}) + \frac{\sum_{i=1}^2 \overline{\mathcal{F}(\mathbf{K}_i)} \mathcal{F}(\mathbf{n})}{\sum_{i=1}^2 |\mathcal{F}(\mathbf{K}_i)|^2}, \quad (16)$$

where $\overline{\mathcal{F}(\cdot)}$ represents the complex conjugate operator. The nulls of one kernel are very likely to be filled by the other's. Therefore, more information of the scene can be captured in the blurry image pair, the deconvolution is jointly invertible and noise sensitivity is significantly improved.

Finally, the challenge of blurry image alignment is overcome by our image capture strategy. The Blind motion deblurring using two images obtained by consecutive shooting is sensitive to the alignment error. Most works have the assumption that the motion deblurring is performed on the well-aligned images. However, accurate automatic alignment of blurry images is quite a challenging task. In [6], the blurry images are aligned manually using Photoshop. Since the images are blurred with different complex motions, even user interactions can not give an accurate result. For our imaging system, the light is split by a beam spliter, and the spatial calibration is performed previously. The two blurry images are captured by a consumer camera simultaneously, so that the color responses and the noise levels are the same. Moreover, when the camera system is placed on a fast-moving platform, such as a moving car or airplane,
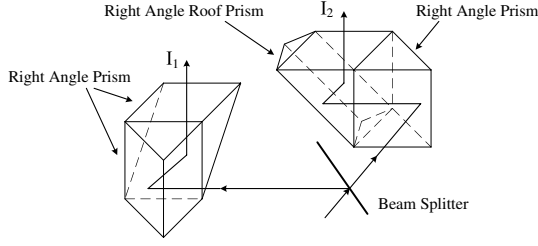
Figure 2: Diagram of the optical path.

Figure 3: Illustration of the image acquisition model.

Figure 4: (a)-(b) Log spectrum of the kernels $\mathbf{K}_1$ and $\mathbf{K}_2$. (c) The spectrum of the combination of two blur kernels.

it can not shoot consecutively to capture two blurry images with the same static scene. If the image capture strategy can be integrated with the camera, such problem can be successfully solved, since the blurred images are taken simultaneously and well-aligned.

2.3. Practical approach

The analysis above indicates that the ability of motion deblurring can be prominently improved by judiciously designing the image capture system. Any mapping f satisfies the designed rules in Eq. (4) and (6) can be used here. For example, all the reflection, translation, rotation, and the combination of them in Euclidean spaces are isometric isomorphisms and also satisfy Eq. (4). In particular, we design a prism system to rotate the image with 90 degrees, so the mapping $f(\cdot)$ is instantiated by $rot90(\cdot)$. Firstly, a beam splitter splits the light into two perpendicular output paths. Then half of the light is transmitted and pass through a right angle prism and a right angle roof prism to rotate 90 degrees with horizontal mirror, the other half is reflected and deviated by other two right angle prisms. The optical path is illustrated in Figure 2. For convenience, the prism system is constructed by the combination of several optical components, which can also be produced as a whole.

Reflection from the prism system returns two images placed side-by-side, while one is a duplicate of the other with 90-degree rotation, as shown in the first column of Figure 3. The two images can be captured by one consumer camera. When the hand-held camera is shaken, the image pair will suffer from the same motion blur. After imaging, the part II of blurred image is rotated 90 degrees inversely. Then $\mathbf{B}_1$ and $rot270(\mathbf{B}_2)$, indicated by the dashed rectangle in Figure 3, are used as the input for the subsequent deblurring process. The two images are split by a beam splitter, so that less calibration is required. The problem is equivalent to the images blurred by two different blur kernels, which have the relationship that $\mathbf{K}_1 = rot90(\mathbf{K}_2)$. As discussed above, the nulls of $|\mathcal{F}(\mathbf{K}_i)|$ are filled by each other, and $\sum_{i=1}^{2} |\mathcal{F}(\mathbf{K}_i)|^2$ contains few small values as shown in Figure 4(c). Moreover, Rav-Acha *et al.* [16] has investigated the relationship between the restoration error and the angle between the directions of two 1**D** motion blurs. They demonstrated that more information of the scene is preserved and better restoration is achieved as the blur kernels are closer to orthogonal.
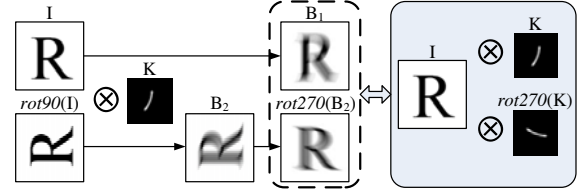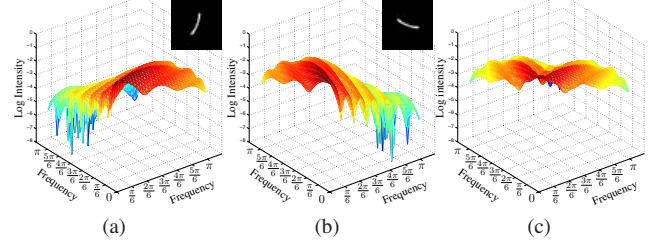
3. Numerical algorithm

In our algorithm, the kernel estimation and image deconvolution of Eq. (12) and (13) are progressively performed with the multi-scale iterative process. We build pyramids $\{\mathbf{B}_1^l\}_{l=1}^{L}$ and $\{\mathbf{B}_2^l\}_{l=1}^{L}$ of the blurry images $\mathbf{B}_1$, $\mathbf{B}_2$ using bicubic down-sampling with scale factor of $\sqrt{2}$. At the top level, $\mathbf{B}_1^L$ and $\mathbf{B}_2^L$ is the full-resolution images. The scale of the blur kernel is also down-sampled by the same manner. At the coarsest level, $\mathbf{K}_1^0$ and $\mathbf{K}_2^0$ are 3×3 kernels. In each level, the estimations are iteratively updated by three steps: salient edges selection, kernel estimation, and non-blind image deconvolution. After $\mathbf{I}$ is estimated at a coarse level, the image is up-sampled to the next level and the sharp edges are selected from it. The blur kernel obtained at the finest level is further refined using ISD method [21]. Finally, the final image deconvolution is performed to obtain the deblurred result. The overall process of our algorithm is illustrated in Figure 5, while the main parts of the proposed method are described in detail in subsequent subsections.

3.1. Salient edges prediction

The work in [23] has discussed the usefulness of edge information and demonstrated that the edges of large-scale objects are significant to the kernel estimation, while the edges smaller than the kernel scale could increase the estimation arbitrary. Similarly, the strong edges of large-scale objects are selected in our algorithm. The sharp edges are firstly selected using the blurred image. For each pixel i in the blurred image $\mathbf{B}^l$, we form a local window $\mathbf{W}_h(i)$ centered at it with size $h \times h$, which is the same as that of the corresponding blur kernel at scale l. In each window, the sum of gradient magnitudes is calculated to formulate the local prior. The usefulness of the edges is measured as

$$r(i) = \frac{\| \sum_{j \in \mathbf{W}_h(i)} \nabla \mathbf{B}(j) \|}{\sum_{j \in \mathbf{W}_h(i)} \| \nabla \mathbf{B}(j) \| + \varepsilon}, \tag{17}$$
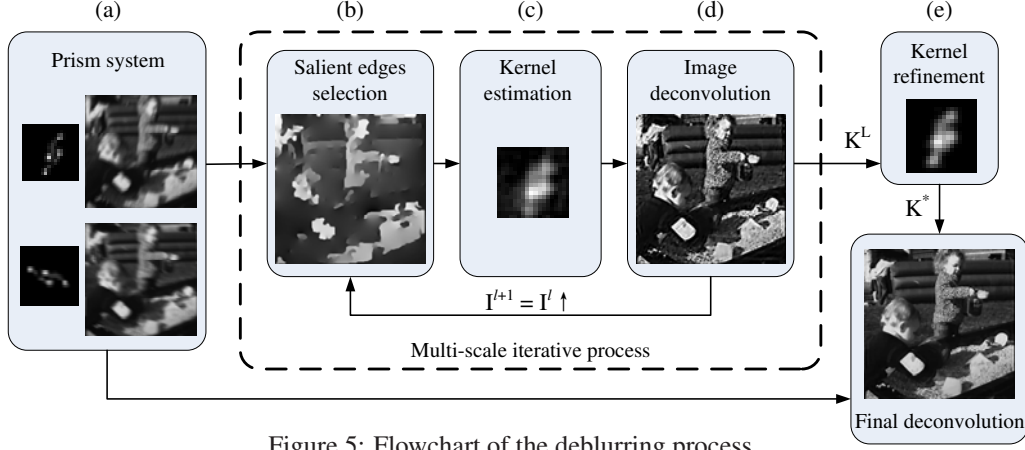
Figure 5: Flowchart of the deblurring process.

where $\nabla\mathbf{B} = (\partial_x\mathbf{B}, \partial_y\mathbf{B})'$ is the image gradient. A large value r implies that the image structure is strong in the local window. When the local region is flat or has fine texture, the gradients will be neutralized by each other, so that the value of $\|\sum_{j\in\mathbf{N}_h(i)}\nabla\mathbf{B}(j)\|$ is small. We construct a matrix $\mathbf{M}_r$ as a binary mask using a unit step function

$$\mathbf{M}_r(i) = u_{\tau_r}(r(i) - \tau_r) = \begin{cases} 1, & r(i) \geqslant \tau_r, \\ 0, & r(i) < \tau_r. \end{cases} \quad (18)$$

For each pixel, $\mathbf{M}_r(i) = 1$ if the r-value is larger than the threshold τ_r, and $\mathbf{M}_r(i) = 0$ is otherwise.

The further selection of sharp edges is determined by the gradient magnitude of shock filtered image. Shock filter [15] is an effective tool to restore strong edges from the blurred signals. Traditional shock filter is sensitive to noise. In our method, the complex shock filter described in [9] is used to select salient sharp edges as

$$\mathbf{I}_t = -\frac{2}{\pi}\arctan(a \cdot Im(\frac{\mathbf{I}}{\theta}))|\nabla\mathbf{I}| + c_1\triangle\mathbf{I} + c_2(\triangle\mathbf{I})_\perp, \quad (19)$$

where $\triangle\mathbf{I}$ and $(\triangle\mathbf{I})_\perp$ are the second-order spatial derivatives with

$$\begin{aligned}\triangle\mathbf{I} &= \partial_{xx}\mathbf{I}(\partial_x\mathbf{I})^2 + 2\partial_{xy}\mathbf{I}\partial_x\mathbf{I}\partial_y\mathbf{I} + \partial_{yy}\mathbf{I}(\partial_y\mathbf{I})^2, \\ (\triangle\mathbf{I})_\perp &= \partial_{xx}\mathbf{I}(\partial_y\mathbf{I})^2 + 2\partial_{xy}\mathbf{I}\partial_x\mathbf{I}\partial_y\mathbf{I} + \partial_{yy}\mathbf{I}(\partial_x\mathbf{I})^2. \end{aligned} \quad (20)$$

Such filtering can produce ideal shock response to a blurred image with noise. In $\mathbf{I}_t$, the image features are recovered from the blurred image. The large values of the gradient magnitude are selected by

$$\nabla\mathbf{I}^s = \nabla\mathbf{I}_t \circ \mathbf{M}_s, \quad (21)$$
$$\mathbf{M}_s(i) = u_{\tau_s}(\|\nabla\mathbf{I}_t\|_2 \circ \mathbf{M}_r - \tau_s), \quad (22)$$

where $\circ$ is the element-wise multiplication operator.

At the coarsest scale, the complex shock filter is firstly applied to the down-sampled blurry image $\mathbf{B}^0$ with Gaussian smoothing. Since the extents of blurs have been narrowed in the low-resolution image, the sharp edges can be

selected more accurately. At the finer scale, the image estimation up-sampled by bicubic interpolation is used as the input for the edge prediction. The predicted edge gradients $\nabla\mathbf{I}^s$ at top level is illustrated in Figure 5(b). Only the strong edges of large-scale objects can be seen in it. For visualization, we use Poisson reconstruction to restore the selected gradients, which is not necessary in the deblurring process.

3.2. Estimating the blur kernel

In our algorithm, the initial kernel estimation is obtained by iteratively performing the salient edges selection, kernel estimation and image deconvolution at each level l. Based on the salient sharp edges $\nabla\mathbf{I}_l^s$, we rewrite Eq. (12) with a Tikhonov regularization term as

$$\begin{aligned}\widehat{\mathbf{K}^l} = \arg\min_K\{&\|\nabla\mathbf{I}_l^s \otimes \mathbf{K} - \nabla\mathbf{B}_1^l\|_2 \\ &+ \|rot90(\nabla\mathbf{I}_l^s) \otimes \mathbf{K} - \nabla\mathbf{B}_2^l\|_2 + \lambda_K\|\mathbf{K}\|_2\}.\end{aligned} \quad (23)$$

The regularizer $\|\mathbf{K}\|_2$ is similar to [7, 23, 25]. For simplicity, we set $\nabla\mathbf{I}_1^s = \nabla\mathbf{I}_l^s$ and $\nabla\mathbf{I}_2^s = rot90(\nabla\mathbf{I}_l^s)$. Let $\mathcal{F}^{-1}(\cdot)$ denote the inverse Fourier transform, Eq. (23) can be calculated efficiently in the Fourier domain by

$$\widehat{\mathbf{K}^l} = \mathcal{F}^{-1}\big(\frac{\Sigma_{i=1}^2\overline{\mathcal{F}(\partial_x\mathbf{I}_i^s)}\mathcal{F}(\partial_x\mathbf{B}_i^l) + \overline{\mathcal{F}(\partial_y\mathbf{I}_i^s)}\mathcal{F}(\partial_y\mathbf{B}_i^l)}{\Sigma_{i=1}^2\mathcal{F}(\partial_x\mathbf{I}_i^s)^2 + \mathcal{F}(\partial_y\mathbf{I}_i^s)^2 + \lambda_K}\big). \quad (24)$$

In the image deconvolution, we constrain the estimated image gradient to be approximate to the blurry image gradient for all the pixels selected in $\nabla\mathbf{I}^s$ in Eq. (13) and get

$$\begin{aligned}\widehat{\mathbf{I}^l} = \arg\min_I\{&\|\mathbf{I} \otimes rot90(\widehat{\mathbf{K}^l}) - \mathbf{B}_1^l\|_2 \\ &+ \|\mathbf{I} \otimes \widehat{\mathbf{K}^l} - rot270(\mathbf{B}_2^l)\|_2 + \lambda_I\|\nabla\mathbf{I} - \nabla\mathbf{I}_l^s\|_2\}.\end{aligned} \quad (25)$$

Here we have $\mathbf{K}_1 = rot90(\widehat{\mathbf{K}^l})$, $\mathbf{K}_2 = \widehat{\mathbf{K}^l}$, $\widetilde{\mathbf{B}}_1 = \mathbf{B}_1^l$ and $\widetilde{\mathbf{B}}_2 = rot270(\mathbf{B}_2^l)$. The solution of Eq. (25) can be expressed as

$$\widehat{\mathbf{I}^l} = \mathcal{F}^{-1}\big(\frac{\Sigma_{i=1}^2\overline{\mathcal{F}(\mathbf{K}_i)}\mathcal{F}(\widetilde{\mathbf{B}}_i) + \lambda_I\Delta}{\Sigma_{i=1}^2\mathcal{F}(\mathbf{K}_i)^2 + \lambda_I(\mathcal{F}(\partial_x)^2 + \mathcal{F}(\partial_y)^2)}\big), \quad (26)$$
$$with \quad \Delta = \overline{\mathcal{F}(\partial_x)}\mathcal{F}(\partial_x\mathbf{I}^s) + \overline{\mathcal{F}(\partial_y)}\mathcal{F}(\partial_y\mathbf{I}^s).$$

277

Figure 6: The Comparison of deblurred results and the close-up views. (a) Two blurred images; (b) The blur kernel, from top to bottom are real kernel, kernels estimated by [17] and our method; (c) Image deblurred by the algorithm proposed in [17]; (d) Image deblurred by our method with the kernel estimated by [17]; (e) Image deglurred by our method; (f) The close-up views of the region in the red rectangle of (c)-(e).

After that, the initial estimation of image and kernel is obtained as shown in Figure 5(c) and (d). To further enhancing the sparsity of the kernel, we apply the ISD method to refine it. In [21], the implementation of ISD, called threshold-ISD, is introduced to efficiently address the ℓ_1 minimization with insufficient measurements. The algorithm forms a partial support $\mathbf{S} = \{j | \mathbf{K}(j) > \tau\}$ based on the current estimation. We replace the regularization term in Eq. (23) by $\sum_{j \in \overline{\mathbf{S}}} |\mathbf{K}(j)|$ and get

$$
\begin{aligned}
\mathbf{K}^* = \arg\min_K \{ & \|\nabla \mathbf{I}_l^s \otimes \mathbf{K} - \nabla \mathbf{B}_1^l\|_2 \\
& + \|rot90(\nabla \mathbf{I}_l^s) \otimes \mathbf{K} - \nabla \mathbf{B}_2^l\|_2 + \beta \sum_{j \in \overline{\mathbf{S}}} |\mathbf{K}(j)| \}.
\end{aligned} \tag{27}
$$

This optimization is efficiently implemented by running Matlab source codes publicly available on their website [20]. Then the large-value elements of kernel are preserved, while the insignificant elements are iteratively suppressed to reduce the noise. Interested readers can refer to [21] for more details.

3.3. Image deconvolution

Given the estimation $\mathbf{K}^*$ of the kernel, we compute the final deblurred result by solving Eq. (13) with TV regularization $E_I(\mathbf{I}) = \|\nabla \mathbf{I}\|_1$, similar to [5].

$$
\begin{aligned}
\mathbf{I}^* = \arg\min_I \{ & \|\mathbf{I} \otimes rot90(\mathbf{K}^*) - \mathbf{B}_1\|_2 \\
& + \|\mathbf{I} \otimes \mathbf{K}^* - rot270(\mathbf{B}_2)\|_2 + \gamma\|\nabla \mathbf{I}\|_1 \}.
\end{aligned} \tag{28}
$$

The deblurred image is recovered by the deconvolution technique proposed in [19]. For color image, the deconvolution is applied on each color channel of the blurry image independently. The final result is shown in Figure 5(e) with the refined kernel. Since the high quality of the kernel estimation, the small edges and textures are restored well and few artifacts are introduced.

4. Experimental evaluation

In this section, we test our method with several synthetic and real blurred images. We assume the range of image intensity to be $[0, 1]$. All the simulated blurred images are generated by the model described in Eq. (3) with zero-mean white Gaussian noise. The parameters of our experiments are set: $\lambda_K = 10$, $\lambda_I = 2e^{-3}$, $\beta = 1$, and $\gamma = 3e^{-4}$. The thresholds τ_r and τ_s are selected as discussed in [7, 23] and gradually decreased by multiplying 0.9 in iterations to introduce more and more edges information. In our experiments, we limited the number of iterations to be eight at each scale. In terms of computation time, the Matlab implementation of our algorithm spends about 29 seconds to restore a 500×500 image blurred with 21×21 kernel on a Windows PC with an Intel 3.0G Hz Core2Duo CPU.

One challenging example of our method is shown in Figure 1. The image with size 350×500 has small depth variation of scene, which is blurred by large-scale and complex motion kernels (95×95) with noise level $\sigma_n = 0.005$. It can by observed that our approach can provide good estimation of the kernel, and then outstanding image restoration is obtained as shown in Figure 1(d). However, such large-scale blur can not be solved by many single-image deblurring method. If such seriously blurry and noisy images are captured by consecutive shooting, it will be quite difficult to perform accurate automatic image alignment, which will further affect the kernel estimation and image restoration by dual-image deblurring algorithm. This experimental result shows the robustness of the proposed approach for real-world image deblurring.

To demonstrate the effectiveness of our method, we compare our deblurred results with the state-of-the-art single-image motion deblurring scheme proposed by Shan *et al*. [17] and dual-image deblurring method proposed by Chen *et al*. [6]. The results of Shan *et al*. [17] are obtained by running the executable program publicly available on their website with manual-adjusted parameters. The deblurred images of Chen *et al*. [6] are given by the authors.
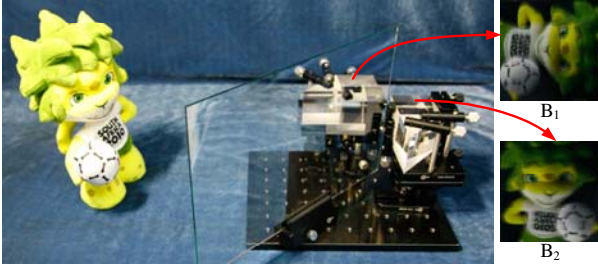
Figure 7: The prototype of image capture system.

Firstly, we compare the kernel estimation with the single-image algorithm [17] as shown in Figure 6(b). The kernel estimated by our method is based on dual blurry images. Since the two kernels are not completely arbitrary but have the relationship of 90-degree rotation, the kernel estimation is more accurate than that provided by [17]. We also compare the image deconvolution using our dual-image method with the single-image approach. In Figure 6(d), the image is restored by our image deconvolution method, where the kernel estimated by [17] are used. From the close-up views in Figure 6(f), we can see that our approach can provide clearer image even the kernel estimation is not accurate enough, as the two input images are excellently complementary to each other.

In Figure 8, the blurred images are generated with the blur kernel used in [13]. The blur kernel and the image restoration are better than that estimated respectively by the method of [17] and [6]. The fragment in the red rectangle contains many characters. Such small objects with salient edges are very challenging to make high-quality reconstruction. Our method provides clearer image with finer textures, while the estimation obtained by [17] has many ringing artifacts along the edges, and there still exists slight blur in the result from the method of [6].

We also tested our approach on real images in Figure 9. The aided prism system used in the experiments is shown in Figure 7, which is put in front of the camera lens and relatively static to the scene. We captured the images by a Canon 5D camera at 2 fps. Our method provides some improvements for the image restoration, which is especially noticeable in the close-up views.

5. Conclusions and future work

In this paper, we elaborately analyzed the capture strategy of well-aligned blurry image pair, and proposed a set of design rules to capture such images. Based on the image pair, dual-image deblurring algorithm is able to give more accurate kernel estimation and obtain higher-quality image restoration. Extensive experimental results have strongly supported the effectiveness of the proposed image capture strategy. While this paper developed and evaluated a practical application, both our theoretical analysis and practical algorithm can be easily extended to other instantiations.

However, the prism system has the limitation that it should be relatively static to the scene. Our future work will focus on the research of other dual-image mapping or multiple-image capture strategy. The aided hardware design is also the research direction under consideration. It is ideal for the future industrial trend to integrate the judicious practical design of image capture system with a consumer camera. Such system can restore the blurry image captured on fast-moving vehicle.

Acknowledgment

References

[1] A. Agrawal, Y. Xu, and R. Raskar. Invertible motion blur in video. *ACM Transactions on Graphics (Proc. SIGGRAPH)*, 28(3):1–8, 2009.

[2] M. Ben-Ezra and S. K. Nayar. Motion-based motion deblurring. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(6):689–698, 2004.

[3] J. Cai, H. Ji, C. Liu, and Z. Shen. Blind motion deblurring from a single image using sparse approximation. In *Proc. CVPR*, pages 104–111, 2009.

[4] J. Cai, H. Ji, C. Liu, and Z. Shen. Blind motion deblurring using multiple images. *Journal of Computational Physics*, 228(14):5057–5071, 2009.

[5] J. Cai, H. Ji, C. Liu, and Z. Shen. High-quality curvelet-based motion deblurring from an image pair. In *Proc. CVPR*, pages 1566–1573, 2009.

[6] J. Chen, L. Yuan, C. K. Tang, and L. Quan. Robust dual motion deblurring. In *Proc. CVPR*, pages 1–8, 2008.

[7] S. Cho and S. Lee. Fast motion deblurring. *ACM Transactions on Graphics (Proc. SIGGRAPH Asia)*, 28(5):1–8, 2009.

[8] R. Fergus, B. Singh, A. Hertzmann, S. T. Roweis, and W. T. Freeman. Removing camera shake from a single photograph. *ACM Transactions on Graphics (Proc. SIGGRAPH)*, 25(3):787–794, 2006.

[9] G. Gilboa, N. Sochen, and Y. Y. Zeevi. Image enhancement and denoising by complex diffusion processes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(8):1020–1036, 2004.

[10] N. Joshi, S. B. Kang, C. L. Zitnick, and R. Szeliski. Image deblurring using inertial measurement sensors. *ACM Transactions on Graphics (Proc. SIGGRAPH)*, 29(4):1–9, 2010.

[11] N. Joshi, R. Szeliski, and D. J. Kriegman. PSF estimation using sharp edge prediction. In *Proc. CVPR*, pages 1–8, 2008.

[12] A. Levin, R. Fergus, F. Durand, and W. T. Freeman. Image and depth from a conventional camera with a coded aper-
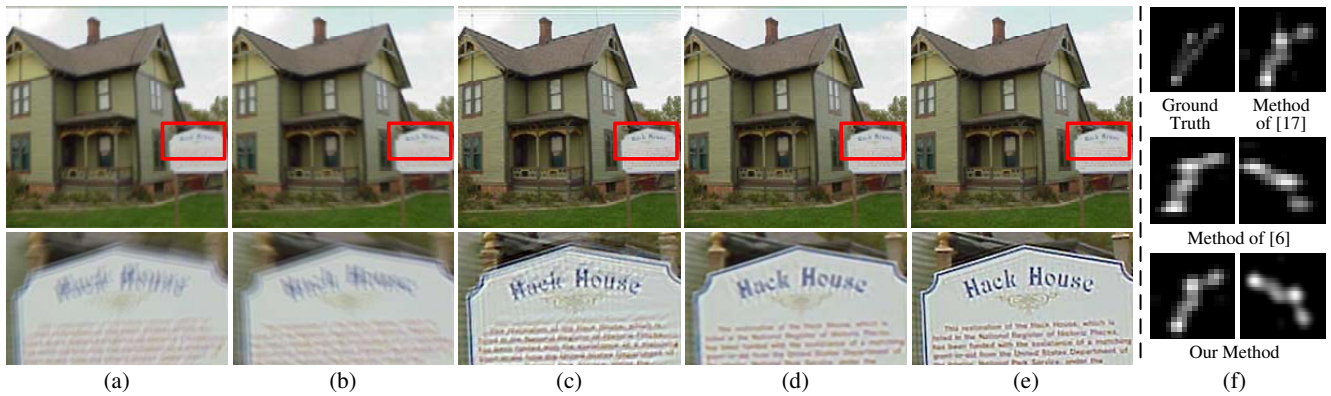
Figure 8: Deconvolution result comparison of synthetic images. (a)-(b) Two blurred images. (c) The deblurred image of (a) using Shan *et al.*'s method [17]. (d) The deblurred image using Chen *et al.*'s method [6]. (e) Deblurred result by our approach. The corresponding close-up views are shown in the bottom line. (f) Visualization of the kernel estimations.
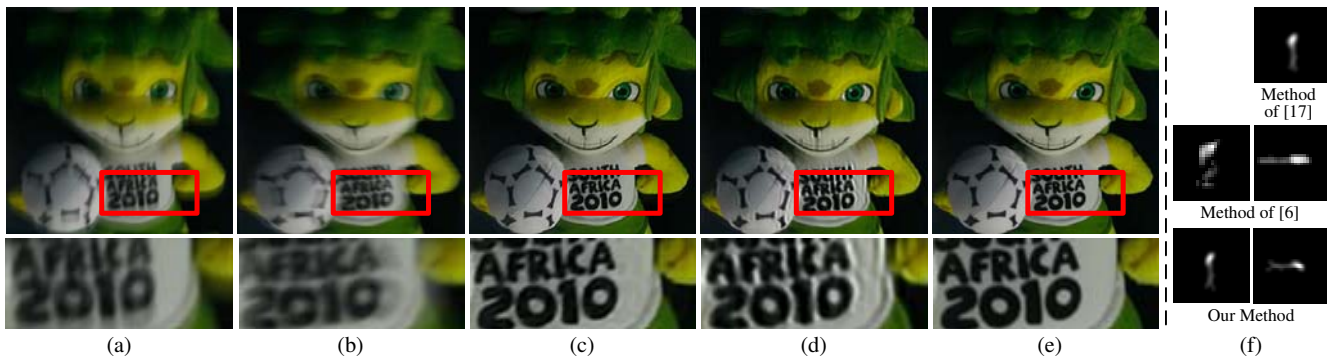

Figure 9: Deconvolution result comparison of real images. (a)-(b) Two blurred images. (c) The deblurred image of (a) using Shan *et al.*'s method [17]. (d) The deblurred image using Chen *et al.*'s method [6]. (e) Deblurred result by our approach. The corresponding close-up views are shown in the bottom line. (f) Visualization of the kernel estimations.

ture. *ACM Transactions on Graphics (Proc. SIGGRAPH)*, 26(3):70–78, 2007.

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

[14] L. B. Lucy. An iterative technique for the rectification of observed distributions. *Astronomical Journal*, 79(6):745–754, 1974.

[15] S. Osher and L. I. Rudin. Feature-oriented image enhancement using shock filters. *SIAM Journal on Numerical Analysis*, 27(4):919–940, 1990.

[16] A. Rav-Acha and S. Peleg. Two motion-based images are better than one. *Pattern Recognition Letters*, 26(3):311–317, 2005.

[17] Q. Shan, J. Jia, and A. Agarwala. High-quality motion deblurring from a single image. *ACM Transactions on Graphics (Proc. SIGGRAPH)*, 27(3):1–10, 2008.

[18] Y.-W. Tai, H. Du, M. S. Brown, and S. Lin. Correction of spatially varying image and video motion blur using a hybrid camera. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(6):1012–1028, 2010.

[19] Y. Wang, J. Yang, W. Yin, and Y. Zhang. A new alternating minimization algorithm for total variation image recon-

struction. *SIAM Journal on Imaging Sciences*, 1(3):248–272, 2008.

[20] Y. Wang and W. Yin. Threshold-ISD software. http://www.caam.rice.edu/~optimization/L1/ISD/, September 2009.

[21] Y. Wang and W. Yin. Sparse signal reconstruction via iterative support detection. *SIAM Journal on Imaging Sciences*, 3(3):462–491, 2010.

[22] N. Wiener. Extrapolation, interpolation, and smoothing of stationary time series. *MIT Press*, 1964.

[23] L. Xu and J. Jia. Two-phase kernel estimation for robust motion deblurring. In *Proc. ECCV*, pages 157–170, 2010.

[24] J. Yang, Y. Zhang, and W. Yin. An efficient TVL1 algorithm for deblurring multichannel images corrupted by impulsive noise. *SIAM Journal on Scientific Computing*, 31(4):2842–2865, 2009.

[25] L. Yuan, J. Sun, L. Quan, and H. Shum. Image deblurring with blurred/noisy image pairs. *ACM Transactions on Graphics (Proc. SIGGRAPH)*, 26(3):1–10, 2007.

[26] L. Yuan, L. Sun, L. Quan, and H. Shum. Progressive interscale and intra-scale non-blind image deconvolution. *ACM Transactions on Graphics (Proc. SIGGRAPH)*, 27(3):1–10, 2008.