Rotational Motion Deblurring of a Rigid Object from a Single Image[*]

Qi Shan, Wei Xiong, and Jiaya Jia
Department of Computer Science and Engineering
The Chinese University of Hong Kong
{qshan,wxiong,leojia}@cse.cuhk.edu.hk

Abstract

Most previous motion deblurring methods restore the degraded image assuming a shift-invariant linear blur filter. These methods are not applicable if the blur is caused by spatially variant motions. In this paper, we model the physical properties of a 2-D rigid body movement and propose a practical framework to deblur rotational motions from a single image. Our main observation is that the transparency cue of a blurred object, which represents the motion blur formation from an imaging perspective, provides sufficient information in determining the object movements. Comparatively, single image motion deblurring using pixel color/gradient information has large uncertainties in motion representation and computation. Our results are produced by minimizing a new energy function combining rotation, possible translations, and the transparency map using an iterative optimizing process. The effectiveness of our method is demonstrated using challenging image examples.

1. Introduction

Motion blur in a natural image is caused by camera shake or object motion during exposure when the shutter speed is relatively slow. The image degradation is usually modeled as the convolution of a clear image with a shift-invariant blur kernel

$$B = S \otimes k + n,$$

where S is the latent unblurred image, k is the point spread function (PSF), and n is the possible additive noise. Normally, the problem of recovering details from a single blurred image is severely ill-posed given the large set of unknowns and complex color information from the blurred image. The solution uniqueness cannot be generally guar-

anteed since the convolution with a blur kernel is non-invertible. To tackle this problem, additional image priors, such as the global gradient distribution from clear images [7], are proposed. Some approaches use multiple images or additional visual cues [2, 20] to constrain the kernel estimation.
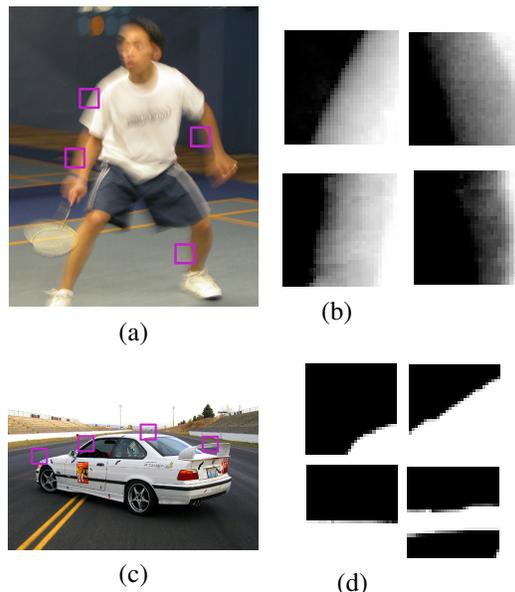
Figure 1. Transparency map. (a) An blurred image. (b) The transparency map in the highlighted regions of (a). (c) An unblurred image. (d) The transparency map in the highlighted regions of (c).

The spatially variant blur kernel estimation is an even more difficult problem. The blur kernel, in this case, may vary in size, shape, and values among pixels [24]. This generality makes it extremely difficult to estimate an appropriate PSF. To make the problem trackable, additional assumptions, such as modeling the general kernel as a weighted sum of several spatially invariant kernels [17], are proposed in previous work. However, due to the complexity of the unknown kernel and the expanded solution space, these methods cannot produce convincing kernel estimation and are easily stuck in local minimum in optimization.

In this paper, we propose a practical algorithm to deblur rigid-body object undergone primarily the rotational motions from a single image. The blur kernel is spatially-variant and unknown. The possible translations during rotation are allowed. Our blur kernel representation is general and is modeled explicitly by a novel motion descriptor. In order to reduce the ambiguities in motion deblurring using complex image structures, we approximate the transparency map physically produced by the blurred object [11] and employ it to robustly estimate the spatially variant kernels.

The transparency map [11] contains fractional pixel values which can be physically interpreted as the time percentage that the foreground pixels are exposed during image capture. Each pixel has a fractional value between 0 and 1. With the fact that the foreground always occludes the background, it further describes the partial occlusion of the background scene by the foreground object during the exposure. So, essentially, the transparency map abstracts object motion information. The pixels in object center have value 1 and the pixels around the boundary of a blurred object may have fractional values. For an unblurred object with solid boundary, each pixel in the image corresponds to either the background or the foreground. So the transparency map can be approximated by binary values consisting of only 0 and 1. One illustration is shown in Figure 1.

Using the transparency map in our motion kernel estimation has the following benefits. First, the transparency map has very simple representation, each value is between 0 and 1, and is not related to the complex image structure. Second, the transparency map has useful physical representation and directly reflects the object motion structures. Third, the latent transparency map after deblurring should contain only binary values for a solid object which greatly reduces the ambiguities in image reconstruction. Moreover, the binary value property in an unblurred transparency map makes it possible to apply an efficient discrete optimization in blur kernel estimation.

By adopting the transparency map, our modeling and computation process are simplified comparing to using only the image color information. In our method, a robust iterative optimization algorithm is proposed to estimate the motion descriptor, which is further applied to recovering the original color image.

The rest of this paper is organized as follows. We first review previous work in Section 2. The motion analysis and descriptor are described in 3. In Section 4, we explain the model and the proposed algorithm. We show the results in Section 5. Finally, we conclude our method and discuss it in Section 6.

2. Previous work

There are two main streams in research categorized by the blur models, i.e., the spatially-invariant and spatially-variant motion deblurring. If the point spread function (PSF) is known, several algorithms such as Wiener filtering, quick Pixon smoothing, Richardson-Lucy deconvolution can be applied to inferring the original clear image.

Deconvolution is also an important topic in astronomical imaging research. Gull discusses in [9] the principle of maximizing the entropy in deconvolving images. Tsumuraya *et al.* [23] apply Lucy's algorithm to reconstruct the deblurred images. Zarowin *et al.* [27] propose a modified version of van Cittert deconvolution (VCD) to relieve the limitation of the traditional method which is only applicable to shapes and related points.

Without knowing the blur kernel, one can recover a general blur image using blind convolution. In [13, 5], methods are proposed to estimate the blur kernel assuming *a priori* knowledge. Two-step approaches are adopted by first estimating the kernel and then deconvolving the image. With the assumption that the unblurred image is two-tone, Li *et al.* [15] propose an iterative approach to refine a deblurring filter until the final result is visually acceptable. In [7], Fergus *et al.* propose an ensemble learning [16] approach with the prior modeling the distribution of image gradient. Most recently, Jia proposes to solve the spatially invariant motion blur from a transparency point of view [11]. Comparing to [11], our goal is to estimate the spatially-VARIANT kernels with entirely different motion models.

In the shift-variant deblurring methods [17, 24, 4], a general PSF is assumed. Calvetti *et al.* [3] assume that the PSF is known, and then the restoration of a blur and noisy image can be accomplished by solving a large linear system. The generalized minimal residual (GMRES) method is used to iteratively solve this linear system. Nagy and Leary [17] model the spatially variant blur kernel as the combination of spatially invariant kernels. It provides a way to simplify the PSF modeling, but it lacks of physical analysis and constraints. Ribaric *et al.* [21] propose a method to restore images blurred by circular motion. This method assumes that the rotational center and the angular velocity are known, which makes it difficult to be applied to natural images.

To provide additional information for image restoration, multiple images or hardware are used to improve the result construction. Without estimating the PSF, Jia *et al.* [12] enhance a short-exposure dark image by transferring the color from a long-exposure blurred image. Yuan *et al.* [26] produce high quality deblurred images using blur/noisy image pair configurations. The ringing artifacts are significantly reduced by the proposed gain-controlled Richardson-Lucy (RL) algorithm. Ben-Ezra *et al.* [3] construct a hybrid camera system to acquire the camera motion tracks for accurately estimating kernels. Raskar *et al.* [19] reduce the ill-posedness of the deblurring problem by rapidly opening and closing the shutter with a pre-specified binary sequence, information lost of high frequencies can be reduced. There

are many previous works in research of image deblurring. We refer the reader to the respective surveys [22, 18, 10].

Comparing to other work, our single image deblurring method provides a novel solution to deblur rotation plus translation motions. Our framework is practical thanks to our kernel modeling, transparency map representation, and the robust iterative optimization.

3. Motion blur descriptor

While the state-of-the-art research on spatially invariant motion deblurring has reached a certain degree of maturity [19, 2, 7, 11], methods dealing with spatially variant blur are still far from practicality. To sufficiently describe object motions, a single blur filter may not be enough. In this section, we propose a motion blur descriptor to model the object rotational motion with possible irregular translations.

An image can be regarded as an integral of the light hitting the image sensor per time unit. The blur effect in an image is caused by the object motion during exposure. If we consider the mixed light emitted or reflected from the foreground and background, the image blur formulation can be very complex. Hence we introduce the transparency map, as described in Section 1, to describe the motion blur.

We denote f as the transparency map of a solid object without motion in the latent (unblurred) image, where $f(x, y) = 0$ if pixel (x, y) is on the background, and $f(x, y) = 1$ if (x, y) belongs to the foreground, as shown in Figure 1. We also denote B as the transparency map where a solid object undergoes motions in a blurred image. Obviously, B can be formulated as an integral of different f after rotation and translation during exposure. Let $Tr(x, y; \theta_t, \Delta x_t, \Delta y_t)$ be the value of transparency map in (x, y) after f rotates with angle θ_t and shifts by $(\Delta x_t, \Delta y_t)$, the blurred transparency map $B(x, y)$ is given by

$$B(x, y) = \frac{1}{T} \int_0^T Tr(x, y; \theta_t, \Delta x_t, \Delta y_t) dt, \quad (1)$$

where T denotes the exposure time. By discretization, $B(x, y)$ can be estimated by

$$\hat{B}(x, y, Q, f) = \frac{1}{T} \sum_{i=0}^{N-1} Tr(x, y; \theta_i, \Delta x_i, \Delta y_i)(t_{i+1} - t_i). \quad (2)$$

The value of N determines the level of accuracy. In the experiments, we found $N = 20$ is sufficient to produce a visually satisfying result. By dividing the exposure time T evenly, the motion descriptor Q contains all the information to describe the object motions

$$Q = \{(\theta_i, \Delta x_i, \Delta y_i) | i \in \{0, 1, \dots, N-1\}\}.$$

3.1. Descriptor analysis

In Eqn. (2), there is one $Tr(x, y; \theta, \Delta x, \Delta y)$ for each time slot, modeling the object motion. Due to the implicit format of these motion parameters, this function cannot be used in efficient optimization. In the following, using the Fourier transform, we derive an explicit representation.

Considering the object motion during image capture, we map each pixel (x, y) in the transformed image $Tr(x, y; f, \theta, \Delta x, \Delta y)$ to one pixel (x', y') in f where

$$[x', y']^T = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} [x, y]^T + [\Delta x, \Delta y]^T.$$

We thus have

$$Tr(x, y; \theta, \Delta x, \Delta y) = f(x \cos(\theta) - y \sin(\theta) + \Delta x,$$
$$x \sin(\theta) + y \cos(\theta) + \Delta y). \quad (3)$$

In order to establish an explicit function of motion descriptors for Tr, we transform the image to the frequency domain. Eqn. (3) can then be written as

$$Tr(\cdot) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} F(u, v) e^{j2\pi(u\Delta x + v\Delta y)}$$
$$e^{j2\pi((ux+vy)\cos(\theta)+(vx-uy)\sin(\theta)))} du dv, \quad (4)$$

where $F(u, v)$ is a Fourier transform of $f(x, y)$ and $j^2 = -1$:

$$F(u, v) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) e^{-j2\pi(ux+vy)} dx dy. \quad (5)$$

Due to the discreteness in both the image domain and the

(a)   (b)
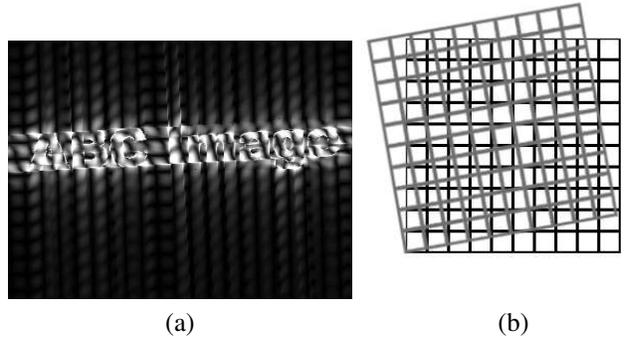
Figure 2. Artifacts introduced by Discrete Fourier Transformation (DFT). (a) Artifacts of image reconstruction after DFT. (b) Grid misalignment after rotation.

frequency domain, we need to sample f and F for further simplification. Directly sampling them using integer coordinates may introduce serious artifacts, as shown in Figure 2. This is because the pixel positions of the rotated image may not fall into the points with integer coordinates

as shown in Figure 2 (b). So the sampling rate should be adjusted to reduce the artifacts. Without losing of generality, we simplify the presentation by assuming the image is square with size $M \times M$. Let $q = 1/M$, by applying the Bluestein's FFT algorithm [1, 6], we rewrite (5) as

$$
\begin{aligned}
F_{(u,v)} &= F(u\kappa_r, v\kappa_r) \\
&\approx \sum_y \sum_x f(x\kappa_s, y\kappa_s) e^{-j2\pi q(u\kappa_r x\kappa_s + v\kappa_r y\kappa_s)} \kappa_s^2 q^2 \\
&= \sum_y \sum_x f(x\kappa_s, y\kappa_s) e^{-j2\pi q(ux+vy)\kappa_r \kappa_s} \kappa_s^2 q^2, \quad (6)
\end{aligned}
$$

where κ_s is the distance between neighboring samples in the image domain and κ_r is the distance in the frequency domain. For simplicity, here we use nearest neighborhood interpolation to get $f(x\kappa_s, y\kappa_s)$. Accordingly, we have

$$
\begin{aligned}
Tr(\cdot) &= \sum_v \sum_u F_{(u,v)} e^{j2\pi (u\Delta x + v\Delta y)\kappa_r q} \\
&\quad e^{j2\pi((ux+vy)\cos(\theta)+(vx-uy)\sin(\theta))\kappa_r q} \kappa_r^2. \quad (7)
\end{aligned}
$$

Directly computing (6) and (7) is computationally inefficient since the time complexity is $\Theta(\chi^2)$, where χ is the number of pixels. So through some algebraic manipulations, we derive

$$
\begin{aligned}
Tr(x,y; \theta, \Delta x, \Delta y) &= \kappa_r^2 \overline{H}(y,x) \sum_v \sum_u (K(u,v) \\
&\quad H(u,v) \overline{H}(u-x, v+y)), \quad (8)
\end{aligned}
$$

where

$$
\begin{aligned}
H(x,y) &= e^{j\pi\kappa_r q((x^2-y^2)\cos(\theta)+2xy\sin(\theta))} \\
\overline{H}(x,y) &= conj(H(x,y)) \\
&= e^{j\pi\kappa_r q((y^2-x^2)\cos(\theta)-2xy\sin(\theta))} \\
K(u,v) &= e^{j2\pi\kappa_r q(u\Delta x + v\Delta y)} F(u,v).
\end{aligned}
$$

The convolution in (7) is substituted by the efficient multiplication in the frequency domain. We verify the correctness of our formulation by using a blur image synthesis shown in Figure 3 using a motion descriptor. The synthesized image in (b) is quite realistic even using the discrete transform.

4. Optimization

Our goal is to restore the latent image f by minimizing the energy function

$$
E_D(Q, f) = \sum_{x,y} (B_{ref}(x,y) - \hat{B}(x,y,Q,f))^2. \quad (9)
$$

where B_{ref} is the input blurred image. It requires that using the estimated descriptor Q and image f, the reconstructed blur image $\hat{B}$ should be similar to the input blur image B_{ref}. Another implicit constraint is that $f(x,y) \in \{0,1\}$ which will be imposed in our discrete optimization process.

Here we naturally assume that $f(x,y)$, Δx_i, Δy_i, and θ_i are all independent. In order to make the problem trackable and to successfully restore the degraded image, the rotation angles $\theta_{t+1} - \theta_t$ in short time intervals are assumed small.

4.1. Parameter initialization

(a)   (b)
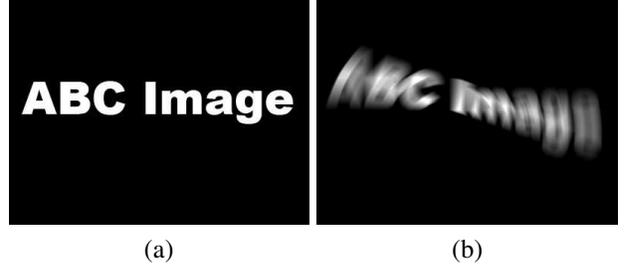
Figure 3. Rotational blur image synthesis. (a) The original image. (b) The blurred image generated using DFT. The naturalness validates the correctness of our motion descriptor.

The transparency map B used in our method describes the transparency of the foreground occluding the background due to object motion. It can be computed using the natural image alpha matting methods. In our approach, the method in [14] is used to compute the fractional transparency map. Since matting methods only use color information to compute the alpha matte, there is no real physical considerations of the background or foreground. So there may exist errors in some pixels. As shown in Section 5, we only need to select and use the pixels with structures in obedience to the human perception instead of all of them.

We first initialize the object rotation center (c_x, c_y). It is common that the blurred color image contains texture information. So the users can interactively select several pairs of points from both the transparency map and the color image indicating the object motion. The pixels in each pair are selected approximately at the start and the end positions of one object point during its movement. In this paper, these points are called characteristic points. One illustration is shown in Figure 4. The estimation of rotation center does not require much accuracy. The errors can be compensated by the estimation of translation Δx_i and Δy_i in the following optimization process.

After selecting the characteristic point pairs, the rotation center can be estimated by solving a least square problem. Assuming each point pair consists of the start point (x_s^i, y_s^i) and the end point (x_e^i, y_e^i), we have

$$
(x_s^i - c_x)^2 + (y_s^i - c_y)^2 = (x_e^i - c_x)^2 + (y_e^i - c_y)^2,
$$

as illustrated in Figure 5. Considering all selected points,

(a)   (b)

Figure 4. Blurred color image. (a) The original blur image. (b) The user selected characteristic points according to the motion patterns of object structure. Each pixel pair roughly indicates the start and end position of one point in motion.
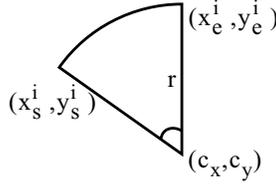
Figure 5. The matched points can be used to identify the rotation center (c_x, c_y).

the rotation center can be computed by solving

$$A_{M \times 2} \cdot [c_x, c_y]^T = B_{M \times 1}, \tag{10}$$

where M is the number of the corresponding point pairs, $A_{i,1} = 1$, $A_{i,2} = \frac{y_e^{(i)} - y_s^{(i)}}{x_e^{(i)} - x_s^{(i)}}$, and $B_i = \frac{x_e^{(i)2} + y_e^{(i)2} - x_s^{(i)2} - y_s^{(i)2}}{2(x_e^{(i)} - x_s^{(i)})}$, $\forall i = 1, 2, 3, \cdots, M$. After c_x and c_y are estimated, we move the coordinate center to (c_x, c_y).

4.2. Iterative optimization

In the following, we propose an iterative optimization to solve all unknowns. The discrete and continuous optimizations are alternatively performed. Our optimization can be divided into three steps, and these three steps are processed iteratively.

Estimating θ_i. In this step, assuming the values of F, Δx_i, Δy_i are fixed, we apply a continuous optimization to compute θ_i by minimizing E_D. Denoting $m_1^{(i)} = (ux + vy)\cos(\theta_i) + (vx - uy)\sin(\theta_i)$ and $m_2^{(i)} = (vx - uy)\cos(\theta_i) - (ux + vy)\sin(\theta_i)$, the first order partial derivatives of E_D can be explicitly given by

$$\frac{\partial E_D}{\partial \theta_i} = -\frac{4\pi \kappa_r^3 q}{N} \sum_{x,y} (\zeta_{x,y} \cdot (\sum_{u,v} F_{(u,v)} j$$

$$e^{j2\pi(u\Delta x + v\Delta y + m_1^{(i)})\kappa_r q} \cdot m_2^{(i)})),$$

where $\zeta_{x,y} = B_{ref}(x,y) - \hat{B}(x,y)$. For small θ_i's, we use the approximation that $\sin\theta_i \approx \theta_i$, $\cos\theta_i \approx 1 - \frac{\theta_i^2}{2}$

to avoid computation of many trigonometric functions. The Hessian matrix can be derived similarly. Using the above representation, θ_i is computed using the conjugate gradient method.

Estimating the latent transparency map f. Given the estimated θ, Δx, and Δy, we compute the binary map f using discrete optimization by minimizing

$$E_D(Q, f) = \sum_p (B_{ref}(p) - \frac{1}{N} \sum_{i=0}^{N-1} Tr(p; \theta_i, \Delta x_i, \Delta y_i))^2,$$

where $p = (x, y)$. Given that the motion parameters are estimated, we can apply inverse motions to the input blur image to derive the latent one. Since Tr denotes the motion of the object in a small time interval, applying (3), we get $f(x_i', y_i') = Tr(x_i, y_i; \theta_i, \Delta x_i, \Delta y_i)$, where $i \in \{0, 1, ..., N-1\}$. Under the rotational motion, x_i and y_i are not guaranteed to be integers. We thus use the bilinear interpolation from the neighboring four pixels with integer coordinates to compute the transparency values:

$$f(x_i', y_i') = \sum_{j=0}^{3} \alpha_i^j f(x_i^j, y_i^j), \ \ x_i^j, y_i^j \in Z. \tag{11}$$

Integrating the above equations, we have

$$(B_{ref}(x, y) - \frac{1}{N} \sum_{i=0}^{N-1} Tr(x_i, y_i; \theta_i, \Delta x_i, \Delta y_i))^2$$

$$= (B_{ref}(x, y) - \frac{1}{N} \sum_{i=0}^{N-1} \sum_{j=0}^{3} \alpha_i^j f(x_i^j, y_i^j))^2$$

$$\propto E_1 + E_2, \tag{12}$$

where

$$E_1 = \frac{2}{N} \sum_{i,j} B_{ref}(x, y) f(x_i^j, y_i^j), \tag{13}$$

and

$$E_2 = \frac{2}{N^2} \sum_{(i_1, j_1) \neq (i_2, j_2)} \alpha_{i_1}^{j_1} \alpha_{i_2}^{j_2} f(x_{i_1}^{j_1}, y_{i_1}^{j_1}) f(x_{i_2}^{j_2}, y_{i_2}^{j_2})$$

$$+ \frac{1}{N^2} \sum_{i,j} (\alpha_i^j)^2 f^2(x_i^j, y_i^j). \tag{14}$$

We build a Markov Random Field (MRF) in the image plane to solve the above energy minimization problem where each pixel is a node. The edges are constructed by connecting node $(x_{i_1}^{j_1}, y_{i_1}^{j_1})$ to $(x_{i_2}^{j_2}, y_{i_2}^{j_2})$ if there exists a corresponding term $f(x_{i_1}^{j_1}, y_{i_1}^{j_1}) f(x_{i_2}^{j_2}, y_{i_2}^{j_2})$ in E_2. All the nodes are 0-1 labeled. Energy E_D is minimized using loopy belief propagation [8][25].

Estimating Δx_i and Δy_i In this stage, we estimate object translations Δx_i and Δy_i. Since θ_i and f are fixed, $Tr(x, y; \theta_i, 0, 0)$ is also known. We build Gaussian image pyramids for $Tr(x, y; \theta_i, 0, 0)$ and B_{ref} respectively. We use a multi-scale approach with these Gaussian image pyramids to compute Δx_i and Δy_i in a top-down manner. Suppose the finest level is 0 and the coarsest level is s, it is obvious that the translation in coarse level is small. Assuming that the translations $\Delta x_i^{(s)}$ and $\Delta y_i^{(s)}$ in level s are smaller than 1 pixel, i.e., $-1 < \Delta x_i^{(s)}, \Delta y_i^{(s)} < 1$, we compute the translation vector in the fine level by

$$
\begin{aligned}
\Delta x_i^{(\gamma)} &= 2\Delta x_i^{(\gamma+1)} + \psi_i^{(\gamma)}(x), \\
\Delta y_i^{(\gamma)} &= 2\Delta y_i^{(\gamma+1)} + \psi_i^{(\gamma)}(y),
\end{aligned} \tag{15}
$$

where $\psi_i^{(\gamma)}(x), \psi_i^{(\gamma)}(y) \in (-1, 1)$. With the constructed Gaussian pyramids, we minimize $E_D^{(\gamma)}$ on each level and propagate the estimated translation vector to the finer one. Combining (3), (9), (15) we minimize

$$
\begin{aligned}
&E_D^{(\gamma)} \\
&= \sum_{x,y}(B_{ref}^{(\gamma)}(x,y) - \frac{1}{N}\sum_i Tr^{(\gamma)}(x,y;\theta_i,\Delta x_i^{(\gamma)},\Delta y_i^{(\gamma)}))^2 \\
&= \sum_{x,y}(B_{ref}^{(\gamma)}(x,y) - \frac{1}{N}\sum_i f^{(\gamma)}(\hat{x}^i + \psi_i^{(\gamma)}(x), \hat{y}^i + \psi_i^{(\gamma)}(y)))^2,
\end{aligned}
$$

where $\hat{x}^i = x\cos(\theta_i) - y\sin(\theta_i) + 2\Delta x_i^{(\gamma+1)}$ and $\hat{y}^i = x\sin(\theta_i) + y\cos(\theta_i) + 2\Delta y_i^{(\gamma+1)}$. Since $\psi_i^{(\gamma)}(x)$ and $\psi_i^{(\gamma)}(y)$ are small values, we apply the Taylor expansion to $f^{(\gamma)}(\hat{x}^i + \psi_i^{(\gamma)}(x), \hat{y}^i + \psi_i^{(\gamma)}(y))$ and derive

$$
\begin{aligned}
E_D^{(\gamma)} &\approx \sum_{x,y}(B_{ref}^{(\gamma)}(\cdot) - \frac{1}{N}\sum_i(f^{(\gamma)}(\hat{x}^i, \hat{y}^i) \\
&+ \frac{\partial f^{(\gamma)}(\hat{x}^i, \hat{y}^i)}{\partial x}\psi_i^{(\gamma)}(x) + \frac{\partial f^{(\gamma)}(\hat{x}^i, \hat{y}^i)}{\partial y}\psi_i^{(\gamma)}(y)))^2.
\end{aligned} \tag{16}
$$

The term $f^{(\gamma)}(\hat{x}^i, \hat{y}^i)$ in (16) in the Gaussian pyramid can be computed by

$$
f^{(\gamma)}(\hat{x}^i, \hat{y}^i) = Tr^{(\gamma)}(x, y; \theta_i, 2\Delta x_i^{(\gamma+1)}, 2\Delta y_i^{(\gamma+1)}).
$$

Our computation starts from the coarsest level. In the optimization in level γ, E_D^γ is minimized by setting the partial derivatives $\frac{\partial E_D^{(\gamma)}}{\partial \psi_i^{(\gamma)}(x)}$ and $\frac{\partial E_D^{(\gamma)}}{\partial \psi_i^{(\gamma)}(y)}$ to zeros and solving an over-determined linear system with $2N$ unknowns: $\psi_i^{(\gamma)}(x)$'s and $\psi_i^{(\gamma)}(y)$'s. The optimization stops when $\Delta x_i^{(0)}$ and $\Delta y_i^{(0)}$ are computed.

4.3. Recover the color image

The above optimization steps iteratively process. In our experiments, they always converge in 10 iterations. To recover the color image, using the transparency map, we separate the blur image into foreground I_f and background I_b.

Then each pixel in I_f can be expressed as a weighted sum of the pixels in the original clear object. Since θ_i, Δx_i, and Δy_i are estimated, the spatially variant blur kernel for each pixel can be uniquely constructed. We write

$$
I_f(p) = R_f(p) \otimes k(p) + n, \tag{17}
$$

where $R_f(p)$ is the corresponding pixel in the latent unblurred object and $k(p)$ is the kernel constructed for pixel p.

We apply the Richardson-Lucy deconvolution method to iteratively restore the unblurred foreground $\hat{R}_f$. Then the final image result is constructed by combining $\hat{R}_f$ and I_b with the recovered transparency map f.

5. Result and analysis

Examples are demonstrated in this section. All images shown contain a blurred object undergoing unknown rotation plus translation motions during image capture.

We first show the doll example in Figure 6. The doll is held by a hand undergoing inconsistent motions for different pixels during image capture as shown in (a). (b) shows the selected characteristic points marked by red crosses in the input image for computing an approximated rotation center. The error introduced in this stage is inevitable but does not greatly influence our optimization. We apply the matting method to compute the blur transparency map with an input of several strokes indicating definite foreground and background pixels. We show the stroke-overlaid image in (c). (d) illustrates the generated alpha matte. Once we estimate all motion parameters, we apply them to constructing spatially variant kernels. The kernels are used to restore the transparency map (e) and latent object (f) using Richardson-Lucy deconvolution. (g) shows the deblurred object composed back to the extracted background I_b. (h) shows the magnified region. The motion blur is removed successfully on the object.

We show another two examples in Figure 7. (a) and (b) show the input blur images with marked characteristic points. These points are selected easily using the recognizable image patterns. (c) and (d) are the computed transparency maps. We apply our iterative optimization to the transparency maps and estimate object motions. Using the constructed blur filter, we deblur the foreground object. The compositions of the deblurred object and the background are shown in (e) and (f).

Our method employs the natural image matting method to compute the transparency map. In a complex scene, existing matting methods may not work well and may generate incorrect alpha values for some pixels. We show in Figure 8 that the errors will not significantly influence our method in the motion estimation since not all transparency pixels are taken into computations. In Figure 8 (a), the hand is
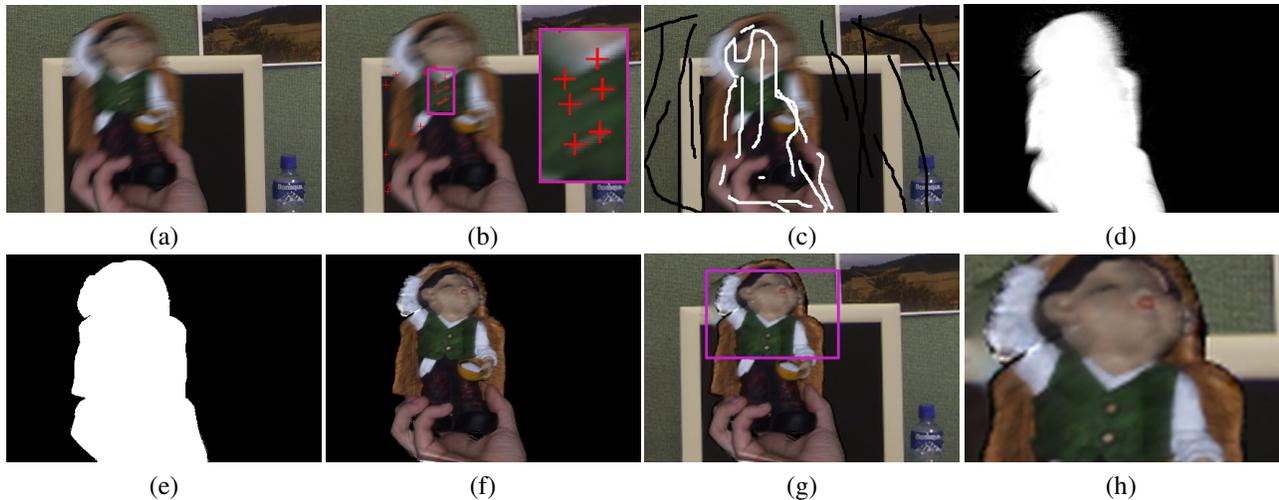
Figure 6. Doll. (a) The input image containing a motion blurred doll. (b) The selected characteristic points are illustrated using red crosses. (c) Strokes specifying the foreground and the background to be used in the matting method. (d) The computed blur transparency map after image matting. (e) The inferred unblurred transparency map by our method. (f) The computed unblurred object on a black background. (g) The unblurred foreground object composed to the original background I_b. (h) The magnified view of (g).

blended to the background. It is difficult to obtain a good alpha matte, as shown in (c). We only use the transparency pixels with less error as illustrated in (d). Applying the proposed optimization, the motion parameters can be successfully estimated. The deblurred transparency map and the composited object to the background are shown in (e) and (f) respectively. It can be observed that since the transparency map contains errors, there are pixels not correctly deblurred. However, most pixels in the result are visually plausible. The patterns on the sleeve are also clear.

6. Conclusion and Discussion

To conclude, we have proposed a novel and highly practical method to restore the degraded image containing a blurred object due to the rotational and translational motions. Instead of using the complex color information in input images, we introduce a transparency map encoding the object motion and apply it to estimating the motion parameters. Our motion descriptor separates the motion blur in different time intervals and the computation is facilitated by Fourier transform. A novel optimization method is proposed to iteratively refine parameter estimation using discrete and continuous optimizations. Our method has been applied to several difficult examples.

In discussion, the quality of our final result is closely related to the quality of the transparency map generated by the matting methods. Our further work includes developing a more reliable transparency map estimation approach using multiple input images. Reducing the manual work, such as the characteristic point selection, will also be studied.

References

[1] D. H. Bailey and P. N. Swarztrauber. The fractional fourier transform and applications. *SIAM Review*, 33(3):389–404, 1991.

[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] D. Calvetti, B. Lewis, and L. Reichel. Restoration of images with spatially variant blur by the gmres method. *Proc. SPIE*, 4116:364–374, 2000.

[4] R. H. Chan, T. F. Chan, L. Shen, and Z. Shen. Wavelet deblurring algorithms for spatially varying blur from high-resolution image reconstruction. *Linear Algebra and its Applications*, pages 139–155, 2003.

[5] M. Chang, A. Tekalp, and A. Erdem. Blur identification using the bispectrum. *IEEE Transactions on Signal Processing*, 39(10):2323–2325, 1991.

[6] R. Cox and R. Tong. Two- and three-dimensional image rotation using the fft. *Image Processing, IEEE Transactions on*, 8(9):1297–1299, 1999.

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

[8] B. J. Frey. *Graphical Models for Machine Learning and Digital Communication*. MIT Press, 1998.

[9] S. F. Gull. *Maximum entropy and bayesian methods*. Kluwer Academic, 1988.

[10] P. C. Hansen, J. G. Nagy, and D. P. O'Leary. *Deblurring Images: Matrices, Spectra, and Filtering*. Society for Industrial and Applied Mathematic, 2006.

[11] J. Jia. Single image motion deblurring using transparency. In *CVPR*, page To appear, 2007.

[12] J. Jia, J. Sun, C.-K. Tang, and H.-Y. Shum. Bayesian correction of image intensity with spatial consideration. In *ECCV*, pages III: 342–354, 2004.

[13] S. K. Kim and J. Paik. Out-of-focus blur estimation and restoration for digital auto-focusing system. *Electronics letters*, 34(12):1217–1219, 1998.

[14] A. Levin, D. Lischinski, and Y. Weiss. A closed form solution to natural image matting. In *CVPR*, pages 61–68, 2006.

(a)　(b)

(c)　(d)

(e)　(f)
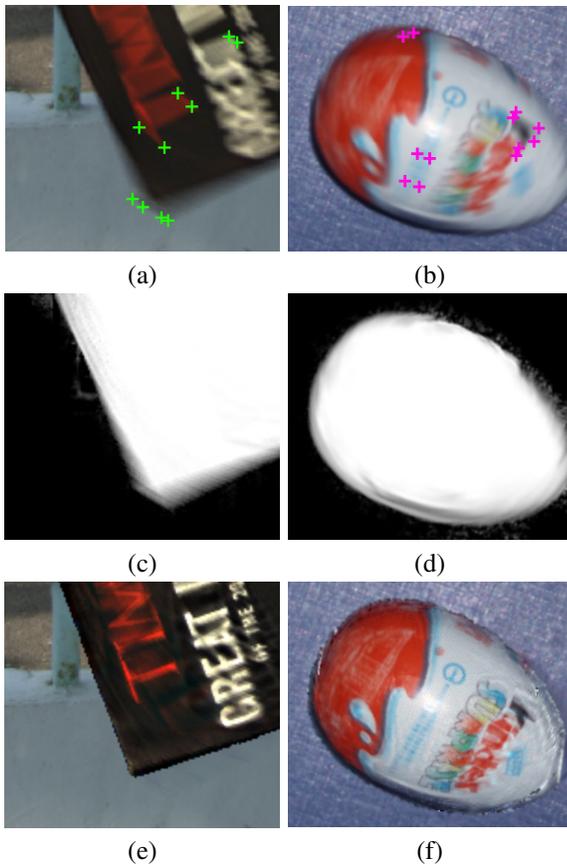
Figure 7. Book and egg chocolate. (a) and (b) are the input blur images with selected characteristic point pairs. (c) and (d) show the computed blur transparency maps. (e) and (f) show the final deblurred image generated by our method.

(a)　(b)

(c)　(d)

(e)　(f)
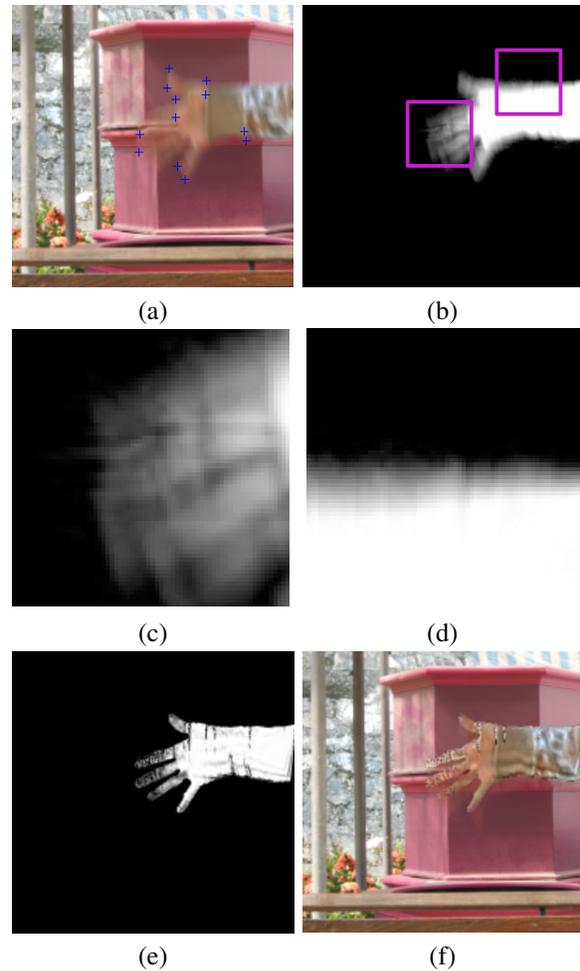
Figure 8. Waving hand. (a) The input blur image with characteristic points. (b) The computed transparency map has errors due to the complex blending between the foreground and the background. We only use visually plausible transparency map pixels in estimating the motion descriptor. (c) The magnified region in (b) with noticeable errors. (d) The magnified region in (b) with less errors. We use these pixels to robustly estimate motion parameters. (e) The final transparency map restored using R-L deconvolution. (f) The final deblurred result.

[15] T.-H. Li and K.-S. Lii. A joint estimation approach for two-tone image deblurring by blind deconvolution. *IEEE Transactions on Image Processing*, 11(8):847–858, 2002.

[16] J. W. Miskin. *Ensemble Learning for Independent Component Analysis*. PhD thesis, Univ. of Cambridge, 2000.

[17] J. G. Nagy and D. P. O'Leary. Restoring images degraded by spatially variant blur. *SIAM Journal on Scientific Computing*, 19(4):1063–1082, 1998.

[18] R. C. Puetter, T. R. Gosnell, and A. Yahil. Digital image reconstruction: Deblurring and denoising. *Annual Review of Astronomy and Astrophysics*, 43:139–194, 2005.

[19] R. Raskar, A. Agrawal, and J. Tumblin. Coded exposure photography: Motion deblurring using fluttered shutter. *ACM Transactions on Graphics*, 25(3):795–804, 2006.

[20] A. Rav-Acha and S. Peleg. Two motion-blurred images are better than one. *Pattern Recogn. Lett.*, 26(3):311–317, 2005.

[21] S. Ribaric, M. Milani, and Z. Kalafatic. Restoration of images blurred by circular motion. *Image and Signal Processing and Analysis, 2000. Proceedings of the First International Workshop on*, pages 53–60, 2000.

[22] J. Starck, E. Pantin, and F. Murtagh. Deconvolution in astronomy: a review. *Astronomical Society of the Pacific*, pages 1051–1069, 2002.

[23] F. Tsumuraya, N. Miura, and N. Baba. Iterative blind deconvolution method using lucy's algorithm. *Astronomy and Astrophysics*, 282:699–708, 1994.

[24] M. Welk, D. Theis, and J. Weickert. Variational deblurring of images with uncertain and spatially variant blurs. *Pattern Recognition*, pages 485–492, 2005.

[25] J. Yedidia, W. Freeman, and Y. Weiss. Constructing free-energy approximations and generalized belief propagation algorithms. *IEEE Transactions on Information Theory*, (73):2282–2312, 2005.

[26] L. Yuan, J. Sun, L. Quan, and H.-Y. Shum. Image Deblurring with Blurred/Noisy Image Pairs. In *SIGGRAPH*, 2007.

[27] C. B. Zarowin. Robust, noniterative, and computationally efficient modification of van cittert deconvolution optical figuring. *J. Opt. Soc. Am. A*, 11:2571–2583, 1994.