

Soft-Segmentation Guided Object Motion Deblurring

Jinshan Pan^{1,2}, Zhe Hu², Zhixun Su^{1,3}, Hsin-Ying Lee⁴, and Ming-Hsuan Yang²

¹Dalian University of Technology ²University of California, Merced

³National Engineering Research Center of Digital Life ⁴University of Southern California

<http://vllab.ucmerced.edu/~jinshan/projects/object-deblur/>

Abstract

Object motion blur is a challenging problem as the foreground and the background in the scenes undergo different types of image degradation due to movements in various directions and speed. Most object motion deblurring methods address this problem by segmenting blurred images into regions where different kernels are estimated and applied for restoration. Segmentation on blurred images is difficult due to ambiguous pixels between regions, but it plays an important role for object motion deblurring. To address these problems, we propose a novel model for object motion deblurring. The proposed model is developed based on a maximum a posteriori formulation in which soft-segmentation is incorporated for object layer estimation. We propose an efficient algorithm to jointly estimate object segmentation and camera motion where each layer can be deblurred well under the guidance of the soft-segmentation. Experimental results demonstrate that the proposed algorithm performs favorably against the state-of-the-art object motion deblurring methods on challenging scenarios.

1. Introduction

Recent years have witnessed significant advances in deblurring [19, 23, 25]. Numerous methods [5, 10, 24, 38, 42, 47, 48, 50] have been proposed to address this problem, but most of them are designed for camera motion blur. Considerably fewer methods have been proposed to remove image blur caused by moving objects, panning cameras, or both.

Object motion blur is caused by relative motion between a camera and objects, which usually results in different blur effects on moving objects and the background (See Figure 1(a)). Uniform deblurring methods (e.g., [5, 47]) cannot be effectively applied to this problem directly as objects in the scene undergo different motion blurs. Although non-uniform deblurring methods (e.g., [46, 48]) consider different blur effects across an image that are caused by camera rotations and translations, they are less effective for abrupt blur changes caused by fast moving objects or panning cameras. Several methods [3, 8, 26, 41, 44] have been proposed to solve the object motion deblurring problem by

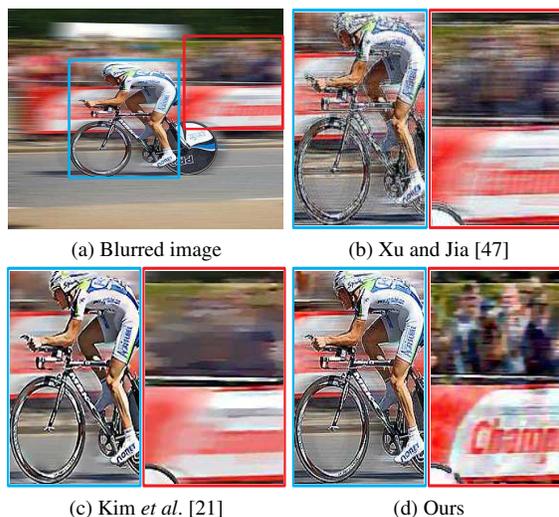


Figure 1. Object motion blur caused by a panning camera.

directly segmenting a blurred image into different regions and deblurring each segmented region. However, it is difficult to identify correct contours of moving objects from a blurred image. The recent method [21] adopts a novel non-local regularization on the residual of estimated results and blurred image to handle object segmentation for dynamic scene deblurring. However, it may not segment the objects undergoing large blurs and affect the deblurred results (See Figure 1(c)). Although segmentation plays a critical role in object motion deblurring, existing methods directly consider it as a pre-processing step and its role in object motion deblurring can be better explored.

In this paper, we propose a novel algorithm for object motion deblurring in which both segmentation as well as deblurring are considered and optimized within one framework. The object layer estimation is achieved by a soft-segmentation method, and the relationship between the soft-segmentation and deblurring is naturally explored and modeled in a maximum a posteriori (MAP) framework. Furthermore, we develop an efficient numerical algorithm to solve the problem. The deblurring component benefits from improving soft-segmentation of the scene, which enables the

proposed method to handle large blur caused by fast moving objects or panning cameras. One challenging example is shown in Figure 1, where the background contains large blur. The proposed algorithm is able to recover the characters on the board in the background whereas the state-of-the-art dynamic scene deblurring method [21] cannot.

2. Related Work and Problem Context

In this section we discuss the most relevant algorithms and put this work in the proper context.

Since blind deblurring is an ill-posed problem, it requires certain assumptions or additional information to constrain the solution space. To solve this issue, numerous regularizations have been developed. One representative work by Chan and Wong [4] uses the total variation to regularize both blur kernels and latent images, and this approach is analyzed in details [34]. Recently, a mixture of Gaussians by Fergus *et al.* [10] is used to approximate the image gradient prior for the latent image and the blur kernels are estimated by a variational Bayesian method. Comprehensive analysis in [29] shows that variational Bayesian based deblurring methods (*e.g.*, [10, 30]) are able to remove trivial solutions in comparison to other approaches with naive MAP formulations. Due to the high computational load of the variational Bayesian inference, some methods improve the MAP based approach by carefully designing image priors and likelihood functions [2, 24, 31, 48, 51] for deblurring.

It has been shown that object boundaries help estimate blur kernels with a transparency (alpha matting) map [18]. The effectiveness of this method hinges on whether a transparency map can be extracted well or not, and improvements have been made [5, 6, 20, 47]. These methods explicitly select sharp edges for kernel estimation and perform well on a recent benchmark dataset [23]. Instead of using priors from natural image statistics, exemplar based methods [13, 33, 42] are proposed to exploit properties for specific object classes or scenes. The aforementioned methods are not formulated for practical camera motion blur, including rotational and translational movements, which lead to spatially variant blur effects.

To deal with spatially variant blur, a general projective motion model is proposed in [45]. Whyte *et al.* [46] simplify this model and solve the deblurring problem in a variational Bayesian framework as [10]. Gupta *et al.* [12] use a motion density function to represent the camera motion trajectory for the non-uniform deblurring. In [39], Shan *et al.* solve the rotational motion blur using a transparency map. Since the optimization steps in the non-uniform deblurring methods are computationally expensive, locally uniform patch-based methods [14, 16] are developed where the deconvolution step can be efficiently computed by the fast Fourier transform. We note that the aforementioned deblurring methods are developed for the camera motion blur and

not effective to account for object motion blur.

Several deblurring methods have been proposed to deal with object motion blur. In [35, 43], hybrid camera systems are designed to acquire additional information from moving objects. Although images can be deblurred well by hybrid camera systems, these methods require specially designed hardware. Levin [26] proposes a new method that first segments blurred regions by comparing likelihoods with a set of one dimensional box filters, and then applies the Richardson-Lucy deconvolution algorithm to each segmented region with its blur kernel. This method is limited by the quality of segmented regions as the segmentation and deblurring are carried out independently. The transparency map [28] is employed by [8, 44] to separate an image into the foreground and background to deal with object motion blur. In [8], Dai and Wu first estimate blur kernels using [1, 7] and then alternatively estimate the transparency map, foreground, and background of latent images. While this method is effective for images with partial blur, the kernel estimation and segmentation processes are independent, which limits the refinement of the segmentation results and accordingly affects the recovered image. Chakrabarti *et al.* [3] use a mixture of Gaussians to model the heavy-tail properties of natural image gradients for deblurring which is able to deal with certain blur (*e.g.*, Gaussian blur with small kernel width). Kim *et al.* [21] alternatively estimate blur kernels and segmentation to handle the dynamic scene deblurring problems. However, this method is less effective for large object motion blur as discussed earlier. In [22], a method based on a local linear motion without segmentation is proposed, which incorporates the optical flow method to guide the blur kernel estimation. Although this method is able to deal with certain object motion blur, the specific assumption on the blur kernel limits the application domains.

We note that Favaro and Soatto [9] develop a unified model to jointly estimate blur and occlusion. However, this method focuses on defocus blur and it has difficulty in handling the blur caused by moving objects.

In this paper, we focus on handling the blur caused by moving objects. Different from existing methods, the proposed algorithm is designed to consider both segmentation and deblurring. We propose a novel formulation that accommodates the soft-segmentation technique to guide the deblurring process in a unified framework.

3. Proposed Algorithm

Since the object blur is mainly caused by moving objects, the blur is not uniform (*e.g.*, the background and foreground in Figure 1 undergo different blurs). Our goal is to split an image into different layers according to moving objects and assume that each layer corresponds to a blur kernel.

We formulate the object motion deblurring problem within a MAP framework. Given a blurred image B , we estimate the latent image I and the blur kernel k ,

$$\begin{aligned}
(I, k) &= \arg \max_{I, k} p(k, I|B) \\
&= \arg \max_{I, k} p(B|k, I)p(k)p(I) \\
&= \arg \max_{I, k} \sum_{i=1}^N p(B, l_i|k, I)p(I)p(k) \\
&= \arg \max_{I, k} \sum_{i=1}^N p(B|l_i, k_i, I)p(l_i|k_i, I)p(I)p(k_i), \tag{1}
\end{aligned}$$

where N denotes the number of segmented layers; l_i is a binary mask for the i -th layer which has the same size as the input image; k_i denotes the blur kernel corresponding to the i -th layer; and $k = \{k_i\}_{i=1}^N$.

For the likelihood $p(B|l_i, k_i, I)$, we assume that pixels of an image are independent. So we have $p(B|l_i, k_i, I) = \prod_{\mathbf{u}} p(B_{\mathbf{u}}|l_{i\mathbf{u}}, k_{i\mathbf{u}}, I_{\mathbf{u}})$, where \mathbf{u} denotes the spatial location of a pixel. The probability $p(B_{\mathbf{u}}|l_{i\mathbf{u}}, k_{i\mathbf{u}}, I_{\mathbf{u}})$ is formulated as the data fitting errors:

$$p(B_{\mathbf{u}}|l_{i\mathbf{u}}, k_{i\mathbf{u}}, I_{\mathbf{u}}) = \begin{cases} \frac{1}{Z_d} \exp(-|(B - I * k_i)_{\mathbf{u}}|) & l_{i\mathbf{u}} = 1, \\ C & l_{i\mathbf{u}} = 0, \end{cases} \tag{2}$$

where Z_d is a normalization term, C is a positive constant, $*$ is a convolution operator, and the Laplacian distribution is used to handle large noise [47]. Based on (2), $p(B|l_i, k_i, I)$ can be equivalently expressed as

$$p(B|l_i, k_i, I) = \frac{1}{Z_d} \exp\left(-\sum_{\mathbf{u}} l_{i\mathbf{u}} |(B - I * k_i)_{\mathbf{u}}|\right). \tag{3}$$

For the prior $p(l_i|k_i, I)$, we introduce an auxiliary segmentation confidence map s_i of the latent image I , which is related to l_i . That is, we set $l_{i\mathbf{u}} = 0$, if $s_{i\mathbf{u}}$ is close to zero and $l_{i\mathbf{u}} = 1$, otherwise. According to the law of total probability, we have

$$\begin{aligned}
p(l_i|k_i, I) &= \sum_{s_i \in \mathcal{S}_i} p(l_i, s_i|k_i, I) \\
&= \sum_{s_i \in \mathcal{S}_i} p(l_i|s_i, k_i, I)p(s_i|I, k_i), \tag{4}
\end{aligned}$$

where \mathcal{S}_i is the space of all possible configurations of s_i .

Since we assume that s_i is a segmentation confidence map of the latent image I , it is independent of the blur kernel k_i . Thus, we have $p(s_i|I, k_i) = p(s_i|I)$ and define the prior $p(s_i|I)$ in this paper as

$$p(s_i|I) = \frac{1}{Z_{s_i}} \exp(-\eta \mathbf{s}_i^T \mathbf{L} \mathbf{s}_i), \tag{5}$$

where Z_{s_i} is a normalization term, η is a weight parameter, \mathbf{s}_i is the vector form of s_i , \mathbf{L} is an Laplacian matrix, and it is defined by $\mathbf{L} = \text{diag}(\mathbf{W}) - \mathbf{W}$, where $\text{diag}(\mathbf{W})$ is a diagonal matrix of \mathbf{W} , and \mathbf{W} is defined by

$$\mathbf{W}_{\mathbf{u}\mathbf{v}} = \exp(-\beta \|I_{\mathbf{u}} - I_{\mathbf{v}}\|^2), \tag{6}$$

where β is a positive weight; \mathbf{u} and \mathbf{v} denote the spatial locations of image pixels. We note that \mathbf{W} is the affinity matrix which is used in normalized cuts [40] and random walk image segmentation method [11]. Levin *et al.* [28] show that the matting Laplacian matrix generates better image segmentation results than those of (6). The affinity matrix in matting is defined by

$$\mathbf{W}_{\mathbf{u}\mathbf{v}} = \sum_{m: (\mathbf{u}, \mathbf{v}) \in \mathbf{w}_m} \frac{1}{\mathcal{C}(\mathbf{w}_m)} \left(1 + \frac{(I_{\mathbf{u}} - \mu_m)(I_{\mathbf{v}} - \mu_m)}{\varepsilon + \sigma_m^2}\right), \tag{7}$$

where \mathbf{w}_m is the m -th patch of I ; $\mathcal{C}(\mathbf{w}_m)$ denotes the number of pixels in \mathbf{w}_m ; μ_m and σ_m are the mean as well as variance of the intensities in \mathbf{w}_m ; and ε is a weight which controls the smoothness of segmentation boundaries. In this paper, we use the matrix (7) to construct \mathbf{L} in (5) (See analysis in Section 6).

The likelihood $p(l_i|s_i, k_i, I)$ measures the similarity of l_i and s_i . Since s_i is independent of k_i , it has similar properties to l_i according to its definition. Therefore, we assume that l_i is independent with respect to k_i and have $p(l_i|s_i, k_i, I) = p(l_i|s_i, I)$ which is defined by

$$p(l_i|s_i, I) = \frac{1}{Z_{l_i}} \exp\left(-\alpha \sum_{\mathbf{u}} D_{\mathbf{u}} (l_{i\mathbf{u}} - s_{i\mathbf{u}})^2\right), \tag{8}$$

where Z_{l_i} is a normalization constant and $D_{\mathbf{u}}$ is a weighted parameter. In this paper, we define $D_{\mathbf{u}}$ which is used in the alpha matting methods [15, 28] as

$$D_{\mathbf{u}} = \sum_{m: (\mathbf{u}, \mathbf{u}) \in \mathbf{w}_m} \frac{1}{\mathcal{C}(\mathbf{w}_m)} \left(1 + \frac{(I_{\mathbf{u}} - \mu_m)^2}{\varepsilon + \sigma_m^2}\right). \tag{9}$$

We note that when an image patch \mathbf{w}_m covers only a smooth region, the value $D_{\mathbf{u}}$ is close to 1. Otherwise, it is larger than 1, which penalizes more on the inconsistency of $l_{i\mathbf{u}}$ and $s_{i\mathbf{u}}$ in (8).

Based on above discussions, the remaining task is to define the priors $p(I)$ and $p(k_i)$ of the latent image I and the blur kernel k_i . We use the sparsity image gradient prior [27] for the latent image I and an Laplacian prior for the blur kernel k_i , which are defined by

$$\begin{aligned}
p(I) &= \frac{1}{Z_I} \exp(-\lambda \phi_I(I)), \\
p(k_i) &= \frac{1}{Z_k} \exp(-\gamma \phi_k(k_i)), \tag{10}
\end{aligned}$$

where $\phi_I(I) = \sum_{\mathbf{u}} (|\partial_x I_{\mathbf{u}}|^{0.8} + |\partial_y I_{\mathbf{u}}|^{0.8})$, $\phi_k(k_i) = \sum_{\mathbf{u}} |k_{i\mathbf{u}}|$; ∂_x and ∂_y denote the differential operators along the x and y directions; and Z_I as well as Z_k are normalization terms; λ and γ are weights.

We take negative log likelihood of (1) and have the pro-

posed deblurring model as follows,

$$\min_{I,k,l,s} \sum_{i=1}^N \sum_{u,v} l_{iu} |(I * k_i - B)_u| + \alpha D_u (l_{iu} - s_{iu})^2 + \gamma |k_{iu}| + \eta \mathbf{W}_{uv} (s_{iu} - s_{iv})^2 + \lambda (|\partial_x I_u|^{0.8} + |\partial_y I_u|^{0.8}). \quad (11)$$

4. Optimization

In this section, we propose an efficient algorithm to solve (11) for object motion deblurring.

4.1. Soft-Segmentation Estimation

Given the estimates of latent image I , blur kernel k_i , and label l_i , the soft-segmentation problem can be modeled as

$$\min_s \sum_i \sum_{u,v} D_u (l_{iu} - s_{iu})^2 + \frac{\eta}{\alpha} \mathbf{W}_{uv} (s_{iu} - s_{iv})^2. \quad (12)$$

We note that the optimization problem with respect to s_i of each layer can be solved separately. For the i -th layer, the problem can be equivalently expressed by

$$\min_{\mathbf{s}_i} (\mathbf{l}_i - \mathbf{s}_i)^\top \mathbf{D} (\mathbf{l}_i - \mathbf{s}_i) + \frac{\eta}{\alpha} \mathbf{s}_i^\top \mathbf{L} \mathbf{s}_i, \quad (13)$$

where \mathbf{l}_i is the vector form of l_i and \mathbf{D} is a diagonal matrix whose element is defined as $\mathbf{D}_{uu} = D_u$. By setting the derivative with respect to s_i to zero, the solution to this optimization problem is given by solving a linear system,

$$\left(\mathbf{D} + \frac{\eta}{\alpha} \mathbf{L} \right) \mathbf{s}_i = \mathbf{D} \mathbf{l}_i. \quad (14)$$

We note that directly solving (14) is computationally expensive due to the large matrix \mathbf{L} . He *et al.* [15] prove that this linear equation can be approximated by a linear edge-preserving filter. In this work, we use the fast approximation algorithm [15] to solve (14) and obtain s_i .

After obtaining s_i , the problem with respect to l_i is

$$\min_{l_i} \sum_{i=1}^N \sum_u l_{iu} |(I * k_i - B)_u| + \alpha D_u (l_{iu} - s_{iu})^2. \quad (15)$$

Note that this is a least-squares problem and the closed-form solution of l_i is

$$l_{iu} = s_{iu} - \frac{1}{2\alpha D_u} |(I * k_i - B)_u|. \quad (16)$$

4.2. Intermediate Latent Image Estimation

With the estimates of k_i and l_i , the latent image estimation can be written as

$$\min_I \sum_{i=1}^N \sum_u l_{iu} |(I * k_i - B)_u| + \lambda (|\partial_x I_u|^{0.8} + |\partial_y I_u|^{0.8}). \quad (17)$$

Algorithm 1 Proposed object motion deblurring algorithm

Input: Blurred image B and the number of layers: N .

Output: Latent image I and blur kernel k_i .

- 1: Initialize I , k_i , s_i , l_i with the results from the coarser level.
 - 2: **for** $t_1 = 1 \rightarrow 3$ **do**
 - 3: solve for s_i by (12);
 - 4: solve for l_i by (15);
 - 5: **for** $t_2 = 1 \rightarrow 5$ **do**
 - 6: solve for k_i by (19);
 - 7: solve for I by (17);
 - 8: **end for**
 - 9: **end for**
-

Since (17) is highly non-convex, we use the iteratively reweighted least squares (IRLS) method [27] to solve it. In each iteration, we need to minimize the weighted quadratic problem,

$$\min_I \sum_{i=1}^N \sum_u l_{iu} \omega_{du} |(I * k_i - B)_u|^2 + \lambda (\omega_u^x |\partial_x I_u|^2 + \omega_u^y |\partial_y I_u|^2), \quad (18)$$

where the weights $\omega_{du} = |(I * k_i - B)_u|^{-1}$, $\omega_u^x = |\partial_x I_u|^{-1.2}$, and $\omega_u^y = |\partial_y I_u|^{-1.2}$ are computed from the results in last iteration.

4.3. Kernel Estimation

In the kernel estimation step, image gradients have been shown to be more effective than the intensities [5, 30, 48]. Thus, we use image derivatives in the data fitting term and remove small gradient values according to [5]. The blur kernel can be estimated by

$$\min_k \sum_{i=1}^N \sum_u l_{iu} \|(\nabla I * k_i - \nabla B)_u\|_1 + \gamma |k_{iu}|. \quad (19)$$

We employ the IRLS method to solve (19).

Similar to the state-of-the-art methods, kernel estimation is carried out in a coarse-to-fine manner using an image pyramid [5] to achieve better performance. Algorithm 1 shows the main steps for the kernel estimation algorithm on one image pyramid level.

Initialization: Since one subproblem of the proposed algorithm involves the soft-segmentation method, the initialization step is important. Similar to [11], we first select seed points as shown in Figure 2(a) and then compute a convex hull to include these points (Figure 2(b)). Another way to initialize l_i is to use rectangular bounding boxes similar to [36] (*e.g.*, the bounding box shown in Figure 2(e)). We show that these two initializations generate similar deblurring results in the following and the supplementary document.

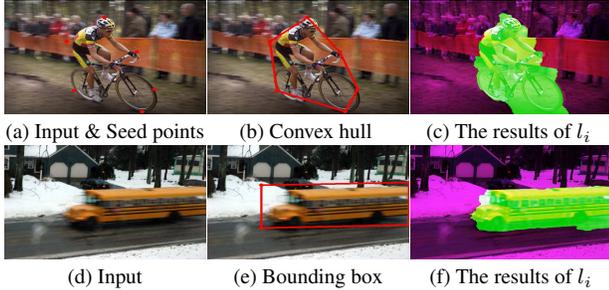


Figure 2. Different initializations in the proposed method.

5. Experimental Results

We present experimental evaluations of the proposed algorithm against several state-of-the-art methods for object motion deblurring. More experimental results and the code can be found at our project website.

Parameter Settings and Implementation Details: In all the experiments, we set $\alpha = 2$, $\lambda = 0.5$, and $\gamma = 0.001$, respectively. The number of layers N is set according to the number of moving objects. As l_i is a binary mask for the i -th layer, we apply the OTSU method to find the threshold for l_i after obtaining l_i by (16). In steps 6 and 7 of Algorithm 1, we estimate the latent image and blur kernel independently at each layer. In (14), we use the estimated RGB image by (17) to compute s_i . Due to separate estimation for each label l_i , the proposed method allows overlapping regions for different layers, which will result in obvious ringing artifacts in the recovered image. To deal with boundaries, we normalize the overlapping regions for different layers so that the sum of each label satisfies $\sum_{i=1}^N l_i = \mathbf{1}$ in the intermediate latent image estimation step, where $\mathbf{1}$ has the same size as each label l_i and its element value is 1.

Quantitative Evaluation: We first evaluate the proposed algorithm using 16 synthetic images. To synthesize blurred images, we use the matting method [28] to separate each clear image into background and foreground regions, to which different blur kernels are applied. Finally, we merge the background and foreground regions using the alpha map to generate the blurred images according to [28] (See Figure 3(a)). We use the PSNR metric to measure the quality of each restored image. Table 1 shows the quantitative evaluation results of each method. The proposed algorithm generates the deblurred images with higher PSNR values¹. One example is shown in Figure 3 for visual comparisons. The deblurred image by the proposed algorithm is clearer (*e.g.*, the vehicle wheel in Figure 3(c)) and the segmentation results are similar to the ground truth.

Real Images: We then evaluate the proposed algorithm using real blurry images. Figure 4(a) shows one example

¹Since the code of existing deblurring methods (*e.g.*, [3, 21, 26, 37]) is not available, we compare the most related method [22] based on our implementation on this dataset.

where the background contains large blur. Because the image blur cannot be described by a uniform blur kernel, state-of-the-art uniform deblurring methods [5, 47] do not perform well on this image. Although the non-uniform deblurring methods [46, 48] are able to deal with the blur caused by camera rotation and translation, they are less effective in handling the abrupt blur changes caused by moving objects. The deblurred image generated by the object motion deblurring method [21] still contains blur effects and the boundaries of the cyclist contain ringing artifacts due to the failure of segmentation. In contrast, the proposed algorithm generates clearer results both in the background and foreground regions. The persons in the background can be recognized and the foreground (*e.g.*, the head in the blue box of Figure 4(g) and (h)) is also comparable to other methods. The results shown in Figure 4(g) and (h) demonstrate that the proposed algorithm is robust to different initializations.

Figure 5(a) shows another example with moving objects and large blur in the background. Since the blur in this image is different at each region, the uniform [5, 47] and non-uniform [46, 48] deblurring methods do not generate clear results. The deblurred images are similar to the blurred image in Figure 5(a). Compared to the camera shake deblurring methods [5, 46, 47, 48] and the recent object motion deblurring approach [21], the proposed algorithm generates a clear image where the type of contents in the background (*e.g.*, people) can be identified.

Comparisons with Segmentation-Free Object Deblurring Methods: Recently, Kim and Lee [22] propose a deblurring method to deal with dynamic scenes without using segmentation. As the blur model is based on local linear motion flow vectors, it is difficult to deal with complex object and camera motions. Figure 6(a) shows an example from [22]. The deblurred image generated by the non-uniform camera shake deblurring method [46] contains ringing artifacts due to the influences of moving objects. Compared with the reported results in [22], the proposed algorithm generates a sharper image with more details (*e.g.*, door, window, and grass regions).

6. Analysis and Discussion

In this section, we present more analysis on how the proposed algorithm performs on object motion deblurring and discuss its connection to the most relevant methods. In addition, we discuss the limitations and extensions of the proposed algorithm.

Compared to the existing MAP-based deblurring methods, the proposed algorithm introduces an additional term l_i , which helps segment an image into different layers. This is mainly because that the sub-problem (12) is the alpha matting method in [15, 28], which provides soft-segmentation of an object in the input image. After optimizing (15), the label l_i is obtained based on the segmentation

Table 1. Quantitative comparisons on the synthetic examples.

	Blur images	Cho and Lee [5]	Xu and Jia [47]	Whyte <i>et al.</i> [46]	Xu <i>et al.</i> [48]	Kim and Lee [22]	Ours
Average PSNRs	22.27	21.40	21.71	13.25	21.90	21.66	22.79

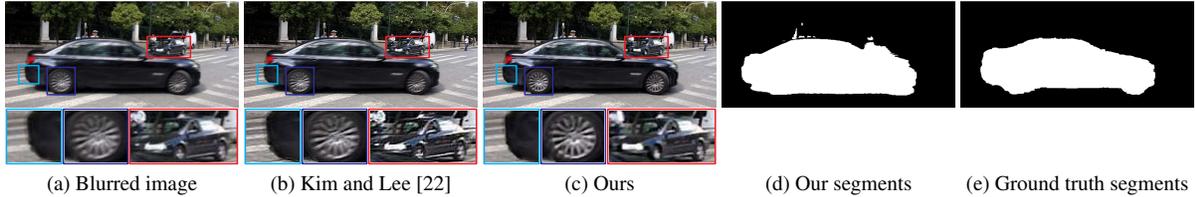


Figure 3. Visual comparisons on an example from the synthetic dataset.



Figure 4. An object motion deblurring example where the background contains large blur.

result s_i . Figure 7(c) shows one intermediate result of l_i .

We note that the work [21] also focuses on the object motion deblurring problem in which regular patches are used at first and refined iteratively. The segmentation results are achieved by solving an optimization problem with non-local regularization on the data fidelity term. Since the data fidelity term is based on the residual of estimated results and a blur image, it is not robust to large motion blur. As shown in Figure 7(b), this method is less effective when segmenting moving objects. The segmented foreground contain not only clear regions but also blurred ones, which affect kernel estimation and lead to the blurry results (See Figure 7(g)). To deal with partial blur, Schelten and Roth [37] segment an image by a variational Bayesian method, which is computationally expensive. Furthermore, the inference step is still based on the residual of estimated results and a blurred image ((14) in [37]), which is not robust to large blur. We also note that the matting method has been used in [17]. However, this method mainly focuses on camera shake re-

moval that takes scene depth into consideration. As the layers for different depth are pre-computed using matting and are fixed during the optimization, the algorithm does not update the segment of moving objects. In addition, global constraints on all the layers in [17] are enforced based on the camera motion which is different from our scenarios. Different from prior work, our method is formulated within a unified probabilistic framework for both blur and soft-segmentation estimations. It explicitly incorporates the segmentation method (*i.e.*, matting method [28]) which relies on the estimated latent image, thereby facilitating the segmentation task (See Figure 7(c)). Moreover, the proposed segmentation problem has a closed-form solution, which can be efficiently solved by [15]. Figure 7(i) shows that our deblurred image recovers fine textures.

We also note that several approaches (*e.g.*, [26, 41]) first detect blur regions and then use existing methods (*e.g.*, [32, 47]) to deblur images. However, these methods are limited by whether image regions can be identified well



Figure 5. An example with a fast moving object and the background contains large blur.

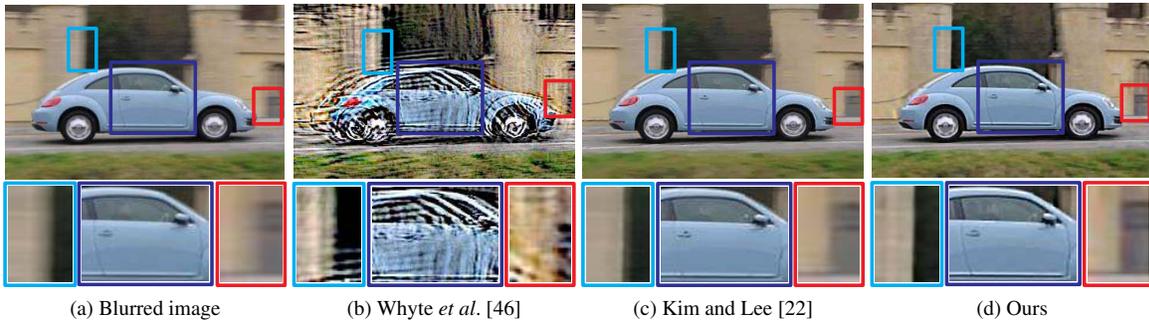


Figure 6. Comparisons with the segmentation-free dynamic scene deblurring method [22].

or not. Figure 8 shows one example from [21]. We compare the blur detection method [41] and the proposed deblurring method without using soft-segmentation. Figure 8(a) shows that the approach in [41] does not always generate clear results. The result in Figure 8(b) suggests that it is not effective to deblur an image with only a pre-detected region. Although the soft-segmentation step using (6) is able to generate segmentation results (See Figure 8(h)), the algorithm using the matting Laplacian matrix generates better results. Thus, we use (7) in the soft-segmentation step. In addition, the segmentation results shown in Figure 8(e)-(g) indicate that the proposed algorithm has good convergence in practice.

The proposed method can also be incorporated with a clustering method [52]. If we solve the sub-problem (12) on a super-pixel level and consider each pixel as a graph node, this becomes the saliency detection problem considered in [49].

Limitations and Extensions. The object motion deblurring method described in Section 3 is based on the assumption that each layer has a uniform blur kernel. This assumption holds for dynamic scenes with the translation motion. However, it is less effective if moving objects involve the rotation and non-rigid motion. For a complex scene, the blur caused by moving objects is often spatially variant [3, 43, 44]. We can use the geometric model of camera motion [45, 46] to approximate the non-uniform object motion blur caused by panning cameras. According to [45, 46], the blur process is modeled by

$$\mathbf{B} = \sum_{j=1}^t w_j K_{\theta_j} \mathbf{I} + \mathbf{e}, \quad (20)$$

where \mathbf{B} , \mathbf{I} , and \mathbf{e} denote the vector forms of the blurred image, latent image, and noise, respectively; $\{\theta_j\}_{j=1}^t$ denote the sampled camera poses; $\{K_{\theta_j}\}_{j=1}^t$ are the warping matrixes corresponding to different sampled camera poses, which transform the latent image \mathbf{I} accordingly; and $\{w_j\}_{j=1}^t$ are weights that satisfy $w_j \geq 0$ as well as

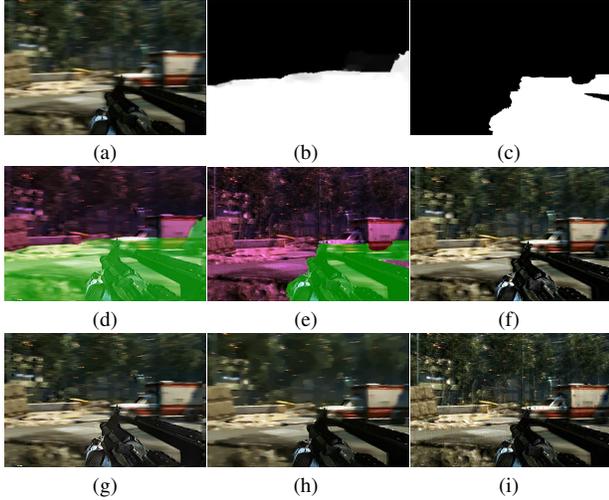


Figure 7. The effectiveness of the proposed soft-segmentation in object deblurring. (a) Blurred image. (b) One intermediate result of [21]. (c) One intermediate result of l_i by the proposed algorithm. (d) The segmentation results of [21]. (e) Our segmentation results. (f)-(h) Results of Xu and Jia [47] and dynamic scene deblurred results [21, 22], respectively. (i) Ours.

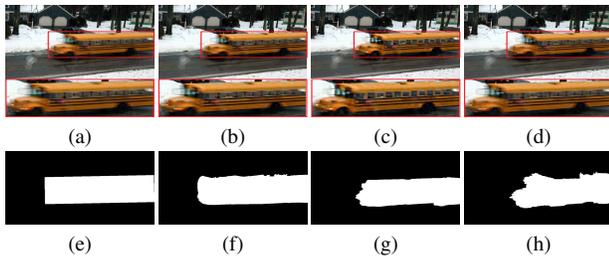


Figure 8. Deblurring methods with pre-detected blur regions. (a) Result of the blur detection method [41]. (b) Result by the proposed algorithm with fixed initialization in (e). (c) Result by the proposed algorithm using (7). (d) Result by the proposed algorithm using (6). (e)-(g) show some segmentation results using (7) over iterations. (h) Estimated segmentation result using (6).



Figure 9. Comparisons with optical flow based object motion deblurring method [7].

$$\sum_j w_j = 1.$$

Based on (20), our non-uniform deblurring method is similar to Algorithm 1, where we only need to replace (16), (17), and (19) with

$$l_{iu} = \alpha s_{iu} + \left| \left(\sum_{j=1}^t w_j K_{\theta_j} \mathbf{I} - \mathbf{B} \right)_u \right| / 2, \quad (21)$$

$$\min_{\mathbf{I}} \sum_{i=1}^N \sum_{\mathbf{u}} l_{iu} \left| \left(\sum_{j=1}^t w_j K_{\theta_j} \mathbf{I} - \mathbf{B} \right)_u \right| + \lambda (|\partial_x \mathbf{I}_u|^{0.8} + |\partial_y \mathbf{I}_u|^{0.8}), \quad (22)$$

and

$$\min_{w_j} \sum_{i=1}^N \sum_{\mathbf{u}} l_{iu} \left| \left(\sum_{j=1}^t w_j K_{\theta_j} (\nabla \mathbf{I}) - \nabla \mathbf{B} \right)_u \right| + \gamma \sum_{j=1}^t |w_j|. \quad (23)$$

We use the fast forward approximation approach with the locally-uniform assumption in [16] to solve the above models.

This extension is able to handle some examples with non-rigid motion to some extent. Figure 9(a) shows a blurred example from [7] where the blur is caused by the non-rigid motion of the hand. Although our deblurred image contains some blur effect (*e.g.*, the area near the little finger), it has fewer ringing artifacts compared to the result generated by [7].

We note that the proposed method can also be applied to deblur images caused by camera rotation and translation. More experimental results are included in the supplemental material.

7. Conclusion

In this paper, we propose an object motion deblurring algorithm within a MAP framework. The proposed method incorporates a soft-segmentation method to take moving objects and background regions into account for kernel estimation. We present an efficient algorithm to solve the proposed model. Experimental results show that the proposed algorithm performs favorably against the state-of-the-art object motion deblurring methods as well as non-uniform deblurring approaches.

We note that the proposed algorithm is based on soft-segmentation and may fail for the images where the blur is caused by both depth variation and moving objects (*e.g.*, Figure 9). Our future work will focus on exploiting the depth information to facilitate the kernel estimation for object motion deblurring.

Acknowledgements: We thank T. H. Kim for providing the deblurred results of his methods [21, 22]. J. Pan is supported by a scholarship from China Scholarship Council. Z. Su is supported by the NSFC (No. 61572099 and 61320106008). Z. Hu and M.-H. Yang are supported in part by the NSF CAREER Grant (No. 1149783), NSF IIS Grant (No. 1152576).

References

- [1] S. Bae and F. Durand. Defocus magnification. *Computer Graphics Forum*, 26(3):571–579, 2007.
- [2] J.-F. Cai, H. Ji, C. Liu, and Z. Shen. Framelet based blind motion deblurring from a single image. *IEEE TIP*, 21(2):562–572, 2012.

- [3] A. Chakrabarti, T. Zickler, and W. T. Freeman. Analyzing spatially-varying blur. In *CVPR*, pages 2512–2519, 2010.
- [4] T. Chan and C. Wong. Total variation blind deconvolution. *IEEE TIP*, 7(3):370–375, 1998.
- [5] S. Cho and S. Lee. Fast motion deblurring. In *SIGGRAPH Asia*, pages 145:1–145:8, 2009.
- [6] T. S. Cho, S. Paris, B. K. P. Horn, and W. T. Freeman. Blur kernel estimation using the radon transform. In *CVPR*, pages 241–248, 2011.
- [7] S. Dai and Y. Wu. Motion from blur. In *CVPR*, 2008.
- [8] S. Dai and Y. Wu. Removing partial blur in a single image. In *CVPR*, pages 2544–2551, 2009.
- [9] P. Favaro and S. Soatto. Seeing beyond occlusions (and other marvels of a finite lens aperture). In *CVPR*, pages 579–586, 2003.
- [10] R. Fergus, B. Singh, A. Hertzmann, S. T. Roweis, and W. T. Freeman. Removing camera shake from a single photograph. In *SIGGRAPH*, pages 787–794, 2006.
- [11] L. Grady. Random walks for image segmentation. *IEEE TPAMI*, 28(11):1768–1783, 2006.
- [12] A. Gupta, N. Joshi, C. L. Zitnick, M. F. Cohen, and B. Curless. Single image deblurring using motion density functions. In *ECCV*, pages 171–184, 2010.
- [13] Y. HaCohen, E. Shechtman, and D. Lischinski. Deblurring by example using dense correspondence. In *ICCV*, pages 2384–2391, 2013.
- [14] S. Harmeling, M. Hirsch, and B. Schölkopf. Space-variant single-image blind deconvolution for removing camera shake. In *NIPS*, pages 829–837, 2010.
- [15] K. He, J. Sun, and X. Tang. Guided image filtering. *IEEE TPAMI*, 35(6):1397–1409, 2013.
- [16] M. Hirsch, C. J. Schuler, S. Harmeling, and B. Schölkopf. Fast removal of non-uniform camera shake. In *ICCV*, pages 463–470, 2011.
- [17] Z. Hu, L. Xu, and M.-H. Yang. Joint depth estimation and camera shake removal from single blurry image. In *CVPR*, pages 2893–2900, 2014.
- [18] J. Jia. Single image motion deblurring using transparency. In *CVPR*, 2007.
- [19] J. Jia. *Mathematical models and practical solvers for uniform motion deblurring*. Cambridge University Press, 2014.
- [20] N. Joshi, R. Szeliski, and D. J. Kriegman. PSF estimation using sharp edge prediction. In *CVPR*, 2008.
- [21] T. H. Kim, B. Ahn, and K. M. Lee. Dynamic scene deblurring. In *ICCV*, pages 3160–3167, 2013.
- [22] T. H. Kim and K. M. Lee. Segmentation-free dynamic scene deblurring. In *CVPR*, pages 2766–2773, 2014.
- [23] R. Köhler, M. Hirsch, B. J. Mohler, B. Schölkopf, and S. Harmeling. Recording and playback of camera shake: Benchmarking blind deconvolution with a real-world database. In *ECCV*, pages 27–40, 2012.
- [24] D. Krishnan, T. Tay, and R. Fergus. Blind deconvolution using a normalized sparsity measure. In *CVPR*, pages 2657–2664, 2011.
- [25] S. Lee and S. Cho. Recent advances in image deblurring. In *SIGGRAPH Asia 2013 Course*, 2013.
- [26] A. Levin. Blind motion deblurring using image statistics. In *NIPS*, pages 841–848, 2006.
- [27] A. Levin, R. Fergus, F. Durand, and W. T. Freeman. Image and depth from a conventional camera with a coded aperture. In *SIGGRAPH*, pages 70–78, 2007.
- [28] A. Levin, D. Lischinski, and Y. Weiss. A closed form solution to natural image matting. In *CVPR*, pages 61–68, 2006.
- [29] A. Levin, Y. Weiss, F. Durand, and W. T. Freeman. Understanding and evaluating blind deconvolution algorithms. In *CVPR*, pages 1964–1971, 2009.
- [30] A. Levin, Y. Weiss, F. Durand, and W. T. Freeman. Efficient marginal likelihood optimization in blind deconvolution. In *CVPR*, pages 2657–2664, 2011.
- [31] Y. Lou, A. L. Bertozzi, and S. Soatto. Direct sparse deblurring. *Journal of Mathematical Imaging and Vision*, 39(1):1–12, 2011.
- [32] L. B. Lucy. An iterative technique for the rectification of observed distributions. *Astronomy Journal*, 79(6):745–754, 1974.
- [33] J. Pan, Z. Hu, Z. Su, and M.-H. Yang. Deblurring face images with exemplars. In *ECCV*, pages 47–62, 2014.
- [34] D. Perrone and P. Favaro. Total variation blind deconvolution: The devil is in the details. In *CVPR*, pages 2909–2916, 2014.
- [35] R. Raskar, A. K. Agrawal, and J. Tumblin. Coded exposure photography: motion deblurring using fluttered shutter. In *SIGGRAPH*, pages 795–804, 2006.
- [36] C. Rother, V. Kolmogorov, and A. Blake. “GrabCut”: interactive foreground extraction using iterated graph cuts. In *SIGGRAPH*, pages 309–314, 2004.
- [37] K. Schelten and S. Roth. Localized image blur removal through non-parametric kernel estimation. In *ICPR*, pages 702–707, 2014.
- [38] Q. Shan, J. Jia, and A. Agarwala. High-quality motion deblurring from a single image. In *SIGGRAPH*, 2008.
- [39] Q. Shan, W. Xiong, and J. Jia. Rotational motion deblurring of a rigid object from a single image. In *ICCV*, pages 1–8, 2007.
- [40] J. Shi and J. Malik. Normalized cuts and image segmentation. *IEEE TPAMI*, 22(8):888–905, 2000.
- [41] J. Shi, L. Xu, and J. Jia. Discriminative blur detection features. In *CVPR*, pages 2965–2972, 2014.
- [42] L. Sun, S. Cho, J. Wang, and J. Hays. Edge-based blur kernel estimation using patch priors. In *ICCP*, 2013.
- [43] 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 TPAMI*, 32(6):1012–1028, 2010.
- [44] Y.-W. Tai, N. Kong, S. Lin, and S. Y. Shin. Coded exposure imaging for projective motion deblurring. In *CVPR*, pages 2408–2415, 2010.
- [45] Y.-W. Tai, P. Tan, and M. S. Brown. Richardson-lucy deblurring for scenes under a projective motion path. *IEEE TPAMI*, 33(8):1603–1618, 2011.
- [46] O. Whyte, J. Sivic, A. Zisserman, and J. Ponce. Non-uniform deblurring for shaken images. *IJCV*, 98(2):168–186, 2012.
- [47] L. Xu and J. Jia. Two-phase kernel estimation for robust motion deblurring. In *ECCV*, pages 157–170, 2010.

- [48] L. Xu, S. Zheng, and J. Jia. Unnatural L_0 sparse representation for natural image deblurring. In *CVPR*, pages 1107–1114, 2013.
- [49] C. Yang, L. Zhang, H. Lu, X. Ruan, and M.-H. Yang. Saliency detection via graph-based manifold ranking. In *CVPR*, pages 3166–3173, 2013.
- [50] H. Zhang, D. P. Wipf, and Y. Zhang. Multi-image blind deblurring using a coupled adaptive sparse prior. In *CVPR*, pages 1051–1058, 2013.
- [51] H. Zhang, J. Yang, Y. Zhang, and T. S. Huang. Sparse representation based blind image deblurring. In *ICME*, pages 1–6, 2011.
- [52] D. Zhou, J. Weston, A. Gretton, O. Bousquet, and B. Schölkopf. Ranking on data manifolds. In *NIPS*, pages 169–176, 2003.