Dynamic Scene Deblurring

Tae Hyun Kim, Byeongjoo Ahn, and Kyoung Mu Lee
Department of ECE, ASRI, Seoul National University, 151-742, Seoul, Korea
{lliger9, bjahn11, kyoungmu}@snu.ac.kr, http://cv.snu.ac.kr

Abstract

Most conventional single image deblurring methods assume that the underlying scene is static and the blur is caused by only camera shake. In this paper, in contrast to this restrictive assumption, we address the deblurring problem of general dynamic scenes which contain multiple moving objects as well as camera shake. In case of dynamic scenes, moving objects and background have different blur motions, so the segmentation of the motion blur is required for deblurring each distinct blur motion accurately. Thus, we propose a novel energy model designed with the weighted sum of multiple blur data models, which estimates different motion blurs and their associated pixel-wise weights, and resulting sharp image. In this framework, the local weights are determined adaptively and get high values when the corresponding data models have high data fidelity. And, the weight information is used for the segmentation of the motion blur. Non-local regularization of weights are also incorporated to produce more reliable segmentation results. A convex optimization-based method is used for the solution of the proposed energy model. Experimental results demonstrate that our method outperforms conventional approaches in deblurring both dynamic scenes and static scenes.

1. Introduction

Blurring artifacts are among the most common flaws in photographs. Camera shake or motion of objects during the time of exposure cause these artifacts under low light conditions. To address this problem, single image deblurring methods which restore a sharp image from a blurred image have been considerable research in the field of computer vision with the recent increased demand for clear images.

In general, the blind deblurring problem that restores the blurry image without knowing the blur kernel is highly ill-posed. So various energy models that are composed of the regularization and data term have been proposed to find the sharp image and the blur kernel jointly, in the form of

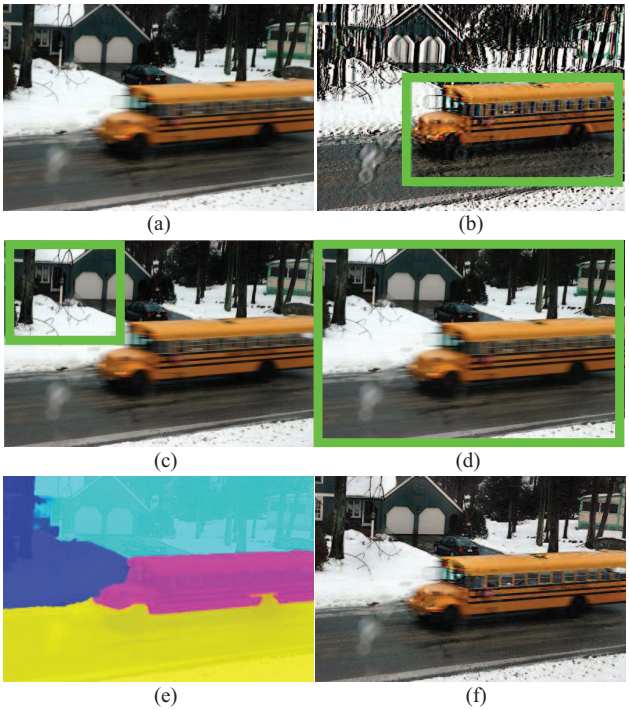


Figure 1. Comparison of deblurring results. The green box illustrates the region used for blur kernel estimation for each image. (a) Input dynamic blurry image of a moving bus. (b)-(c) Deblurring results of Xu et al. [19]. (d) Deblurring result of Whyte et al. [18]. (e) Our motion blur segmentation result. Each color denotes a blur kernel and a map of its associated weight variable. (f) Our deblurring result. Note that both the bus and background region are restored significantly better than those in (b)-(d).

$$E = E_{data}(\mathbf{L}, \mathbf{K}, \mathbf{B}) + E_{reg}(\mathbf{L}, \mathbf{K}), \tag{1}$$

where $\mathbf{L}$ and $\mathbf{B}$ denote the vector form of the latent and blurred images, respectively. The matrix $\mathbf{K}$ denotes the blur kernel whose row vector corresponds to the blur kernel placed at each pixel location. The data term E_{data} measures the data fidelity and the regularization term E_{reg} enforces the smoothness constraint to the latent image as well as to the blur kernel.

Depending on the type of the blur kernel, blind deblur-

ring approaches can be categorized into two types. One is the uniform kernel approach, which assumes the blur kernel is spatially invariant, and the other is the non-uniform approach, which assumes the blur kernel varies spatially.

If we assume that the blur kernel is shift invariant and uniform over the entire image [4, 15], it is possible to restore the latent image quickly with the aid of a fast Fourier transform (FFT) and parallel processing [20, 2, 19]. However the assumption of shift invariant blur kernel does not hold good when there exists a rotational movement of the camera or a moving object in the shaken image.

To alleviate these limitations of uniform motion blur assumption, several non-uniform kernel based methods are proposed. In particular, recent approaches focus on modeling the rotation of camera as well as translation [5, 7, 18], and they obtained promising results in the deblurring of *static* scene.

However, problems still remain in more general settings where not only camera shake but also moving objects exist. For example, in Fig. 1, restoring the image with the uniform blur kernel that has been estimated from the moving bus raises a severe artifacts in the background region (Fig. 1(b)). And also the uniform kernel estimated from the background fails deblurring the bus (Fig. 1(c)). Note that even the state-of-the-art non-uniform blur kernel method [18] which can deblur rotational camera shake does not restore the moving bus either (Fig. 1(d)).

So, the dynamic scene deblurring problem is deeply challenging. Thus far, only a limited amount of research has been done on this problem [11, 6, 8]. However, much of this work is still in a nascent stage and even the hardware assisted method can not handle this problem well [9].

Levin [11] proposed a sequential two-stage approach to solve this problem. She argued that moving objects and background should be handled with different blur kernels to remove the artifacts from deblurring with an inaccurate blur kernel. To begin with, she segmented blur motions by comparing likelihoods with a set of given one dimensional box filters, then applied the Richardson-Lucy deconvolution algorithm to each segmented region with its corresponding box filter. For the first time, she approached this challenging problem with a simple and intuitive way. However, the kernel for the segmentation is limited to the box filters and thus the poor segmentation results could cause undesirable artifacts since the segmentation-stage and the deblurring-stage are separated. Harmeling et al. [6] proposed a method that restores overlapping patches of the blurred image. This approach could handle smoothly varying blur kernels but could not handle the abrupt change of the blur kernel near the boundary of moving objects since they did not segment the motion blurs. More recent work of Ji et al. [8] is based on the interpolation of initially estimated kernels and showed much better results by reducing errors from inaccurate blur kernels, but it also could not overcome the motion boundary problems as in [6].

In principle, the dynamic scene deblurring problem also requires the segmentation of differently blurred regions. In this work, we address the problem of estimating latent image as well as different blur motions and their implicit (soft) segmentations. To the best of our knowledge, this is the first dynamic deblurring work that can estimate these variables jointly. In our framework, we propose a new energy model including multiple blur kernels and their associated pixel-wise weights. The weight of a kernel takes high value he kernel gives high data fidelity. At the same time, the blur kernels are estimated from the pixels whose associated weights have high values. Therefore, locally varying weight information allow us to segmentation the blur motions. In addition, we add non-local regularization to the weight variables to enforce the smoothness in segmentation.

In this study, we introduce a more general and new deblurring framework that can adaptively combine different blur models to estimate the spatially varying blur kernels. Also, as illustrated in Fig. 1(e)-(f), we provide the segmentation of the motion blur as well as better deblurring results. Note that since our framework is general in nature, any blur models and optimization method can be incorporated. We demonstrate the effectiveness of our new deblurring framework by the test results on very challenging images on which conventional techniques break down.

2. Dynamic Scene Deblurring Model

In our dynamic scene deblurring model, we assume the existence of various blur motions. So we have to find both blur kernels and their corresponding blur regions. Also, we do not restrict the types of blur kernels, so we employ both the uniform kernels, which are simple and fast, and the non-uniform kernels which can handle camera rotation.

As there are multiple blur kernels in a dynamic scene, each blurred pixel should be restored from one of them and each kernel should be estimated from its related pixels. For this, we introduce pixel-wise weight variables. A pixel-wise weight variable is associated with a blur kernel and it gains high values on the pixels related with the blur kernel. Therefore, the weight variables imply the segmentation of motion blur. The proposed energy model is given by,

$$E = E_{data}(\mathbf{L}, \mathbf{W}, \mathbf{K}, \mathbf{B}) + E_{reg}(\mathbf{L}, \mathbf{W}, \mathbf{K}). \qquad (2)$$

The set $\mathbf{K} = \{\mathbf{K}_i\}$ denotes a set of N blur kernels, and the set $\mathbf{W} = \{\mathbf{W}_i\}$ means a set of N weight variables where $i = 1, 2, \ldots, N$. A weight vector $\mathbf{W}_i$ is associated with the corresponding blur kernel $\mathbf{K}_i$.

Compared with the conventional model in (1), our new energy model involves additional weight variables and multiple blur kernels, so it becomes a more complex and chal-

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

Figure 2. Multiple blur kernel models give a much better result without additional process to remove ringing artifact. (a) A shaken image of static scene. (b) Deblurring result of Shan et al. [15]. Severe ringing artifacts are observed near edges. (c) Deblurring result of our method with one uniform kernel. It shows serious ringing artifacts near edges. (d) Deblurring results of our method with six uniform blur kernels and weight variables. Ringing artifacts are significantly reduced and the texture of the wallpaper is restored well. (e) Illustration of six weight variables. Slightly different blur kernels are estimated and the result shows that it is hard to estimate a perfect uniform kernel due to unexpected blur effects.

lenging problem. Note that the conventional model is a special case of our model. Since (2) is more general, it could also provide reliable results even for static scenes.

Note that even in the case of a static scene with only translational camera shake, the captured real image may contain various blur motions because depth variation or radial distortion may generate unexpected blur effects. Since the proposed model employs multiple blur kernels, it could handle this problem complementary and produce much better results than the conventional methods. In Fig. 2, six similar but slightly different uniform blur kernels and their associated six blur regions are jointly estimated and a sharper deblurred image with less ringing artifacts is obtained by our method.

2.1. Adaptive Blur Model Selection

In this section, we propose a data term that adaptively selects and fuses proper blur models among candidate models. For this, we adopt a strategy that chooses the locally (pixel-wise) best suited model by measuring the data fidelities and gives a high value to the associated weight variable. At the same time, to obtain correct blur kernels, it is required that each blur kernel is estimated from pixels whose associated weight variable shows high values. In this way, the set $\mathbf{W}$ segments the motion blurs by selecting the locally best suited data model. The data term of the proposed new energy model is formulated by a weighted sum of the multiple data models with some constraints as follows, and minimizing it is equal to select locally best data model.

$$E_{data}(\mathbf{L},\mathbf{W},\mathbf{K},\mathbf{B}) = \lambda \sum_{i=1}^{N} \sum_{\partial_*} \|\mathbf{W}_i^{\frac{1}{2}} \odot (\mathbf{K}_i \partial_* \mathbf{L} - \partial_* \mathbf{B})\|^2 \tag{3}$$

where N is the number of maximal blur models in the scene and λ is the parameter adjusting the scale of our data term and the continuous weight vector is constrained to be (pixel-wise) $\mathbf{W}_i \succeq \boldsymbol{0}$ and $\sum_{i=1}^{N} \mathbf{W}_i = \boldsymbol{1}$. The operator $\odot$ means the element-wise (Hadamard) product of two vectors and the operator $\partial_* \in \{\partial_x, \partial_y\}$ denotes the partial derivative in horizontal and vertical directions [2]. To reduce ringing artifacts, we also use gradient maps, but we do not use brightness map or second order gradient maps unlike conventional methods [15, 2, 19]. Despite this, we can obtain satisfying results by means of multiple blur models and reduce the computational cost.

2.2. Regularization

As dynamic scene deblurring is a highly ill-posed problem, regularization enforcing the smoothness of variables is necessary to obtain a reliable solution. In our energy model, three primal variables are the latent image $\mathbf{L}$, the set of blur kernel matrices $\mathbf{K}$ and the set of the weight variables $\mathbf{W}$, and each has different kinds of regularization as

$$E_{reg}(\mathbf{L},\mathbf{W},\mathbf{K}) = E_{reg}(\mathbf{L}) + E_{reg}(\mathbf{W}) + E_{reg}(\mathbf{K}), \tag{4}$$

and the details of which are described in the following sections.

2.2.1 Regularization on L

We design the latent image to be sharp in edge regions and smooth in flat regions to suppress noise. For this purpose many researchers have studied various priors of the latent image and it is known that l_p norm on the gradient map with $0.7 \leq p \leq 1$ could capture the statistics of natural images with heavy tailed distribution [12, 10, 5]. However, conventional optimization algorithms with sparse norm less than

$p < 1$ are hard to optimize and require additional computational efforts. Thus, our model adopts the total variation model used in [19] as the prior of the latent image, as follows:

$$E_{reg}(\mathbf{L}) = |\nabla \mathbf{L}|. \quad (5)$$

2.2.2 Regularization on W

We assumed that a blurry object can be restored by one of the various blur models, and the motion blur does not change abruptly except on the boundary of a moving object. So, the locally varying weight variable should be segmented and we adopt non local regularization for this purpose.

Non-local regularization is widely used in computer vision and have come into the spotlight lately [17, 16]. The formulation incorporated in our deblurring model is,

$$E_{reg}(\mathbf{W}) = \sum_{i=1}^{N} \sum_{\mathbf{x}} \sum_{\mathbf{y} \in \mathcal{N}(\mathbf{x})} g(\mathbf{x}, \mathbf{y}) |\mathbf{W}_i(\mathbf{x}) - \mathbf{W}_i(\mathbf{y})|, \quad (6)$$

where $\mathcal{N}(\mathbf{x})$ denotes neighboring pixels of $\mathbf{x}$ and the function $g(\mathbf{x}, \mathbf{y})$ is a non-local similarity map which is used to define the mutual support between the pixels at positions $\mathbf{x}$ and $\mathbf{y}$. Similar to the work in [17] the non-local similarity map between two neighboring pixels is defined as

$$g(\mathbf{x}, \mathbf{y}) = e^{-(\frac{\|\mathbf{x}-\mathbf{y}\|}{\sigma_D})^2} \cdot e^{-(\frac{\mathbf{L}_0(\mathbf{x}) - \mathbf{L}_0(\mathbf{y})}{\sigma_I})^2}, \quad (7)$$

where the parameters σ_D and σ_I are used to adjust the slope of the non-local similarity map and the given latent image $\mathbf{L}_0$ can be obtained from the initial of each level in the coarse to fine approach or previous result in the iterative optimization procedure. To reflect the properties of the weight variable $\mathbf{W}_i$, with similar values between neighboring pixels but discontinuity on the boundary of moving objects, it is necessary to use a model that could give sparsity on the difference of weights between neighboring pixels. For this reason, our regularization of weight variables is also based on the total variation. Note that, in contrast to the result from weight regularization with only four neighbors, the result from regularization with dozens of neighbors is much better in both motion blur segmentation and deblurring as shown in Fig. 3.

2.2.3 Regularization on K

As we use both uniform and non-uniform kernels in our blur models, two different regularization models are required.

First, if the blur kernel matrix consists of uniform blur kernel, we use Tikhonov regularization which is typically used in other methods of uniform blur kernel regularization due to its simplicity [20, 2, 19]. By using this regularization, we can have a smooth kernel. The energy function for regularization on an uniform kernel $\mathbf{K}_i$ is formulated by

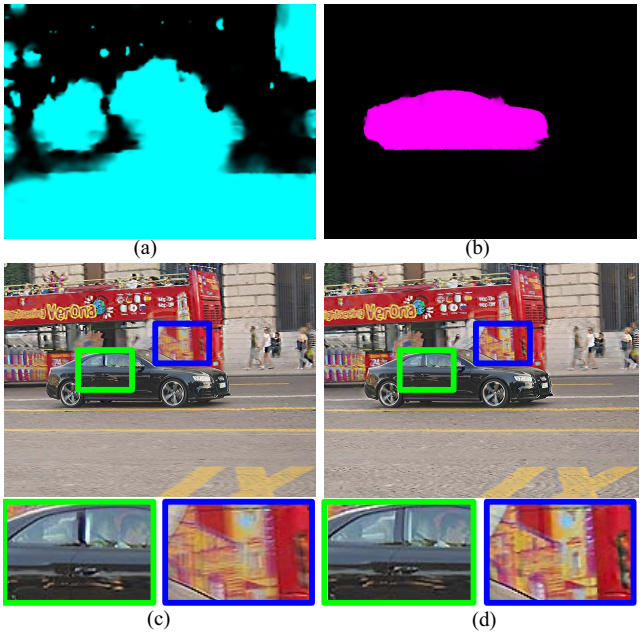


Figure 3. Comparison of weight maps corresponding to the car, and the deblurring results with varying number of neighboring pixels in non-local regularization. (a) 4 neighbor pixels are used, giving a noisy and uneven result. (b) 80 neighbor pixels are used and the weight variable has high values in the exact area of the car. (c)-(d) Deblurring results with weight variables in (a) and (b) respectively. Figure (d) shows a visually more satisfactory result.

$$E_{reg}(\mathbf{K}_i) = \beta \|\mathbf{k}_i\|^2, \quad (8)$$

where $\mathbf{k}_i$ is a vector form of the uniform kernel $\mathbf{K}_i$ and the parameter β controls the influence of regularization on $\mathbf{k}_i$.

Secondly, for a non-uniform blur kernel $\mathbf{K}_i$, we also use Tikhonov regularization but in a different manner from the case of uniform blur kernel because a non-uniform kernel is estimated in a different way. To be specific, similar to [5, 7], the non-uniform kernel matrix $\mathbf{K}_i$ is made by restricting it to linear combinations of the several basis kernels as

$$\mathbf{K}_i = \sum_{m=1}^{M} \mu_{m,i} \mathbf{b}_m, \quad (9)$$

where the vector $\mathbf{b}_m$ is the m_{th} basis kernel induced by a possible camera shake and M denotes the total number of basis kernels. $\mu_{m,i}$ is the coefficient of basis kernel $\mathbf{b}_m$ which satisfies $\mu_{m,i} \geq 0$ and $\sum_{m=1}^{M} \mu_{m,i} = 1$. Since the non-uniform kernel is determined by $\mathbf{u}_i = [\mu_{1,i}, \mu_{2,i}, \cdots, \mu_{M,i}]^T$, we regularize $\mathbf{u}_i$ instead of the $\mathbf{K}_i$ itself. Then the energy function for regularization of the non-uniform kernel is formulated by

$$E_{reg}(\mathbf{K}_i) = \gamma \|\mathbf{u}_i\|^2, \quad (10)$$

where the parameter γ adjusts the scale of regularization on $\mathbf{u}_i$.

3. Optimization

The proposed dynamic scene deblurring model introduced in the previous section and the final objective function is as follows:

$$\begin{aligned}\min_{\mathbf{L},\mathbf{W},\mathbf{K}} \lambda \sum_{i=1}^{N}\sum_{\partial_*} \|\mathbf{W}_i^{\frac{1}{2}} \odot (\mathbf{K}_i\partial_*\mathbf{L} - \partial_*\mathbf{B})\|^2 + \\ |\nabla\mathbf{L}| + \sum_{i=1}^{N}\sum_{\mathbf{x}}\sum_{\mathbf{y}\in\mathcal{N}(\mathbf{x})} g(\mathbf{x},\mathbf{y})|\mathbf{W}_i(\mathbf{x}) - \mathbf{W}_i(\mathbf{y})| + \\ \beta \sum_{\substack{i=1,\\ \mathbf{K}_i:uniform}}^{N} \|\mathbf{k}_i\|^2 + \gamma \sum_{\substack{i=1,\\ \mathbf{K}_i:non-\\ uniform}}^{N} \|\mathbf{u}_i\|^2,\end{aligned} \tag{11}$$

where $\mathbf{W}_i(\mathbf{x}) \geq 0$ and $\sum_{i=1}^{N}\mathbf{W}_i(\mathbf{x}) = 1$. Though our final objective function is not a jointly convex problem, three sub-problems with respect to $\mathbf{L}$, $\mathbf{W}$ and $\mathbf{K}$ are convex. Therefore, instead of using complex optimization such as sampling based technique for global optimum, we propose an iterative optimization method similar to [2, 19, 15] for easier inference. By alternatively optimizing each subproblem in an iterative process, we can efficiently estimate the $\mathbf{L}$, $\mathbf{W}$ and $\mathbf{K}$ and can obtain successful results. Each proposed subproblem can be modeled as a convex function, and we adopt the first-order primal-dual algorithm [1] to solve each problem.

3.1. Sharp Image Restoration

Sharp image restoration methods are widely researched in both non-blind and blind deblurring methods and some fast solutions are available with the aid of FFT. The update procedure of $\mathbf{L}$ by the first-order primal-dual algorithm is

$$\begin{cases}\mathbf{q}^{n+1} = \frac{\mathbf{q}^n + \sigma_L\mathbf{S}\mathbf{L}^n}{\max(\mathbf{1}, \mathbf{q}^n + \sigma_L\mathbf{S}\mathbf{L}^n)} \\ \mathbf{L}^{n+1} = \arg\min_{\mathbf{L}} \frac{(\mathbf{L} - (\mathbf{L}^n - \tau_L\mathbf{S}^T\mathbf{q}^{n+1}))^2}{2\tau_L} + \\ \qquad\qquad \lambda\sum_{i=1}^{N}\sum_{\partial_*} |\mathbf{W}_i^{\frac{1}{2}} \odot (\mathbf{K}_i\partial_*\mathbf{L} - \partial_*\mathbf{B})\|^2,\end{cases} \tag{12}$$

where $n \geq 0$ means iteration number, $\mathbf{q}$ denotes the dual variable of $\mathbf{L}$ defined on the vector space and $\mathbf{S}$ is a continuous linear operator that calculates the difference between two pixels. The update steps σ_L and τ_L control the convergence rate as defined in [1]. Initially $\mathbf{q}^0 = \mathbf{0}$, $\mathbf{L}^0$ is equal to the blurred image $\mathbf{B}$. In particular, since the primal update for $\mathbf{L}$ in (12) is a quadratic form, we have adopted the Landweber method [3] to solve with FFT similar to [20].

3.2. Weight Estimation

The use of non-local regularization and constraints on weight variables makes it hard to infer, but with the aid of convexity, we also adopt the first-order primal-dual algorithm and the update step is given by,

$$\begin{cases}\mathbf{r}_{i,\mathbf{y}}^{n+1}(\mathbf{x}) = \min(g(\mathbf{x},\mathbf{y}), \max(-g(\mathbf{x},\mathbf{y}), \\ \qquad\qquad \mathbf{r}_{i,\mathbf{y}}^{n}(\mathbf{x}) + \sigma_{\mathbf{w}}(\mathbf{Z}_{\mathbf{y}}\mathbf{W}_i^n)(\mathbf{x})) \\ \mathbf{W}_i^{n+1} = \mathbf{W}_i^n - \tau_{\mathbf{w}}(\sum_{\mathbf{y}} \mathbf{Z}_{\mathbf{y}}^T\mathbf{r}_{i,\mathbf{y}}^{n+1} + \\ \qquad\qquad \lambda\sum_{\partial_*}(\mathbf{K}_i\partial_*\mathbf{L} - \partial_*\mathbf{B}) \odot (\mathbf{K}_i\partial_*\mathbf{L} - \partial_*\mathbf{B})) \\ \mathbf{W}^{n+1} = \Pi_{\mathbf{W}}(\mathbf{W}^{n+1}),\end{cases} \tag{13}$$

where $\mathbf{r}_{i,\mathbf{y}}$ is a dual variable defined on the vector space, and update steps $\sigma_{\mathbf{w}}$ and $\tau_{\mathbf{w}}$ control the convergence rate as defined in [1]. $\mathbf{Z}_{\mathbf{y}}$ is a continuous linear operator that calculates the difference between two neighboring pixels at $\mathbf{x}$ and $\mathbf{y}$. Since $\mathbf{W}$ has some constraints, $\mathbf{W}_i(\mathbf{x}) \geq 0$ and $\sum_{i=1}^{N}\mathbf{W}_i(\mathbf{x}) = 1$, the orthogonal projection $\Pi_{\mathbf{W}}$ projects $\mathbf{W}$ onto the unit simplex [14]. This projection converges with N iterations at most and the detail is in Algorithm 1.

Algorithm 1 The algorithm of projection onto unit simplex

1: $T = \{1, ..., N\}$
2: $\mathbf{W}_i(\mathbf{x}) \leftarrow \mathbf{W}_i(\mathbf{x}) - (\sum_i \mathbf{W}_i(\mathbf{x}) - 1)/|T|$, if $i \in T$
3: $T \leftarrow T - \{i\}$, if $\mathbf{W}_i(\mathbf{x}) < 0$
4: $\mathbf{W}_i(\mathbf{x}) \leftarrow 0$, if $i \notin T$
5: Repeat steps 2-4 until $\sum_{i=1}^{N}\mathbf{W}_i(\mathbf{x}) = 1$, for all $\mathbf{x}$.

For weight estimation, it is important to set proper initial value for $\mathbf{W}$. The initial weight map is designed so that at least one of the segments for uniform kernels can cover the moving object. However since we don't know the position and the size of it, we use many overlapping segments to cover it initially. If the moving object occupies large part or has strong edges in an initial segment, then the blur kernel of the moving object can be roughly estimated from that segment. Then, by iterations, both the accuracies of the blur kernel and segment increase. An example of using 6 uniform blur kernels and 1 non-uniform blur kernel and their corresponding initial weight maps is illustrated in Fig. 4. We observed empirically that different settings of initial weight map do not change the results significantly. Thus, we used the same initial segmentations as in Fig. 4 and set the initial values of 1/3 for each segment for all experiments.

3.3. Blur Kernel Estimation

The proposed method includes multiple blur models and a blur model could be either uniform and non-uniform kernel. Therefore, we have to estimate both uniform and

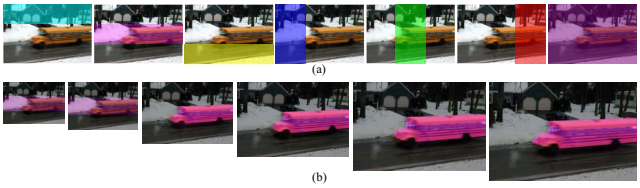


Figure 4. (a) An example of the initial set-up of weight variables. The six columns on the left illustrate the initial weight variables corresponding to six uniform models, and the right most column shows the initial weight variable corresponding to a non-uniform model. (b) Change of a weight variable from coarse to fine level. The distribution of a weight variable gradually changes and finally fits on the moving bus.

non-uniform kernels. The blur kernel estimation methods for both approaches have been widely studied in blind deblurring methods, but ours is somewhat different because the proposed model includes additional weight variables. Since proper initial value for blur kernel is also important, the method guiding the latent image using prediction step [2, 19] is widely used. So, we adopt the predicted gradient maps $\{\mathbf{p}_x, \mathbf{p}_y\}$ defined in [2], instead of using latent image itself for accurate kernel estimation.

3.3.1 Uniform Kernel Estimation

For $\mathbf{L}$ and $\mathbf{W}$ being fixed, our energy model for uniform kernel $\mathbf{K}_i$ is quadratic and the solution can be easily obtained. The quadratic objective function with some constraints on the uniform kernel is given by

$$\min_{\mathbf{k}_i} \lambda((\mathbf{P}_x\mathbf{k}_i - \partial_x\mathbf{B})^T diag(\mathbf{W}_i)(\mathbf{P}_x\mathbf{k}_i - \partial_x\mathbf{B}) + (\mathbf{P}_y\mathbf{k}_i - \partial_y\mathbf{B})^T diag(\mathbf{W}_i)(\mathbf{P}_y\mathbf{k}_i - \partial_y\mathbf{B})) + \beta\|\mathbf{k}_i\|^2. \tag{14}$$

Note that matrices $\mathbf{P}_x$ and $\mathbf{P}_y$ consist of $\mathbf{p}_x$ and $\mathbf{p}_y$, respectively and the vector form of uniform kernel $\mathbf{k}_i$ is used where elements of $\mathbf{k}_i$ are larger or equal to zero and their sum is one. To help understand this and represent it as a quadratic form we introduce $diag(\mathbf{W}_i)$ which is a diagonal matrix whose diagonal entries are the elements of $\mathbf{W}_i$. As this problem is convex, we can use any quadratic programming methods to solve it, and we have adopted the Landweber method [3] to iteratively minimize with FFT for reducing computations.

3.3.2 Non-Uniform Kernel Estimation

Since non-uniform kernel $\mathbf{K}_i$ is a weighted sum of M basis kernels and the blurry image $\mathbf{B}$ is equal to $\mathbf{K}_i\mathbf{L}$, we can derive an equation,

$$\mathbf{K}_i\mathbf{L} = \mathbf{A}\mathbf{u}_i, \tag{15}$$

where the matrix $A = [\mathbf{b}_1\mathbf{L}, \mathbf{b}_2\mathbf{L}, \cdots, \mathbf{b}_M\mathbf{L}]$. Therefore, the minimization on the coefficient vector $\mathbf{u}_i$ for non-uniform kernel is given by

$$\min_{\mathbf{u}_i} \lambda(\|\mathbf{W}_i^{\frac{1}{2}} \odot (\mathbf{A}_x\mathbf{u}_i - \partial_x\mathbf{B})\|^2 + \|\mathbf{W}_i^{\frac{1}{2}} \odot (\mathbf{A}_y\mathbf{u}_i - \partial_y\mathbf{B})\|^2) + \gamma\|\mathbf{u}_i\|^2, \tag{16}$$

where $\mathbf{A}_x$ and $\mathbf{A}_y$ are derived from $\mathbf{K}_i\mathbf{p}_x = \mathbf{A}_x\mathbf{u}_i$ and $\mathbf{K}_i\mathbf{p}_y = \mathbf{A}_y\mathbf{u}_i$, respectively. Since this energy function can also be represented as the quadratic form, we can find an optimal $\mathbf{u}_i$ by quadratic programming. To be specific, the quadratic programming is formulated as

$$\min_{\mathbf{u}_i} \frac{1}{2}\mathbf{u}_i^T\mathbf{H}\mathbf{u}_i + \mathbf{f}^T\mathbf{u}_i, \tag{17}$$

where

$$\begin{cases} \mathbf{H} = \mathbf{A}_x^T diag(\mathbf{W}_i)\mathbf{A}_x + \mathbf{A}_y^T diag(\mathbf{W}_i)\mathbf{A}_y + \frac{\gamma}{\lambda}\mathbf{I}, \\ \mathbf{f} = \mathbf{A}_x^T diag(\mathbf{W}_i)\mathbf{B}_x + \mathbf{A}_y^T diag(\mathbf{W}_i)\mathbf{B}_y. \end{cases} \tag{18}$$

The minimization is performed by the interior point method and we can obtain the non-uniform blur kernel matrix as $\mathbf{K}_i = \sum_{m=1}^{M} \mathbf{u}_i(m)\mathbf{b}_m$ from (9).

3.4. Overall Procedure

In the previous sections, we introduce the efficient minimization methods for $\mathbf{L}$, $\mathbf{W}$ and $\mathbf{K}$, respectively. However, there exist many unknown variables in our model and the traditional iterative optimization is prone to be stuck in local minimum. To alleviate this problem, we adopt coarse to fine approach like most recent blind deconvolution algorithms [13, 19, 2], and the overall procedure of our dynamic scene deblurring is in Algorithm 2.

Algorithm 2 The overall procedure of the proposed dynamic scene deblurring algorithm

Input: A blurry image $\mathbf{B}$
Output: $\mathbf{L}$, $\mathbf{W}$ and $\mathbf{K}$

```
Build an image pyramid, which has 5 levels, with a scale factor of 0.5
for t = 1 to 3 do
    Update K with the predicted gradient maps {p_x, p_y}. (Sec. 3.3)
    for n = 1 to 30 do
        Continuous optimization of L. (Sec. 3.1)
        Continuous optimization of W. (Sec. 3.2)
    end for
end for
Propagate variables to the next pyramid level if exists.
Repeat steps 2-9 from coarse to fine pyramid level.
```

4. Experimental Results

Although many parameters are used in our experiments, fortunately most of them are reliable and less sensitive to various blurry images except the parameter λ. Since we do

not estimate the noise level and the blur strength of the input image, λ that adjusts the influence of data term should be tuned from the statistics of the input image. It ranges from 50 to 500 and it has a low value when the noise level is high or the blur is severe. The other parameters are fixed and we use six uniform kernel models and one non-uniform kernel model, so $N = 7$ in all experiments. By setting N as large as possible, we can handle various kinds of blur motions but it raises costs and thus we determined the number of models empirically and fixed it. Note, however, that the numbers of segmented regions in the final results in Fig. 5 are less than 7 and adaptive to each image. This is due to our sparsity priors to the weight variables. The initial values of the seven weight variables are set as illustrated in Fig. 4(a) and the value of each area is set $\frac{1}{3}$. We use 80 neighbors of a pixel in a 9×9 patch for non-local regularization and the parameters are $\sigma_D = 40$, $\sigma_I = \frac{25}{255}$, $\beta = 10\lambda$, $\gamma = 1000\lambda$.

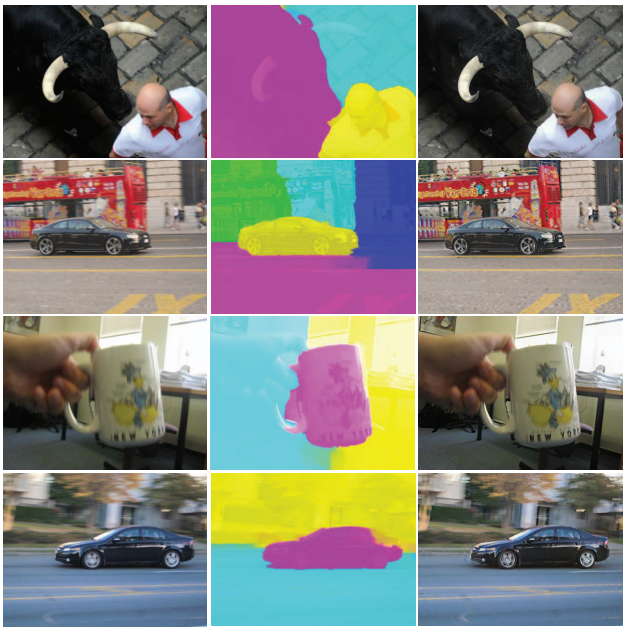
The framework of our method in Algorithm 2 is designed for gray image restoration. However, the estimated set of blur kernels $\mathbf{K}$ from a gray image is also used for deblurring the corresponding color image by applying the sharp image restoration step introduced in Section 3.1 for each color channel.

In Fig. 5, the motion blur segmentation and deblurring results of real dynamic scenes are shown. We observe that substantial improvements are achieved in the hair of the running bull and the letter on the bus. Also we compared our results with dynamic scenes to conventional methods. As shown in Fig. 6, there are serious artifacts near the boundaries of moving objects in the results of other methods, while our method gives relatively clean results with the aid of various blur models and motion blur segmentations.

We also compared our deblurring results on static scenes with conventional methods in Fig. 7 and Fig. 8. The blurry input Picasso image in Fig. 7 is degraded by an uniform kernel. Although it is possible to obtain a sharp image with methods based on uniform kernel, our model shows much better result in reducing ringing artifacts and restoring details well. In addition, the blurry input used in Fig. 8 has non-uniform blur motion which is generated by rotational camera shake. In this case, one of seven blur models which corresponds to the non-uniform blur kernel gained almost all weights and still showed competitive result compared to the state-of-the-art non-uniform kernel based methods.

5. Discussion and Conclusion

We proposed a novel single image deblurring framework that can handle multiple moving objects in the scene as well as camera shake. By introducing multiple blur models and their locally varying weight variables which favor the blur models giving better data fidelity, we could also obtain the segmented blur region as well as restored image. We demonstrated the superiority of our method over conventional methods in dynamic scene deblurring as well as in static scene cases. The future challenges and remaining problems are determining the number of moving objects via non-parametric methods. Since the number of moving objects is unknown in a real image, we should set a large number of blur models. Another problem is the run-time of our method. Due to multiple blur models and additional weight variables, computational costs increase. Thus, our future works will include developing an efficient optimization method and parallel implementation using GPGPU.



(a) (b) (c)

Figure 5. Deblurring results of real dynamic scenes. (a) Blurry images. (b) Our motion blur segmentation results. (c) Our deblurring results.

Acknowledgments

This research was supported in part by the National Research Foundation of Korea (NRF) grant funded by the Ministry of Science, ICT & Future Planning (MSIP) (No. 2009-0083495).

References

[1] A. Chambolle and T. Pock. A first-order primal-dual algorithm for convex problems with applications to imaging. *Journal of Mathematical Imaging and Vision*, 40(1):120–145, May 2011. 5

[2] S. Cho and S. Lee. Fast motion deblurring. In *SIGGRAPH*, 2009. 2, 3, 4, 5, 6

[3] H. Engl, M. Hanke, and A. Neubauer. *Regularization of Inverse Problems*. Mathematics and Its Applications. Springer, 1996. 5, 6

[4] R. Fergus, B. Singh, A. Hertzmann, S. T. Roweis, and W. Freeman. Removing camera shake from a single photograph. In *SIGGRAPH*, 2006. 2

[5] A. Gupta, N. Joshi, L. Zitnick, M. Cohen, and B. Curless. Single image deblurring using motion density functions. In *ECCV*, 2010. 2, 3, 4, 8

[6] S. Harmeling, H. Michael, and B. Schoelkopf. Space-variant single-image blind deconvolution for removing camera shake. In *NIPS*, 2010. 2

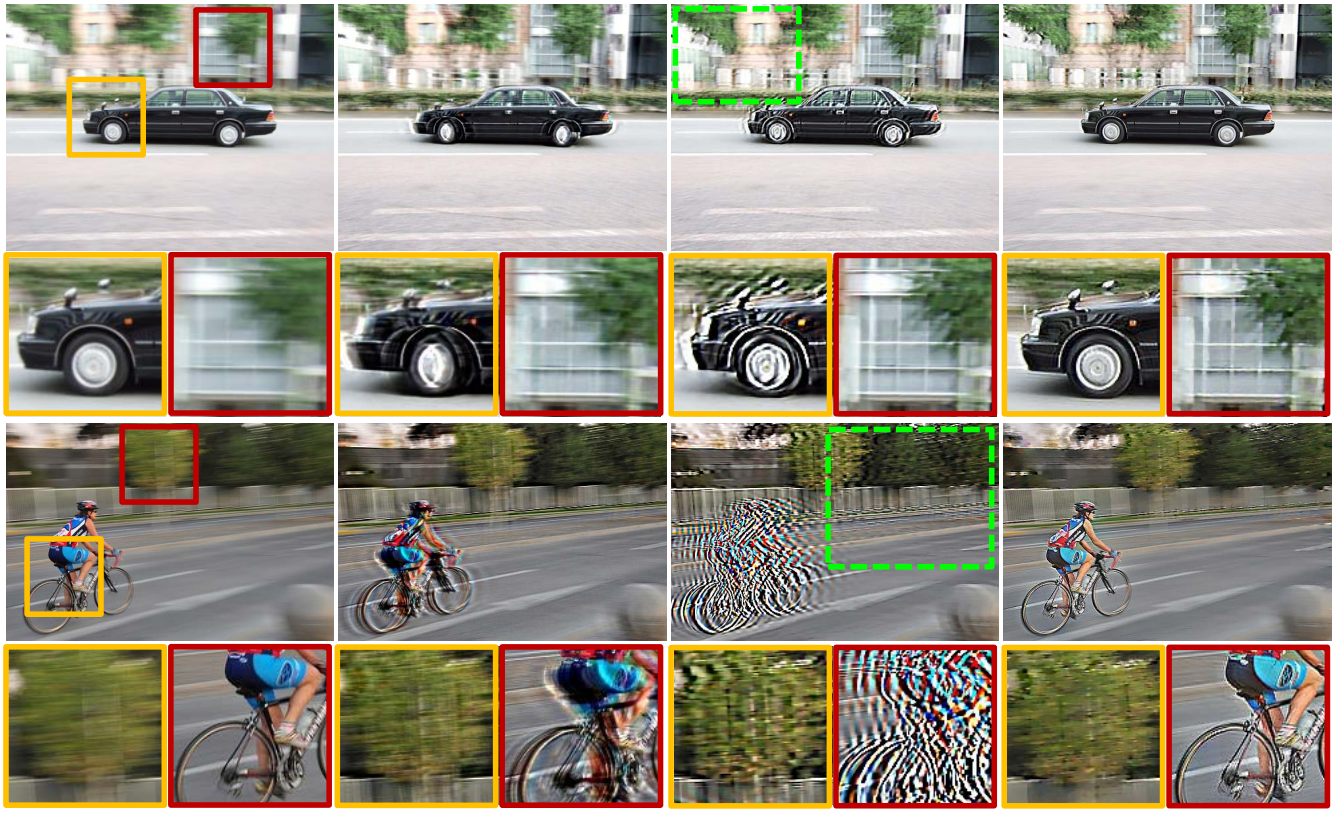


Figure 6. Comparison of dynamic scene deblurring results. (a) Blurry images of real dynamic scenes. (b) Deblurring results of Whyte et al. [18]. (c) Deblurring results of Xu et al. [19]. Dashed green boxes in the figures denote the regions used for estimating uniform blur kernels and used for restoring the background regions. (d) Our results.

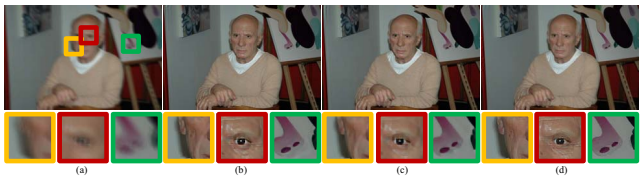


Figure 7. Comparison of static scene deblurring. (a) Blurry Picasso image. Synthetic uniform kernel is used to blur the Picasso image. (b) Result of Shan et al. [15]. (c) Result of Xu et al. [19]. (d) Our result.

Figure 8. Comparison of static scene deblurring. Magazine image is blurred by rotational camera shake and requires non-uniform blur kernel to be restored. (a) Blurry Magazine image. (b) Result of Hirsch et al. [7]. (c) Result of Gupta et al. [5]. (d) Our result.

[7] M. Hirsch, C. J. Schuler, S. Harmeling, and B. Scholkopf. Fast removal of non-uniform camera shake. In *ICCV*, 2011. 2, 4, 8

[8] H. Ji and K. Wang. A two-stage approach to blind spatially-varying motion deblurring. In *CVPR*, 2012. 2

[9] N. Joshi, S. B. Kang, C. L. Zitnick, and R. Szeliski. Image deblurring using inertial measurement sensors. In *SIGGRAPH*, 2010. 2

[10] D. Krishnan and R. Fergus. Fast image deconvolution using hyper-laplacian priors. In *NIPS*, 2009. 3

[11] A. Levin. Blind motion deblurring using image statistics. In *NIPS*, 2006. 2

[12] A. Levin and Y. Weiss. User assisted separation of reflections from a single image using a sparsity prior. *IEEE Trans. Pattern Analysis Machine Intelligence*, 29(9):1647–1654, 2007. 3

[13] A. Levin, Y. Weiss, F. Durand, and W. T. Freeman. Understanding and evaluating blind deconvolution algorithms. In *CVPR*, 2009. 6

[14] C. Michelot. A finite algorithm for finding the projection of a point onto the canonical simplex of rn. *J. Optim. Theory Appl.*, 50(1):195–200, July 1986. 5

[15] Q. Shan, J. Jia, and A. Agarwala. High-quality motion deblurring from a single image. In *SIGGRAPH*, 2008. 2, 3, 5, 8

[16] Y.-W. Tai and S. Lin. Motion-aware noise filtering for deblurring of noisy and blurry images. In *CVPR*, 2012. 4

[17] M. Werlberger, T. Pock, and H. Bischof. Motion estimation with non-local total variation regularization. In *CVPR*, 2010. 4

[18] O. Whyte, J. Sivic, A. Zisserman, and J. Ponce. Non-uniform deblurring for shaken images. *International Journal of Computer Vision*, 98(2):168–186, 2012. 1, 2, 8

[19] L. Xu and J. Jia. Two-phase kernel estimation for robust motion deblurring. In *ECCV*, 2010. 1, 2, 3, 4, 5, 6, 8

[20] L. Yuan, J. Sun, L. Quan, and H.-Y. Shum. Image deblurring with blurred/noisy image pairs. In *SIGGRAPH*, 2007. 2, 4, 5