Changsoo Je, Hyeon Sang Jeon, Chang-Hwan Son, and Hyung-Min Park.
Disparity-Based Space-Variant Image Deblurring.
Signal Processing: Image Communication, Volume 28, Issue 7, pp. 792-808, August, 2013.
http://dx.doi.org/10.1016/j.image.2013.04.005

Disparity-Based Space-Variant Image Deblurring

Changsoo Je[a], Hyeon Sang Jeon[a,b], Chang-Hwan Son[a], Hyung-Min Park[a,*]

[a] *Department of Electronic Engineering, Sogang University, 35 Baekbeom-ro, Mapo-gu, Seoul 121-742, Republic of Korea*
[b] *Telecom Service Development Team2, SK C&C, SK u-Tower, 25-1 Jeongja-dong, Bundang-gu, Sungnam-si, Gyeonggi-do 463-844, Republic of Korea*

Abstract

Obtaining a good-quality image requires exposure to light for an appropriate amount of time. If there is camera or object motion during the exposure time, the image is blurred. To remove the blur, some recent image deblurring methods effectively estimate a point spread function (PSF) by acquiring a noisy image additionally, and restore a clear latent image with the PSF. Since the groundtruth PSF varies with the location, a blockwise approach for PSF estimation has been proposed. However, the block to estimate a PSF is a straightly demarcated rectangle which is generally different from the shape of an actual region where the PSF can be properly assumed constant. We utilize the fact that a PSF is substantially related to the local disparity between two views. This paper presents a disparity-based method of space-variant image deblurring which employs disparity information in image segmentation, and estimates a PSF, and restores a latent image for each region. The segmentation method firstly over-segments a blurred image into sufficiently many regions based on color, and then merges adjacent regions with similar disparities. Experimental results show the effectiveness of the proposed method.

Keywords: Image deblurring, space-variant deblurring, disparity, segmentation, point spread function, deconvolution

*Corresponding author
Email addresses: vision@sogang.ac.kr (Changsoo Je), hsjeon@sk.com (Hyeon Sang Jeon), chson@sogang.ac.kr (Chang-Hwan Son), hpark@sogang.ac.kr (Hyung-Min Park)

1. Introduction

Imaging a scene from a fixed camera viewpoint is generally required to capture a clear image. However, a camera often moves with respect to the target scene during a single exposure time. This is mostly the way: without using a tripod, taking a picture is frequently affected by a hand tremor. A camera image is obtained as the integration of the input intensity of electromagnetic radiation for an exposure time. The motion of the camera during the exposure time makes *motion blur* in the image. Under a poor lighting environment, a camera image is typically captured by a long exposure time, and a hand tremor is likely to cause a serious sequential rigid motion of the sensor array, resulting in a highly motion-blurred image. In many applications, motion blur is undesirable, and researchers have investigated methods of *deblurring* which restore a clear *latent image* from a *blurred image*. Deblurring algorithms mostly consists of two main steps, (a) estimation of a blur kernel, namely, *point spread function* (PSF) corresponding to the camera or object motion, and (b) latent image reconstruction based on the estimated PSF.

Many deblurring algorithms [1, 2] take a single blurred image as the input, and they generally has two major disadvantages, longer computational times and lower accuracies compared to multi-image-based algorithms, due to the deficiency of information. Recently the advances in CMOS sensors have made possible rapid capture of images, and multi-image-based deblurring [3, 4, 5, 6, 7, 8, 9] has become more feasible.

The hybrid imaging method [8] measures the camera motion from a sequence of images, and uses the motion to compute a PSF. However, using special hardware is the major disadvantage of the method. Double-image-based approaches have been proposed, and many of them employ a short-exposure (dark or noisy) image in addition to a blurred image [3, 4, 5, 6, 7]. Jia et al. used an under-exposed (dark) image and a blurred image to produce an optimal color mapping function [6]. Without PSF estimation, a high-quality image is obtained by the color mapping function. However, the dark images they used were not noisy. Other important methods using a short-exposure image estimate a PSF mostly using a *noisy image*, and restore a latent image with the PSF [3, 4, 5, 7]. Actually, their work can be considered as either deblurring of a blurred image using a noisy image or denoising of the noisy image using the blurred image. The (blurred or noisy) images of different exposure times and possibly different ISOs can be subsequently captured in

the exposure bracketing mode of a camera[1], prohibiting serious geometric transformation between any pair of images [3].

Most of existing methods assume two kinds of blur models. One is the space-invariant model in which the PSF is identical for a whole image, and the other is the space-variant model in which the PSF can vary with the location. Since an object scene generally has its spatial variation of the distance from the sensor array, the motion of features in a camera image varies with the spatial location, and space-invariant models cannot handle the spatial differences of the blur.[2] Šorel and Flusser [9] considered blur in the pinhole camera model depending on the depth of scene and the camera motion, and estimated a depth-dependent (space-variant) PSF for a piecewise planar scene from multiple blurred images. However, the method requires specification of a region of approximately constant depth as a user input, and the camera motion is assumed to be an arbitrary curve parallel to the image plane without any rotations. To deal with general camera motion, a blockwise approach for PSF estimation has been proposed [4]. However, the block to estimate PSF is a straightly demarcated rectangle which is different from the shape of an actual region where the PSF can be properly assumed constant. In case there exists significant spatial variation of the groundtruth PSF in a block, the blockwise PSF still cannot reliably handle the spatial variation. If a sufficiently small block size is chosen to properly resolve the spatial variation, features such as edges and corners in each block will be reduced, and PSFs will not be reliably estimated for many blocks.

A latent image is estimated by deconvolution of a blurred image with a PSF. For the image restoration, methods based on total variation (TV) have been often used [11, 4]. The TV-based methods have some advantages in that they finds a better optimized solution compared to the Richardson-Lucy (RL) algorithm [12] (which is employed in [3]), and they generally suppress the noise and sharpen edges. Unfortunately, the objective function

[1]Some non-SLR compact digital cameras as well as many DSLR cameras have an exposure bracketing function.

[2]Actually, an important portion of the PSF is produced by many factors that are independent of the relative motion between the scene and sensor [10], such as defocus, diffraction, spherical and chromatic aberrations, and wave aberrations. Those factors are mainly characterized by the geometry of the aperture, lens, and mirror in the imaging system. The total PSF caused by motion as well as those factors, is spatially variant in general. This paper discusses on the space-variant total PSF.

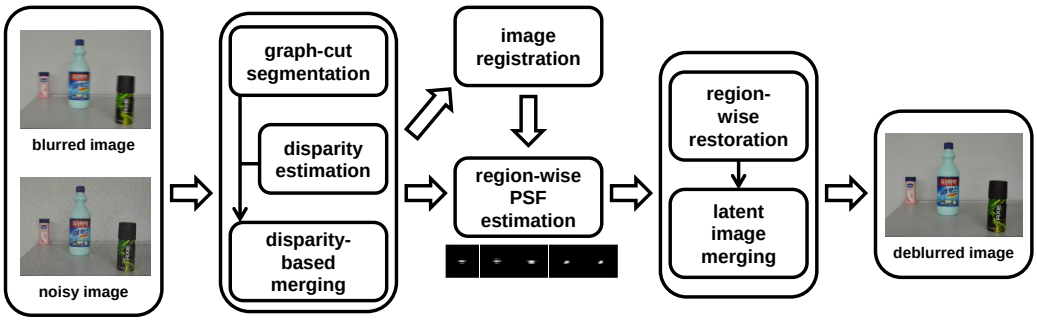


Figure 1: **Overall procedure of the disparity-based deblurring method.**

of the TV-based methods has an L_1 regularization term, and the solution cannot be generally expressed in a closed form. While the space-domain iterative method based on conjugate gradient (CG) [13] does not usually converge rapidly, the half-quadratic iterative approach based on hyper-Laplacian distribution priors [14] is very efficient by using the fast Fourier transform (FFT).

Disparity[3] (or binocular disparity) is the locational difference of a scene point in the images from two distinct viewpoints, and is one of the most important depth cues in human and machine vision. *Motion parallax* is the parallax achieved by motion of a viewpoint, and can be equivalently interpreted as disparity. In general, the magnitude of motion parallax is proportional to the amount of motion of the viewpoint and the inverse of the depth (the distance between the viewpoint and object). Similarly, the magnitude of disparity is proportional to the distance between the two viewpoints and the inverse of the depth (the distance between the viewpoints and object). The motion-blurred image can be represented as the temporal integral of a latent image with time-varying motion parallax. That is, the blur kernel corresponds to the temporal integral of a neutral scene point with the time-varying motion parallax. In case of deblurring using at least two images, we can estimate the disparity between any pair of images, and the disparity can be considered as a discrete version of the time-varying motion parallax, and therefore we assume regions of similar disparities are likely to have similar blur kernels.

In this paper, we propose a disparity-based regional PSF estimation and

[3]In this paper, we consider disparity basically as two-dimensional vector(s).

restoration algorithm (see Fig. 1). It uses a pair of images from an object scene: a blurred image captured with long exposure and possibly low ISO, and a noisy image with short exposure and possibly high ISO. We assume there is disparity between the two images caused by camera motion which also causes the blur of the blurred image. Therefore, we utilize the fact that a PSF is substantially related to the local disparity between two views: a pair of blurred and noisy images. We first over-segment the blurred image into regions based on color (RGB intensities) using the graph-cut method [15, 16, 17, 18]. Then, we detect corners of the blurred image with the Harris detector [19], and estimate the disparities of the corners between the two images by the subpixel hierarchical block matching [20, 21, 22, 23, 24]. The disparity of each segmented region is determined as the vector median[4] of the disparities of the corners in the segmented region. Finally, if adjacent regions have similar disparities, they are merged into one region. Meanwhile the homography between the two images is estimated from the block matching result, and the noisy image is registered to the blurred image according to the homography.[5] Now a PSF can be estimated for each region using the blurred image and the registered version of the noisy image. Then, we run region-wise image restoration using the half-quadratic iterative method [14] with the hyper-Laplacian distribution prior to model the gradients in natural-scene images. Finally, the region-wise restoration results are combined to form a whole latent image.

The rest of this paper is organized as follows. Section 2 describes multi-image-based image deblurring, and Section 3 presents disparity-based image segmentation. Regional PSF estimation is presented in Section 4, and image reconstruction is described in Section 5. Experimental results are shown in Section 6, and Section 7 concludes this paper.

[4]The vector median is defined in Section 3.

[5]Homography is suitable not for a general 3D scene, but for a planar scene. However, as demonstrated earlier, a pair of blurred and noisy images we use for deblurring, are captured intentionally without serious geometric transformation between the pair. Therefore, a singe homography matrix is a practically good model to represent the geometric transformation between the two images. Moreover, we assume homography is appropriate for typical deblurring using multiple images captured from very similar viewpoints.

Figure 2: **Image blur model.**

2. Multi-Image-Based Image Deblurring

In this section, we demonstrate a model of motion blur, and describe existing methods of multi-image-based space-invariant and space-variant image deblurring. Moreover, we discuss related problems of the existing methods which motivated our work in this paper.

Under poor lighting, an image sensor should be exposed for a sufficiently long time to take sufficient amount of photons. During the long time the image sensor is likely to move by a hand tremor, which makes the image blurred. Generally, the blurred image is assumed to be the convolution of a latent sharp image and a PSF which is the spatial distribution of the temporal density of the displacements from a location caused by the motion during the exposure time. Figure 2 illustrates an image blur model which is expressed as a relation between a blurred image, latent image, PSF, and noise. The relation can be given as [1]

$$B = X \otimes K + e, \tag{1}$$

where $\otimes$ denotes the convolution, and B, X, K, and e are the blurred image, latent image, PSF, and sensing noise, respectively.

Image deblurring is basically a blind deconvolution which is an ill-posed problem. To solve it, the MAP (maximum a posteriori) method has been frequently used. Bayesian rule gives [1]

$$p(X, K|B) \propto p(B|X, K)p(X)p(K), \tag{2}$$

where $p(B|X, K)$, $p(X)$, and $p(K)$ are the likelihood and priors of X and K, respectively. In the blind deconvolution using a single image, many iterations of PSF estimation and image reconstruction are required, and it just achieves relatively low accuracy.

Yuan et al. have proposed an image deblurring method using a noisy image and a blurred image whose exposure times and ISOs are different from

each other [3]. It estimates a more accurate PSF, and thus more stably reconstructs a latent image compared to methods using only a single blurred image. In their method, a single PSF is estimated, and the image reconstruction is globally performed by the PSF. The estimation of a PSF K can be given by the following equation,

$$\hat{K} = \arg\min_{K} \left\{ ||B - K \otimes N||^2 + R_K(K) \right\}, \tag{3}$$

where the first term on the right side of the above equation is a fidelity term, and N is the noisy image registered to B. R_K is a regularization term which minimizes the estimation error caused by over-fitting due to the fidelity term. After estimation of a PSF, a latent image is reconstructed by RL method [12]. The PSF is assumed spatially invariant. However, a PSF is spatially variant in general.

Šorel and Šroubek proposed a blockwise space-variant deblurring algorithm using noisy and blurred images [4]. The whole image is segmented into blocks, and PSF estimation is performed independently for each block as

$$\hat{K}_{ij} = \arg\min_{K} \left\{ ||B_{ij} - K \otimes N_{ij}||^2 + R_K(K) \right\}, \tag{4}$$

where i and j represent the 2D indices of a block. Each block image is reconstructed by a TV method including regularization as L_1 norm. It results in stronger edges compared to RL methods. Finally, all the reconstructed block images are concatenated into one image.

If there exists significant spatial variation of the groundtruth PSF in a block corresponding to a straightly demarcated rectangle, the PSF cannot be reliably estimated. Therefore, it is desirable to estimate a blur kernel for each region in which the blur kernel can be properly assumed invariant. To segment the blurred image into such space-invariant regions, we use disparity which is closely related to the blur kernel as discussed in Section 1.

3. Disparity-Based Image Segmentation

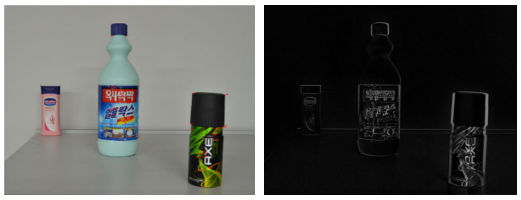
This section presents a method of image segmentation based on disparity between two input images. The method divides the blurred image into regions so that the PSF in each region can be properly assumed space-invariant. We use the blurred image in segmentation since it has less noise than the noisy image.

Roughly, the disparity between the two images varies with the depth of the scene point from the camera. For translational motion, the nearer the scene point is from the camera, the bigger the disparity is, and vice versa. Considering arbitrary motion, however the disparity variation cannot be properly assumed only coherent with the depth. Instead, it is also significantly affected by rotation, and thus somehow more complicated. Actually, the disparity rather corresponds to the motion parallax, the relative motion of the scene point in the camera image. In addition, a PSF corresponds to the temporal integral of a neutral scene point with the time-varying motion parallax. Therefore, we consider disparity-based segmentation a very effective strategy to handle space-variant PSFs. However, since disparity bears a discrete time-segment version of the motion corresponding to a PSF, even two regions of the same disparity can have substantially different PSFs. Hence, we use color as well as disparity for proper segmentation.

Our disparity-based segmentation consists of color-based over-segmentation using the graph-cut method, disparity estimation, and disparity-based merging (see Fig. 1). To estimate disparities, we first detect corner features in the blurred image by the Harris detector [19]. We use corners as features to avoid aperture problem in using edges as features. We set 5×5 pixels for the window size of the detector. Based on the detected corners the disparity is estimated by the subpixel hierarchical block matching [20, 21, 22, 23, 24] in which the subpixel localization is performed by parabolic curve fitting [25]. We use the 2D vector form of disparity rather than its magnitude, to utilize its full dimensional information. Figure 3 shows a disparity estimation result. The more distant from the camera the scene region, the smaller the disparity, and vice versa (although this is not so rigorous as previously discussed). The image is initially over-segmented by the graph-cut method [15, 16, 17, 18] based on color. Then, we selectively merge some adjacent regions based on their disparities. Figure 4 shows the process of image segmentation. We compute the vector median of the disparities in each segmented region. For a set of vectors, we define the vector median as the vector whose each component is given as the median of the corresponding components. For example, the vector median in 2D, as in our application, is given as

$$\text{median}\{\mathbf{v}_i\} = (\text{median}\{v_{1,i}\}, \text{median}\{v_{2,i}\})\,, \tag{5}$$

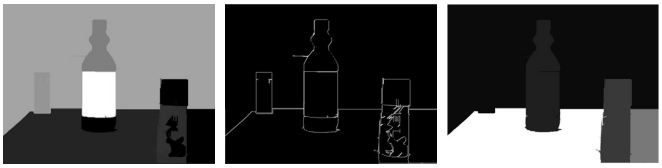
where i denotes the vector index, and $v_{1,i}$ and $v_{2,i}$ represent respectively the first and second components of the i-th vector, $\mathbf{v}_i$ among a set of vectors $\{\mathbf{v}_i\}$. Finally, we merge adjacent regions with the similar median disparities into



(a) (b)

(c) (d) (e)

Figure 3: **Disparity estimation result.** (a) Vector disparities at corners, (b) gray levels representing pixelwise disparity magnitude, and (c), (d), and (e) vector disparities of blocks.



(a) (b) (c)

Figure 4: **Disparity-based segmentation.** (a) Over-segmentation by graph-cut, (b) boundary pixels, and (c) merging regions with similar disparities.

one in which the similarity is measured as the Euclidean distance between their vector-median disparities, and then is moderately thresholded.

4. Regional PSF Estimation

After segmentation, we execute image registration, and then perform PSF estimation for each region (see Fig. 1). For accurate estimation of a regional PSF, we use the blurred image and the registered version of the noisy image. In Section 3 we carried out the subpixel hierarchical block matching, and the matching result is used to estimate the homography between the pair of input images by RANSAC (RANdom SAmple Consensus) [26]. According to the estimated homography, we generate the registered version of the noisy image. For PSF estimation we use the Tikhonov regularization method [27] which is also employed in [3]. In PSF estimation of each region, we use the gradient of the corresponding gray image for the fidelity term since image blurring (or deblurring) can be properly considered as the change of image gradients. The estimation of a regional PSF is given as

$$\hat{K}_s = \arg\min_{K} \left\{ \sum_{\xi \in \{x,y\}} \left(\frac{\lambda}{2} ||\partial_\xi B_s - K \otimes \partial_\xi N_s||^2 + ||\partial_\xi K||^2 \right) \right\} \tag{6}$$

where s is the index of a region, and ∂_x and ∂_y represent the partial differential operators with respect to horizontal and vertical directions, respectively. Since we use L_2 norm in PSF estimation, we can get the closed-form solution, and FFT can be used for fast computation.

An estimated PSF may be unreliable if the edge information in the corresponding segmented region is not sufficient for reliable estimation. Therefore, reliability of an estimated PSF should be checked. As the reliability measure, we use kurtosis [28] which indicates sparseness of the PSF. The normalized kurtosis for zero-mean data η, is given by

$$\tilde{\kappa}(\eta) = \frac{E(\eta^4)}{\{E(\eta^2)\}^2} - 3, \tag{7}$$

where $E(\cdot)$ denotes the expectation or mean. Since we usually use a sufficiently broad window for a PSF compared to the typical spread range of a PSF, a PSF tends to be relatively sparse, and hence corresponding to relatively high kurtosis [29, 14, 30]. In case of unreliable PSF estimation, the

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

Figure 5: **Estimated regional PSFs.**

PSF is likely not to be sparse. Due to the lack of sufficient features containing blur information, some unreliable PSFs may accidentally have extremely high kurtosis since such a PSF is similar to the Dirac delta function, and thus achieves only a tiny effect of deblurring. Therefore, we remove a PSF with the kurtosis outside of a predetermined interval considered as the kurtosis range of reliable PSFs, and replace it with one of another region with a similar disparity. In our experiments, a reliable PSF of 33×33 pixels usually has kurtosis from 60 to 150. Figure 5 shows the result of region-wise PSF estimation for a pair of images of 1365×1024 pixels. The PSFs are reliably estimated, and they are usually distributed densely around the center. Their kurtoses are 80, 79, 60, 96, and 106, respectively. In our implementation, we set the kurtosis range of reliable PSFs of 33×33 pixels to that from 30 to 180. The kurtoses for the PSFs shown in Fig. 5 are all in-between this range.

5. Image Reconstruction

In this section, we describe image reconstruction using the blurred image and the estimated regional PSFs. We first restore a latent image of each segmented region by performing deconvolution of each region of the blurred image with the corresponding PSF, and then merge all the reconstructed regions into one (see Fig. 1). For each regional image reconstruction, we use an FFT-based method of hyper-Laplacian regularization ($p(x) \propto \exp(-k|x|^{\alpha})$ where $0 < \alpha \leq 1$) [14], and we choose L_1 regularization ($\alpha = 1$). Hence the problem is given as

$$\min_{X,w} \left\{ \frac{\lambda}{2} ||B_s - K_s \otimes X||^2 + \sum_q \left(\frac{\beta}{2} ||\nabla_q X - w_q||^2 + ||w_q|| \right) \right\} \qquad (8)$$

where q is the pixel index, $\nabla_q X \equiv (\partial_x X(x_q, y_q), \partial_y X(x_q, y_q))^{\mathrm{T}}$ is the gradient of X at pixel q, w_q is the vector of two auxiliary variables at pixel q, and β

is the scaling for TV regularization. The above equation is obtained by substituting $\alpha = 1$ (Laplacian) to the corresponding equation in [14]. According to the deconvolution results in [14], $\alpha = 2/3$ (a strict hyper-Laplacian case) usually attains the highest SNR. Nevertheless, $\alpha = 1$ is still a good choice in that it results in relatively high SNR and performs faster. The objective function is minimized by iterative estimation of X and w. From [31], subproblems of w is given as

$$\hat{w}_{s,q} = \max\left\{||\nabla_q X_s|| - \frac{1}{\beta}, 0\right\} \frac{\nabla_q X_s}{||\nabla_q X_s||}. \tag{9}$$

The above FFT-based deblurring is fast, and the use of L_1 norm for regularization preserves strong edges.

Performing deconvolution of an extremely smooth region with the estimated PSF may cause undesirable artifact without apparent enhancement. Protecting such smooth regions have been investigated in [32, 33] for image restoration and sharpening, respectively.

6. Experimental Results

We have tested our algorithm for captured real images as well as artificially generated images, and compared the results of our method with those of the space-invariant method of Yuan et al. [3] and of the block-based method of Šorel and Šroubek [4]. We have used a DSLR camera (Nikon D7000) and two compact digital cameras (Samsung VLUU ST5000 and Canon Digital IXUS 110 IS) to capture real images.

6.1. Results for artificially generated images

We made an artificial blurred image from a fairly-captured input image of 1365×1024 resolution (with exposure time of 1/50 second, relative aperture of f/9, and ISO 1600 from Nikon D7000) and spatially-different artificial blur-kernels. Figure 6(a) shows the blurred image. We first manually segmented the captured image into six regions by considering the regional depths of the scene. Figure 6(b) shows the six regions and the corresponding blur-kernels. For region-wise blurring, we employed `conv2` in MATLAB which performs 2D convolution, then merged all the blurred regions into one, and filtered it with a small Gaussian kernel to decrease the artifact on the boundary of regions. To make a noisy image, we added 5% of Gaussian noise to the original input image. From the pair of blurred and noisy images, PSFs and a latent image

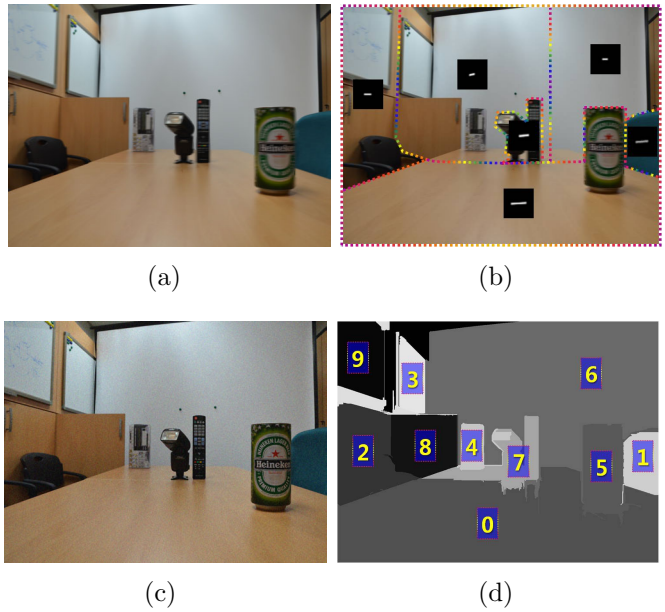


Figure 6: **Artificially generated blurred and noisy images.** (a) Blurred image, (b) regional PSFs, (c) noisy image, and (d) the segmentation result by the proposed method.

are estimated by each method. Figure 6(d) shows the segmentation result by the proposed method, and each region is labeled with a number.

Figure 7 shows the PSFs estimated from the artificially generated blurred and noisy images by the three methods: (a) Yuan et al. [3], (b and c) Šorel and Šroubek [4], (d) our method, and (e) the groundtruth, where (c), (d), and (e) are the PSFs corresponding to the ten labeled regions ($0 \sim 9$) in order. As shown in Fig. 7, the PSFs estimated by our algorithm are more similar to the groundtruth PSFs compared to those by the other two methods.

Figures 8 and 9 compare the latent images estimated by the three methods from the artificially generated blurred and noisy images. In Fig. 9, the objects of upper images is relatively close to the camera, those of middle-row images is more distant from the camera, and those of lower images is more distant than that of middle-row images, as shown in Fig. 6(a). The method of Yuan et al. uses a single global PSF for the whole image without consideration of the spatial difference, so results in much artifact. The method of Šorel and Šroubek attains better results than that of Yuan et al. However, for some blocks including much different disparities simultaneously, PSFs cannot be

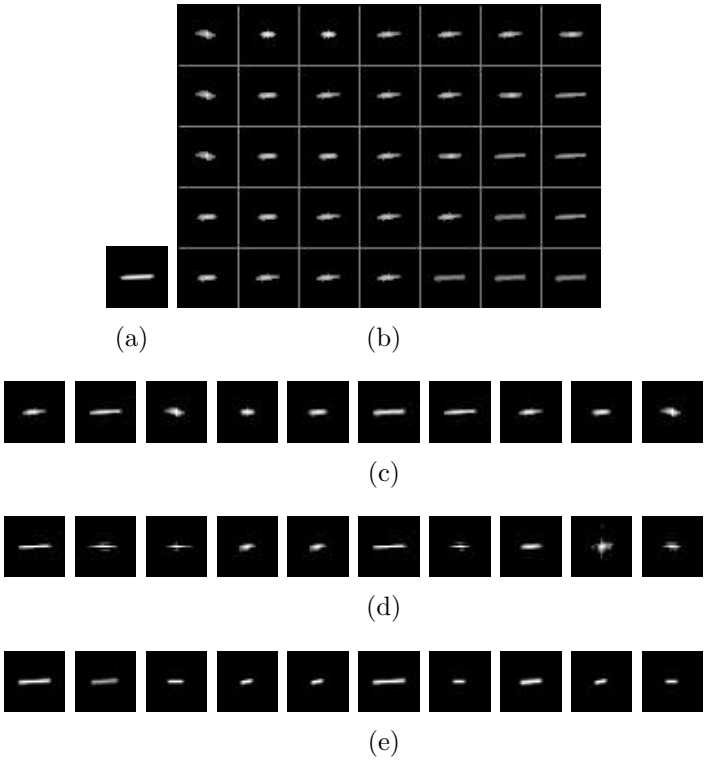
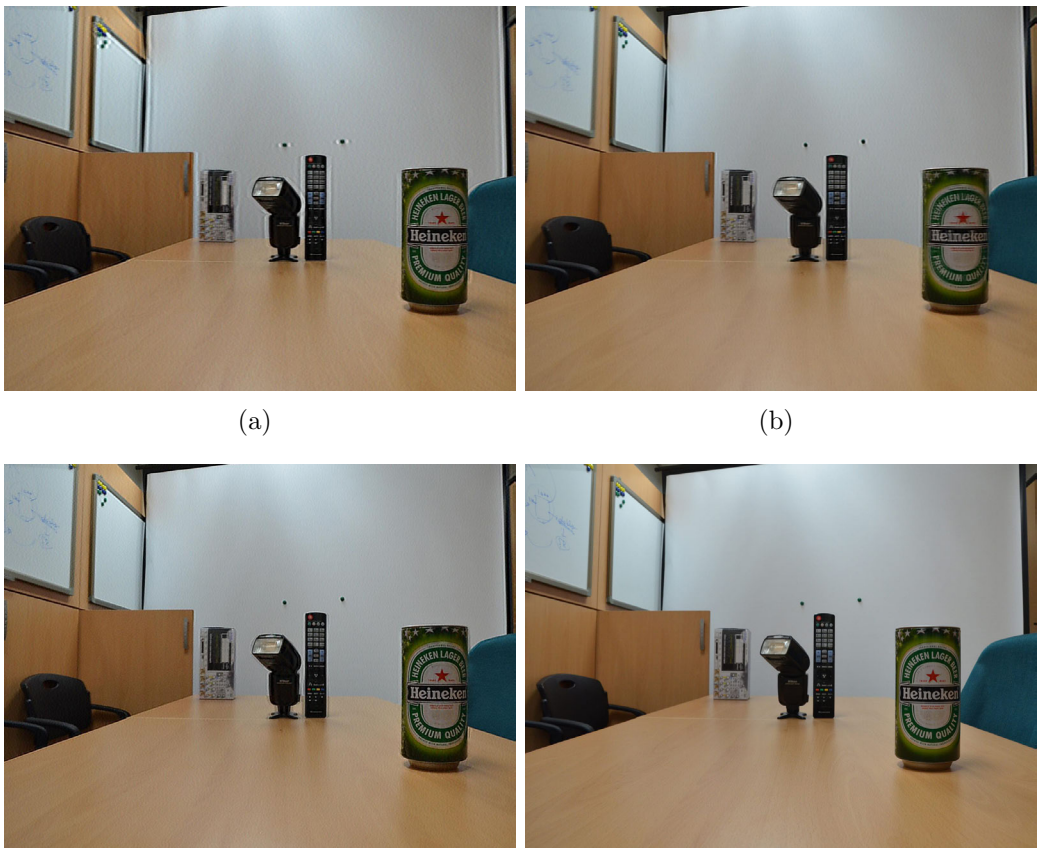


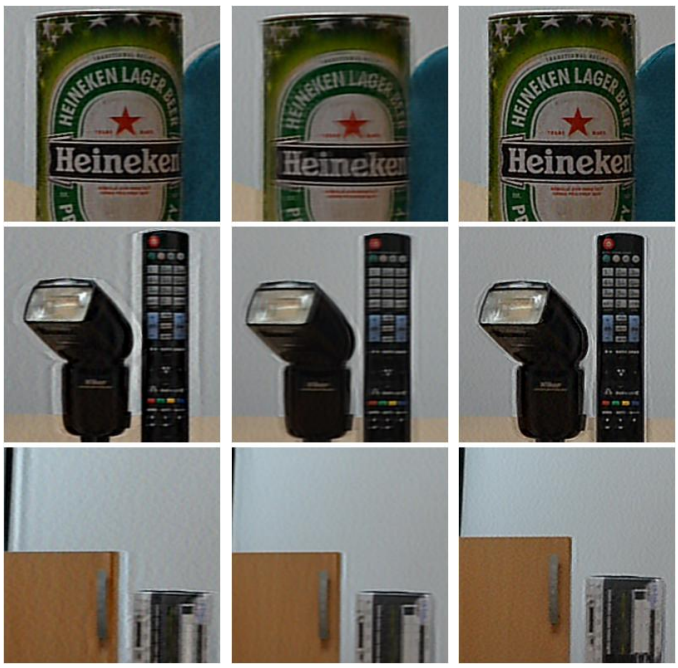
Figure 7: **Comparison of PSFs estimated by three methods from artificially generated blurred and noisy images in Fig. 6.** (a) PSF by space-invariant method of Yuan et al. [3], (b) PSFs by block based method of Šorel and Šroubek [4], (c) PSFs in (b) corresponding to the labeled regions in Fig. 6(d), (d) PSFs by our method, and (e) groundtruth PSFs.



(a) (b)

(c) (d)

Figure 8: **Comparison of the latent images estimated by three methods from artificially generated blurred and noisy images in Fig. 6.** (a) Method of Yuan et al. [3], (b) method of Šorel and Šroubek [4], (c) our method, and (d) the groundtruth.



(a) (b) (c)

Figure 9: **Regional comparison of the latent images estimated by three methods from artificially generated blurred and noisy images in Fig. 6.** (a) Method of Yuan et al. [3], (b) method of Šorel and Šroubek [4], and (c) our method.

Table 1: Comparison of each channel's PSNRs of the images reconstructed by the three methods from artificially generated blurred and noisy images in Fig. 6 [dB].

Method	R-channel	G-channel	B-channel
Space-invariant [3]	31.54	31.70	31.47
Blockwise [4]	33.41	33.99	33.29
Our method	34.87	35.05	34.90

Table 2: Exposure times, f-numbers, and ISO settings of real images.

	Blurred image	Noisy image
Three objects (Fig. 10)	1/15s, f/8, ISO 800	1/125s, f/11, ISO 12800
Parking lot (Fig. 14)	1/15s, f/4.5, ISO 100	1/125s, f/6.3, ISO 3200
Kitchen (Fig. 15)	1/10s, f/3.3, ISO 400	1/40s, f/3.3, ISO 1600
Desks (Fig. 16)	1/15s, f/4.5, ISO 100	1/15s, f/4.5, ISO 100
Cups (Fig. 23)	1s, f/2.8, ISO 80	1/30s, f/2.8, ISO 1600

The kitchen images are acquired by Samsung ST5000, the cups images by Canon IXUS 110 IS, and the others by Nikon D7000. The noisy image of desks is made with additional 5% of Gaussian noise.

reliably estimated, and the image restoration results of those blocks are not so desirable compared to those by our method.

Table 1 compares each channel's PSNRs (peak signal-to-noise ratios) of the latent images reconstructed by the three methods. As shown in the table, the reconstructed image by our algorithm has higher PSNRs than that by the other methods. That is caused by the fact that our algorithm estimates a more accurate PSF for each region segmented by the disparity.

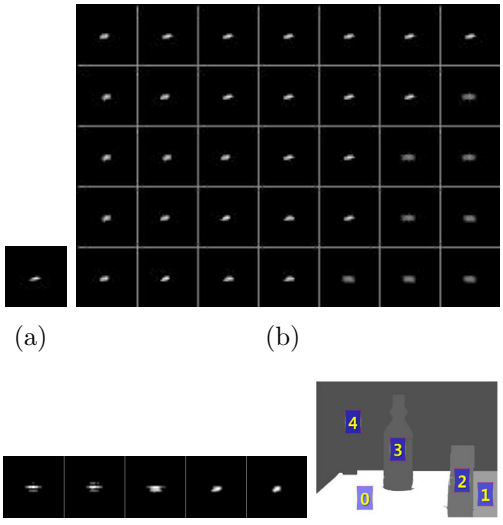
6.2. Results for captured real images

We have captured a pair of images whose resolution is 1365×1024: a blurred image under long exposure and low ISO and a noisy image under short exposure and high ISO. They are sequentially captured under a usual hand tremor in the bracketing mode of the camera. Figure 10 shows the pair of blurred image and registered version of noisy image. Their exposure times, f-numbers, and ISO settings are shown in Table 2 (row of three objects).

Figure 11 shows the results of PSF estimation from the pair of real motion-blurred and noisy images by the three methods: (a) Yuan et al. [3], (b) Šorel and Šroubek [4], (c) our method, and (d) the segmentation result by our method. Each PSF of (c) is corresponding to the labeled region ($0 \sim 4$) of

(a) (b)

Figure 10: **Sequentially captured pair of blurred and noisy images of three objects.** (a) Blurred image and (b) registered version of noisy image.



(a) (b)

(c) (d)

Figure 11: **Comparison of PSFs estimated by three methods from sequentially captured real images in Fig. 10.** (a) Method of Yuan et al. [3], (b) method of Šorel and Šroubek [4], (c) our method, and (d) segmentation result by our method.

(d) in order. As shown in (b) and (c), the PSF varies with its location, and it is highly related to its disparity. In the result by Šorel and Šroubek [4], for any block including dissimilar disparities together, the PSF could not be properly estimated. In such a case, the PSF is estimated mainly based on a region with relatively large gradients. For example, some blocks partially includes the wall and an object. In the scenes the wall is relatively distant from the camera while the object is relatively close to the camera. Since the wall has less gradients, the PSF is estimated mainly based on the close object which has larger gradients. Therefore, the shape of the estimated PSF is somehow wide due to the gradients of the close object pixels. On the other hand, the result by our method has the regional PSFs coherent with their disparity.

Figures 12 and 13 compare the latent images estimated by the three methods from the sequentially captured images. The result by our method show the best qualities among the results by the three methods.

Figures 14, 15, and 16 shows other sequentially captured pairs of blurred and noisy images of a typical parking lot, kitchen, and desks, respectively. The blur and depth are more complex than those of the previous three objects scene. Their exposure times, f-numbers, and ISO settings are shown in Table 2 (rows of parking lot, kitchen, and desks). The deblurring results by the three methods are shown in Figs. 17, 18, 19, 20, 21, and 22. Note that the result by our method still exhibits the highest qualities and least undesirable artifacts.

Figure 23 shows a substantially degraded result from images of cups. In this example, the camera is quite close to the scene compared to the depth range of the scene. As a result, the perspectivity is highly affected by camera motion, both the motion in the exposure time of the blurred image and the motion between the pair of images. Therefore, a single homography model is not sufficient for this example. For large and complicated camera motion, region-wise homography (or affine) transformation can be more appropriate.

7. Conclusion

We presented a disparity-based deblurring algorithm using a pair of noisy and blurred images. Our algorithm adequately segments the image into regions by initial graph-cut over-segmentation based on color, and disparity-based merging. For each region a PSF is estimated, and a regional latent image is restored. Finally the restored regional images are merged into a

(a)

(b)

(c)

Figure 12: **Comparison of latent images estimated by three methods from sequentially captured real images in Fig. 10.** (a) Method of Yuan et al. [3], (b) method of Šorel and Šroubek [4], and (c) our method.

(a) (b) (c)

Figure 13: **Regional comparison of latent images estimated by three methods from sequentially captured real images in Fig. 10.** (a) Method of Yuan et al. [3], (b) method of Šorel and Šroubek [4], and (c) our method.

(a) (b)

Figure 14: **Sequentially captured pair of blurred and noisy images of a parking lot.** (a) Blurred image and (b) registered version of noisy image.

(a) (b)

Figure 15: **Sequentially captured pair of blurred and noisy images of a kitchen.** (a) Blurred image and (b) registered version of noisy image.

(a) (b)

Figure 16: **Sequentially captured pair of blurred and noisy images of desks.** (a) Blurred image and (b) registered version of noisy image.

(a)

(b)

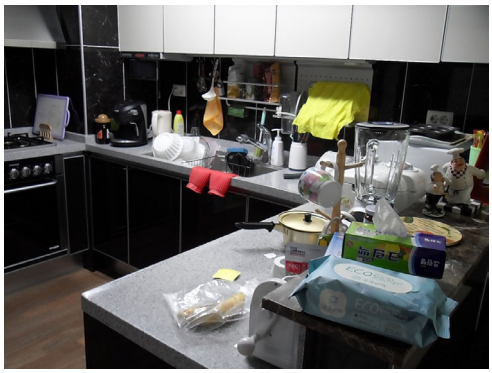
(c)

Figure 17: **Comparison of latent images estimated by three methods from sequentially captured real images in Fig. 14.** (a) Method of Yuan et al. [3], (b) method of Šorel and Šroubek [4], and (c) our method.

(a) (b) (c)

Figure 18: **Regional comparison of latent images estimated by three methods from sequentially captured real images in Fig. 14.** (a) Method of Yuan et al. [3], (b) method of Šorel and Šroubek [4], and (c) our method.

(a)



(b)

(c)

Figure 19: **Comparison of latent images estimated by three methods from sequentially captured real images in Fig. 15.** (a) Method of Yuan et al. [3], (b) method of Šorel and Šroubek [4], and (c) our method.

(a) (b) (c)

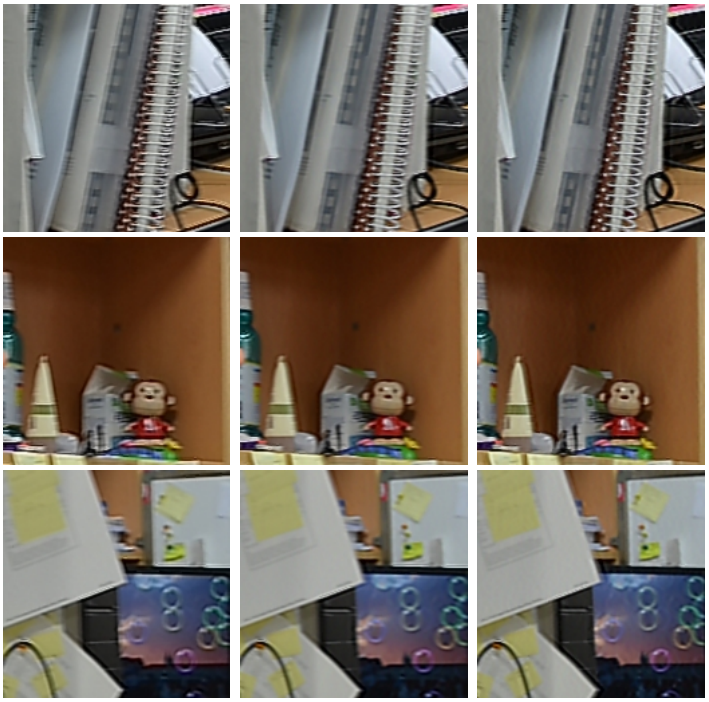
Figure 20: **Regional comparison of latent images estimated by three methods from sequentially captured real images in Fig. 15.** (a) Method of Yuan et al. [3], (b) method of Šorel and Šroubek [4], and (c) our method.

(a)

(b)

(c)

Figure 21: **Comparison of latent images estimated by three methods from sequentially captured real images in Fig. 16.** (a) Method of Yuan et al. [3], (b) method of Šorel and Šroubek [4], and (c) our method.



(a) (b) (c)

Figure 22: **Regional comparison of latent images estimated by three methods from sequentially captured real images in Fig. 16.** (a) Method of Yuan et al. [3], (b) method of Šorel and Šroubek [4], and (c) our method.

(a) (b)

(c)

Figure 23: **Degraded result from images of cups.** (a) Blurred image, (b) registered version of noisy image, and (c) the debulering result by our method. The camera settings are shown in Table 2 (row of cups).

latent image. The experimental results of artificial and real sets of blurred and noisy images have shown that our algorithm attains better qualities than the two existing distinguished methods. The proposed method is particularly useful for images with high variation of disparity.

Acknowledgment

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (No. 2010-0025703 and No. 2012R1A1A2009138), and partially by Samsung Electronics Co., Ltd.

References

[1] Q. Shan, J. Jia, A. Agarwala, High-quality motion deblurring from a single image, ACM Trans. Graph. 27 (2008) 73:1–73:10.

[2] S. Cho, S. Lee, Fast motion deblurring, ACM Trans. Graph. 28 (2009) 145:1–145:8.

[3] L. Yuan, J. Sun, L. Quan, H.-Y. Shum, Image deblurring with blurred/noisy image pairs, ACM Trans. Graph. 26 (2007) 1–10.

[4] M. Šorel, F. Šroubek, Space-variant deblurring using one blurred and one underexposed image, in: Proc. IEEE Int. Conf. Image Process., pp. 157–160.

[5] S.-H. Lee, H.-M. Park, S.-Y. Hwang, Motion deblurring using edge map with blurred/noisy image pairs, Optics Communications 285 (2012) 1777 – 1786.

[6] J. Jia, J. Sun, C.-K. Tang, H.-Y. Shum, Bayesian correction of image intensity with spatial consideration, in: T. Pajdla, J. Matas (Eds.), Computer Vision - ECCV 2004, volume 3023 of *Lecture Notes in Computer Science*, Springer Berlin / Heidelberg, 2004, pp. 342–354.

[7] S. H. Lim, D. A. Silverstein, Method for deblurring an image, US Patent Application, pub. no. US 2006/0187308 A1 (2006).

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

[9] M. Šorel, J. Flusser, Space-variant restoration of images degraded by camera motion blur, Image Processing, IEEE Transactions on 17 (2008) 105 –116.

[10] W. J. Smith, Modern optical engineering: the design of optical systems, McGraw-Hill, New York, NY, 1966.

[11] A. Chambolle, P.-L. Lions, Image recovery via total variation minimization and related problems, Numerische Mathematik 76 (1997) 167–188.

[12] W. H. Richardson, Bayesian-based iterative method of image restoration, J. Opt. Soc. Am. 62 (1972) 55–59.

[13] A. Levin, R. Fergus, F. Durand, W. T. Freeman, Image and depth from a conventional camera with a coded aperture, ACM Trans. Graph. 26 (2007).

[14] D. Krishnan, R. Fergus, Fast image deconvolution using hyper-Laplacian priors, in: Proc. Neural Inf. Process. Syst., pp. 1033–1041.

[15] D. M. Greig, B. T. Porteous, A. H. Seheult, Exact maximum a posteriori estimation for binary images, Journal of the Royal Statistical Society Series B Methodological 51 (1989) 271–279.

[16] Y. Boykov, V. Kolmogorov, An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision, Pattern Analysis and Machine Intelligence, IEEE Transactions on 26 (2004) 1124 –1137.

[17] Y. Boykov, O. Veksler, R. Zabih, Fast approximate energy minimization via graph cuts, in: Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on, volume 1, pp. 377 –384 vol.1.

[18] P. Felzenszwalb, D. Huttenlocher, Efficient graph-based image segmentation, International Journal of Computer Vision 59 (2004) 167–181.

[19] C. Harris, M. Stephens, A combined corner and edge detector, in: Proc. Alvey Vision Conf., pp. 147–151.

[20] P. Dani, S. Chaudhuri, Automated assembling of images: image montage preparation, Pattern Recognit. 28 (1995) 431–445.

[21] P. Thevenaz, U. Ruttimann, M. Unser, Iterative multi-scale registration without landmarks, in: Proc. Int. Conf. Image Process., volume 3, pp. 228–231.

[22] R. Wong, E. Hall, Sequential hierarchical scene matching C-27 (1978) 359–366.

[23] A. Koschan, V. Rodehorst, K. Spiller, Color stereo vision using hierarchical block matching and active color illumination, in: Proc. Int. Conf. Pattern Recognit., volume 1, pp. 835–839.

[24] C. Je, H.-M. Park, Optimized hierarchical block matching for fast and accurate image registration, Signal Processing: Image Communication (in press, available at `http://dx.doi.org/10.1016/j.image.2013.04.002`).

[25] M. Debella-Gilo, A. Kb, Sub-pixel precision image matching for measuring surface displacements on mass movements using normalized cross-correlation, Remote Sens. Environ. 115 (2011) 130–142.

[26] M. A. Fischler, R. C. Bolles, Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography, Commun. ACM 24 (1981) 381–395.

[27] A. N. Tikhonov, V. Y. Arsenin, Solutions of ill-posed problems, V. H. Winston and Sons., 1977.

[28] A. Hyvärinen, J. Karhunen, E. Oja, Independent Component Analysis, Wiley Interscience, 2001.

[29] R. Fergus, B. Singh, A. Hertzmann, S. T. Roweis, W. T. Freeman, Removing camera shake from a single photograph, ACM Trans. Graph. 25 (2006) 787–794.

[30] D. Krishnan, T. Tay, R. Fergus, Blind deconvolution using a normalized sparsity measure, in: Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on, pp. 233 –240.

[31] Y. Wang, J. Yang, W. Yin, Y. Zhang, A new alternating minimization algorithm for total variation image reconstruction, SIAM J. Img. Sci. 1 (2008) 248–272.

[32] C. Wang, Z. Liu, Total variation for image restoration with smooth area protection, J. Signal Process. Syst. 61 (2010) 271–277.

[33] A. Polesel, G. Ramponi, V. Mathews, Image enhancement via adaptive unsharp masking, Image Processing, IEEE Transactions on 9 (2000) 505 –510.