DOI: 10.2352/ISSN.2470-1173.2016.2.VIPC-033

Motion deblurring for depth-varying scenes

Ruiwen Zhen, Robert Stevenson; University of Notre Dame, Notre Dame, IN, 46556

Abstract

Camera motion blur generally varies across the image plane. In addition to camera rotation, scene depth is also an important factor that contributes to blur variation. This paper addresses the problem of estimating the latent image of a depth-varying scene from a blurred image caused by camera in-plane motion. To make this depth-dependent deblurring problem tractable, we acquire a small sequence of images with different exposure settings along with inertial sensor readings using a smart phone. The motion trajectory can be roughly estimated from the noisy inertial measurements. The short/long exposure settings are arranged in a special order such that the structure information preserved in short-exposed images is employed to compensate the trajectory drift introduced by the measurement noise. Meanwhile, these short-exposed images could be regarded as a stereo pair which provide necessary constraints for depth map inference. However, even with ground-truth motion parameters and depth map, the deblurred image may still suffer from ringing artifacts due to depth value ambiguity along objects boundaries resulting from camera motion. We propose a modified deconvolution algorithm that searches the "optimal" depth value in a neighborhood for each boundary pixel to resolve ambiguity. Experiments on real images validate that our deblurring approach achieves better performance than existing state-of-the-art methods on a depth-varying scene.

Introduction

Motion blur due to camera shake is a common artifact in digital images, especially when using hand-held cameras in low-light conditions. The problem of removing the motion blur has received a lot of attention in recent years.

Traditionally, the blurred image is modeled as a latent image convolved with a blur kernel. Since the underlying blur kernel is unknown in practice, image deblurring, also known as blind deconvolution, is an ill-posed problem and additional information is needed to constrain the solution space. Most single image deblurring algorithms incorporate image or kernel prior knowledge. Multiple images and hardware have also been considered. For example, Yuan et al. [21] used a blurred/noisy image pair to make kernel estimation easier and Ben-Ezra et al. [12] proposed a hybrid imaging system where the attached low-resolution video camera can record motion trajectory. A common assumption in these algorithms is that the blur kernel is spatially-invariant. However, the blur kernel generally varies across the image plane due to camera rotation, electronic rolling shutter mechanism and scene depth variation.

Recently Tai et al. [16] introduced a novel model, referred to as projective motion blur model, which interprets the blurred image as an integration of the scene under a sequence of projective motions during the exposure time. This model is able to describe 6D camera motion with global homography transformations. Whyte et al. [18] substituted this model into the existing blind deblurring algorithms to recover 3D camera rotations. Concurrently, Gupta et al. [3] described similar work to explain in-plane rotation and translation. Hirsch et al. [5] and Hu et al. [7] developed fast algorithms to reduce computational loads. Joshi et al. [8] integrated inexpensive and lightweight inertial sensors with a DSLR camera to measure camera motion and calculate the homography parameters based on rigid body kinematics.

None of the above methods ever consider the effect of scene depth variation on blur kernels. The existence of depth factor can significantly complicate the deblurring problem. First, depth map acquisition generally requires special hardware, such as stereo-camera or depth sensor. In addition, a global motion descriptor is not sufficient to describe scene depth variation, so the advantages brought by the projective motion blur model will no longer exist. What's more, depth discontinuity leads to abrupt blur kernel changes along the boundaries. Small errors of depth estimation at the boundary pixels can be magnified as ringing artifacts in the deblurred image.

limited effort has been devoted to depth-aware deblurring. Li et al. [10] extended the hybrid camera with an additional low-resolution video camera so that two low-resolution cameras form a stereo pair and provide a low-resolution depth map. Levin et al. [9] designed a coded aperture to recover both latent image and layered depth map with the help of user-drawn strokes in some cases. Xu et al. [19] inferred depth from two blurred images captured by a stereo camera and proposed a hierarchical estimation framework to remove motion blur caused by in-plane translation. Sorel et al. [15] also discussed depth effect on translational blur, but they estimated the relative depth map using the constraints from multiple images rather than stereo configuration. The algorithm of Paramanand et al. [13] allows for more complex camera motion, in-plane rotation as well as translation. However, the scene is restricted to a bilayer scene (a foreground plus a background). Hu et al. [6] proposed to jointly estimate depth layers and remove non-uniform blur caused by in-plane motion from a single blurred image. While the unified framework is promising, user input for partitioning depth layers is required and potential depth values should be known in advance.

In this paper, we present a multi-image motion deblurring approach that takes advantage of the inertial data and a special combination of blurred/noisy images to restore scene depth as well as latent image. A simple application on smart phones was designed such that it can capture a burst of three images with different exposure settings and record inertial data simultaneously. The first and last images are taken with a short exposure time and thus appear to be sharp and noisy while the intermediate image taken with a long exposure time is motion blurred. As in our previous work [23], we used the sharp image edges preserved in the denoised versions of the first and last images to remove the drift in the motion trajectory estimated from noisy inertial measure-

DOI: 10.2352/ISSN.2470-1173.2016.2.VIPC-033

ments. The depth map is obtained by considering the denoised images as if they were captured by a stereo camera. Although the baseline between the two images is not fixed and unknown, we can approximate it using the recorded inertial data.

In addition, it is found that the depth value along sharp transitions in the depth map is always ambiguous for a blurred image, since camera motion causes one pixel move across depth discontinuity region. Therefore, a simple extension of the Projective Motion Richardson-Lucy algorithm via replacing a constant depth value with pixel-wise depth values may result in severe ringing artifacts in the case of large depth variation. To reduce the ringing artifacts, we attempt to search an optimal depth value from several candidates at the pixel locations of interest for each homography transformation and substitute these updated depth maps in the Projective Motion Richardson-Lucy algorithm. The effectiveness of our modified algorithm has been demonstrated by both synthetic and real images.

Depth-Aware Blur Model

In this section, we present the depth-ware projective motion blur model which describes camera motion and scene depth explicitly.

In a pinhole camera, a point light source in the scene will be projected to different locations of the sensor if the camera moves during the exposure time. The resulting trajectory left on the sensor is the point spread function (PSF) of that point. Let $\boldsymbol{x}$ represent the initial projection of scene point $\boldsymbol{X}$, then the projection at time t is related to $\mathbf{x}$ by a homography transformation $\mathbf{H}(t)$,

$$\boldsymbol{x}'(t) = \mathbf{H}(t)\boldsymbol{x} = \mathbf{K}(\mathbf{R}(t) + \frac{1}{d}\mathbf{T}(t)\mathbf{N}^T)\mathbf{K}^{-1}\boldsymbol{x} \tag{1}$$

where $\mathbf{R}(t)$ and $\mathbf{T}(t)$ are 3×3 rotation matrix and 3×1 translation vector respectively, $\mathbf{N}$ is the normal vector to the image plane and $\mathbf{K}$ represents the camera intrinsic matrix. In the previous work, the depth value d is assumed constant, but for depth-varying scenes, d is a function of pixel location $\boldsymbol{x}$. In this case, the homography matrix $\mathbf{H}$ depends on both camera motion and scene depth map D, denoted by $\mathbf{H}(t,D)$. This equation also implies that only translational blur is affected by depth variation and its size is inversely proportional to depth value. Therefore, the objects that are close to the camera sensor are likely to appear more blurry than those at a distance.

According to the projective motion blur model, the blurred image is generated by integrating all homography-transformed versions of the sharp image over the exposure time T, i.e.,

$$g(\boldsymbol{y}) = \int_0^T f(\mathbf{H}^{-1}(t,D)\boldsymbol{y})dt + n \tag{2}$$

where g, f, n denote the blurred image, latent image and additive Gaussian noise. Note that the mapping that maps any pixel location $\boldsymbol{y}$ in the blurred image g to its original location in the sharp image f is the inverse homography matrix $\mathbf{H}^{-1}(t,D)$. In discrete form, the integral is replaced by a weighted sum over a set of camera poses,

$$g(\boldsymbol{y}) = \sum_{i=1}^{N} w_i f(\mathbf{H}_i^{-1}(D)\boldsymbol{y}) + n \tag{3}$$

The weights $w_i, i = 1\ldots N$ satisfy $\sum_i w_i = 1$ and represent the proportion of exposure time spent at each camera pose.

General camera motion has 6 degrees, consisting of 3 rotations and 3 translations, as shown in Eq. (1). In this work, we only consider in-plane motion (include in-plane translations and rotation) for simplicity. This assumption has been proved valid in many scenarios since small out-of-plane rotations can be approximated by in-plane translations [3]. The homography $\mathbf{H}_i(D)$ then simplifies to

$$\mathbf{H}_i(D) = \mathbf{K}\left(\begin{bmatrix}\cos(\theta_i) & -\sin(\theta_i) & T_{ix}/D(\boldsymbol{y})\\ \sin(\theta_i) & \cos(\theta_i) & T_{iy}/D(\boldsymbol{y})\\ 0 & 0 & 1\end{bmatrix}\right)\mathbf{K}^{-1} \tag{4}$$

Where θ_i, T_{ix}, T_{iy} respectively represent the rotation angle around z-axis (optical axis) and x-y plane translations. With our setting, these unknown motion parameters can be roughly estimated from measured inertial data: θ_i is a single integration of z-axis angular velocity from gyroscope, and T_{ix}, T_{iy} are computed by integrating x- and y-axis accelerometer data twice. The specific details can be found in our previous work [23]. In addition, since inertial sensors are uniformly sampled, the weights w_i are equal to $\frac{1}{N}$.

Motion Deblurring

Given the image sequence f_1, f_2, f_3 captured by the smart phone camera and the simultaneously recorded inertial data $\{\boldsymbol{a}_i = [a_{ix}, a_{iy}, a_{iz}], \boldsymbol{\omega}_i = [\omega_{ix}, \omega_{iy}, \omega_{iz}], i = 1,\ldots,M\}$, our objective is to estimate the latent image f from the blurred observation f_2. In this process, three problems need to be addressed: estimating depth map, compensating temporally-increasing motion drift induced by inertial measurement noise, and generating a deblurred image with less ringing artifacts. Our solutions to these problems are explained in the following subsections.

Disparity Estimation

Due to fast shutter speed and high ISO value, the captured images f_1 and f_3 are noisy and partially lost correct colors, but the sharp image structures can be well preserved after denoising process. This allows the denoised versions of f_1 and f_3 to work as a stereo-like pair from which a coarse disparity map is estimated. The disparity map is then converted to a depth map via the inverse relation,

$$D = \frac{f * b}{Disp} \tag{5}$$

where f, b and $Disp$ respectively denote the camera focal length, baseline and estimated disparity map. Since this stereo-like pair is not captured by a real stereo camera, baseline means the distance between the two camera views from which the pair of images are taken. The baseline is not fixed and unknown because of arbitrary in-plane camera motion.

Fig. 1 presents an example to illustrate each step in the disparity map estimation. Initially we apply the state-of-the-art CBM3D denoising algorithm [1] on f_1 and f_3 to obtain the denoised images f_{1D} and f_{3D}. In order to compute the disparity map, for each pixel in the reference image (f_{1D}), the corresponding pixel in the target image (f_{3D}) has to be found. Most disparity estimation methods (stereo dense matching) require the camera motion that relates one image to the other is parallel to x-axis of the camera sensor so that the search of the corresponding pixel is limited to the pixels on the same horizontal line in the target image.

DOI: 10.2352/ISSN.2470-1173.2016.2.VIPC-033

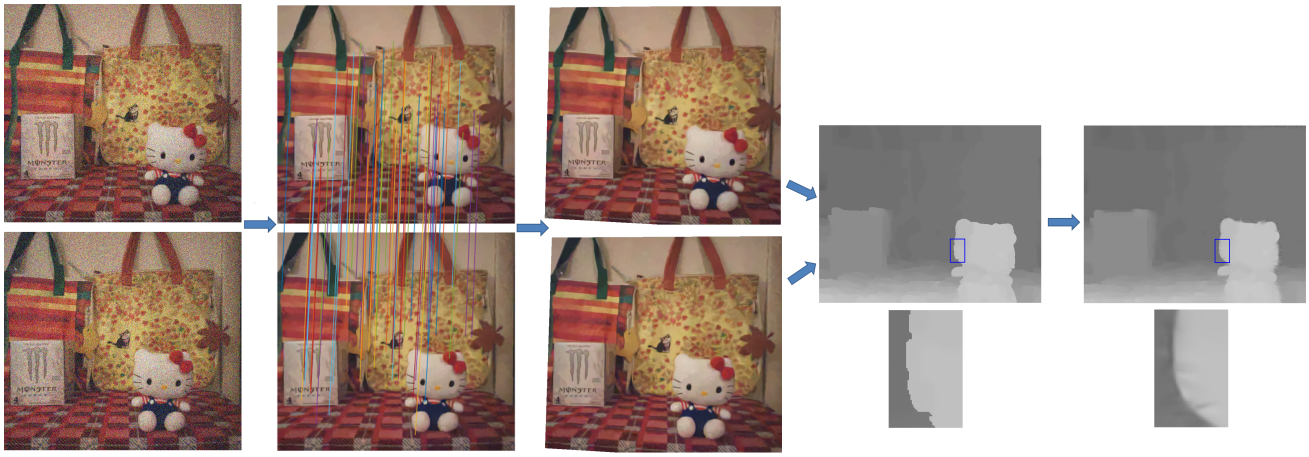


Figure 1. *Disparity map estimation. From left to right, they are noisy images, denoised images marked by point correspondences, rectified images, initially estimated disparity map, final disparity map after inverse transformation and guided filtering, respectively.*

Therefore, it is necessary to perform epipolar rectification which transforms f_{1D} and f_{3D} to make the corresponding points have the same vertical coordinates. For our uncalibrated case, we need to identify a set of point correspondences between f_{1D} and f_{3D} first. SIFT is adopted to find potential matches and RANSAC is applied to discard outliers [17]. Then the fundamental matrix $\mathbf{F}$ can be computed from these matches and Quasi-Euclidean epipolar rectification [2] is used to produce a pair of rectified images f_{1R} and f_{3R} that have small distortion compared to original images.

A coarse disparity map is generated by taking the overlapping area of the rectified images as input to the two-step stereo matching algorithm [11]. To fill occlusions and remove outliers, the generated disparity map is post-processed by a left-to-right cross-checking, a weighted median filter and a subpixel correction. Note that this disparity map is estimated with respect to the left rectified image f_{1R}. The inverse projective transformation from rectification should be applied on the disparity map to align it with the original denoised image f_{1D}. Sometimes the region boundary in the estimated disparity map is not well handled due to low contrast of denoised images. We employ the guided filter [4] to preserve boundaries and the guidance image is obtained by enhancing f_{1D} such that its histogram matches that of the blurred image f_2.

Motion Drift Compensation

In the case of in-plane camera motion, a sequence of motion parameters $\{[T_{ix}, T_{iy}, \theta_i], i = 1, \ldots, M\}$ can be computed from the measured inertial data $\{[a_{ix}, a_{iy}, \omega_{iz}], i = 1, \ldots, M\}$. However, a common issue with using inertial sensors for camera motion estimation is noise amplification. The noise present in the inertial measurements, especially accelerometer data, could lead to significant drift in the estimated motion trajectory and degrade the final deblurring result. Based on our previous work [23], we address this problem for depth-varying scenes by exploiting the underlying relation between the captured images.

The first image f_1 is related to the last image f_3 by the fundamental matrix $\mathbf{F}$, from which the rotation matrix $\mathbf{R}_F$ and translation vector $\mathbf{T}_F$ at the end of the camera motion can be retrieved up to scale and a four-solution ambiguity. The solution ambiguity is resolved by testing a single point to determine if it is in front of both camera views, but the scale factor is unknown and approximated by

$$s = \sqrt{T_{Mx}^2 + T_{My}^2}. \tag{6}$$

Given the desired end point $\{[T_{Fx}, T_{Fy}, \theta_F]\}$ extracted from $\mathbf{R}_F$ and $\mathbf{T}_F$, a less drifted trajectory is searched for such that the accelerations and angular velocities interpreting it matches the measurements in the least-squares sense,

$$\begin{aligned} \underset{\{T_{ix}, T_{iy}, \theta_i\}}{\operatorname{argmin}} \sum_{i=1}^{M} & \left\| \left[\frac{d^2 T_{ix}}{dt^2} \quad \frac{d^2 T_{iy}}{dt^2} \quad C\frac{d\theta_i}{dt} \right] - \left[a_{ix} \quad a_{iy} \quad C\omega_{iz} \right] \right\|^2 , \\ & + \lambda \left\| \left[\frac{T_{Mx}}{s} \quad \frac{T_{My}}{s} \quad C\theta_M \right] - \left[T_{Fx} \quad T_{Fy} \quad C\theta_F \right] \right\|^2 \end{aligned} \tag{7}$$

where C is used to balance translation in unit m with rotation in unit rad and λ is a regularization parameter. The minimization is performed using the simplex method.

In most time, the above constraint alone cannot totally remove the drift. To correct the remaining drift and estimate accurate motion blur kernel for f_2, we regard the denoised image f_{1D} as a good estimate of the latent image f that undergos camera motion $\{[a_{ix}, a_{iy}, \omega_{iz}], i = k, \ldots, k+N-1\}$ to generate the blurred image f_2. Accordingly, we formulate the following objective function,

$$\underset{\{T_{ix}, T_{iy}, \theta_i\}}{\operatorname{argmin}} \sum_{\partial_* \in \{\partial_x, \partial_y\}} \left\| \partial_* f_2(\boldsymbol{y}) - \frac{1}{N} \sum_{i=k}^{k+N-1} \partial_* f_{1D}(\mathbf{H}_i^{-1}(D)\boldsymbol{y}) \right\|^2 \tag{8}$$

The depth map D is computed from the disparity map using Eq. (5) where the baseline b is approximated by s. As the translations T_{ix} and T_{iy} estimated from Eq. (7) are required to be consistent with scale factor s, using s as the baseline will cancel out the bad influence brought by inaccurate depth value when calculating the homography matrix $\mathbf{H}_i(D)$ as Eq. (4). The algorithm to

DOI: 10.2352/ISSN.2470-1173.2016.2.VIPC-033

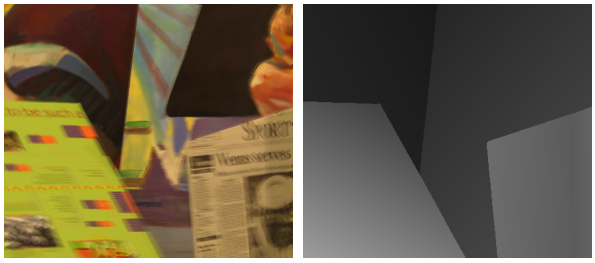
optimize Eq. (8) can be found in [23]. Due to the additional computational expense incurred by spatially-varying scene depth, the locally-uniform approximation technique [5] is adopted to speed up optimization.

Image Restoration

To recover the latent image from the blurred observation f_2, we solve the objective function

$$\underset{f}{\operatorname{argmin}} \left\| f_2(\boldsymbol{y}) - \frac{1}{N} \sum_{i=k}^{k+N-1} f(\mathbf{H}_i^{-1}(D)\boldsymbol{y}) \right\|^2 + \beta \|\nabla f\|^{0.8} \quad (9)$$

which imposes the sparse prior to regularize the solution. The Projective Motion Richardson-Lucy algorithm [16] can be used to perform the optimization. Fig. 2(a)-(c) respectively show a synthesized blurry image caused by in-plane camera motion (mainly large translation), its ground-truth disparity map [14] and the restored image. It is easy to see that the result is not satisfying as expected. It suffers from ringing artifacts along the region boundary where the depth values of two sides have a large difference.



(a) (b)

(c) (d)

Figure 2. *Comparing deblurring results on a synthetic example (a) synthesized blurry image (b) ground-truth disparity map (c) deblurred image by Projective Motion Richardson-Lucy algorithm (d) deblurred image by our algorithm*

One reason behind this problem is attributed to the depth value ambiguity that occurs near the boundary pixels. When the shutter is open, a boundary pixel in the latent image may move from one region with depth value d_1 to its adjacent region with depth value d_2 under camera motion. Theoretically each point on its trajectory is generated by a homography matrix depending on depth value d_1, but in the implementation some homographies are computed using d_2. According to the previous work, to estimate the value of each transformed image at the pixel location $\boldsymbol{y}$, the corresponding pixel $\boldsymbol{x}$ in the latent image f is found by applying an inverse homography matrix $\mathbf{H}_i^{-1}(D(\boldsymbol{y}))$ on $\boldsymbol{y}$, i.e.,

$$\boldsymbol{x} = \mathbf{H}_i^{-1}(D(\boldsymbol{y}))\boldsymbol{y}. \quad (10)$$

However, in terms of the blurring process, the correct inverse mapping should be

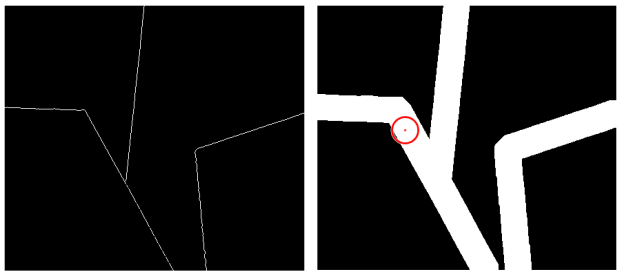
$$\boldsymbol{x} = \mathbf{H}_i^{-1}(D(\boldsymbol{x}))\boldsymbol{y}. \quad (11)$$

When $D(\boldsymbol{x})$ is close to $D(\boldsymbol{y})$, the values of $\boldsymbol{x}$ estimated from Eq. (10) and Eq. (11) are similar and thus no visible artifacts can be seen in the restored image. But if translation exists and the difference between $D(\boldsymbol{x})$ and $D(\boldsymbol{y})$ is large enough, severe ringing artifacts will be generated, as shown in Fig. 2(c).

Since $\boldsymbol{x}$ appears on both sides of Eq. (11), it is difficult to find a solution directly from it. To address this issue, we propose to search an "optimal" depth value from a set of sampled depth values for each pixel near the boundaries. First, discontinuities in the depth map is detected using a sobel filter, followed by a dilation operation which expands the discontinuous areas to include most depth-ambiguous pixels. The size of the structure element is chose to be the maximum kernel size. Fig. 3 shows the result of the synthesized example. For each pixel in the expanded area, a set of depth values is created between the maximum and minimum depth values in its neighborhood, marked by a red circle in Fig. 3(b). The depth values of the pixels in other areas are left unchanged. In practice, we resort to constructing the set of disparity values first instead of depth values because the depth value is continuous while the disparity value has already been quantized in most time. At each camera pose, the optimal depth value assigned for computing the homography matrix that maps the pixel location $\boldsymbol{y}$ on the current image plane to its original location $\boldsymbol{x}$ is determined by

$$\underset{d_j \in \Lambda}{\operatorname{argmin}} |d_j - D(\mathbf{H}_i^{-1}(d_j)\boldsymbol{y})| \quad (12)$$

where Λ represents the sampled set of depth values. Since searching only involves motion parameters of each camera pose and has no relation with image intensities, it is a one-pass process.



(a) (b)

Figure 3. *(a) binary image that indicates depth discontinuities (b) binary image after dilation*

Let $\hat{D}_i$ denote the new depth map updated for camera pose i. The Projective Motion Richardson-Lucy algorithm is modified as following:

$$f^{t+1}(\boldsymbol{y}) = f^t(\boldsymbol{y}) + \frac{1}{N} \sum_{i=k}^{k+N-1} E^t(\mathbf{H}_i(D)\boldsymbol{y}) + \beta \nabla(\|\nabla f^t\|^{0.8})$$

DOI: 10.2352/ISSN.2470-1173.2016.2.VIPC-033

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

Figure 4. *Depth-aware motion deblurring on a real image with small depth variation (a) blurred image (b) deblurring result of [22] (c) deblurring result of [20] (d) our restored image*

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

Figure 5. *Depth-aware motion deblurring on a real image with large depth variation (a) blurred image (b) deblurring result of [20] (c) our restored image using Projective Motion Richardson-Lucy algorithm (d) final restored image using modified Richardson-Lucy algorithm*

(13)

$$E^t(\boldsymbol{y}) = f_2(\boldsymbol{y}) - \frac{1}{N}\sum_{i=k}^{k+N-1} f^t(\mathbf{H}_i^{-1}(\hat{D}_i)\boldsymbol{y}), \tag{14}$$

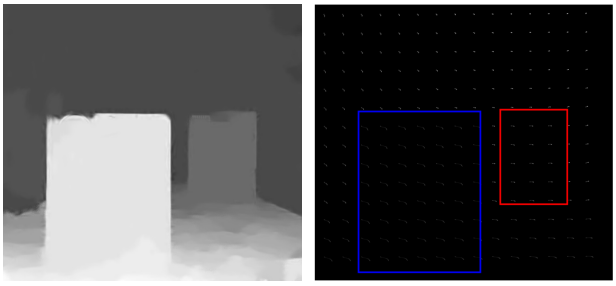
Note that new depth maps are only used in inverse mapping. In forward mapping, we still use the original depth map. Fig. 2(d) is the deblurred image produced by our modified algorithm. Compared to Fig. 2(c), ringing artifacts have been significantly reduced.

Experiments

In this section, we evaluate the proposed deblurring approach on real images and compare its performance against several state-of-the-art methods.

The real images along with inertial sensor measurements are acquired by our designed application running on Google Nexus 5 [23]. The exposure time/ISO values for the captured three images are respectively configured to $\frac{1}{330}s/800$, $0.3s/100$ and $\frac{1}{330}s/800$. The 3-axis accelerometer and 3-axis gyroscope are polled at the fastest rate approximately 200Hz. We measured the standard deviation of the measurement noise, $0.005rad/s$ for gyroscope and $0.023m/s^2$ for accelerometer. In all experiments, the parameters $\lambda = 100$ and $\beta = 0.25$ are used.

Fig. 4(a) shows a real blurred image due to hand-shake and the noisy images captured before and after it have been given in Fig. 1. Though the scene has small depth variation, the state-of-the-art deblurring methods based on constant depth does not work well. Fig. 4(b) is the deblurring result generated by Zhang's method [22] that accepts multiple images as input and assumes the blur kernel is spatially-invariant. Xu's algorithm [20] can handle spatially-varying blur caused by high-degree camera motion. However, constant depth assumption makes their algorithm produce correct rotation component but incorrect translation component, resulting in unsatisfying deblurred image Fig. 4(c). In comparison, scene details at different depths are well recovered in our restored image, as shown in Fig. 4(d), since the proposed approach takes depth factor into account.



(a) (b)

Figure 6. *(a) estimated disparity map (b) estimated spatially-varying blur kernels for example in Fig. 5*

In the previous example, no visible artifacts can be seen along depth discontinuities in the deblurred image generated by the traditionally Projective Motion Richardson-Lucy algorithm. To present the advantages of our modified deconvolution algorithm,

DOI: 10.2352/ISSN.2470-1173.2016.2.VIPC-033

we consider a scene with large depth variation shown in Fig. 5. Obviously, the box in the front is more blurred than the box in the back. Fig. 5(b) and (c) are respectively the deblurring result of Xu's algorithm and our approach. We magnified the text regions to better visualize the difference. Due to the same motion parameters and different deconvolution algorithms, Fig. 5(c) and (d) are the same everywhere except the boundary regions inside the magenta rectangle. Fig. 5(c) is restored using the traditional algorithm while (d) using the modified algorithm. As shown in the close-up, the ringing artifacts caused by large depth step between the box in the front and the background is alleviated by our modified algorithm.

Conclusion

In this work, we have developed an approach for removing camera motion blur from blurred images taken of scenes with depth variation. Our approach uses popular smart phones to capture image data in conjunction with inertial data. No special hardware, such as stereo camera, is required. The relation between captured images as well as between images and inertial measurements are fully utilized to infer depth map and estimate accurate motion trajectory. We also presented modification to Projective Motion Richardson-Lucy algorithm to reduce ringing artifacts along boundaries in the restored image. In the case of scenes with narrow depth range, our approach is still applicable.

There are two directions of future work. One is to improve depth map estimation in the second-pass using the deblurred image. Another is the extension to general camera motion.

References

[1] Kostadin Dabov, Alessandro Foi, Vladimir Katkovnik, and Karen Egiazarian. Color image denoising via sparse 3d collaborative filtering with grouping constraint in luminance-chrominance space. In *Image Processing, 2007. ICIP 2007. IEEE International Conference on*, volume 1, pages I–313. IEEE, 2007.

[2] Andrea Fusiello and Luca Irsara. Quasi-euclidean uncalibrated epipolar rectification. In *Pattern Recognition, 2008. ICPR 2008. 19th International Conference on*, pages 1–4. IEEE, 2008.

[3] Ankit Gupta, Neel Joshi, C Lawrence Zitnick, Michael Cohen, and Brian Curless. Single image deblurring using motion density functions. In *Computer Vision–ECCV 2010*, pages 171–184. Springer, 2010.

[4] Kaiming He, Jian Sun, and Xiaoou Tang. Guided image filtering. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 35(6):1397–1409, 2013.

[5] Michael Hirsch, Christian J Schuler, Stefan Harmeling, and Bernhard Scholkopf. Fast removal of non-uniform camera shake. In *Computer Vision (ICCV), 2011 IEEE International Conference on*, pages 463–470. IEEE, 2011.

[6] Zhe Hu, Li Xu, and Ming-Hsuan Yang. Joint depth estimation and camera shake removal from single blurry image. In *Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on*, pages 2893–2900. IEEE, 2014.

[7] Zhe Hu and Ming-Hsuan Yang. Fast non-uniform deblurring using constrained camera pose subspace. In *BMVC*, pages 1–11, 2012.

[8] Neel Joshi, Sing Bing Kang, C Lawrence Zitnick, and Richard Szeliski. Image deblurring using inertial measurement sensors. In *ACM Transactions on Graphics (TOG)*, volume 29, page 30. ACM, 2010.

[9] Anat Levin, Rob Fergus, Frédo Durand, and William T Freeman. Image and depth from a conventional camera with a coded aperture. In *ACM Transactions on Graphics (TOG)*, volume 26, page 70. ACM, 2007.

[10] Feng Li, Jingyi Yu, and Jinxiang Chai. A hybrid camera for motion deblurring and depth map super-resolution. In *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, pages 1–8. IEEE, 2008.

[11] Mikhail G Mozerov and Joost van de Weijer. Accurate stereo matching by two-step energy minimization. *Image Processing, IEEE Transactions on*, 24(3):1153–1163, 2015.

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

[13] Chandramouli Paramanand and Ambasamudram Narayanan Rajagopalan. Non-uniform motion deblurring for bilayer scenes. In *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*, pages 1115–1122. IEEE, 2013.

[14] Daniel Scharstein and Richard Szeliski. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *International journal of computer vision*, 47(1-3):7–42, 2002.

[15] Michal Šorel and Jan Flu. Space-variant restoration of images degraded by camera motion blur. *Image Processing, IEEE Transactions on*, 17(2):105–116, 2008.

[16] YW Tai, P Tan, and MS Brown. Richardson-lucy deblurring for scenes under projective motion path. *IEEE transactions on pattern analysis and machine intelligence*, 2010.

[17] A. Vedaldi and B. Fulkerson. VLFeat: An open and portable library of computer vision algorithms, 2008.

[18] Oliver Whyte, Josef Sivic, Andrew Zisserman, and Jean Ponce. Non-uniform deblurring for shaken images. *International journal of computer vision*, 98(2):168–186, 2012.

[19] Li Xu and Jiaya Jia. Depth-aware motion deblurring. In *Computational Photography (ICCP), 2012 IEEE International Conference on*, pages 1–8. IEEE, 2012.

[20] Li Xu, Shicheng Zheng, and Jiaya Jia. Unnatural l0 sparse representation for natural image deblurring. In *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*, pages 1107–1114. IEEE, 2013.

[21] Lu Yuan, Jian Sun, Long Quan, and Heung-Yeung Shum. Image deblurring with blurred/noisy image pairs. *ACM Transactions on Graphics (TOG)*, 26(3):1, 2007.

[22] Haichao Zhang, David Wipf, and Yanning Zhang. Multi-image blind deblurring using a coupled adaptive sparse prior. In *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*, pages 1051–1058. IEEE, 2013.

[23] Ruiwen Zhen and Robert Stevenson. Multi-image motion deblurring aided by inertial sensors. *Journal of Electronic Imaging, 2015 submitted*.