Video deblurring and super-resolution technique for multiple moving objects

Takuma Yamaguchi[1], Hisato Fukuda[2], Ryo Furukawa[3], Hiroshi Kawasaki[4], and Peter Sturm[5]

[1] Research & Development Center, The Nippon Signal Co., LTD
1836–1 Oaza Ezura, Kuki, Saitama 346–8524 Japan
ymgc-tkm@signal.co.jp
[2] Saitama University, Department of Information and Computer Science
255 Shimo-okubo, Sakura-ku, Saitama, 338-8570, Japan
fukuda@cv.ics.saitama-u.ac.jp
[3] Faculty of Information Sciences,Hiroshima City University
3–4–1 Ozuka-higasi , Asaminami-ku, Hiroshima, 731–3194 Japan
ryo-f@hiroshima-cu.ac.jp
[4] Faculty of Engineering, Kagoshima University
1–21–24 Korimoto , Kagoshima City, Kagoshima, 890–0065 Japan
kawasaki@ibe.kagoshima-u.ac.jp
[5] INRIA Grenoble – Rhone-Alpes
655 Avenue de l'Europe, 38330 Montbonnot St Martin, France
peter.sturm@inrialpes.fr

Abstract. Video camera is now commonly used and demand of capturing a single frame from video sequence is increasing. Since resolution of video camera is usually lower than digital camera and video data usually contains a many motion blur in the sequence, simple frame capture can produce only low quality image; image restoration technique is inevitably required. In this paper, we propose a method to restore a sharp and high-resolution image from a video sequence by motion deblur for each frame followed by super-resolution technique. Since the frame-rate of the video camera is high and variance of feature appearance in successive frames and motion of feature points are usually small, we can still estimate scene geometries from video data with blur. Therefore, by using such geometric information, we first apply motion deblur for each frame, and then, super-resolve the images from the deblurred image set. For better result, we also propose an adaptive super-resolution technique considering different defocus blur effects dependent on depth. Experimental results are shown to prove the strength of our method.

1 Introduction

Demand for retrieving a high quality single image from video sequence is increasing, such as surveillance and handheld video capture and so on. Since image quality of video camera is usually lower than digital camera, simple frame capture is often insufficient for actual purpose. Although the main reason of the low quality of video data is a low resolution of video camera, motion blur is another important reason of degradation; it commonly occurs because video usually captures moving object, whereas, still camera mainly captures static scene only. Another problem on quality of video data is narrow depth of field; it is also common because video camera requires high frame-rate with fast shutter speed, resulting in wide aperture. Because of the narrow depth of field, the scene other than target object is blurred by defocus blur. Thus, simple frame capture can produce only low quality image and image restoration technique is inevitably required.

To deal with the problem mentioned above, hybrid camera systems are proposed [1, 2]. However, since those systems require additional sensors, the systems become complicated and the technique cannot be applied for common video data. On the other hand super-resolution technique using several input frames are proposed. However, most of them does not consider motion blur and only several papers take the problem into account; they treat motion blur as noise [3]. Therefore, quality of image restoration is limited.

In this paper, we propose a method to restore a sharp and high-resolution image from a video sequence by applying a motion deblurring technique for each frame followed by super-resolution technique for multiple frames. To conduct a motion deblur from an image, motion information is required. Since typical device of deblurring techniques is a still camera, they assume long exposure time and complicated camera motion; thus, sophisticated blind kernel estimation technique is usually required. To the contrary, with video camera, motion is usually small and simple for each frame. One important problem for video is that several objects move independently. In our method, by taking account of such feature of video camera, we propose a motion deblurring technique using optical flow of the scene with scene segmentation technique.

In terms of super-resolution of the image sequence, sub-pixel registration is required and it is usually difficult to achieve with blurry image. Since motion blur is reduced by our method in the first step, the problem is greatly reduced. In addition, since the scene contains several independently moving objects, segmentation and area based registration for each segment is required; it is efficiently solved by our pixel-based plane approximation technique. Further, image quality is further improved by considering the different defocus blur for each segment dependent on different depth with our adaptive super-resolution technique.

2 Related work

In terms of deblurring techniques for motion blur, since the blur is a convolution process, restoration technique has been proposed as a deconvolution technique for known kernel [4, 5]. If the kernel is unknown, such condition is common for usual photos, the problem is ill-conditioned and it cannot be solved without additional information [6]. For simple and straight-forward solution, an additional sensor is used to estimate the blur kernel [1, 7]. Recently, blind deconvolution techniques using the information of natural scene, *i.e.*, "heavy tailed distribution in the gradients" are proposed [8–11]. We also use the same knowledge to estimate the motion blur kernel.

Generally, the main reason of motion blur is assumed to be a camera motion, such as camera shake, thus, previous technique usually uses a single blur kernel for deblurring. Currently several researches are proposed considering object motion in the scene [12, 7]. In addition, more general cases, such as an independent blur kernel for each depth of an object is proposed [13]. We also estimate independent blur kernel for each segment.

In terms of super-resolution techniques, reconstructing a high-resolution image from multiple low-resolution images is intensively researched [14–16]. In those techniques, it is assumed that scenes are either static or dynamic, but consist of single depth or planar objects with little motion, and the camera is also assumed to be static. With such assumptions, registration between frames can be simplified and it can be done with sufficient accuracies with 2D affine or homography transformation. However, for applying techniques to more general purposes, it is necessary to allow 3D scenes containing multiple independently moving objects, non-rigid motion objects (e.g. cloths), etc. With existing super-resolution techniques, it is difficult to achieve this, because of significant appearance changes caused by objects' motion and viewpoint changes. To perform super-resolution for such objects or scenes, 3D information should be considered. Tung *et al.* [17] have applied super-resolution technique to construct a high-resolution 3D video. However, the technique is based on approximating 3D objects by triangular patches, and thus, accurate and dense 3D data is required; it cannot be easily acquired in general.

The technique to achieve both motion deblur and super-resolution is also proposed by Tai *et al.* [2]. The central idea is similar to ours, however, the method to estimate the motion of the scene is totally different; we estimate it only from video data, whereas Tai *et al.* use additional device as hybrid system.

3 Algorithm overview

A simple solution to restore the images that are degraded by blur kernels per each frame and object is to prepare each kernel for calculation. However, the considered input is a video sequence captured by a handheld camera, and thus, such blur kernels are not usually given. In this paper, since the input is a video sequence, we estimate those blur kernels for each segmented region of objects in the scene; those regions are detected by segmentation using optical flow field.

In terms of motion deblurring, we assume that the blur of the region to be combination of motion blur and defocus blur, where defocus blur is constant for each region. With such video data, feature points are also blurry because of motion blur and it is difficult to achieve high accuracy to detect them, however, optical flow field can be accurately acquired with area based method. Therefore, we use the optical flow field to estimate motion blur kernel.

On the other hand, restoration of low resolution image with defocus blur has been researched for long time, typically via super-resolution techniques; it is known that the quality is low if only a single image is used, and thus, many techniques using multiple images and MAP estimation are proposed to achieve reasonable results [15, 16]. To super-resolve images from low-resolution and blurred images, sub-pixel registration is required. In previous methods, where the scene is assumed to be a single plane, accurate registration can be easily achieved. However, natural scenes consist of multiple dynamic 3D objects, and thus, achieving an accurate and robust registration is not easy. In this paper, we propose a plane based registration method to achieve sub-pixel accuracy for registration of all the pixels in the images.

As already described, we assume that image blur to be combination of motion and defocus blur. Since we assume that there are several objects at different depths in the scene, all objects do not suffer from the same defocus blur. Therefore, we propose an *adaptive deblurring method* to change kernels for defocus blur adaptively for each object. Certainly, estimating blur kernels for each object is not easy, therefore, for simplicity, we assume in this paper that the kernels of defocus blur can be described as one-parameter point spread functions (Bessel function). Since we consider moving objects in the scene, defocus blur kernels may vary for each frame. However, since we use between 20 and 40 frames for super-resolution, *i.e.*, just 1 to 2 seconds of video, we assume that large changes of defocus blur kernels are unlikely, and thus, we use the same kernel for the process. Actual algorithm is as follows.

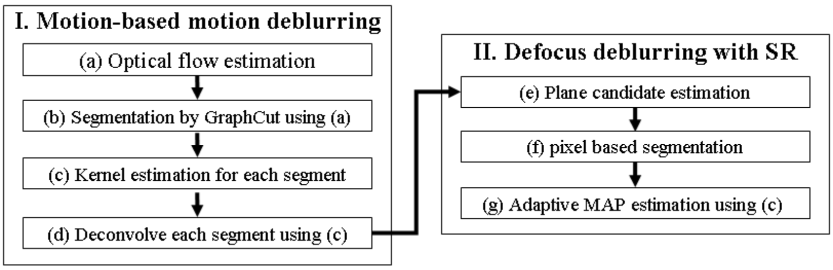


Fig. 1. Flow of the deblurring process.

First, we estimate optical flow field of the input image sequence by using block matching technique. (Fig. 1(a)). The optical flow field is segmented by graph cut method, where each of the regions include almost constant flow vectors (Fig. 1(b)). Then, an initial blur kernel is estimated for each segments (Fig. 1(c)). Simple super-resolution may not bring good results when the input images contain motion blur, because it is difficult to achieve high accuracy registration with such blurry images. Therefore, motion deblurring technique is applied before a super-resolution (Fig. 1(d)). Finally, these frame-wise deblurred results are further improved by using super-resolution technique, simultaneously improving the resolution and defocus blur (Fig. 1 II).

4 Motion deblurring for multiple moving objects

In this paper, we estimate the motion of each region by segmenting the optical flow field, and use the flow vectors for the regions to estimate motion blur kernels. For simplicity, we model image blur as a convolution of a line-shaped motion blur kernel and one-parameter isotropic defocus blur kernel. Certainly, a line-shaped motion blur kernel sometimes results in an insufficient quality, especially for a large camera motion (*e.g.*, severe ringing effects), however, camera motion is usually small and simple in our research, because all images are captured by video camera where a shutter speed is usually faster than 1/60 to keep 30 fps, and such simple kernel can achieve enough restoration in reality.

As the line-shaped motion blur kernel estimation, we use a direction of optical flow for its direction, and the knowledge of the derivatives histogram of natual scene to estimate the scaling parameter; note that such scaling parameter estimation is currently common and used by several research groups [18, 10]. The actual kernel estimation proceeds as follows:

1. The optical flow field is estimated for all the images based on block matching.
2. Input images are segmented into regions, each of which has almost constant motion vectors.
3. For each region, a line-shaped motion blur kernel is estimated from optical flow and the derivatives histogram of the image.
4. Motion blur is reduced by deconvolution algorithm by using the line-shaped blur kernels

These processes are explained in the following sections in detail.

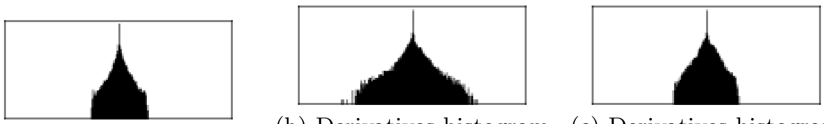
4.1 Segmentation of blurry image sequence

An input data of the proposed method is a captured sequence of images. The image may be captured by a static or moving camera. The captured scene may include multiple objects that may be static or moving. Therefore, segmentation for each object is required. Since input image is blurry, feature based method may not work, and thus, area based approach is used. In this paper, optical flow field is obtained by pyramid based block matching method. Then, multi-value graph-cut method is applied to those flow field. In our implementation, we put a large value on a direction rather than a length of the optical flow for data-term of graph-cut from our experience of several experiments. We also assume only 3 to 5 segments in the scene for fast calculation.

4.2 Blur kernel estimation using optical flow

For each extracted region, blur kernel is estimated. In our research, we assume that the shape of the motion blur kernel to be linear as mentioned above. We use a direction of optical flow for its direction, and estimate the scaling parameter by using the knowledge of the derivatives histogram of natural scene; *i.e.*, the derivatives histograms of the scene for all directions are usually the same in natural scene. Therefore, actual algorithm is as follows.

First, we calculate the derivatives histogram along optical flow vector direction. Then, we add blur to the perpendicular direction by changing the kernel size so that the both derivative histograms become similar. Fig.2(a) and (b) show the both derivatives histogram along optical flow direction and its perpendicular direction. Fig.2(c) shows the derivatives histogram along the perpendicular direction after applying the estimated blur kernel. We can clearly see that the shapes of Fig.2(a) and (c) look similar.



(a) Derivatives histogram along optical flow direction

(b) Derivatives histogram along perpendicular direction to the optical flow

(c) Derivatives histogram along the perpendicular direction after applying motion blur

Fig. 2. motion blur kernel estimation.

4.3 Motion deblurring for each segment

In terms of deconvolution algorithm, several techniques exist; Iterative Back Projection[5] is applied in our approach.

5 Super-resolution technique for multiple depth

We adopt a multi-frame super-resolution technique to restore both low-resolution and defocus blur. To realize an efficient removal of defocus blur, we first carry out a piecewise planar segmentation of the scene, in order to accomplish accurate registration and set appropriate blur kernels dependent on depth in 3D scenes. The segmentation algorithm basically consists of two steps; (1) plane candidate generation by using feature tracking results and (2) pixel-based segmentation by minimizing re-projection errors. For super-resolution, we use a MAP image reconstruction formulation with the registration result for each segment.

5.1 Estimating candidate planes based on feature point tracking

A number of studies have already been reported related to the extraction of planes from the scene for the purpose of 3D reconstruction [19–21]. In these studies, planar areas are extracted as patches by clustering feature points. However, in practice, it is often difficult to perform an accurate plane-based approximation because individual feature point tracks are easily affected by outliers, the aperture problem and view-dependent appearance changes, even if the global ego-motion estimation is accurate. In addition, since features are often not detected along object boundaries, patch creation is another difficult problem.

In this paper, we propose a pixel-based plane estimation which is more suitable than a patch-based technique. More specifically, instead of dividing the scene into patches, candidate planes are first extracted, each of which is defined by a group of tracked feature points included in a single plane. To achieve good results, a sufficient number of candidate planes should be extracted to approximate the 3D scene. A simple solution is to extract as many planes as possible from all combinations of the feature points. On the other hand, the smaller the number of candidate planes, the more efficient the computation. Therefore, we propose an efficient method to reduce the number of candidate planes to approximate the 3D scene by using the knowledge that neighboring feature points usually belong to the same plane.

Our candidate plane estimation method is described in Algorithm 1. First, corresponding feature points between input frames are computed. Then, an initial candidate plane which described by feature point tracks is generated. Using the tracks, the homography matrices between the base frame and the other frames are calculated. Next, the candidate plane is updated. Feature points whose evaluation values are less than the threshold value (0.2 pixel in our case), are added to the plane. We use the average of the re-projection errors of all the corresponding points as the evaluation value. And then, the homography matrix calculation and updating the candidate plane are iterated until the feature point tracks on the plane are converged. Repeating this manner, candidate planes describing the scene are obtained. Fig. 3 shows an example for the generation of three groups, where the black points represent the feature points which are already calculated or assigned to some planes, and the white points represent unselected and unlabeled points.

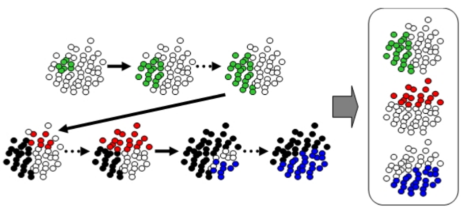


Fig. 3. Candidate plane detection.

5.2 Pixel-based segmentation by minimization of re-projection errors

Since the candidate planes (groups of feature points each of which is included in a single plane) extracted by the aforementioned method are represented as groups

Algorithm 1 Candidate plane estimation.

X is defined as the set of all corresponding feature point tracks across input frames.
$P(x)$ is defined as a predicate that is true if point track x is not selected and unlabeled.
while $\exists x \in X; P(x)$ **do**
  Select a feature point track $a(\subseteq \{x \in X; P(x)\})$ and the k nearest neighbors $b(\subseteq X)$ (in this paper $k := 7$).
  $A^{(0)} := \phi$, $A^{(1)} := a \cup b$, $i := 1$
  while $A^{(i)} \neq A^{(i-1)}$ **do**
    Compute the homography matrix $\mathcal{H}$ of $A^{(i)}$ for each frame.
    $A^{(i+1)} := \phi$
    for $\forall y \in X$ **do**
      if Adequateness of $\mathcal{H}$ for $y \geq$ *threshold* **then**
        $A^{(i+1)} := A^{(i+1)} \cup y$
      end if
    end for
    $i := i + 1$
  end while
  $A^{(i)}$ is a group of feature point tracks residing in the same plane.
end while

of feature points rather than explicit patches, the dense pixel correspondence is not yet determined at this stage. Since transformation parameters of each candidate plane between frames are calculated in the previous step, pixel-based correspondences can be estimated by assigning each pixel to one of the candidate planes by minimizing the re-projection error using the parameters.

In this paper, the homography matrices obtained from the candidate planes are used as transformation parameters. Then, the differences of intensity for each pixel from a reference frame to all other frames are computed, the average of the differences is stored for each plane, and the pixel is assigned to the plane for which that average is the smallest. The actual calculation is as follows. We denote the number of input frames as N, the homography matrix (as obtained from the i-th candidate plane, *i.e.*, the i-th group of feature points) between the reference frame and the k-th frame as $\mathcal{H}_i^{(k)}$, and the respective intensity levels of arbitrary points in the reference frame and the k-th frame as $I(\cdot)$ and $I^{(k)}(\cdot)$, respectively. Then the following equation is obtained for each pixel in the reference frame.

$$\hat{i}_{\boldsymbol{p}} = \arg\min_i \left[\frac{\sum_{k=1}^{M} \left\{ I(\boldsymbol{p}) - I^{(k)}(\mathcal{H}_i^{(k)}\boldsymbol{p}) \right\}^2}{M} \right] \quad (1)$$

Here, $M(\leq N)$ denotes the number of frames for which the pixels were effective before the projection (in other words, the pixels were within the image), and $\boldsymbol{p}$ represents a coordinate vector. By finding the minimum projection difference, each pixel is assigned to plane $\hat{i}_{\boldsymbol{p}}$. Note that since we can reject pixels whose difference measure is large, our method can handle occlusions. The process is shown in Fig. 4.

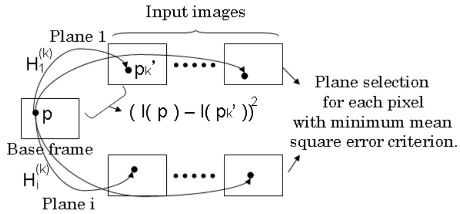


Fig. 4. Plane selection for each pixel.

5.3 Adaptive SR by MAP estimation

We use a maximum a posteriori (MAP) image reconstruction formulation for multi-frame super-resolution as follows:

$$\hat{X} = \underset{X}{\operatorname{argmin}} \left[\sum_{k=1}^{N} ||D_k H_k F_k X - Y_k||_2^2 + \lambda ||\Gamma X||_2^2 \right] \tag{2}$$

where F_k is the geometric motion operator between the high-resolution (HR) frame X and the kth low-resolution (LR) frame Y_k, H_k is the defocus blur matrix representing the camera's point spread function and D_k stands for the decimation matrix (F_k is previously estimated, see Sec. 5.2). $||\Gamma X||^2$ is the Tikhonov regularization cost function and λ is the regularization parameter. Generally, a high-pass operator is used as Γ; we use the Laplacian.

If we assume that all the decimation operations are the same (i.e. $\forall k, D_k = D$) and all the blur operations are the same (i.e. $\forall k, H_k = H$), (2) may be written as

$$\hat{X} = \underset{X}{\operatorname{argmin}} \left[\sum_{k=1}^{N} ||DF_k HX - Y_k||_2^2 + \lambda ||\Gamma X||_2^2 \right]. \tag{3}$$

We decompose this minimization problem into the following two separate steps, as suggested in [3].

1. Compute a defocus blurred HR image $\hat{Z}(= H\hat{X})$ from the LR images.
2. Estimate the HR image $\hat{X}$ from the defocus blurred HR image $\hat{Z}$.

In this paper $\hat{Z}$ is calculated by solving the following minimization problem:

$$\hat{Z} = \underset{Z}{\operatorname{argmin}} \left[\sum_{k=1}^{N} ||DF_k \hat{Z} - Y_k||_2^2 \right]. \tag{4}$$

In the deblurring step, the deblurred HR image $\hat{X}$ is obtained through the following formulation:

$$\hat{X} = \underset{X}{\operatorname{argmin}} \left[||W(HX - Z)||_2^2 + \lambda ||\Gamma X||_2^2 \right]. \tag{5}$$

where W is a diagonal matrix, each of whose diagonal values equals the number of measurements for one pixel. With this formulation, different blur kernels can be set to each pixel.

6 Experiments

6.1 Evaluation of the method using real-data

To test the effectiveness of the method, we conducted experiments using motorized stage. In this data, the scene consists of two planes with texture as shown in Fig.5(a). We set the nearest plane to be in focus and the other plane undergo a depth-dependent defocus blur by the camera aperture. We moved the two objects with different speed and different direction by two different motorized stages. The super-resolution image (SRI) with our method is shown in Fig.5(i). We can still observe small ringing effects remaining near edges, however, strong motion blur is removed and super-resolution is successfully conducted.

Next, we apply the technique to curved surfaces. The result is shown in Fig.6; in Fig.6(c), we can see that the scene is successfully segmented into several planes to approximate the curved surfaces. In Fig.6(h), we can clearly see that the motion blur is removed and super-resolution is successfully applied even if the shape has no planer area.

6.2 Handheld video data scene

In this experiment, we conducted an experiment using a handheld video camera as shown in Fig.7(a). The motion deblurred image with our method is shown in Fig.7(f). We can see that motion blur was successfully removed. The result of plane segmentation applied on the motion deblurred image sequence and the final super-resolved image are shown in Fig.7(c) and (h). Even for such natural sequence captured by handheld video, each plane was successfully segmented and super-resolution is successfully achieved.

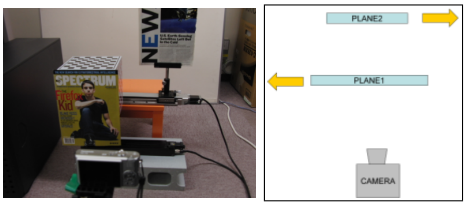
The super-resolution image without motion deblurring is shown in Fig.7(g). We can clearly see that our method gives the best restoration.

6.3 Multiple moving objects captured by static camera

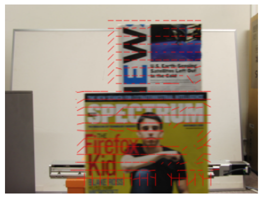
Finally, we conducted the same experiment with static camera and multiple moving objects. Fig.8(a) shows example and optical flows of the input data. Fig.8(f) shows a motion deblurred image and Fig.8(h) shows the final result by applying adaptive MAP estimation on the motion deblurred images. Fig.8(g) shows the result of simple super-resolution and we can confirm that our method achieved the best restoration.

7 Conclusion

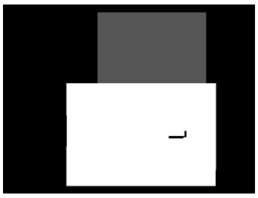
In this paper, we propose a method to restore a sharp and high-resolution image from video data captured by a handheld camera in which both independent motion and defocus blur are observed. The method is based on a motion deblurring technique using estimated blur kernels for each frame and object and



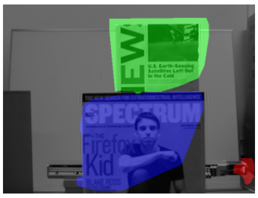
(a)Image capturing scene



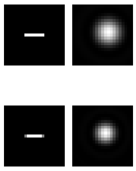
(b)Optical flows



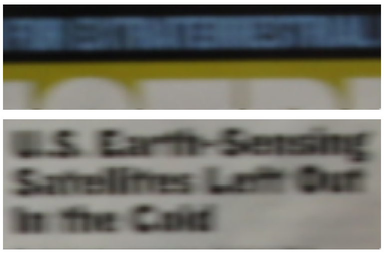
(c)Segmentation with GC



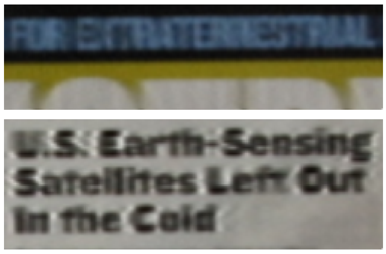
(d)Detected planes for SR



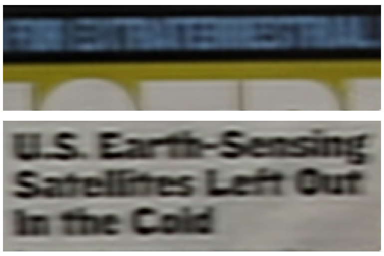
(e) Blur kernels for each segment (left:motion blur kernel, right:defocus blur kernel)



(f)Input image



(g)Motion deblurred image

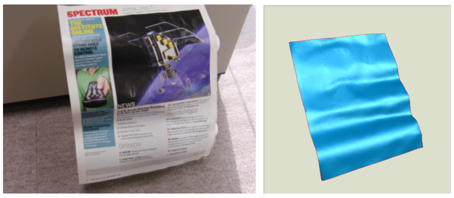


(h)SRI without motion deblurring

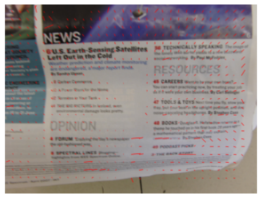
(i)Proposed method

Fig. 5. Multiple object motion by motorized stage.

super-resolution technique with adaptive defocus blur kernel. A motion blur kernel is efficiently estimated by using optical-flow and natural scene statistics and motion blur is reduced by a deconvolution algorithm. A defocus blur is removed by an adaptive MAP estimation technique with pixel-wise plane segmentation method. We conducted several experiments using real data which successfully show the effectiveness of our method. Extended research on deforming object with independent motion blur is our next step.

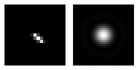


(a)Image capturing scene

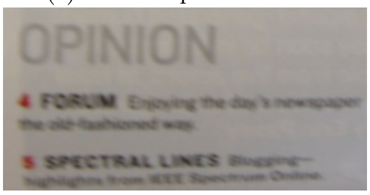


(b)Optical flows

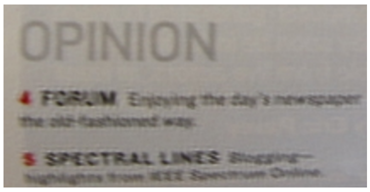
(c)Detected planes for SR



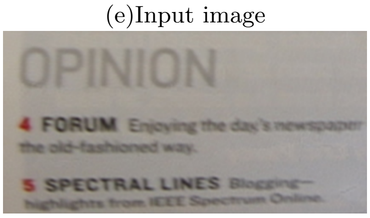
(d)Blur kernels (left:motion blur kernel, right:defocus blur kernel)



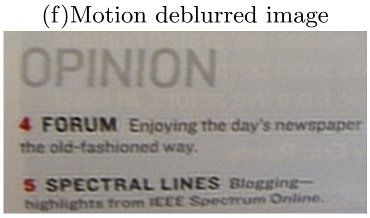
(e)Input image



(f)Motion deblurred image



(g)SRI without motion deblurring



(h)Proposed method

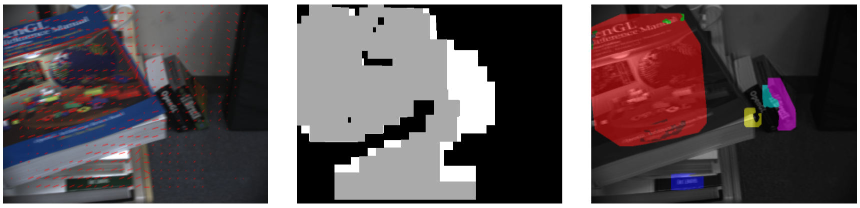
Fig. 6. Motion deblur and super-resolution for curved surface.

Acknowledgment

This work was supported in part by SCOPE No.101710002 and Grant-in-Aid for Scientific Research No.072103013, No.19700098 and 21700183 in Japan.

References

1. Ben-Ezra, M., Nayar, S.: Motion Deblurring using Hybrid Imaging. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Volume I. (2003) 657–664
2. Tai, Y.W., Du, H., Brown, M.S., Lin, S.: Image/video deblurring using a hybrid camera. In: CVPR. (2008)
3. Farsiu, S., Robinson, M.D., Elad, M., Milanfar, P.: Fast and robust multiframe super resolution. IEEE Trans. on Image Processing **13** (2004) 1327–1344
4. Lucy, L.: An iterative technique for the rectification of observed distributions. J. of Astronomy **79** (1974) 745–754



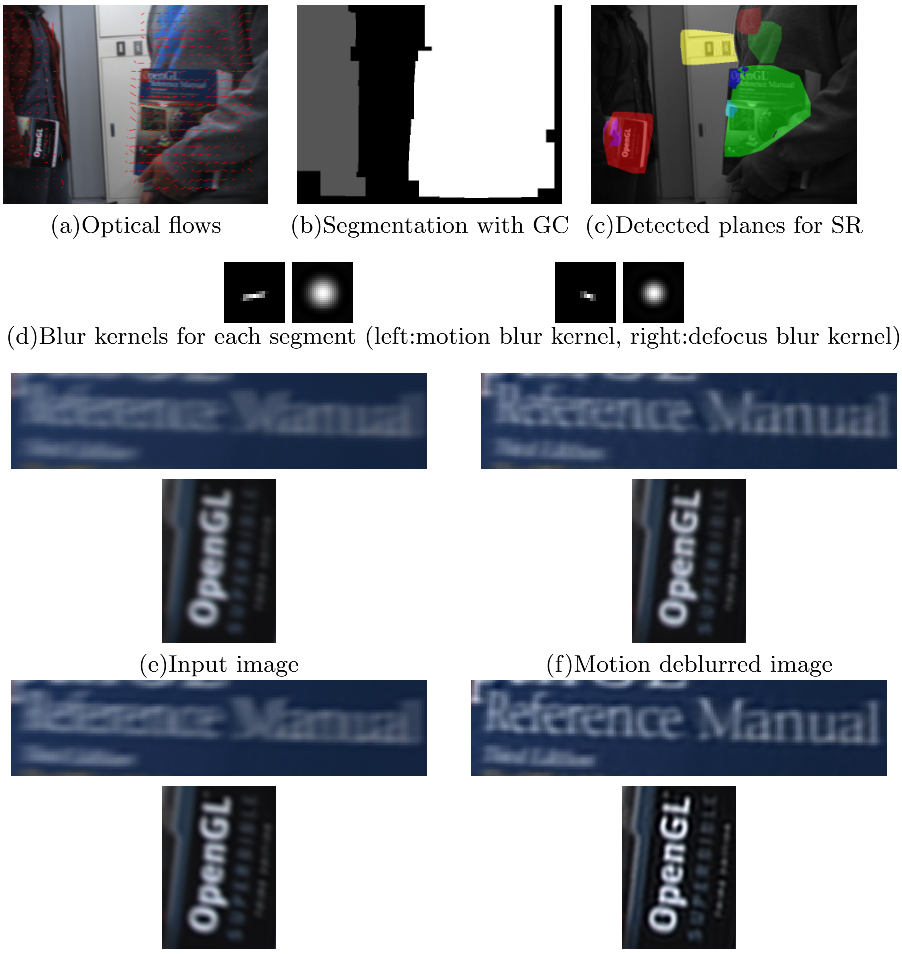
(a)Optical flows (b)Segmentation with GC (c)Detected planes for SR

(d)Blur kernels for each segment (left:motion blur kernel, right:defocus blur kernel)

(e)Input image (f)Motion deblurred image (g)SRI without motion deblurring (h)Proposed method

Fig. 7. Video data capture by handheld camera.

5. Irani, M.I., Peleg, S.: Improving resolution by image registration. CVGIP **53** (1991) 231–239
6. Levin, A., Weiss, Y., Durand, F., Freeman, W.: Understanding and evaluating blind deconvolution algorithms. In: CVPR. (2009) 1964–1971
7. Li, F., Yu, J., Chai, J.: A hybrid camera for motion deblurring and depth map super-resolution. In: CVPR. (2008)
8. Sroubek, F., Gabriel, C., Flusser, J.: A unified approach to superresolution and multichannel blind deconvolution. IEEE Trans. on Image Processing **16** (2007) 2322–2332
9. Fergus, R., Singh, B., Hertzmann, A., Roweis, S.T., Freeman, W.T.: Removing camera shake from a single photograph. In: SIGGRAPH. (2006) 787–794
10. Shan, Q., Jia, J., Agarwala, A.: High-quality motion deblurring from a single image. ACM Transactions on Graphics (SIGGRAPH) (2008)
11. Levin, A., Fergus, R., Durand, F., Freeman, W.: Image and depth from a conventional camera with a coded aperture. In: SIGGRAPH. (2007)
12. Levin, A., Sand, P., Cho, T.S., Durand, F., Freeman, W.T.: Motion-invariant photography. In: SIGGRAPH. (2008) 1–9
13. Cho, S., Matsushita, Y., Lee, S.: Removing non-uniform motion blur from images. In: ICCV. (2007) 1–8
14. Farsiu, S., Elad, M., Milanfar, P.: A practical approach to superresolution. In: Visual Communications and Image Processing. Volume 6077. (2006)
15. Katsaggelos, A.K., Molina, R., Mateos, J.: Super Resolution of Images and Video. Morgan & Claypool Publishers (2006)



(a)Optical flows (b)Segmentation with GC (c)Detected planes for SR

(d)Blur kernels for each segment (left:motion blur kernel, right:defocus blur kernel)

(e)Input image (f)Motion deblurred image

(g)SRI without motion deblurring (h)Proposed method

Fig. 8. Video data captured by static camera.

16. Park, S.C., Park, M.K., Kang, M.G.: Super-resolution image reconstruction: a technical overview. IEEE Signal Processing Magazine **20** (2003) 21–36
17. Tung, T., Nobuhara, S., Matsuyama, T.: Simultaneous super-resolution and 3d video using graph-cuts. In: CVPR. (2008)
18. Levin, A.: Blind motion deblurring using image statistics. In: NIPS. (2006)
19. Zucchelli, M., Santos-Victor, J., Christensen, H.: Multiple plane segmentation using optical flow. In: BMVC. (2002) 313–322
20. Dick, A., Torr, P., Cipolla, R.: Automatic 3d modelling of architecture. In: BMVC. (2000) 372–381
21. Bartoli, A.: Piecewise planar segmentation for automatic scene modeling. In: CVPR. (2001) 283–289