MOTION BLUR FOR MOTION SEGMENTATION

C. Paramanand A. N. Rajagopalan

Department of Electrical Engineering
Indian Institute of Technology Madras, Chennai, India

ABSTRACT

In this paper, we develop a method for motion segmentation using blur kernels. A blur kernel represents the apparent motion undergone by a scene point in the image plane. When the relative motion between the camera and scene is not restricted to fronto-parallel translations, the shape of the blur kernels can vary across image points. For a dynamic scene, we effectively model motion blur using transformation spread functions (TSFs) which represent the relative motions. Given a set of blur kernels that are estimated at different points across an image, we develop a method to segment them according to their relative motion. We initially group the blur kernels based on their 'compatibility'. We refine this initial segmentation by jointly estimating the TSF and removing the outliers.

Index Terms— Motion segmentation, motion blur, point spread function, dynamic scene

1. INTRODUCTION

In many applications, we commonly encounter dynamic scenes wherein the camera as well as the objects move independently. Separating the objects according to their motion and estimating the motion parameters is an important task for scene understanding. The standard approach to this problem is based on separating point correspondences obtained from multiple views or a sequence of images. This problem has been widely studied in the literature [1, 2, 3].

In this paper, our objective is to use motion blur as a cue for motion segmentation. While capturing an image of a dynamic scene, due to the averaging effect of light at the camera sensors, the occurrence of motion blur is quite common. Typically, in point correspondence-based methods, the apparent displacement of scene points due to relative motion is significantly large. Our technique can serve as an alternative to the standard approach for scenarios wherein the relative motion between the camera and the scene is comparatively much smaller. Consider the blurred image shown in Fig. 1 (a). The scene consisted of a static background and two moving objects. One of the object was moving horizontally while the other was moving vertically. Also, there was camera translation along the axial direction because of which even the background is blurred. Point spread functions (PSFs) that were

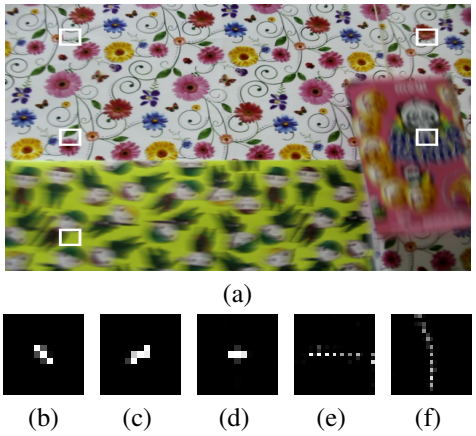


Fig. 1. (a) Blurred observation. (b)-(d): PSFs from the background. PSFs at the (e) yellow and (f) pink objects.

estimated at three points of the background (marked in Fig. 1 (a)) are shown in Figs. 1 (b)-(d). Even though the three points were under the influence of the same camera motion, we observe that the PSFs appear quite different. The blur kernels shown in Fig. 1 (e) and (f) correspond to the two moving objects and appear according to their relative motions. We seek to exploit the motion information present in blur kernels as a cue for segmentation. Specifically, we address the following tasks: (1) Given PSFs estimated at different point across the image, segment them according to their relative motion. (2) Estimate the parameters for each class of relative motion.

Related to our work, techniques exist that infer motion information from blurred images. The method in [4] uses a motion blurred observation in addition to two unblurred observations to aid the optical flow estimation. Caglioti and Guisti [5] propose a technique based on alpha matte to infer object motion during exposure. Methods exist that identify blur parameters in a motion blurred observation [6, 7, 8]. The method in [9] classifies regions in an image according to the effect of blurring. Favaro and Soatto [10] estimate the motion field, depth-map and the restored image from two motion blurred observations which are captured with different exposure times. It must be noted that, none of these methods consider the fact that the same relative motion (such as camera rotation) can result in different blur kernels at different points in the image. In [11], Dai and Wu propose a method

for determining the motion parameters from blur. Their blur model includes rotation and affine motion, which effectively represents non-uniform blur. They also extend their method to perform motion segmentation. However, this technique is applicable only when the alpha matte can be reliably estimated.

We propose a motion segmentation technique that uses blur kernels determined at different points of an image. We follow such an approach because PSFs can be estimated quite accurately and easily from small image patches. Although in this work, we use an unblurred-blurred image pair and estimate PSFs locally from image patches, our method is applicable even when blur kernels can be determined from a single image. In the experimental section, we show the applicability of our method on a single blurred observation. One can comfortably capture an unblurred-blurred image pair by suitably altering the exposure duration and ISO settings [12, 13]. With the use of a hybrid camera [14, 15], it is straightforward to obtain such a pair. An unblurred-blurred pair has also been used in [4] and [16].

Recent techniques [16, 17, 18] model a motion blurred image as a weighted average of transformed instances of the original image. We refer to the weights of the transformation as the transformation spread function (TSF) [17]. When there are camera rotations or axial motion, the PSF varies at every image point. In contrast, a single TSF represents blurring adequately. In the case of a dynamic scene, we model blurring by representing each relative motion with a TSF. We consider the relative motion to be composed of in-plane rotations and translations along all three directions. Since camera out-of-plane rotations can be approximated by translations [16, 18], our relative motion model represents blurring due to camera shake quite well.

A PSF depicts the motion undergone by a scene point and is analogous to a point trajectory. We propose a criterion to deduce whether two image points underwent the same set of transformations by comparing their PSFs. Based on this criterion, we obtain an initial segmentation on the given set of blur kernels. Each segment corresponds to a class of relative motion. Although the segmentation of blur kernels obtained using our constraint is quite accurate, there is a small possibility of the presence of outliers. When the number of blur kernels available is sufficiently large, we propose to simultaneously remove the outliers and estimate the TSF using a RANSAC-based approach.

2. IMAGE MODEL

Consider a dynamic scene wherein both the camera and scene objects are free to move independently. With the camera center as the reference, let the number of distinct relative motions be N_m. Let f denote the image of a scene captured without any relative motion between the scene points and the camera. Let f^τ denote the intensities of f within the region corresponding to the τth motion ($\tau = 1 \ldots N_m$) of the scene. Due to the dynamic nature of the scene, the blurred observation of the scene g, contains different regions blurred in different ways. Let g^τ denote the intensities in the region of the blurred image corresponding to τth motion. If the relative motion is not restricted to fronto-parallel translations, each region g^τ will be a space-variantly blurred instance of f^τ. To represent this space-variant blur, we separately model each of the N_m regions using the notion of transformation spread function (TSF) [16, 17, 18]. g^τ is related to f^τ as

$$g^\tau(\mathbf{x}) = \int_{\lambda \in \mathbf{T}} \omega^\tau(\lambda)\, f^\tau(H_\lambda(\mathbf{x}))\, d\lambda \tag{1}$$

where $\omega^\tau : \mathbf{T} \rightarrow \mathbb{R}_+$, is called the *transformation spread function* (TSF) and maps the set of all possible transformations $\mathbf{T}$ to non-negative real numbers. $H_\lambda(\mathbf{x})$ denotes the image coordinates when a homography H_λ is applied on the point $\mathbf{x}$. The TSF ω^τ depicts the τth *relative* motion between the camera and the scene. i.e., for each transformation $\lambda \in \mathbf{T}$, the value of the TSF $\omega^\tau(\lambda)$ denotes the fraction of the total exposure duration for which the relative motion caused a transformation H_λ on the image coordinates (within f^τ).

The blurred segment g^τ can also modeled with a space-variant PSF h^τ as

$$g^\tau(\mathbf{x}) = f^\tau *_v h^\tau(\mathbf{x}) = \int f^\tau(\mathbf{x} - \mathbf{u})\, h^\tau(\mathbf{x} - \mathbf{u}, \mathbf{u})\, d\mathbf{u} \tag{2}$$

where $h^\tau(\mathbf{x}, \mathbf{u})$ denotes the blur kernel at the image point $\mathbf{x}$ as a function of the independent variable $\mathbf{u}$. The PSF represents the displacements undergone by a point light source at $\mathbf{x}$ and can be written [17] in terms of the TSF as

$$h^\tau(\mathbf{x}, \mathbf{u}) = \int_{\lambda \in \mathbf{T}} \omega^\tau(\lambda)\, \delta\left(\mathbf{u} - \left(H_\lambda^{-1}(\mathbf{x}) - \mathbf{x}\right)\right) d\lambda \tag{3}$$

where δ indicates the 2D Dirac Delta function.

Following our discussion in section 1, We consider the relative motion to be composed of in-plane rotations and 3D translations. Consequently, the homography H_λ will be composed of 2D translations, in-plane rotations and scaling (similarity transformation). For practical implementation, we discretize our model. Let $h^\tau(i, j, ;)$ denote the discrete blur kernel at a pixel $\mathbf{p} = (i, j)$. The set $\mathbf{T}$ is sampled to get N_T different transformations (as done in [16, 18]). For $l = 1 \ldots N_T$, let $\omega^\tau(l)$ denote the weight corresponding to the lth transformation of $\mathbf{T}$ and (i_l, j_l) denote the co-ordinates of the point when a transformation H_l^{-1} is applied on $\mathbf{p}$. The discrete form of the relationship in Eqn. (3) is given by

$$h^\tau(i, j; s, t) = \sum_{l=1}^{N_T} \omega^\tau(l)\, \delta_d(s - (i_l - i), t - (j_l - j)) \tag{4}$$

where δ_d denotes the 2D Kronecker delta function. When (i_l, j_l) take non-integer values, we assign values to the pixels neighboring (i_l, j_l) by bilinear interpolation principle.

3. BLUR-BASED SEGMENTATION

The first step of our algorithm is to obtain PSFs at different locations in the image. Given an unblurred-blurred image pair, we initially estimate N_p PSFs (denoted by $h_1 \ldots h_{N_p}$) at different points ($\mathbf{p}_1 \ldots \mathbf{p}_{N_p}$) across the image. Within small image patches of f and g, we assume blur to be space-invariant. We solve for the PSFs by using a least-squares approach with a sparsity prior. The task is to segment the N_p estimated blur kernels according to their respective relative motions. We initially propose a blur compatibility criterion to compare two blur kernels. Using this condition, we obtain an initial segmentation of PSFs. For each set, we estimate a TSF from the blur kernels. In scenarios where the number of blur kernels is significantly large, we can remove the outliers using a RANSAC-based approach [19].

3.1. Blur compatibility criterion

Consider a blur kernel h_a at a point $\mathbf{p}_a = (i, j)$. We refer to a set of transformations $\mathbf{\Lambda}_a$ as the support of h_a. $\mathbf{\Lambda}_a$ is given by $\mathbf{\Lambda}_a = \{\lambda : h_a(i, j; i_\lambda - i, j_\lambda - j) > 0\}$, where (i_λ, j_λ) denotes the co-ordinates of the point when a transformation H_λ^{-1} is applied on $\mathbf{p}_a$. $\mathbf{\Lambda}_a$ contains all possible transformations from $\mathbf{T}$ that shift the point $\mathbf{p}_a$ to a position at which the blur kernel h_a has a positive entry. Since many transformations can induce the same shift at a point, $\mathbf{\Lambda}_a$ will contain many transformations including those present in the true TSF.

Given two blur kernels, h_a and h_b corresponding to locations $\mathbf{p}_a$ and $\mathbf{p}_b$, respectively, we determine the supports of the blur kernels $\mathbf{\Lambda}_a$ and $\mathbf{\Lambda}_b$. If $\mathbf{p}_a$ and $\mathbf{p}_b$ underwent the same relative motion, the blur kernels h_a and h_b would be related to a single TSF, and there would be a common set of transformations between $\mathbf{\Lambda}_a$ and $\mathbf{\Lambda}_b$ that would include the support of the TSF. We evaluate the intersection of the two supports $\mathbf{\Lambda}_{ab} = \mathbf{\Lambda}_a \cap \mathbf{\Lambda}_b$. We apply the transformations in $\mathbf{\Lambda}_{ab}$ on $\mathbf{p}_a$ and $\mathbf{p}_b$ to get two sets of displacements $\hat{h_a}$ and $\hat{h_b}$, respectively. We define the blur compatibility criterion as follows: the locations of positive entries of h_a and h_b should be included within $\hat{h_a}$ and $\hat{h_b}$, respectively. If h_a (or h_b) has positive entries at locations other than those in $\hat{h_a}$ (or $\hat{h_b}$), then we can conclude that there are no common transformations between $\mathbf{\Lambda}_a$ and $\mathbf{\Lambda}_b$ that can generate both the blur kernels. This would imply that $\mathbf{p}_a$ and $\mathbf{p}_b$ underwent different motions. It must be noted that, there is a rare possibility that even though two blur kernels satisfy the blur compatibility criterion, they could be due to different relative motions. However, if two points undergo the same relative motion, the blur kernels at the points must satisfy the compatibility criterion.

3.2. Segmenting blur kernels

Out of the N_p blur kernels, we randomly select a blur kernel and test it for compatibility with the remaining $N_p - 1$ blur kernels. A set Ω_1 containing blur kernels that are compatible with the randomly selected PSF is obtained. From the remaining blur kernels (which do not belong to Ω_1), we select another PSF and check for its compatibility with all the other $N_p - 1$ kernels. Repeating the process, we obtain sets of blur kernels denoted by $\Omega_1 \ldots \Omega_{N_m}$. Note that there is a possibility that a blur kernel can belong to more than one of the sets (as mentioned in section 3.1).

For a set of blur kernels corresponding to the same relative motion, we estimate its TSF (which characterizes the relative motion) from the PSFs. In Eqn. (4), we see that each component of a blur kernel $h^\tau(i, j; m, n)$, is a weighted sum of the components of the TSF ω^τ. Consequently, a blur kernel h_a^τ (at a point $\mathbf{p}_a$) can be expressed as $h_a^\tau = M_a \omega^\tau$, where M_a is a matrix whose entries are determined by the location of the blur kernel and the bilinear interpolation coefficients. Suppose that we have N_h PSFs corresponding to the τth motion. By stacking all of these as a vector $\overline{h}$, and suitably concatenating the matrices M_a for $a = 1 \ldots N_h$, we get

$$\overline{h} = M\omega^\tau \tag{5}$$

A TSF is typically sparse due to the fact that during image capture, very few transformations would have occurred out of all possible transformations. From the blur kernels $\overline{h}$, we estimate a sparse TSF by minimizing the following cost.

$$\underset{\omega^\tau}{\operatorname{argmin}} \left\{ \left\| \overline{h} - M\omega^\tau \right\|_2^2 + \lambda_s \left\| \omega^\tau \right\|_1 \right\} \tag{6}$$

3.3. Removal of outliers

For a given set of blur kernels, we remove the outliers using a RANSAC-based approach [19]. In each iteration of RANSAC, we estimate the TSF from four randomly chosen blur kernels by solving Eqn. (6). To test the fitted model and form the consensus set (which consists of the likely inliers), for each blur kernel of the set h_a, we generate a blur kernel $\hat{h}_a$ from the estimated TSF (using Eqn. 4) and compare the two kernels. Following [20], we evaluate the correlation between h_a and $\hat{h}_a$ as the similarity measure. If h_a and $\hat{h}_a$ are close to each other, the blur kernel h_a is included in the consensus set. We stop the iterations when the consensus set becomes larger than a fraction of the total number of blur kernels in the set.

4. EXPERIMENTAL RESULTS

In our experiments, to obtain an unblurred-blurred image pair of a scene, we capture a sequence of images. Following [10], we consider one of the frames as the unblurred reference image, and average a few frames to obtain a blurred observation.

For the scene shown in Fig. 1, we estimated forty blur kernels at regular intervals across the image. Using our kernel compatibility criterion, we obtained three different sets of blur kernels. In Fig. 2, the locations of the blur kernels of the

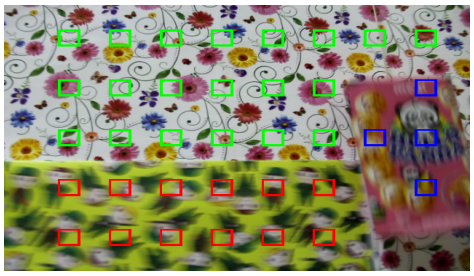
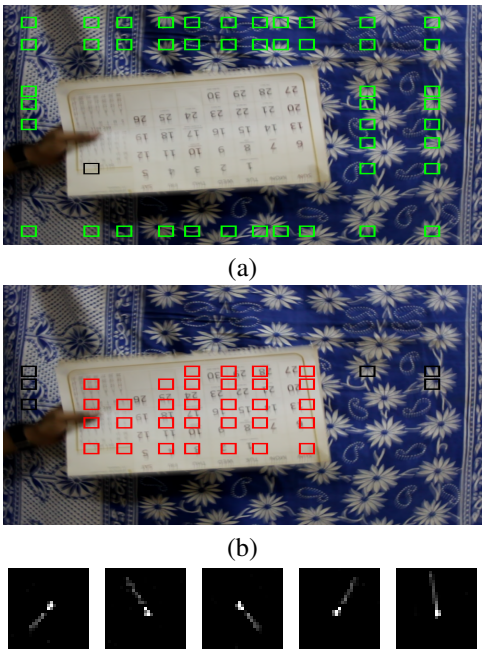


Fig. 2. Segmentation of blur kernels.

three sets are marked in different colours. Although the PSFs shown in Fig. 1 (b)-(d) vary in shape, since they were part of the same relative motion, our method correctly grouped them into the same segment. Some of the blur kernels which were estimated at the boundary of two objects were in error and did not belong to any of the three sets (since the criterion was not satisfied). In this experiment, we did not apply RANSAC-based outlier removal because the number of kernels in two of the segments was small. Also, the result in Fig. 2 shows that there was no need for the removal of outliers.

In our next experiment, the scene (shown in Fig. 3) consisted of an object which was vertically moving against a static background. During image capture, the camera underwent both in-plane rotation and translations resulting in space-variant blur. This can be seen from the first four blur kernels in the third row of Fig. 3, which were estimated at the four corners of the background. The fifth PSF in the third row of Fig. 3 corresponds to the vertically moving object. We estimated blur kernels at 88 different image locations and segmented them using our compatibility criterion. Although the segmentation was largely correct, there were some outliers. The final segmentation was obtained after removing the outliers. Fig. 3 (a) shows the locations of blur kernels of one of the sets marked in green and the location of an outlier (which was correctly detected) marked in black. In Fig. 3 (b), we see the locations of the other set of blur kernels marked in red and its outliers in black.

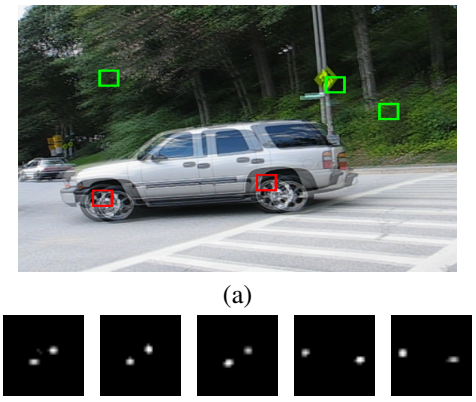
In our next experiment, the blurred observation shown in Fig. 4 (a) was obtained by averaging frames of one of the image sequences from [3]. To demonstrate our method on a single blurred observation, in this experiment we did not use an unblurred image. Five blur kernels were estimated from the blurred patches (at locations marked in Fig. 4 (a)) using the technique in [21]. The first three PSFs of the second row of Fig. 4 are from the background while remaining two are from the foreground. Based on the blur compatibility criterion, the five PSFs were correctly segmented as shown Fig. 4 (a). Although, in this experiment, we used only one observation, user assistance was needed to select patches that led to accurate PSF estimation.



(a)

(b)

Fig. 3. (a) and (b) positions of segmented blur kernels and the outliers marked on the blurred observation.



(a)

Fig. 4. (a) Blurred image. Second row: first three PSFs are from the backround and the next two are from the foreground.

5. CONCLUSIONS

We developed a technique for motion segmentation through blur kernels. Although the shape of PSFs vary when the relative motion is not confined to in-plane translations, our blur compatibility criterion allows for correct segmentation of the PSFs. We estimate the TSF for each relative motion by exploiting its relationship with the PSFs and remove the outliers. A possible extension would be to develop a single image-based method by using colour information for segmentation. We would like to incorporate a more general model for relative motion in future. Our approach can act as a stepping stone towards deblurring of dynamic scenes which is quite challenging.

6. REFERENCES

[1] J.P. Costeira and T. Kanade, "A multibody factorization method for independently moving objects," *International Journal of Computer Vision*, vol. 29, no. 3, pp. 159–179, 1998.

[2] K. Kanatani, "Motion segmentation by subspace separation and model selection," in *Proc. ICCV*, 2001.

[3] R. Tron and R. Vidal, "A benchmark for the comparison of 3-d motion segmentation algorithms," in *Proc. CVPR*, 2007.

[4] A. Sellent, M. Eisemann, B. Goldlucke, D. Cremers, and M. Magnor, "Motion field estimation from alternate exposure images," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 33, no. 8, pp. 1577–1589, 2011.

[5] V. Caglioti and A Giusti, "On the apparent transparency of a motion blurred object," *International Journal of Computer Vision*, vol. 86, no. 2, pp. 243–255, 2008.

[6] A. Chakrabarti, T. Zickler, and W. T. Freeman, "Analyzing spatially-varying blur," in *Proc. CVPR*, 2010.

[7] A. Levin, "Blind motion deblurring using image statistics," in *Proc. NIPS*, 2006.

[8] P. Trouve, F. Champagnat, G. Le Besnerais, and J. Idier, "Single image local blur identification," in *Proc. ICIP*, 2011.

[9] R. Liu, Z. Li, and J. Jia, "Image partial blur detection and classification," in *Proc. CVPR*, 2008.

[10] P. Favaro and S. Soatto, "A variational approach to scene reconstruction and image segmentation from motion-blur cues," in *Proc. CVPR*, 2004.

[11] S. Dai and Y. Wu, "Motion from blur," in *Proc. CVPR*, 2008.

[12] M. Sorel and F. Sroubek, "Space-variant deblurring using one blurred and one underexposed image," in *Proc. ICIP*, 2009.

[13] L. Yuan, J. Sun, L. Quan, and H. Y. Shum, "Image deblurring with blurred/noisy image pairs," in *Proc. SIGGRAPH*, 2007.

[14] M. Ben-Ezra and S. Nayar, "Motion-based motion deblurring," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 26, no. 6, pp. 689–698, 2004.

[15] Y. Tai, H. Du, M. S. Brown, and S. Lin, "Correction of spatially varying image and video blur using a hybrid camera," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 32, no. 6, pp. 1012–1028, 2010.

[16] O. Whyte, J. Sivic, A. Zisserman, and J. Ponce, "Non-uniform deblurring for shaken images," *Intl. Jrnl. Comp. Vis.*, 2011.

[17] C. Paramanand and A. N. Rajagopalan, "Inferring image transformation and structure from motion-blurred images," in *Proc. BMVC*, 2010.

[18] A. Gupta, N. Joshi, L. Zitnick, M. Cohen, and B. Curless, "Single image deblurring using motion density functions," in *Proc. ECCV*, 2010.

[19] M. A. Fischler and R. C. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," *Commun. ACM*, vol. 24, no. 6, pp. 381–395, 1981.

[20] Z. Hu and M. Yang, "Good regions to deblur," in *Proc. ECCV*, 2012.

[21] L. Xu and J. Jia, "Two-phase kernel estimation for robust motion deblurring," in *Proc. ECCV*, 2010.