Inferring Image Transformation and Structure from Motion-Blurred Images

Paramanand Chandramouli
paramanand@gmail.com

A. N. Rajagopalan
raju@ee.iitm.ac.in

Image Processing and Computer Vision Lab
Department of Electrical Engineering
Indian Institute of Technology Madras
Chennai, India

While capturing images of a scene, the relative motion between the camera and the scene during exposure leads to motion-blur in images. The convolution model for motion-blur is applicable only when the camera motion is restricted to in-plane translations. Blur arising due to rotation and out-of-plane translation of camera cannot be modeled using convolution with a single blur kernel [3, 5, 6]. In this paper, we model the motion-blurred image as the weighted average of geometrically transformed versions of the reference image. The model implicitly accounts for the space-variant nature of blurring that occurs due to unrestricted camera motion. In many applications, motion-blur can occur while capturing scenes with depth variations from a moving camera. During the camera motion, the apparent movement of the scene points in the image is related to the shape of the scene [1]. Consequently the extent of blurring at a point is governed by both the camera motion and the scene structure.

Initially, we consider scenes having constant depth and develop an algorithm to estimate the transformations undergone by the reference image during exposure. We next consider images of scenes with depth variations. Based on the estimated transformations, we relate the depth at a scene point to the blurred image intensity through the point spread function. Depth estimation is posed as a state estimation problem and is solved using an unscented Kalman filter (UKF).

A motion-blurred image g can be related to the reference image f through the space-variant point spread function (PSF) h as

$$g(x,y) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x-s,y-t)\,h(x-s,y-t;s,t)\,ds\,dt \qquad (1)$$

For the case of in-plane camera translations, the PSF remains constant at all the image points. However, when the camera motion is not restricted to in-plane translations, the PSF can vary at every image point. We model the motion-blurred image in terms of the reference image using a function which we call as *transformation spread function* (TSF).

Let $\mathbf{T}$ denote the set of possible geometric transformations the image points can undergo during the exposure. We define the transformation spread function $h_T : \mathbf{T} \rightarrow \Re_+$ as a mapping from the set of transformations to the set of nonnegative real numbers. $h_T(T_\lambda)$ denotes the fraction of the exposure duration for the image transformation T_λ. The blurred image is modeled as the weighted sum of the transformed reference images.

$$g = \sum_{T_\lambda \in \mathbf{T}} h_T(T_\lambda)\, f_{T_\lambda} \qquad (2)$$

where f_{T_λ} denotes the reference image warped by the transformation T_λ. Let $h(i,j,;)$ denote the discrete PSF at the image point (i,j). Let (i_λ, j_λ) denote the co-ordinates of the point when a transformation T_λ is applied on (i,j). We can obtain the PSF at each pixel (i,j) from TSF (which is common for all the image points) as

$$h(i,j;m,n) = \sum_{T_\lambda \in \mathbf{T}} h_T(T_\lambda)\,\delta_d\left(m-(i-i_\lambda),\, n-(j-j_\lambda)\right) \qquad (3)$$

where δ_d denotes the 2D Kronecker delta function.

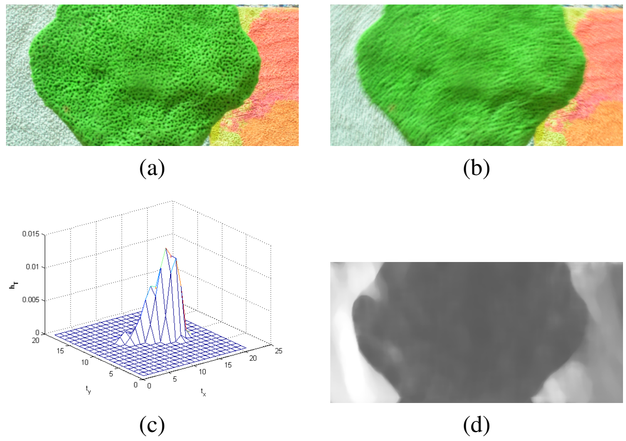
In scenes with depth variations, the extent of blurring at a point depends both on the scene structure and the camera motion. Blur can serve as a depth cue when there is a translational component in the camera motion. Suppose we know the reference TSF h_{T_o} (determined at depth d_o), the PSF at every point can be related to its depth value. The blurred image g in terms of the reference image and the space-variant PSF is given by

$$g(i,j) = \sum_{m,n} f(i-m, j-n)\, h(i-m, j-n; m,n) \qquad (4)$$

Initially, we determine the TSF h_{T_o} from a patch f_o of the reference image f and the corresponding patch g_o of g using the proposed TSF estimation technique. We consider it as the reference TSF and assume that its depth is d_o. We estimate the relative depth $k(i,j) = d(i,j)/d_o$ at every (i,j) with the knowledge of f, g and h_{T_o} by posing it as a recursive state estimation problem [4]. The measurement model is

$$\mathbf{g}(i,j) = \mathbf{H}_{i,j}(k(i,j)) + e(i,j) \qquad (5)$$

where $\mathbf{g}(i,j)$ denotes blurred image pixels corresponding to all three channels and the nonlinear operator $\mathbf{H}_{i,j}$ is in terms of f and h_{T_o}. At each pixel (i,j), the state mean and covariance are predicted through the system model. From these, the observation moments are obtained through unscented transformations [2]. The Bayesian estimate of the state mean is updated based on the observation through the UKF. The updated mean is regarded as the relative depth $k(i,j)$.



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

Figure 1: (a) Reference image. (b) Blurred observation. (c) Estimated TSF at $t_\theta = 3^\circ$ and $t_s = 1$ (d) Estimated depth map.

The reference image of the scene consisting of three objects at different depths is shown in Fig. 1 (a). In Fig. 1 (d), we observe that the objects that are near the camera are correctly assigned a lesser relative depth value than those that are farther.

In this paper, we have developed a technique to determine the image transformations and extract depth information from motion-blurred images. The camera motion was restricted to in-plane rotations and translations. However, the framework can be extended to allow out-of-plane camera rotations.

[1] D. J. Heeger and A. D. Jepson. Subspace methods for recovering rigid motion. *International Journal of Computer Vision*, 7(2):95–117, 1992.

[2] S. Julier and J. Uhlmann. A new extension of the Kalman filter to nonlinear systems. *The 11th International Symposium on Aerospace/Defense Sensing, Simulation and Controls*, pages 182–193, 1997.

[3] C. Mei and I. Reid. Modeling and generating complex motion blur for real-time tracking. *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8, 2008.

[4] C. Paramanand and A. N. Rajagopalan. Unscented transformation for depth from motion-blur in videos. *IEEE Workshop on Three Dimensional Information Extraction for Video Analysis and Mining in conjunction with CVPR 2010.*

[5] Y. Tai, H. Du, M. Brown, and S. Lin. Image/video deblurring using a hybrid camera. *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8, 2008.

[6] O. Whyte, J. Sivic, A. Zisserman, and J. Ponce. Non-uniform deblurring for shaken images. *IEEE Conference on Computer Vision and Pattern Recognition*, 2010.