2014 22nd International Conference on Pattern Recognition

Inferring Plane Orientation from a Single Motion Blurred Image

M.Purnachandra Rao, A.N. Rajagopalan
Department of Electrical Engineering
Indian Institute of Technology Madras, Chennai, India
Email: mpurna2u@gmail.com, raju@ee.iitm.ac.in

Guna Seetharaman
Information Directorate
AFRL/RIEA, Rome NY, USA
Email: guna@ieee.org

***Abstract*—We present a scheme for recovering the orientation of a planar scene from a single translationally-motion blurred image. By leveraging the homography relationship among image coordinates of 3D points lying on a plane, and by exploiting natural correspondences among the extremities of the blur kernels derived from the motion blurred observation, the proposed method can accurately infer the normal of the planar surface. We validate our approach on synthetic as well as real planar scenes.**

I. Introduction

An extensively researched area in computer vision is the recovery of 3D structure from image intensities. Well-known cues for depth recovery include disparity [1], optical flow [2], texture [3], [4], [5], shading [6], [7], defocus blur [8] and motion blur [9], [10], [11], [12], to name a few. While estimation of 3D depth/shape has been of general interest, there have also been works targeting the special case of inferring planar 3D geometry (such as the Manhattan model). This is due to the fact that the world around us can, in many cases, be modeled as being piecewise planar. Approximating a 3D scene with planes (where possible) has tremendous advantage in terms of reducing the computational complexity. Estimation of surface normals of a scene/object plays a crucial role in identifying the 3D geometry/shape of that scene/object. The elegant homography relationship between two images (original and transformed due to relative motion between camera and scene) holds for scene points lying on a plane in the 3D world.

Estimating a plane involves finding its surface normal and the perpendicular distance from the center of the camera to the plane. The relevance of this problem is evident from the many works that exist in the literature. Brown et al. [5] assume textural isotropy and use the foreshortening of texture due to an inclined plane as a cue to estimate the orientation of the plane. Super et al. [3], [4] assume a homogeneously textured input image and use the local variations of spatial frequencies to compute the orientaion of the planar surface. Clark et al. [14] implemented a technique to recover the orientation of text planes using perspective geometry. In [13], Farid et al. reveal the fact that the projection of a planar texture having random phase leads to higher-order correlations in the frequency domain, and these correlations are proportional to the orientaion of the plane. Greinera et al. [15] have proposed a method to determine the surface normal using projective geometry and spectral analysis. Hwa et al. [1] discuss a method to estimate depth information using disparity obtained from stereo images. Behzad et al. [2] have developed a technique to estimate depth information from optical flow. Haines et al. [16] describe a technique that makes use of prior training data gathered in an urban environment to classify planar/non-planar surfaces and to compute the orientaion of the planes.

We propose to use motion blur as a *cue* to estimate the *orientation* of a planar scene given a *single* motion blurred image of the plane. Fig. 1 depicts an example in which the surface normal of the plane is rotated 30 degrees with respect to the optical axis. Note that the degree of blur in the left-half of the image is higher as compared to the right-half of the plane. Usually, blurring is considered as a nuisance whose effect needs to be removed. However, works do exist that, in fact, use blur (optical/motion) as a cue to infer valuable information such as the depth of the scene and relative motion of the camera with respect to the scene. In [9], an unblurred-blurred image pair was used to estimate depth from a translationally blurred image. Sorel et al. [17] derive the relationship among point spread functions (PSFs) at different depths and use it to restore space-variantly blurred images. Xu et al. in [18] used a motion blurred stereo-pair to estimate depth information which is subsequently used for space-variant deblurring. Lin et al [10] describe a method to estimate depth from a single motion blurred image. They assume the scene to consist of a single fronto-parallel plane and calculate the distance of the plane from the camera. Zheng et al. [12] have recently proposed a method to estimate depth from a single motion blurred image but it is designed specifically for low-light conditions.

Fig. 1. A motion blurred inclined plane.

To the best of our knowledge, the only method to estimate plane orientation using blur as a cue is the recent work by McCloskey et al. [19] who have proposed a method based on blur gradients to evaluate the planar orientaion (slant and tilt angles) from a single image using *optical blur* as a cue. They exploit the relationship between blur variations for the equifocal (fronto-parallel scene) plane and a plane's tilt and

1051-4651/14 $31.00 © 2014 IEEE
DOI 10.1109/ICPR.2014.364

slant angles. For a fronto-parallel scene, all the pixels in the image have the same amount of blur. In the case of an inclined plane, the amount of blur varies inversely with depth. The user has to manually mark a patch of interest for which slant and tilt angles are estimated. Their work assumes a homogeneously textured observation.

In this paper, we propose an interesting approach (a first of its kind) to determine the surface normal of a plane from a *single motion-blurred* image. We exploit the homography relation that exists in the image domain under camera motion to determine the surface normal. For a planar scene, the blurred image can be represented as a weighted average of warped versions of the unblurred image. This representation helps in characterizing the space-variant blur by a set of global homographies. We extract patches from the image and estimate blur kernels at these patches. Using the correspondences among the extremities of blur kernels at different locations, we set up a system of linear equations that are solved to yield the surface normal.

The paper is organized as follows. In section II, we describe motion blur model for planar scenes. In section III, we discuss the proposed method to estimate the surface normal from point correspondences across blur kernels. In section IV, we validate our approach on several synthetic and real examples, and conclude with section V.

II. Planar motion blur

Motion blur in an image is due to relative motion between camera and scene during exposure time. Since the camera sensor sees different scene points at different instants of time within the exposure window, these intensities get averaged resulting in a blurred image. Let g be the blurred image captured by a camera with exposure time E_t, and let f be the original image (without camera shake). During the exposure time, f may have undergone a set of transformations due to relative motion between the camera and the scene. The transformed image at time instant τ can be explained using homography H_τ as $g_\tau(H_\tau(\mathbf{x})) = f(\mathbf{x})$ where $\mathbf{x}$ represents pixel coordinates. Therefore, the blurred image can be modeled as the average of transformed versions of f during the exposure time E_t. The blurred image intensity at a location $\mathbf{x}$ can then be expressed as

$$g(\mathbf{x}) = \frac{1}{E_t}\int_0^{E_t} f(H_\tau^{-1}(\mathbf{x}))\; d\tau$$

The homography relation in the image domain holds only for the set of scene points lying on a plane. The homography at time instant τ is given by $H_\tau = K\left(R_\tau + \mathbf{t}_\tau \frac{\mathbf{n}^T}{d}\right)K^{-1}$ where $K = \begin{bmatrix} q & 0 & 0 \\ 0 & q & 0 \\ 0 & 0 & 1 \end{bmatrix}$, with q being the focal length of the camera. Here R_τ denotes the rotation matrix at time instant τ and is a combination of the rotational matrices about the X, Y and Z axes, and d is the perpendicular distance from the center of the camera to the plane and is a constant for the entire plane. Here, $\mathbf{t}_\tau = [T_{X_\tau} T_{Y_\tau} T_{Z_\tau}]^T$ represents the 3D translation vector at time τ and $\mathbf{n} = [N_X\ N_Y\ N_Z]^T$ denotes the surface normal of the planar scene. Following [11], [17], in our analysis, we assume that the motion blur is due to camera translations only. Therefore, R_τ is a 3×3 identity matrix (I) and the homography simplifies to

$$H_\tau = K\left(I + \mathbf{t}_\tau \frac{\mathbf{n}^T}{d}\right)K^{-1}. \tag{1}$$

III. Normal from point-correspondences

The aim of our work is to use a single motion blurred image to estimate the surface normal of a planar scene. It is straightforward to show that the blur kernel centered at location $\mathbf{x}$ can be written as

$$h(\mathbf{x}, \mathbf{u}) = \frac{1}{E_t}\int_0^{E_t} \delta(\mathbf{u} - (H_\tau(\mathbf{x}) - \mathbf{x}))\; d\tau \tag{2}$$

i.e., the PSF represents the displacements undergone by an image point due to a set of motion transformations. The blur kernel induced will ideally consist of impulses at the corresponding shifts, and the weight of the impulse will be governed by the fraction of the exposure time spent in that homography/pose. For a fronto-parallel scene i.e., when $\mathbf{n} = [0\ 0\ 1]^T$, the blur induced would be space-invariant when camera undergoes only in-plane translations in the $xy-$plane. This is because for some transformation $\mathbf{t} = [T_{X_\tau}\ T_{Y_\tau}\ 0]^T$ and $\mathbf{n} = [0\ 0\ 1]^T$, we obtain

$$\begin{bmatrix} x_\tau \\ y_\tau \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & \frac{qT_{X_\tau}}{d} \\ 0 & 1 & \frac{qT_{Y_\tau}}{d} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \tag{3}$$

Clearly, the displacements in x and y directions are a constant (independent of the spatial location) and equal $\frac{qT_{X_\tau}}{d}$ and $\frac{qT_{Y_\tau}}{d}$, respectively. However, for a general inclined plane, the blur induced would be space-variant (due to change in depth of the scene) even for pure in-plane translational motion. Corresponding to this situation, we will have (for $\mathbf{t} = [T_{X_\tau}\ T_{Y_\tau}\ 0]^T$)

$$\begin{bmatrix} x_\tau \\ y_\tau \\ 1 \end{bmatrix} = \begin{bmatrix} 1 + N_X\frac{T_{X_\tau}}{d} & N_Y\frac{T_{X_\tau}}{d} & qN_Z\frac{T_{X_\tau}}{d} \\ N_Y\frac{T_{Y_\tau}}{d} & 1 + N_Y\frac{T_{Y_\tau}}{d} & qN_Z\frac{T_{Y_\tau}}{d} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \tag{4}$$

Note that the displacements along x and y are no longer a constant and, in fact, vary as a function of the spatial location of the image point. Since our interest is in estimating the surface normal $\mathbf{n} = [N_X\ N_Y\ N_Z]^T$ (and not the camera motion per se), we rewrite equation (4) as

$$\begin{bmatrix} x_\tau \end{bmatrix} = \begin{bmatrix} x & y & 1 \end{bmatrix} \begin{bmatrix} 1 + N_X\frac{T_{X_\tau}}{d} \\ N_Y\frac{T_{X_\tau}}{d} \\ qN_Z\frac{T_{X_\tau}}{d} \end{bmatrix} \tag{5}$$

and

$$\begin{bmatrix} y_\tau \end{bmatrix} = \begin{bmatrix} x & y & 1 \end{bmatrix} \begin{bmatrix} N_X\frac{T_{Y_\tau}}{d} \\ 1 + N_Y\frac{T_{Y_\tau}}{d} \\ qN_Z\frac{T_{Y_\tau}}{d} \end{bmatrix} \tag{6}$$

In equations (5) and (6), assuming point correspondences between $(x,\ y)$ and $(x_\tau,\ y_\tau)$ to be known, the unknowns are $N_X,\ N_Y,\ N_Z,\ T_{X_\tau},\ T_{Y_\tau}$ and d, and these appear in the right-most column vector. Note that the ratio $\frac{T_{X_\tau}}{d}$ (or $\frac{T_{Y_\tau}}{d}$)

is a common scale factor multiplying the normal and hence need not be estimated. At first glance, it might appear that one can enforce unit norm on the normal to reduce the unknowns by one. However, we refrain from doing so since we lose the elegance of the linear equations (5) and (6) in the process. Thus, there are effectively three unknowns $(N_X,\ N_Y,\ N_Z)$ that are to be estimated. Hence, we need at least three point correspondences to solve this problem. If we can find point displacements at other locations in the image corresponding to the *same motion* $[T_{X_\tau}\ T_{Y_\tau}\ 0]^T$, then it should theoretically be possible to determine the unknowns. This, in fact, forms the basic premise for our method.

As discussed earlier in equation (2), the PSF or blur kernel encapsulates the displacements of pixels under the influence of camera motion. Thus, if we can establish point-correspondences (all influenced by the same motion) across atleast three blur kernels, then we can solve for the surface normal. However, because the blur kernel estimation itself is prone to small errors, it is only prudent that we use as many correspondences as possible. Note that we need to identify corresponding points among the PSFs with respect to the same *homography*. On this issue, we wish to point out an interesting fact that a natural correspondence exists among the extremities of blur kernels (i.e., non-zero impulses at maximum distance from the origin of the PSF and on either side of the origin) across the image. We could potentially use the left (or right) extremity of the blur kernel in equation (5) or (6). Although it might appear that one can then solve for the normal, there is an ambiguity issue which we wish to highlight. Since the blur kernels are estimated independently across the image, there is a possibility of incurring spatial shifts in the PSFs when employing any blind deblurring method. A blurred patch b can be represented as convolution of latent patch l and blur kernel h i.e., $b = l * h$. Note that a shifted version (translational shift along x and y directions) of the true h also satisfies the convolution relation because $b(\mathbf{x}) = l(\mathbf{x} - \mathbf{s_0}) * h(\mathbf{x} + \mathbf{s_0})$. The shift introduced in the blur kernel is equivalently compensated in the latent image. Hence, if we choose only one extremity from the blur kernels, the surface normal cannot be estimated correctly due to possible misalignment errors. In order to resolve this issue, we choose the displacement *between the extremities* for computing correspondenes since this displacement is independent of any shift in the blur kernel.

From equation (5), the extremity of a PSF (say h_1) due to translation (say T_{X_p}) can be expressed as

$$\begin{bmatrix} x_{l_1} \end{bmatrix} = \begin{bmatrix} x_1 & y_1 & 1 \end{bmatrix} \begin{bmatrix} 1 + N_X \frac{T_{X_p}}{d} \\ N_Y \frac{T_{X_p}}{d} \\ qN_Z \frac{T_{X_p}}{d} \end{bmatrix} \tag{7}$$

where $(x_1\ y_1)$ is the spatial-location of the origin of h_1. Similarly, the $x-$coordinate of the right-extreme point of h_1 due to another translation (say T_{X_q}) will be

$$\begin{bmatrix} x_{r_1} \end{bmatrix} = \begin{bmatrix} x_1 & y_1 & 1 \end{bmatrix} \begin{bmatrix} 1 + N_X \frac{T_{X_q}}{d} \\ N_Y \frac{T_{X_q}}{d} \\ qN_Z \frac{T_{X_q}}{d} \end{bmatrix} \tag{8}$$

Subtracting equation (7) from equation (8), we get

$$\begin{bmatrix} x_{\triangle_1} \end{bmatrix} = \begin{bmatrix} x_1 & y_1 & 1 \end{bmatrix} \begin{bmatrix} N_X \frac{T_{X_q} - T_{X_p}}{d} \\ N_Y \frac{T_{X_q} - T_{X_p}}{d} \\ qN_Z \frac{T_{X_q} - T_{X_p}}{d} \end{bmatrix} \tag{9}$$

where $x_{\triangle_1}$ indicates the difference between the $x-$coordinates of the two extreme points of the blur kernel h_1. If we can determine M such PSFs in the given blurred image, then we have a set of M (≥ 3) linear equations given by

$$\begin{bmatrix} x_{\triangle_1} \\ x_{\triangle_2} \\ . \\ . \\ x_{\triangle_M} \end{bmatrix} = \begin{bmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ . & . & 1 \\ . & . & 1 \\ x_M & y_M & 1 \end{bmatrix} \begin{bmatrix} N_X \frac{T_{X_q} - T_{X_p}}{d} \\ N_Y \frac{T_{X_q} - T_{X_p}}{d} \\ qN_Z \frac{T_{X_q} - T_{X_p}}{d} \end{bmatrix} \tag{10}$$

where $(x_i\ y_i)$ represents the spatial-location of the origin of the i^{th} PSF. Note that $\frac{T_{X_q} - T_{X_p}}{d}$ is a constant that multiplies every component of $\mathbf{n}$ and hence need not be estimated. Therefore, one can solve equation (10) using least-squares to infer the surface normal.

(a) (b)

Fig. 2. (a) Motion blurred planar scene. (b) PSFs at different locations of (a).

The procedure explained above, in fact, is equally applicable to extreme points along the y direction too. Since our scheme relies on pixel motion, we propose to use $x_{\triangle_i}$ or $y_{\triangle_i}$, whichever is higher in magnitude. Note that the fronto-parallel plane is a special case of our formulation in that the PSFs will be identical at all locations i.e, $x_{\triangle_i} = k\ \forall i$ from which the solution can be inferred as $\mathbf{n} = [0\ 0\ 1]^T$.

Consider Fig. 2(a) in which we have shown an inclined plane with surface normal $[0\ 0.5\ 0.8660]$. We applied a set of homographies arising from camera motion and generated this blurred image. We choose eight random locations in the blurred image and their corresponding PSFs are shown in Fig. 2(b). Note that the blur is space-varying, as expected. Due to translational motion of the camera, the PSFs vary with the spatial location of the patch. A patch closer to the camera contains more blur as compared to a patch farther away from the camera. To determine the extremities of a PSF, we calculate the row sum and column sum of the PSF and choose the positions of the first and last non-zero values of the PSF as extreme points. These points are indicated by red (left-most) and green (right-most) pixels. Pixels with the same color constitute point correspondences. Therefore, all the red (green) points correspond to the same homography.

A. PSF estimation

Although our interest is not in estimating camera motion, we need to determine PSFs at different spatial locations in the

blurred image. There exist several methods [20] - [24] in the literature for blur kernel estimation. We used an off-the-shelf blind motion deblurring technique [20] to estimate the blur kernel for a selected patch. Estimating the PSF from a single motion blurred image is a very ill-posed problem since there exist many possible combinations of PSF and latent image that can lead to the same blurred image. Hence, blind motion deblurring methods typically impose priors on the PSF and the latent image. The method of [20] reveals that strong edges need not always lead to accurate PSF estimation and employs a two-phase approach to estimate PSF. In the first phase, the authors define a metric to identify useful edges. These edges are considered to estimate a coarse blur kernel. In the second phase, an iterative support detection method is used (instead of hard-thresholding) to estimate the sparse blur kernel. The method executes fast and the accuracy of PSF estimation is also quite satisfactory [20].

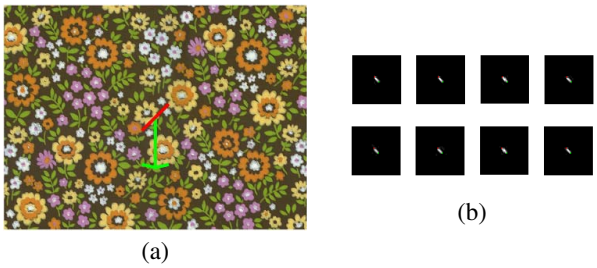
IV. Experiments

In this section, we validate the proposed method with examples, both synthetic and real. Since both PSF estimation as well as extreme point detection can involve small errors, we propose to use about 8 point-correspondences (instead of the minimum of 3) in equation (10) for robustness against noise. For the synthetic case, we choose focal length $q = 1200$ pixels which is a practical value. For these experiments, we assumed a surface normal, applied a set of homographies (camera translations) on an unblurred textured image, and computed the weighted average of the transformed images to yield the blurred observation. For the real case, the focal length (usually in mm) is gathered from the meta-data itself, and is converted into pixels using the sensor dimensions and the resolution of the image. The value of d in equation (10) is the same for all the points lying on the plane and it can be any constant (other than zero). In this work, we are interested only in the orientation of plane (and not in d which is embedded in the constant that multiplies $\mathbf{n}$ in equation (10)).

A. Synthetic case

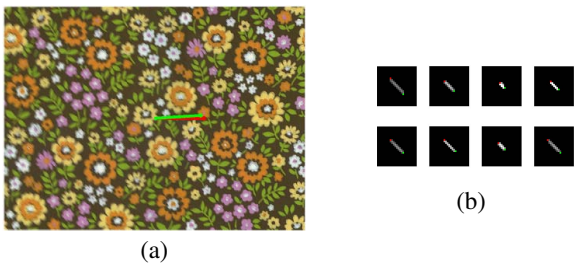
In the first example, we assumed a fronto-parallel planar scene ($\mathbf{n} = [0\ 0\ 1]^T$). We applied a set of translations along both x and y directions and the blurred image thus obtained is shown in Fig. 3(a). Due to the fronto-parallel nature of the scene, all the 3D points are at the same distance from the camera and experience identical blur. We randomly select eight (spatially well-separated) patches in Fig. 3(a) and estimate their PSFs using [20]. These PSFs are shown in Fig. 3(b) and, as expected, have the same form. The extreme points in each PSF are detected as discussed earlier in section III and the displacements between these points is substituted in equation (10) to solve for the normal. The estimated normal turned out to be $\hat{\mathbf{n}} = [0\ 0.0632\ 0.998]$ which is quite close to the true normal.

Next, we use the same image as in the earlier example but assume an inclined plane with normal $\mathbf{n} = [-0.6428\ 0\ 0.7660]$. Following the procedure outlined earlier for the fronto-parallel case, a blurred observation (Fig. 4(a)) was generated using a set of transformations for the camera motion. We randomly selected eight patches (each of size 120×120 pixels) and the corresponding PSFs estimated using [20] are shown in Fig. 4(b). Note that the blur kernel is space-variant, as expected. The extreme point correpondences among the blur kernels have also been indicated in Fig. 4(b). From the displacements of the extremities, the normal was estimated using equation (10) by employing only the $x-$translations. The result was $\hat{\mathbf{n}} = [-0.6007\ 0.0370\ 0.7986]$ which is close to the actual normal. The angular error between the actual (red arrow) and the estimated normal (green arrow) is only 3.7 degrees as depicted in Fig. 4(a).



(a) (b)

Fig. 3. (a) A fronto-parallel scene with translational blur. (b) PSFs estimated using [20] at random locations in (a).



(a) (b)

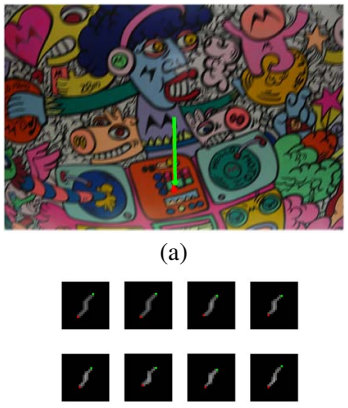
Fig. 4. (a) Inclined plane with motion blur. (b) PSFs estimated at different locations in (a).

Yet another synthetic example is shown in Fig. 5(a) with the true normal being $\mathbf{n} = [0\ \ -0.5\ \ 0.8660]$. The PSFs are shown in Fig. 5(b) and the normal estimated from the extremities of the PSFs was found to be $\hat{\mathbf{n}} = [-0.0061\ \ -0.4484\ \ 0.8938]$. The true and estimated normals shown in Fig. 5(a) have an angular deviation of 3.9 degrees which can be attributed to small errors in estimating the PSFs and their extreme points.

B. Real case

We used a Canon 60D camera to capture real data. The sensor width of the camera was 23.2 mm and the spatial resolution was 720×480 pixels. For the real experiments, we employed a translational stage to induce translational motion blur along both x and y directions.

In the first example, we captured a translationally blurred fronto-parallel textured board (Fig. 6(a)). Akin to the synthetic case, we choose eight different patches (again of size 120×120 pixels) and determine the PSFs corresponding to the center of these patches using [20]. The estimated PSFs are shown in Fig. 6(b) with extreme points marked. The normal estimated using equation (10) was found to be $\hat{\mathbf{n}} = [0\ 0\ 1]$. Since we



(a)

(b)

Fig. 5. (a) A blurred inclined plane, and (b) PSFs obtained at eight different locations in (a).

(a)

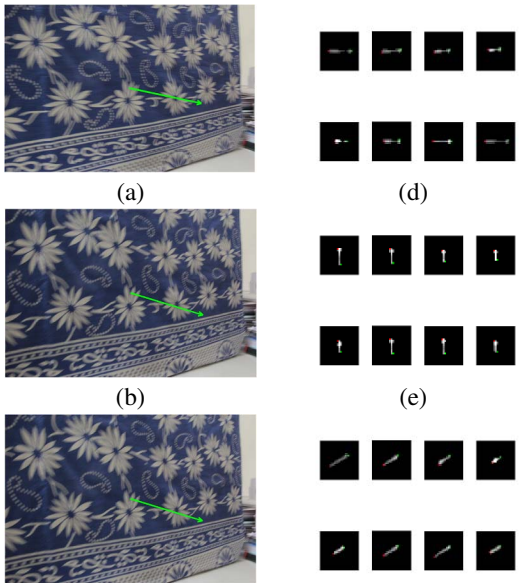
(b)

Fig. 6. (a) Fronto-parallel blurred image. (b) PSFs estimated at different locations in (a).

know apriori that the scene is fronto-parallel, we can conclude that the estimated normal is indeed correct.

Next, we captured a blurred image of an inclined plane as shown in Fig. 7(a). One can visually perceive the space-variant nature of the blur in this image. We randomly picked eight patches and the estimated PSFs are shown in Fig. 7(d). The extreme points in each PSF are represented with red (left-most) and green (right-most) colors. By following the same procedure discussed in the earlier experiments, the surface normal was found to be $\hat{\mathbf{n}} = [-0.4601\ 0.0970\ 0.8825]$. Since, this is a real example, we do not know the true normal. We ascertain the correctness of the estimated normal by capturing blurred images of the *same plane* but with two different camera translations. Ideally, the estimated normals should be identical irrespective of the camera motion. We captured two more blurred images with different in-plane translations and these are shown in Figs. 7(b)-(c), with their corresponding PSFs (Figs. 7(e)-(f)). The estimated normals were found to be $\hat{\mathbf{n}}$= $[-0.4886\ \ -0.051\ 0.871]$ and $[-0.4601\ 0.1400\ 0.8767]$ respectively. Note that the estimated normals in all the three cases are quite close to one another reaffirming the correctness of our procedure. Furthermore, we physically measured the orientation of the plane and found it to be 30 degrees. This is indeed close to the value of 28 degrees obtained using the proposed method.

We show another real example in Fig. 8(a). We captured an outdoor ground-plane with the optical axis of the camera



(a)

(b)

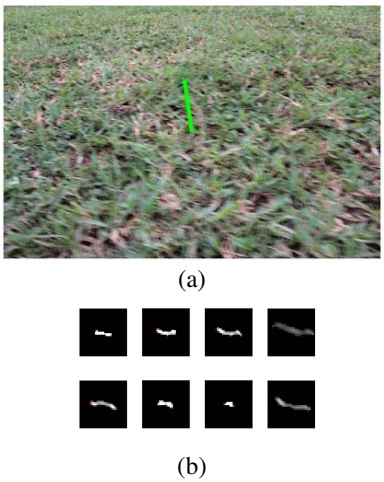
(c)

(d)

(e)

(f)

Fig. 7. (a)-(c) Blurred images of an inclined plane for different camera translations. (d)-(f) PSFs corresponding to figures.(a)-(c), respectively.



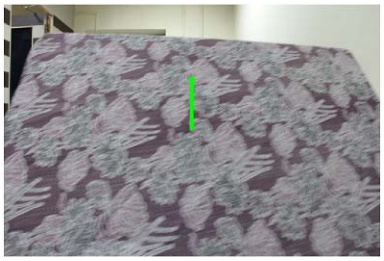
(a)

(b)

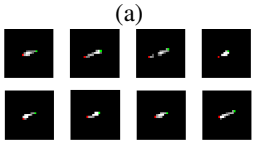
Fig. 8. (a) Planar surface with motion blur. (b) PSFs estimated at different locations in (a).

approximately parallel to the ground-plane. The focal length was 18 mm. From Fig. 8(a), we observe that the image has significant variations in blur. The lower portion of the image (close to the camera) has more blur compared to the upper portion (far from the camera). To recover surface normal, we selected eight patches such that the patches were spreadout across the image. Their estimated PSFs are shown in Fig. 8(b). From the PSFs, we can infer that the translation is more prevalant along the $x-$direction. The detected extremities in each PSF are also indicated in Fig. 8(b). By following the procedure discussed earlier, the normal was computed as $[0.1373\ 0.9800\ 0.1442]$. Because the optical axis was not exactly parallel to the plane, the resultant angle turns out to be 81 degrees which is as expected.

In the final example, a planar scene was imaged as shown in Fig. 9(a). The bottom of the plane is closest to the camera while the top edge of the plane is the farthest. The focal length of the camera (18 mm) was obtained from the image meta-data. Using sensor dimensions and image size, the focal length translates to 581 pixels. We randomly picked eight patches throughout the image and their corresponding PSFs are shown in Fig. 9(b). After substituting the focal length and displacements of each PSF in equation (10), the computed surface normal was found to be $[0 \ -0.3713 \ 0.9151]$ and is indicated by a green arrow.



(a)



(b)

Fig. 9. (a) A planar surface with translational motion blur. (b) Blur kernels extracted at different spatial locations in (a).

V. Conclusions

We proposed a scheme, which we believe is the first of its kind, to estimate planar orientaion from a single motion blurred image. We revealed the underlying relationship between the surface normal of a planar scene and the induced space-variant nature of blur due to translational motion. Exploiting correspondences among the extreme points of the PSFs, we constructed a set of linear equations whose solution yields the surface normal. The method was validated on synthetic as well as real images.

Since our method can explain planes, it offers a convenient platform for attempting restoration of piecewise planar motion-blurred scenes. As future work, we plan to relax the translational constraint and extend our method to the case of general camera motion. Yet another direction to pursue is to automatically segment and estimate surface normals when the scene consists of multiple planes.

Acknowledgments

A part of this work was supported by a grant from the Asian Office of Aerospace Research and Development, AOARD/AFOSR. The support is gratefully acknowledged. The results and interpretations presented in this paper are that of the authors, and do not necessarily reflect the views or priorities of the sponsor, or the US Air Force Research Laboratory.

References

[1] L.S Hwa, and S. Sharma, "Real-time disparity estimation algorithm for stereo camera systems", *IEEE Transactions on Consumer Electronics*, vol. 57, no. 3, 1018–1026, 2011.

[2] S. Behzad, and M.K. Brown, "Robust depth estimation from optical flow", *in Proc. International Conference on Computer Vision*, 1988.

[3] B.J Super, and A.C Bovik, "Planar surface orientation from texture spatial frequencies", *Pattern Recognition* vol. 28, no. 5, 729-743, 1995.

[4] B.J Super, A.C Bovik, "Shape from texture using local spectral moments", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 17, no. 4, 333-343, 1995.

[5] L.G Brown, H. Shvayster, "Surface orientation from projective foreshortening of isotropic texture autocorrelation", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, no. 6, 584-588, 1990.

[6] R.T. Frankot and R. Chellappa, "A method for enforcing integrability in shape from shading", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 10, no. 7, 439-451, 1988.

[7] A. Pentland, "Shading into texture", *Artificial Intelligence*, vol. 29, no. 2, 147-170, 1986.

[8] S. Chaudhuri and A.N Rajagopalan, "Depth from defocus: A real aperture imaging approach", *Springer-Verlag*, New York, March 1999.

[9] C. Paramanand and A.N. Rajagopalan, "Inferring image transformation and structure from motion-blurred images", *in Proc. British Machine Vision Conference (BMVC)*, 2010.

[10] H.Y. Lin and C.H. Chang, "Depth recovery from motion blurred images", *in Proc. International Conference on Pattern Recognition (ICPR)*, 2006.

[11] J.S Fox, "Range from translational motion blurring", *in proc. Computer Vision and Pattern Recognition*, 360-365, 1988.

[12] Y. Zheng, S. Nobuhara, and Y. Sheikh, "Structure from motion blur in low light", *in Proc. Computer Vision and Pattern Recognition (CVPR)*, 2569-2576, 2011.

[13] H. Farid and J. Kosecka, "Estimating planar surface orientaion using bispectral analysis", *IEEE Transactions on Image Processing*, vol. 16, no. 8, 2154-2160, 2007.

[14] P. Clark, and M. Majid, "Estimating the orientation and recovery of text planes in a single image", *in Proc. British Machine Vision Conference (BMVC)*, 2001.

[15] T. Greinera, Shivani G. Rao and Sukhendu Das, "Estimation of orientation of a textured planar surface using projective equations and separable analysis with M-channel wavelet decomposition", *Pattern Recognition*, vol 43, Jan 2010.

[16] O. Haines and A. Calway, "Detecting planes and estimating their orientation from a single image", *in Proc. British Machine Vision Conference (BMVC)*, 2012.

[17] M. Sorel, and J. Flusser, "Space-variant restoration of images degraded by camera motion blur", *IEEE Transactions on Image Processing*, vol. 17, no. 2, 105-116, 2008.

[18] L. Xu and J. Jia, "Depth-aware motion deblurring", *in Proc. International Conference on Computational Photography (ICCP)*, 2012.

[19] S. McCloskey and M. Langer, "Planar orientation from blur gradients in a single image", *in Proc. Computer Vision and Pattern Recognition (CVPR)*, 2009.

[20] Xu Li, and J. Jia, "Two-phase kernel estimation for robust motion deblurring", *in Proc. European Conference on Computer Vision (ECCV)*, Springer Berlin Heidelberg, 157-170, 2010.

[21] R. Fergus, B. Singh, A. Hertzmann, S.T. Roweis, and W.T. Freeman, "Removing camera shake from a single photograph", *in ACM Transactions on Graphics (TOG)*, vol. 25, no. 3, pp. 787-794, 2006.

[22] Q. Shan, J. Jia, and A. Agarwala, "High-quality motion deblurring from a single image", *in ACM Transactions on Graphics*, vol. 27, Aug. 2008.

[23] O. Whyte, J. Sivic, A. Zisserman, and J. Ponce, "Non-uniform deblurring for shaken images", *in Proc. Computer Vision and Pattern Recognition (CVPR)*, Jun. 2010.

[24] A. Gupta, N. Joshi, L. Zitnick, M. Cohen, and B. Curless, "Single image deblurring using motion density functions", *in Proc. European Conference on Computer Vision (ECCV)*, 2010.