BMVC2000

Detecting Multiple Texture Planes using Local Spectral Distortion

Eraldo Ribeiro* and Edwin R. Hancock
Department of Computer Science,
University of York, York Y01 5DD, UK

Abstract

This paper presents a spectral method for identifying multiple texture planes. We commence by showing how pairs of spectral peaks can be used to make direct estimates of the slant and tilt angles. Our method commences by computing the affine distortion matrices for pairs of corresponding spectral peaks. We show that the directions of the eigenvectors of the affine distortion matrices can be used to estimate local planar pose. The leading eigenvector points in the tilt direction and the direction of the second eigenvector can be used to estimate the slant direction. By searching for clusters in the computed slant and tilt angles, we locate multiple texture planes. We apply the method to a variety of real-world and synthetic imagery.

1 Introduction

The identification of perspective planes from texture information is a key problem in shape-from-texture. Broadly speaking the problem can be approached in either a global or a local manner. The global approach involves identifying vanishing points from either texture-gradient or spectral information. This method invariably revolves around the structural analysis of pre-segmented texture primitives [1, 2, 3]. While providing geometrically intuitive means of recovering the parameters of perspective pose, these methods are not well suited to the analysis of scenes containing multiple planes. The second approach is to use measures of spectral distortion to make local estimates of the pose parameters. These Fourier domain methods revolve around simplifying the perspective geometry to obtain locally affine measures of texture distortion [4, 5, 6, 7]. While being more suitable for the analysis of scenes containing multiple planes or curved surfaces, these local methods rely on numerical optimisation and do not return closed-form pose estimates.

Since it represents one of the few attempts at texture-plane segmentation reported in the literature, we pause to consider Krumm and Shafer's [8] method in more detail. The method uses spectral back-projection to estimate local surface orientation. This is done by recovering the parameters of affine projection which numerically minimise the sum-of-squares difference between local spectra. Once the pose parameters are to hand, then each local power-spectrum is projected onto the front-parallel plane. Texture planes are segmented using a dendrogram-based clustering method.

*Supported by CAPES-BRAZIL, under grant: BEX1549/95-2

BMVC 2000 doi:10.5244/C.14.11

Our aim in this paper is to develop a more direct method for detecting multiple perspective planes in textured images. We follow Krumm and Shafer [8] by assuming that the local texture variations due to the perspectivity can be approximated in an affine manner. In the Fourier domain, this means that the local frequency content of the texture also undergoes affine distortion. Our measure of texture variation are the affine distortion matrices between corresponding spectral peaks at different locations on the image plane. We present an analysis which shows how the eigenvectors of the spectral distortion matrices can be used to make closed-form estimates of the slant and tilt angles of local texture planes. This result applies under the assumption that the underlying texture is homogeneous.

Once we have pose estimates to hand, we can proceed to develop an algorithm for multiple plane detection. This is a two-step process. The first step is to improve the consistency of the orientation field through the use of local contextual information. Here we use a robust smoothing method which has proved successful in the shape-from-shading domain [9, 10] to improve the organisation of the needle map. The second step is to locate planes by clustering the local slant and tilt angles. This is done by using the Gustavson-Kessel fuzzy c-means algorithm [11].

2 Spectral distortion under perspectivity

This paper is concerned with recovering a dense map of surface orientations for surfaces which are uniformly painted with periodic textures. Our approach is a spectral one which is couched in the Fourier domain. We make use of an analysis of spectral distortion under perspective geometry extensively developed by Super and Bovik, Krumm and Shafer, and, by Malik and Rosenholtz among others. This analysis simplifies the full perspective geometry of texture planes using a local affine approximation. Suppose that $\mathbf{U}_t$ represents the frequency vector associated with a spectral peak detected at the point with position-vector $\mathbf{X}_t$ on the texture plane. Further, let $\mathbf{U}_i$ and $\mathbf{X}_i$ represent the corresponding frequency vector and position vector when the texture plane undergoes perspective projection onto the image plane. If the perspective projection can be locally approximated by an affine distortion $T_A(\mathbf{X}_i)$, then the relationship between the texture-plane and image-plane frequency vectors is $\mathbf{U_i} = T_A(\mathbf{X}_i)^{-T}\mathbf{U_t}$. The affine distortion matrix is given by

$$T_A(\mathbf{X}_i) = \frac{\Omega}{hf\cos\sigma}\begin{bmatrix} x_i \sin\sigma + f\cos\tau\cos\sigma & -f\sin\tau \\ y_i\sin\sigma + f\sin\tau\cos\sigma & f\cos\tau \end{bmatrix} \tag{1}$$

where f is the focal length of the camera, σ is the slant angle of the texture-plane, τ is the tilt angle of the texture-plane and $\Omega = f\cos\sigma + \sin\sigma\,(x_i\cos\tau + y_i\sin\tau)$.

In this paper we aim to show how an affine analysis of local spectral distortion can be used to recover the parameters of planar pose in closed form. Although affine analysis has previously been used by Krumm and Shafer in their work on shape-from-texture [8], the recovery of pose parameters has relied on exhaustive numerical search and spectral back-projection. In this paper, we demonstrate that the eigenvectors of the affine transformation matrix can be used to directly estimate the slant and tilt angles.

2.1 Spectral homographies

Our aim is to make local estimates of the slant and tilt angles using the observed distortions of the texture spectrum across the image plane. To do this we measure the affine distortion between corresponding spectral peaks. To commence, consider the point S on the texture plane. We sample the texture projection at two neighbouring points A and B laying on the image plane. The co-ordinates of the two points are respectively $\mathbf{X}_A = (x, y)^T$ and $\mathbf{X}_B = (x + \Delta x, y + \Delta y)^T$ where Δx and Δy are the image-plane displacements between the two points. Figure 1 illustrates the idea.

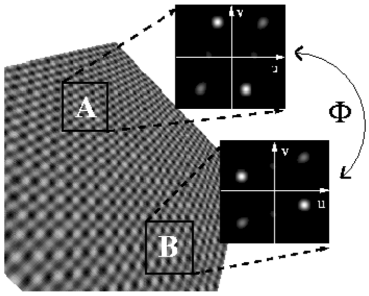


Figure 1: Affine distortion between two neighbour local spectra

Consider a local-planar patch on the observed texture surface which has a spectral peak with frequency vector $U_S = (u_s, v_s)^T$. Suppose that we sample the corresponding image plane spectral peak at two distinct locations X_A and X_B where the measured frequency vectors are respectively $U_A = (u_A, v_A)^T$ and $U_B = (u_B, v_B)^T$. Provided that the two image-plane locations belong to the same planar region of the observed texture, then the image plane peak frequency vectors are related to the texture-surface frequency vector via the equations $U_A = (T_A(X_A)^{-1})^T U_S$ and $U_B = (T_A(X_B)^{-1})^T U_S$, where $T_A(X_A)$ is the local affine projection matrix of the planar surface patch at the point A and $T_A(X_B)$ is the corresponding affine projection matrix at the point B. As a result, the frequency vectors for the two corresponding spectral peaks on the image-plane are related to one-another via the local spectral distortion

$$U_B = \left(T_A(X_A) T_A(X_B)^{-1} \right)^T U_A \tag{2}$$

As a result, the texture-surface spectral distortion matrix $\Phi = (T_A(X_A)T_A(X_B)^{-1})^T$ is a 2x2 matrix. This matrix relates the affine distortion of the image plane frequency vectors to the local planar pose parameters. Substituting for the affine approximation to the perspective transformation from Equation (2), the required matrix is given in terms of the slant and tilt angles as

$$\Phi = \frac{\Omega(\mathbf{A})}{\Omega^2(\mathbf{B})} \begin{bmatrix} \Omega(\mathbf{A}) + \Delta y \sin\sigma \sin\tau & -\Delta y \sin\sigma \cos\tau \\ -\Delta x \sin\sigma \sin\tau & \Omega(\mathbf{A}) + \Delta x \sin\sigma \cos\tau \end{bmatrix} \tag{3}$$

where $\Omega(A)= f\cos\sigma + \sin\sigma\,(x\cos\tau + y\sin\tau)$ and $\Omega(B)= f\cos\sigma + \sin\sigma\,((x+\Delta x)\cos\tau + (y+\Delta y)\sin\tau)$. The above matrix represents the distortion of the spectrum sampled at the location B with respect to the sample at the location A. In the next Section we show how to solve directly for the parameters of planar pose, i.e. the slant and tilt angles, using the eigen-structure of the transformation matrix Φ.

2.2 Eigenstructure of the spectral homography matrix

In the previous section, we showed that if local planarity is assumed then neighbouring power spectra can be related to one-another by means of an affine transformation matrix. The assumption of local planarity implies that the neighbouring spectra will be distorted by the same slant and tilt angles. In this Section we present the key theoretical contribution of this paper, which is to show how the eigenvectors of the affine distortion matrix can be used to estimate the slant and tilt angles. To commence, let us consider the eigenvector equation for the distortion matrix Φ, i.e. $\Phi\mathbf{w}_\lambda = \lambda\mathbf{w}_\lambda$, where $\lambda = (\lambda_1, \lambda_2)$ are the eigenvalues of the distortion matrix Φ and $\mathbf{w}_\lambda$ are the corresponding eigenvectors. Since Φ is a 2x2 matrix the two eigenvalues are found by solving the quadratic eigenvalue equation $det[\Phi - \lambda I] = 0$ where I is the 2x2 identity matrix. The explicit eigenvalue equation is $\lambda^2 - Trace(\Phi)\lambda + det(\Phi) = 0$, where $Trace(\Phi)$ and $det(\Phi)$ are the trace and determinant of Φ. Substituting for the elements of the transformation matrix Φ, we have $\lambda^2 - [\frac{\Omega(\mathbf{A})}{\Omega(\mathbf{B})} + \frac{\Omega^2(\mathbf{A})}{\Omega^2(\mathbf{B})}]\lambda + [\frac{\Omega(\mathbf{A})}{\Omega(\mathbf{B})} \times \frac{\Omega^2(\mathbf{A})}{\Omega^2(\mathbf{B})}] = 0$. The two eigenvalue solutions of the above quadratic equation are $\lambda_1 = \frac{\Omega^2(\mathbf{A})}{\Omega^2(\mathbf{B})}$ and $\lambda_2 = \frac{\Omega(\mathbf{A})}{\Omega(\mathbf{B})}$. The corresponding eigenvectors are $\mathbf{w}(\lambda_1) = [\mathbf{w}_x(\lambda_1), \mathbf{w}_y(\lambda_1)]^T$ and $\mathbf{w}(\lambda_2) = [\mathbf{w}_x(\lambda_2), \mathbf{w}_y(\lambda_2)]^T$. More explicitly, the eigenvectors are $\mathbf{w}(\lambda_1) = [1, \tan\tau]^T$ and $\mathbf{w}(\lambda_2) = [1, -\frac{\Delta x}{\Delta y}]^T$. As a result we can directly determine the tilt angle from the vector components of the eigenvector associated with the eigenvalue λ_1. The tilt angle is given by

$$\tau = \arctan(\frac{\mathbf{w}_y\,(\lambda_1)}{\mathbf{w}_x\,(\lambda_1)}) \tag{4}$$

The intuitive justification for this result is that under perspective projection the only direction which remains invariant at all locations on the image plane is the tilt direction. As a result, a frequency vector which is aligned in the tilt direction will maintain a constant angle, but it will change in magnitude according to the position on the image plane. In other words, the tilt direction is an eigenvector of the local affine transformation.

Once the tilt angle has been obtained, we recover the slant angle by solving the equation $\lambda_2 = \frac{\Omega(\mathbf{A})}{\Omega(\mathbf{B})} = \frac{f\cos\sigma + \sin\sigma(x\cos\tau + y\sin\tau)}{f\cos\sigma + \sin\sigma(x+\Delta x)\cos\tau + (y+\Delta y)\sin\tau)}$. The solution is

$$\sigma = \arctan\left[\frac{f\,(\lambda_2 - 1)}{(y(1-\lambda_2) - \lambda\Delta y)\sin\tau + (x(1-\lambda_2) - \lambda_2\Delta x)\cos\tau}\right] \tag{5}$$

3 Computing local planar orientation

In this Section we explain how to recover local planar surface orientation using our affine distortion method. The first step in orientation recovery is to estimate the affine distortion matrix which represents the transformation between different local texture regions on the image plane. These image texture regions are assumed to belong to a single local planar patch on the curved texture surface. We do this by selecting pairs of neighbouring points on the image plane. At each point there may be several clear spectral peaks. Since the affine distortion matrix Φ has four elements that need to be estimated, we need to know the correspondences between at least two different spectral peaks at the different locations. Suppose that $U_1^{p_1} = (u_1^{p_1}, v_1^{p_1})^T$ and $U_1^{p_2} = (u_1^{p_2}, v_1^{p_2})^T$ represent the frequency vectors for two distinct spectral peaks located at the point with co-ordinates $\mathbf{X}_1 = (x_1, y_1)^T$ on the image plane. The frequency vectors are used to construct the columns of a 2x2 spectral measurement matrix $V_1 = (\mathbf{U}_1^{p_1}|\mathbf{U}_1^{p_2})$. Further, suppose that $U_2^{p_1} = (u_2^{p_1}, v_2^{p_1})^T$ and $U_2^{p_2} = (u_2^{p_2}, v_2^{p_2})^T$ represent the frequency vectors for the corresponding spectral peaks at the point $\mathbf{X}_2 = (x_2, y_2)^T$. The corresponding spectral measurement matrix is $V_2 = (\mathbf{U}_2^{p_1}|\mathbf{U}_2^{p_2})$. Under the affine model presented in Section 2, the spectral measurement matrices are related via the equation $V_2 = \Phi V_1$. As a result the local estimate of the spectral distortion matrix is $\Phi = (V_1^T)^{-1} V_2$. If correspondences between more than two spectral peaks are available then the affine distortion matrix can be obtained through least-squares estimation. If V_1 and V_2 are 2xM spectral measurement matrices for M corresponding spectral peaks, then the least squares estimate which minimises the quantity $(V_2 - \Phi V_1)^T (V_2 - \Phi V_1)$ is $\Phi = (V_1^T V_1)^{-1} V_1^T V_2$.

In practice, we only make use of the most energetic peaks appearing in the power spectrum. That is to say we do not consider the detailed distribution of frequencies. Our method requires that we supply correspondences between spectral peaks so that the distortion matrices can be estimated. We use the energy amplitude of the peaks is establish the required correspondences. This is done by ordering the peaks according to their energy amplitude. The ordering of the amplitudes of peaks at different image locations establishes the required spectral correspondences.

In order to improve the orientation estimates returned by eigenvectors of the affine distortion matrix, we use the robust vector smoothing method developed by Worthington and Hancock [10] for shape-from-shading.

4 Experiments

We have experimented with both synthetic surfaces with known ground truth and real-world images. The former are used to assess the accuracy of the method, while we use the latter to demonstrate the practical utility of the method.

4.1 Real world data

We commence with experiments on real-world images which illustrate how our method can be used to identify multiple texture planes. Here we use the Gustavson-Kessel [11] fuzzy clustering method to locate clusters in the distribution of needle-map directions after robust smoothing has been applied. Each cluster is taken to represent a distinct plane. The mean surface normal direction for the cluster represents the orientation of the segmented plane. In Figure 2 we show the result of applying the new texture plane location method to some real world image data. The images in column a) are the original images. Column b) shows the needle-maps, i.e. the local orientation information, extracted using our new method. In column c) we show the distribution of needle-map directions with the detected clusters marked as circles. Here the surface normal data is visualised on a unit sphere. The polar angles defining the hemisphere are the slant and tilt directions of the surface normals. The cluster centres detected by Gustavson-Kessel method are indicated by circles, whose radius indicates the cluster-variance. Finally, column d) shows the segmentations obtained when the pixels in the original image are labelled according to the cluster which has the largest fuzzy membership weight.

The image in the top row is synthetic and represents three distinct planes. These are cleanly segmented from the image data. The second image is obtained by placing a table-cloth over two sides of a rectangular box. Again, the two planes are cleanly segmented. The image shown in the third row is more complicated. Here the table-cloth is folded over the corner of a rectangular box. All three planes are located. Finally, the bottom row shows the result obtained when the method is applied to an image of a folded piece of wallpaper. Here the two planes are well located, However, there are segmentation errors associated with the location of the fold.

4.2 Sensitivity analysis

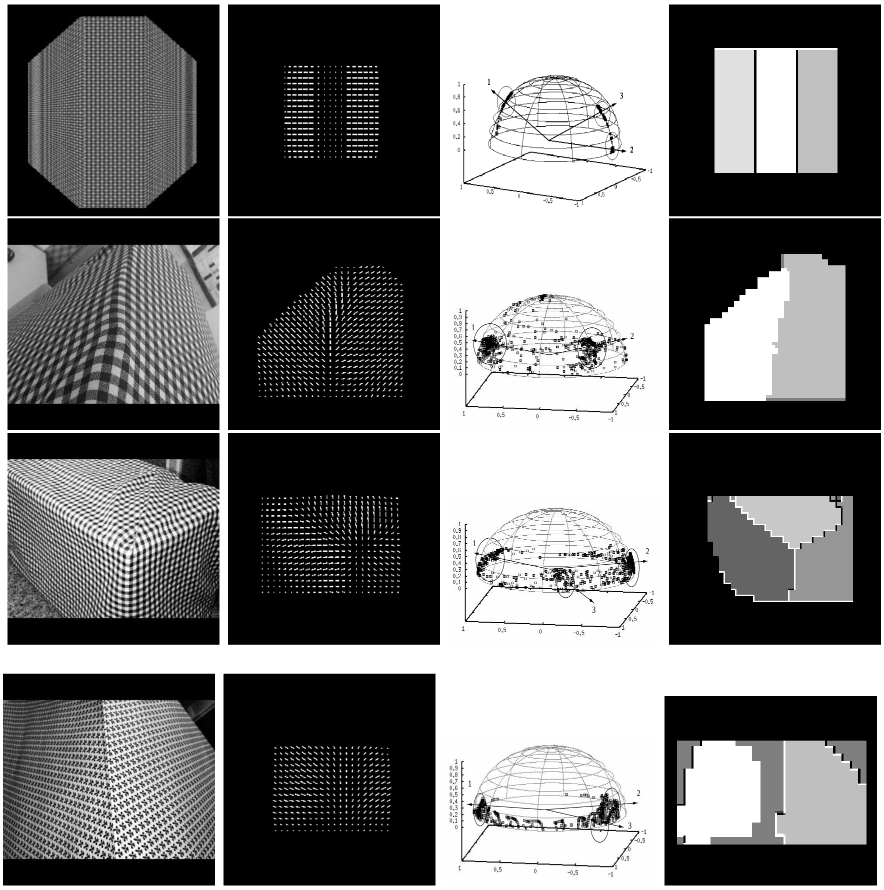
Next we turn our attention to measuring the sensitivity of the the method on synthetic images. We commence by assessing the ability of the method to recover reliable slant and tilt information. Here we have generated synthetic regularly textured surfaces. We use the ground truth surface normals to investigate the systematics of the errors in the estimated needle-map.

In Figure 3 we show scatter plots of the estimated slant and tilt angles versus their ground truth values. In each case there is a clear regression line.

We have also measured the sensitivity of our method to imperfections in the regularity in the texture. Our method is based on the assumption that local texture planes are homogeneously painted with regular textures. We have investigated with the effect of disturbing the regularity of the texture in both the frequency domain and the spatial domain. In the frequency domain, we have added white noise (i.e. noise with a uniform frequency distribution) of increasing amplitude to the texture. Figure 4 shows the effect of adding increased amounts of white noise to a sinusoidal texture. The perspective projections of the randomised textures with slant and tilt angles of 45 degrees are shown in Figure 5. In the spatial domain, we have investigated the effect of randomising the positions of texture primitives.

We do this by adding a random displacement sampled from a two-dimensional circularly symmetric Gaussian distribution of zero mean and known variance. In Figure 6 we show the effect of increasing the variance on the distribution of texture primitives. The perspective projections of these textures are shown in Figure 7.

In Figure 8 we show the effect of the spatial domain and frequency domain noise on the recovered slant and tilt angles. Here we have projected the noise corrupted textures onto a plane whose slant and tilt angles are both equal to 45



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

Figure 2: 3D-plane segmentation. (a) original images; (b) estimated needle map; (c) spherical clusters; (d) segmented planes

degrees. Figure 8(a) shows the measured slant and tilt angles when the sinusoidal textures shown in Figures 4 and 5 are subjected to frequency domain noise. Here the method degrades once the amplitude of the white noise exceeds 12% of the amplitude of the original sinusoidal frequency. Figure 8(b) repeats this experiment for the spatial-domain texture shown in Figures 6 and 7. The plot shows the slant and tilt error as a function of the standard-deviation of the random spatial displacement error. Here the method recovers good estimates of the slant and tilt parameters provided that the ratio of the standard deviation of the displacement error to the inter-point distance does not exceed 10%.

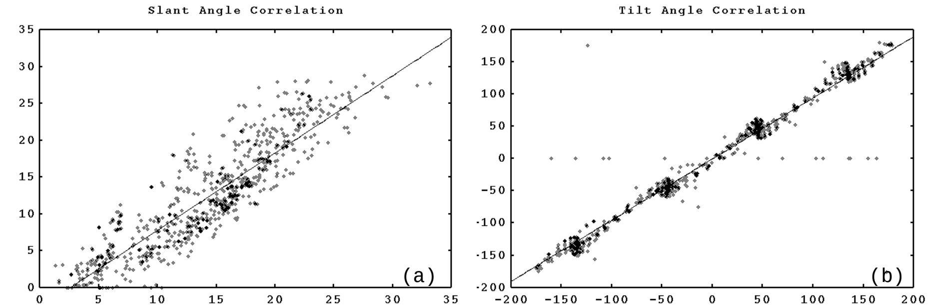


Figure 3: Scatter Correlation Plots (Robust Kernel). (a) Slant Angle Correlation($\mu = 0.91$); (b) Tilt Angle Correlation($\mu = 0.96$). Where μ is the linear correlation coefficient

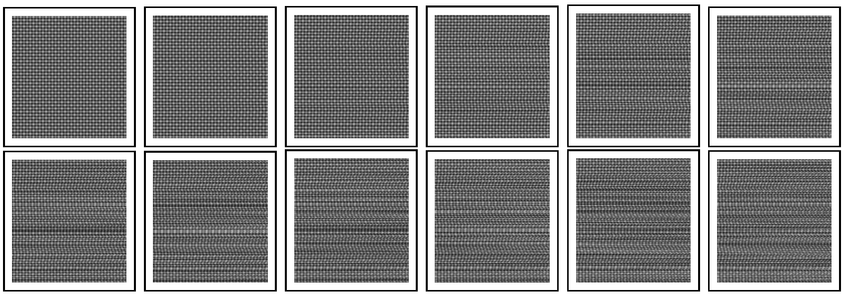


Figure 4: Sinusoid images with increasing spectral frequency irregularity

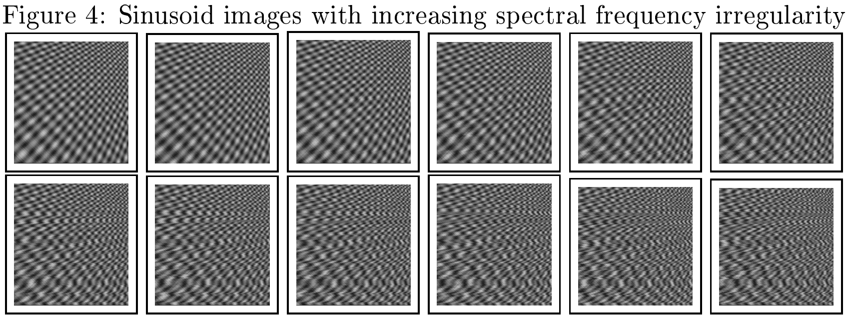


Figure 5: Projected sinusoid images

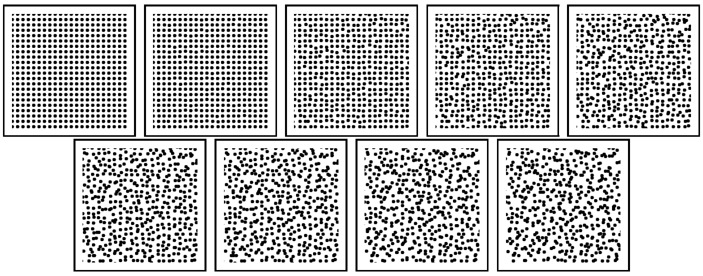


Figure 6: Circular texture images with increasing spatial location irregularity.

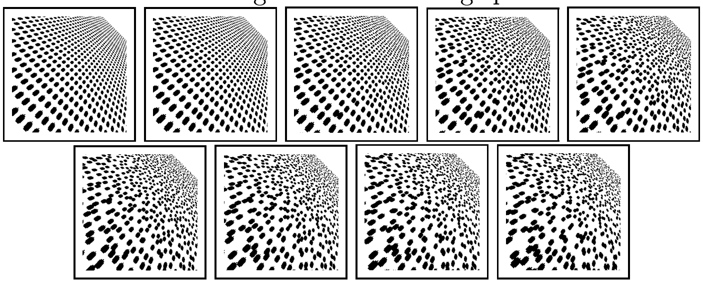
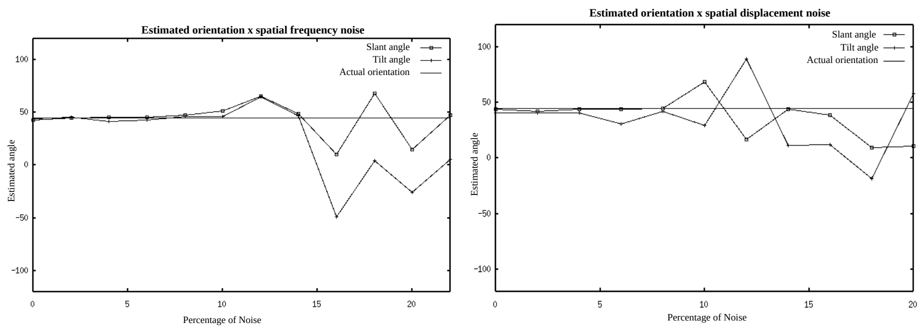


Figure 7: Projected circular texture images.



(a) (b)

Figure 8: Slant-tilt error plots. (a) Increasing spectral frequency noise; (b) Increasing spatial location noise

5 Conclusions

We have presented a new method for locating multiple planes in textured imagery. The method uses the eigenvectors of the local spectral distortion matrix to make closed-form estimates of the slant and tilt angles of planar pose. By clustering the resulting surface normals we locate individual texture planes. The method is demonstrated on both real-world and synthetic data. Real world examples indicate that it is capable of cleanly segmenting planar textures. The synthetic data shows that the method is robust to the presence of significant irregularities in the textures.

Our future plans revolve around using the eigen-structure of the spectral distortion matrices to recover topographic information from curved surfaces. This study is in hand and the results will be reported in due course.

References

[1] K. Ikeuchi. Shape from regular patterns. *Artificial Intelligence*, 22:49–75, 1984.

[2] John Aloimonos and M.J. Swain. Shape from texture. *Biological Cybernetics*, 58(5):345–360, 1988.

[3] K. Kanatani and T. Chou. Shape from texture: General principle. *Artificial Intelligence*, 38:1–48, 1989.

[4] John Krumm and Steve A. Shafer. Shape from periodic texture using spectrogram. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 284–289, 1992.

[5] B.J. Super and A.C. Bovik. Planar surface orientation from texture spatial frequencies. *Pattern Recognition*, 28(5):729–743, 1995.

[6] Jonas Garding. Shape from texture for smooth curved surfaces. In *European Conference on Computer Vision*, pages 630–638, 1992.

[7] J. Malik and R. Rosenholtz. A differential method for computing local shape-from-texture for planar and curved surfaces. In *IEEE Conference on Vision and Pattern Recognition*, pages 267–273, 1993.

[8] John Krumm and Steve A. Shafer. Texture segmentation and shape in the same image. In *IEEE International Conference on Computer Vision*, pages 121–127, 1995.

[9] P. L. Worthington and E.R. Hancock. Needle map recovery using robust regularizers. *Image and Vision Computing*, 17(8):545–559, 1998.

[10] P.L. Worthington and E.R. Hancock. New constraints on data-closeness and needle map consistency for shape-from-shading. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 21(12):1250–1267, December 1999.

[11] J. C. Bezdek. *Pattern Recognition with Fuzzy Objective Algorithms*. Plenum Press, 1981.