A Scale Adaptive Method for Estimating the Perspective Pose of Texture Planes

ERALDO RIBEIRO* EDWIN R. HANCOCK

Department of Computer Science,
University of York, York Y01 5DD, UK
eraldorj@minster.york.ac.uk

Abstract. Multiple vanishing point detection provides the key to recovering the perspective pose of textured planes. If vanishing points are to be detected from spectral information then there are two computational problems that need to be solved. Firstly, the search of the extended image plane is unbounded, and hence the location of vanishing points at or near infinity is difficult. Secondly, correspondences between local spectra need to be established so that vanishing points can be triangulated. In this paper we offer a way of overcoming these two difficulties. We overcome the problem of unbounded search by mapping the information provided by local spectral moments onto the unit-sphere. According to our representation, the position and direction of each local spectrum maps onto a great circle on the unit-sphere. The need for correspondences is overcome by accumulating the great circle intercepts. Vanishing points occur at local accumulator maxima on the unit sphere. To improve the accuracy of the recovered perspective pose parameters for highly slanted planes, we use an adaptive spectral window. This selects the window so as to reduce spectral defocusing by minimising the determinant of the spectral covariance matrix. We experiment with the new shape-from-texture technique on both synthetic and real world data. Here it proves to be an accurate and robust means of estimating perspective pose.

keywords: Shape-from-texture, unit-sphere accumulation, adaptive spectral analysis, texture analysis, planar surface recovery.

1 Introduction

The perspective foreshortening of surface patterns is an important cue for the recovery of surface orientation from 2D images [1, 2]. Broadly speaking there are two routes to recovering the parameters of perspective projection for texture patterns. The first of these is to estimate the texture gradient [3, 4]. Geometrically, the texture gradient determines the tilt direction of the plane in the line-of-sight of the observer and its magnitude determines the slant angle of the plane. A more direct and geometrically intuitive alternative route to the local slant and tilt parameters of the surface is to estimate the whereabouts of vanishing points [5, 6, 7]. Provided that two or more vanishing points are available, then planar surface orientation can be directly determined.

Unfortunately the location of vanishing point from texture distribution is not itself a straightforward task. If direct analysis is being attempted in the spatial domain, then the tractability of the problem hinges on the regularity and structure of the texture primitives [5, 6]. Moreover, multiple vanishing point detection may be even more elusive. It is for this reason that frequency domain analysis offers an attractive alternative [8, 9, 10]. The main reason for this is that the analysis of spectral moments can provide a convenient means of identifying the individual radial patterns associated with multiple vanishing points.

In this paper we make use of an interesting spectral property which provides a direct route to vanishing point location via the use of frequency domain information [11]. The observation is a simple one. At each point on the image plane, the spectral angle points in the direction of a vanishing point. Lines that radiate from a vanishing point therefore connect points of uniform spectral angle.

The idea underpinning our method is to exploit the unit sphere representation of the image plane [7] by relating the uniform spectral angle property in the image plane to great circles on the unit-sphere. This representation bounds the search space for vanishing points, since parallel lines meet at opposite poles of the unit-sphere. Additionally, it also solves the problem of finding correspondences between the spectral components related to each vanishing point.

Unfortunately, before reliable local spectra can be estimated, there needs to be an estimate of the local distortion of the texture so that the size of the spectral window can be set. However, this local distortion is, after all, the ultimate goal of perspective pose recovery. The main problem stems from the fact that if the window size is incorrectly set then the local estimate of the texture spectrum becomes defocussed. This defocusing in turn leads to poor estimation of perspective pose parameters when texture planes are subjected to severe perspective foreshortening. Distant texture

*Suported by Fundação Coordenação de Pessoal de Nível Superior (CAPES-BRAZIL), under grant: BEX1549/95-2

elements appear smaller while closer ones appear bigger. In order to accurately quantify the information provided by this texture gradient we must be able to locally adapt the size of the spectral window.

Most the spectral shape-from-texture methods opt to use a spectral window of fixed size. In other words, the accuracy of their perspective pose estimates are likely to be limited by defocusing. There are two exceptions. Garding and Lindeberg [12] address the scale problem employing a Gaussian scale-space decomposition locally over the structural primitives. Stone and Isard [13] have a method which interleaves the adjustment of local filters for adaptive scale edge detection and the estimation of planar orientation in an iterative feedback loop.

We use a spectral version of the idea proposed by Leung and Malik [14] for locating repeated elements in an image by measuring their affine similarities. The idea has also be exploited by Schaffalitzky and Zisserman [15]. Our method uses the local spectral distortion in order to compute the affine similarities and estimate the best local scale for the spectral descriptor.

2 Perspective Modelling

We commence by reviewing the projective geometry for the perspective transformation of points on a plane [8, 9]. Specifically, we are interested in the perspective transformation between the object-centred co-ordinates of the points on the texture plane and the viewer-centred co-ordinates of the corresponding points on the image plane. Suppose that the texture plane is a distance h from the camera which has focal length $f < 0$. Consider two corresponding points that have co-ordinates $\mathbf{X_t} = (x_t, y_t)^T$ on the texture plane and $\mathbf{X_i} = (x_i, y_i)^T$ on the image plane. The perspective transformation between the two co-ordinate systems is

$$\mathbf{X_i} = T_p \mathbf{X_t} \tag{1}$$

We represent the orientation of the viewed surface plane using the slant σ and tilt τ angles. This parametrisation is a natural way to model local surface orientation. For a given plane, the slant is the angle between viewer line of sight and the normal vector of the plane. The tilt is the angle of rotation of the normal vector around the line of sight axis. The elements of the transformation matrix T_p can be computed using the slant angle σ and tilt angle τ in the following manner

$$T_p = \frac{f\cos\sigma}{h - x_t \sin\sigma} \begin{bmatrix} \cos\tau & -\sin\tau \\ \sin\tau & \cos\tau \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{\cos\sigma} \end{bmatrix} \tag{2}$$

The perspective transformation in Equation 2 represents a non-linear geometric distortion of a surface texture pattern onto an image plane pattern. Unfortunately, the non-linear nature of the transformation makes Fourier domain analysis of the texture frequency distribution somewhat intractable. In order to proceed we therefore derive a local linear approximation to the perspective transformation. However, it should be stressed that the global quilting of the local approximations preserves the perspective effects required for recovering shape-from-texture. With this linear model, the perspective distortion can be represented by T_p^* which is the linear version of the perspective transformation given by Equation 2. We linearize T_p using a first-order Taylor formula. Let (xo_t, yo_t, h) be the origin or expansion point of the local coordinate system for the resulting affine transformation. This origin projects to the point (xo_i, yo_i, f) on the image plane. We denote the corresponding local coordinate system on the image plane by $\mathbf{X_i'} = (x_i', y_i', f)$ where $x_i = x_i' + xo_i$ and $y_i = y_i' + yo_i$. The linearised version of T_p in equation 2 is obtained through the Jacobian $J(.)$ of $\mathbf{X_i}$ where each partial derivative is calculated at the point $\mathbf{X_i'} = \mathbf{0}$. After the necessary algebra, the resulting linear approximation is

$$T_p^* = \frac{\Omega}{hf\cos\sigma} \begin{bmatrix} xo_i \sin\sigma + f\cos\tau\cos\sigma & -f\sin\tau \\ yo_i \sin\sigma + f\sin\tau\cos\sigma & f\cos\tau \end{bmatrix} \tag{3}$$

where $\Omega = f\cos\sigma + \sin\sigma\,(xo_i\cos\tau + yo_i\sin\tau)$. Hence, T_p^* depends on the expansion point (xo_i, yo_i) which is a constant. The transformation T_p^* in Equation 3 operates from the texture plane to the image plane. The model is similar to the scaled orthographic projection [16].

The net effect of the global perspective transformation is to distort the viewer-centred texture pattern in the direction of vanishing points on the image plane. In order to recover the parameters of perspective pose, we need to know the position of at least two different vanishing points in the image plane. Suppose that the two points are $\mathbf{V_1} = (x_{v1}, y_{v1})^T$ and $\mathbf{V_2} = (x_{v2}, y_{v2})^T$. Let the normal-vector to the texture plane be $N = (p, q, 1)$. The resulting normal vector components p and q are found by solving the system of simultaneous linear equations

$$\begin{bmatrix} x_{v1} & y_{v1} \\ x_{v2} & y_{v2} \end{bmatrix} \begin{bmatrix} p \\ q \end{bmatrix} = - \begin{bmatrix} f \\ f \end{bmatrix} \tag{4}$$

The solution parameters, p and q are $p = f\frac{y_{v1} - y_{v2}}{x_{v1}y_{v2} - x_{v2}y_{v1}}$ and $q = f\frac{x_{v2} - x_{v1}}{x_{v1}y_{v2} - x_{v2}y_{v1}}$. The slant and tilt angles are computed using $\sigma = \arccos\left(\frac{1}{\sqrt{p^2+q^2+1}}\right)$ and $\tau = \arctan\left(\frac{q}{p}\right)$. If more than two vanishing points are available, then the recovery of perspective pose parameters is over-constrained and can be effected by least-squares estimation.

3 Projective Distortion of the Power Spectrum

The Fourier transform provides a representation of the spatial frequency distribution of a signal. The novel contri-

bution in this section we show how local spectral distortion resulting from our linear approximation of the perspective projection of a texture patch can be computed using an affine transformation of the Fourier representation. We will commence by using an affine transform property of the Fourier domain [17]. This property relates the linear effect of an affine transformation A in the spatial domain to the frequency domain distribution. Suppose that $F(.)$ represents the Fourier transform of the image. Furthermore, let $\mathbf{X}$ be a vector of spatial co-ordinates and let $\mathbf{U}$ be the corresponding vector of frequencies. According to Bracewell *et al* [17], the distribution of image-plane frequencies $\mathbf{U_t}$ resulting from the Fourier transform of the affine transformation $\mathbf{X_i} = A\mathbf{X_t} + B$ is given by

$$F(\mathbf{U_i}) = \frac{1}{|det(A)|} e^{2\pi j \mathbf{U_t}^T A^{-1} \mathbf{B}} F[A^{-T}\mathbf{U_t}] \qquad (5)$$

In our case, the affine transformation is T_p^* as given in Equation 3 and there are no translation coefficients, i.e., $\mathbf{B} = \mathbf{0}$. As a result Equation 5 simplifies to:

$$F(\mathbf{U_i}) = \frac{1}{|det(T_p^*)|} F[T_p^{*-T}\mathbf{U_t}] \qquad (6)$$

In other words, the effect of the affine transformation of coordinates T_p^* induces an affine transformation T_p^{*-T} on the texture-plane frequency distribution. The spatial domain transformation matrix and the frequency domain transformation matrix are the inverse transpose on one-another.

We will consider here only the affine distortion over the frequency peaks, i.e., the energy amplitude will not be considered in the analysis. For practical purposes we will use the local power spectrum as the spectral representation of the image. This describes the energy distribution of the image as a function of its frequency content. In this way we will ignore complications introduced by phase information. Using the power spectrum, small changes in phase due to translation will not affect the spectral information and hence Equation 6 will hold. The power spectrum representation of an image $f(\mathbf{X_t})$ may be defined as the Fourier transform of the autocorrelation function of the image.

In order to obtain a smooth spectral response we use the Blackman-Tukey power spectrum estimator. This is the frequency response of the windowed autocorrelation function. We employ a triangular smoothing window $w(\mathbf{X})$ [18] due to its stable spectral response. The spectral estimator is then

$$P(\mathbf{U_i})^{BT} = \mathcal{F}\{c_{xx}(\mathbf{X_i}) \times w(\mathbf{X_i})\} \qquad (7)$$

Where c_{xx} is the estimated autocorrelation function of the image patch. Our overall goal is to consider the effect of perspective transformation on the power-spectrum. In practice, however, we will be concerned with semi-periodic textures in which the power spectrum is strongly peaked. In this case we can confine our attention to the way in which the dominant frequency components transform. According to our affine approximation and Equation 6, the way the Fourier domain transforms locally is governed by

$$\mathbf{U_i} = T_p^{*-T}\mathbf{U_t} \qquad (8)$$

This spectral property has also been exploited by Rosenholtz and Malik [10] in their work on local shape-from-texture.

4 Local Spectral Frequency Under Unit-Sphere Mapping

In this section we follow Barnard [7] and model the image plane in terms of spherical coordinates by projecting it onto a unit-sphere centred at the optical centre. This projection simplifies the representation of the perspectivity of the texture plane and the search for its vanishing points. The main advantage is that unlike the image plane, the unit-sphere is a closed space parametrised the two angles of azimuth and zenith or elevation. Spherical projections of the image plane have been exploited by several authors [19, 7]. However, they have employed structural representations of texture. Instead, we use the local spectral frequency to model texture.

Figure 1 illustrates the projection geometry. The unit-sphere is placed at the focal point and the image plane lies at a distance f along the optical axis. For each point on the image plane, the position (x_i, y_i) and the measured spectral angle α_i specify the equation of a line. The orientation $\theta = \alpha_i$ and the normal distance is given by $r_\theta = \sqrt{x_i^2 + y_i^2}\cos(\phi_i - \alpha_i)$ (where $\phi_i = \arctan\frac{y_i}{x_i}$). Each such

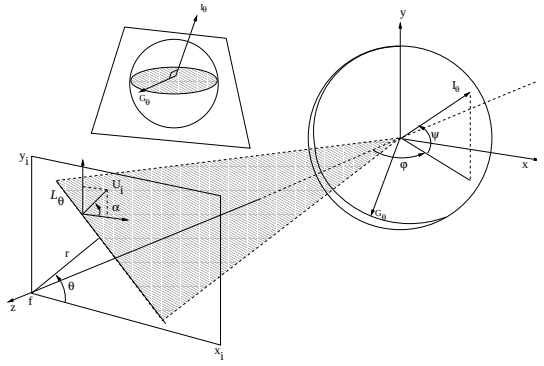


Figure 1: Projecting lines from the image plane onto the unit-sphere.

line, $\mathcal{L}_\theta$ projects onto a great circle on the unit-sphere. The great circles are constructed by intersecting the sphere by the plane that both contains the line $\mathcal{L}_\theta$ and the center of the sphere. Let $\mathbf{G}_\theta$ be a vector that points from the centre of the unit-sphere to a point on the corresponding great circle. If

the point has azimuth φ and elevation ψ, then the vector is given by $\mathbf{G}_\theta = (\sin\varphi\cos\psi, \sin\psi, \cos\varphi\cos\psi)$. Suppose that I_θ is the normal vector of the plane which contains the great circle. Figure 1 shows in the geometry of the plane normal vector I_θ and the great circle generator vector G_θ. Since the two vectors are perpendicular to one-another they satisfy the condition that $\mathbf{G}_\theta.\mathbf{I}_\theta = 0$ The azimuth and elevation angles of points on the great-circle are related to the normal distance parameters of the straight-line in the following manner

$$f\left(\sin\theta.\tan\psi + \cos\theta\sin\varphi\right) = r_\theta\cos\varphi \qquad (9)$$

5 Accumulation on the unit-sphere

Having established the relationship between spectral angle and vanishing point location, we are now in a position to develop an accumulation algorithm on the unit-sphere. We exploit the following two properties to map the search for vanishing points on to the unit-sphere:

PROPERTY 1 (SPECTRAL FREQUENCY ANGLE CONSTANCY). If $\mathcal{L}_\theta$ is a line radiating from a vanishing point on the image plane, then every local spectral distribution taken at points belonging to $\mathcal{L}_\theta$ will have a constant spectral angle α. Conversely, each spectral angle α estimated from a local frequency distribution on the image plane specifies the equation of a line $\mathcal{L}_\theta$ which radiates from a corresponding vanishing point [11].

We now exploit Property 1 to directly relate the local spectral angle to great circles on the unit-sphere. Using the equality between the angles θ and α and using the expression for a great circle in Equation 9, we find

$$\psi = -\arctan\frac{f\cos\alpha\sin\varphi - r_\theta\cos\varphi}{f\sin\alpha} \qquad (10)$$

PROPERTY 2 (FROM SPECTRAL FREQUENCY ANGLES TO GREAT CIRCLES). Each spectral angle α estimated from the local frequency distribution centred at a point on the image plane maps to a great circle on the unit-sphere. When several great circles intercept on the unit-sphere, then the corresponding image-plane spectra will have originated from a common vanishing point.

To compute the spectral angle distribution, we require a way of sampling the local power spectrum. In particular we need a sampling procedure which provides a means of recovering the angular orientation information residing in the peaks of the power-spectrum. We accomplish this by simply searching for local maxima over a filtered representation of the local power spectrum. Since we are interested in the angular information rather than the frequency contents of the power spectrum, we ignore the very low frequency components of the power-spectrum since these mainly describe micro-texture patterns or very slow energy variations. Providing that we have at least two representative spectral peaks we can directly generate line directions according to the angular constancy property. We can use as many distinct spectral components as we can estimate. However, a two component decomposition is sufficient for our purposes. We extract angular decompositions for the local power spectra at several locations on the image plane. Using Equation 10 we accumulate evidence for the intersections of great circles on the unit-sphere. To do this we quantise the unit-sphere into accumulator cells of approximately equal area. Each great circle is traced across the unit-sphere and the vote count is incremented each time it crosses a new accumulator cell. Vanishing points are located in cells which have accumulated local voting maxima. Once two or more intersection points are located, then the perspective pose of the plane can then be determined as described in Section 2.

6 Adapting the Scale of the Local Spectral Descriptor

As pointed out earlier, the selection of the physical scale for local spectral descriptors is critical to accurate shape from texture. To illustrate this point, Figure 2 shows a doubly sinusoidal texture viewed under perspective projection. The perspective distortion of the texture is controlled by the slant-angle of the texture-plane, which in this case is in the horizontal direction. The main feature to note is that the spatial frequency of the texture increases in the slant direction. To accurately estimate the frequency content using local spectral descriptors, it is important that the local sampling window is adapted to avoid distortion of the measurements. If the size of the window is fixed, then the computed descriptors will not accurately reflect the local spectral content at the relevant point on the image plane. Moreover, if the window is too large then the texture spectra will be defocussed. The top row shows the spectra estimated with a fixed window, while the lower row shows the spectra obtained if an optimally sized sampling window is used. The main feature to note is that the spectra estimated with the fixed size window are blurred. The spectra estimated with the optimally sized window, on the other hand, are crisp. In order to adapt the size of the spectral window and hence improve the quality of the local spectral information we use a spectral version of the idea proposed by Leung and Malik [14]. The process is based on the assumption that the distortion between two patches in the image plane can be approximated by an affine transformation [8, 10]. This affine transformation can be also computed in the spectral domain by using the Bracewell theorem [17] which relates the spectral distortion of local patches under affine transformation. The idea is basically to track the spectral pattern at

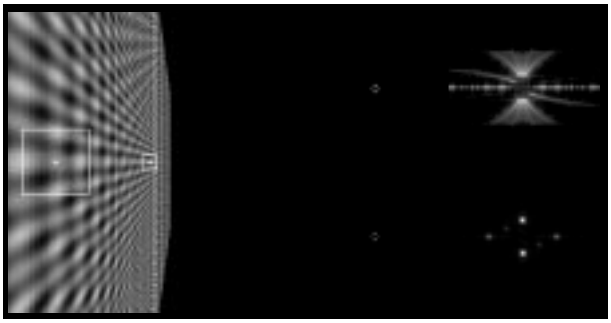


Figure 2: Projected artificial texture with squares showing the window sizes employed by the spectral estimator together with the power spectrum response using a fixed data window and our adaptive window.

different locations over the image plane allowing the scale to adapt by minimising the affine error between the two local spectra. The process consists of the following steps: (1) given an initial spectrum, find the affine transformation between two local spectra for several scales up to a maximum scale allowed; (2) estimate the backprojection affine error for each scale; (3) find the scale that gives the minimum error between the spectra after backprojection.

The local spectra are assumed to be related by a 2x2 affine transformation matrix Φ. The transformation governing the affine projection of the spectral component U_1 is

$$\mathbf{U}_1' = \Phi \mathbf{U}_1 \tag{11}$$

Our aim is to track the spectra across the image plane so as to minimise the degree of defocusing. Suppose U_1 is a crisply estimated spectra and that $U_2(s)$ is a neighbouring but defocussed spectrum estimated with a window size s. Our tracking process consists of interleaved iterations. First, we find the least squares affine parameter matrix $\Phi_{1,2}$ between the two spectra which satisfies the condition

$$\Phi_{12} = \arg\min_{\Phi} (U_2(s) - \Phi U_1)^T (U_2(s) - \Phi U_1) \tag{12}$$

The solution to this overdetermined problem is

$$\Phi = \left(\mathbf{U}_1^T \mathbf{U}_1\right)^{-1} \mathbf{U}_1^T \mathbf{U}_2(s) \tag{13}$$

Once the estimates of the affine parameters are to hand, then we can adjust the size of the spectral window s, so that the Euclidean distance between the estimated spectra $U_2(s)$ and projected spectra $\Phi_{12} U_1$ is minimised. The optimal scale satisfies the condition

$$s = \arg\min_{s} ||U_2(s) - \Phi_{1,2} U_1|| \tag{14}$$

The affine projection and scale re-estimation steps are interleaved and iterated to convergence. It is important to point out that we need at least two different spectral components to allow the transformation matrix be estimated. This is not a difficult requirement to be achieved in most of nearly regular textures. This assumption is only not valid for pure directional line textures or very random texture where the detection of spectral peaks is very difficult to be achieved. Moreover, we consider here only the affine distortion over the most energetic frequency peaks. The energy amplitude of the peaks is considered in the analysis as a guide to provide the correspondence between them.

7 Experiments

In this section we provide some results which illustrate the accuracy of the planar pose estimation achievable with our shape-from-texture algorithm. This evaluation is divided into two parts. We commence by considering textures with known ground-truth slant and tilt. This part of the study is based on projected Brodatz textures [20]. The second part of our experimental study focuses on real texture planes where the ground truth is unknown.

7.1 Synthetic texture planes

In Figure 3, we have taken three different texture images from the Brodatz album and have projected them onto planes of known slant and tilt. The textures are regular real textures of almost uniform element distribution. Superimposed on the projected textures are the estimated lines radiating from the corresponding vanishing points as given by our algorithm. Figure 4 shows the back-projection of the images onto the recovered texture plane. In most cases there is little residual perspective distortion. The main feature to note is that the method performs well even when the texture plane is highly inclined.

7.2 Real World Examples

This part of the experimental work focuses on real world textures with unknown ground-truth. The textures used in this study are two views of a brick-wall, a York pantile roof and the lattice casing enclosing a PC monitor. The images were collected using a Kodak DC210 digital camera and are shown in Figure 5. There is some geometric distortion of the images due to camera optics. This can be seen by placing a ruler or straight-edge on the brick-wall images and observing the deviations along the lines of mortar between the bricks.

Superimposed on the images are the lines radiating from the vanishing points. In the case of the brick-wall images these closely follow the mortar lines. In Figure 6 we show the back-projection of the textures onto the fronto-parallel plane using the estimated orientation angles. In the case of the brick-wall, any residual skew is due to error in

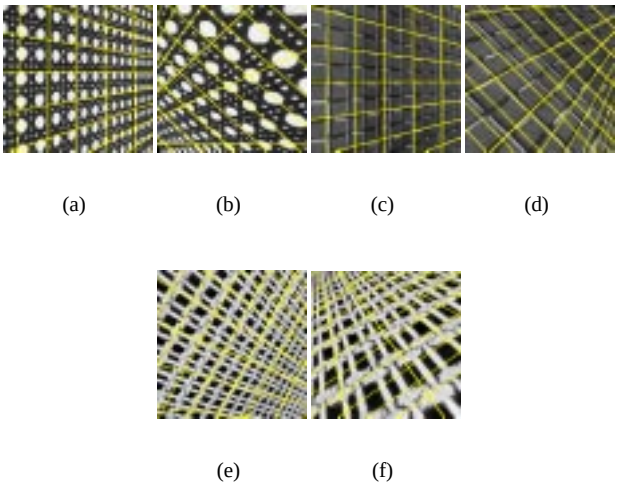


Figure 3: *Brodatz textures. (a)-(b) D101; (b)-(c) D1; (d)-(e) D20.*

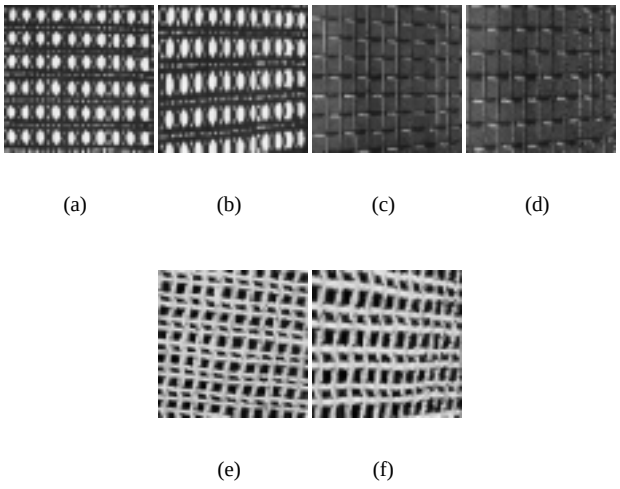


Figure 4: *Back-projected Brodatz textures. (a)-(b) D101; (b)-(c) D1; (d)-(e) D20.*

the estimation of the slant and tilt parameters. It is clear that the slant and estimates are accurate but that there is some residual skew due to poor tilt estimation.

The final set of real world experiments focusses on recovering the vanishing points for large man-made planar objects. The images are provided by indoor views of our lab and views of skyscrapers collected from the Internet. These images contain both texture information and bounding rectangles that can be used to confirm by visual inspection the direction of the vanishing point. These lines are not used in the analysis. For instance in the indoor scenes there are rectangular beams and gantries which converge to the vanishing point. In the skyscraper images, the sides of the planar faces can be used to confirm the accuracy of the vanishing point.

In Figure 7 we show the reconstructed perspective planes superimposed on the original images. There are several examples of both the indoor scenes and the skyscraper images. The textures used in the analysis are quite varied. For instance, in the indoor scenes there are rectangular ventilation ducts on the roofing and webbing on the gantry. The skyscraper textures are largely composed of rectangular window patterns. The planes are visualised by drawing radial lines from the vanishing point. In all cases the reconstructed radial lines from the vanishing point match the bounding rectangles. In other words, the reconstructions are good.

8 Conclusions

We have described an algorithm for estimating the perspective pose of textured planes by projecting spectral information onto the unit-sphere. We exploit the fact that the local spectral components are oriented in the direction of vanishing points. As a result each estimated spectral component can be mapped onto a great circle of the unit-sphere. Vanishing points are characterised by locations at which several great circles intercept. Based on this observation, we pose the problem of estimating perspective pose as that of searching for accumulator cells of maximum contents on the unit-sphere.

To overcome the problem of defocusing, we also introduce a technique for adaptively setting the size of the local spectral window for estimating perspective pose from frequency information.

The method is illustrated to operate effectively on both synthetic imagery with known ground truth and on a wide variety of real-world textured planes. One advantage of the method is that it does not rely on potentially unreliable estimates of texture gradient to constrain the tilt angle.

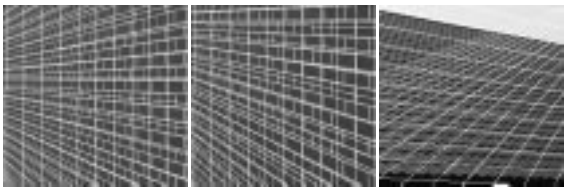
Our future plans revolve around using the method to estimate shape from the texture distribution of curved objects. Suffice to say that studies aimed at addressing this topic are in hand and will be reported in due course.

References

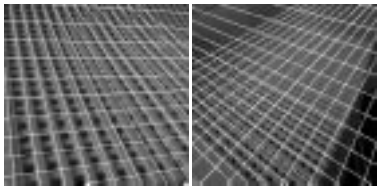
[1] Marr D. *Vision: A computational investigation into the human representation and processing of visual information*. Freeman, 1982.

[2] Gibson J. J. *The perception of the visual world*. Houghton Miffin, Boston, 1950.

[3] Cutting E.J. and Millard R.T. Three gradients and the perception of flat and curved surfaces. *Journal of Experimental Psychology:general*, 113(2):198–216, 1984.

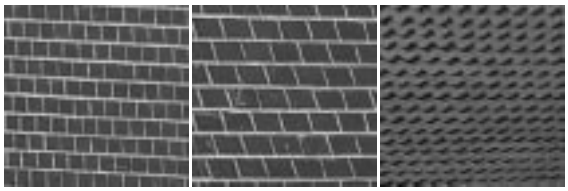


(a) Brick Wall (b) Brick Wall (c) Roof

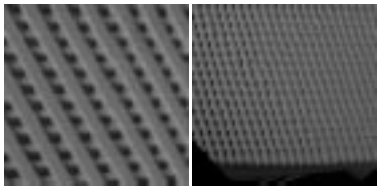


(d) PC casing (e) PC casing

Figure 5: *Outdoor texture images.*



(a) Brick wall (b) Brick wall (c) Roof



(d) PC casing (e) PC casing

Figure 6: *Back-projected Outdoor images.*

[4] Stevens K.A. On gradients and texture "gradients". *Journal of Experimental Psychology:general*, 113(2):217–220, 1984.

[5] Kender J.R. Shape from texture: an aggregation transform that maps a class of texture into surface orientation. In *6th IJCAI, Tokyo*, pages 475–480, 1979.

[6] Kwon J.S., Hong H.K., and Choi J.S. Obtaining a 3-d orientation of projective textures using a morphological method. *Pattern Recognition*, 29:725–732, 1996.

[7] Barnard S.T. Interpreting perspective images. *Artificial Intelligence*, 21:435–462, 1983.

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

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

[10] Malik J. and Rosenholtz R. Recovering surface curvature and orientation from texture distortion: a least squares algorithm and sensitive analysis. *Lectures Notes in Computer Science - ECCV'94*, 800:353–364, 1994.

[11] Ribeiro E. and Hancock E.R. 3-d planar orientation from texture: estimating vanishing point from local spectral analysis. In *IX British machine Vision Conference*, pages 326–335, Sep. 1998.

[12] Garding J. and Lindeberg T. Direct computation of shape cues using scale-adapted spatial derivatives operators. *International Journal of Computer Vision*, 17(2):163–191, 1994.

[13] Stone J.V. and Isard S.D. Adaptive scale filtering: A general method for obtaining shape from texture. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 17(7):713–718, 1995.

[14] Leung T. and Malik J. Detecting, localizing and grouping repeated scenes elements from a image. In *ECCV, LNCS 1064*, pages 546–555, 1996.

[15] Schaffalitzky F. and Zisserman A. Geometric grouping of repeated elements within images. In *British Machine Vision Conference*, pages 13–22, 1998.

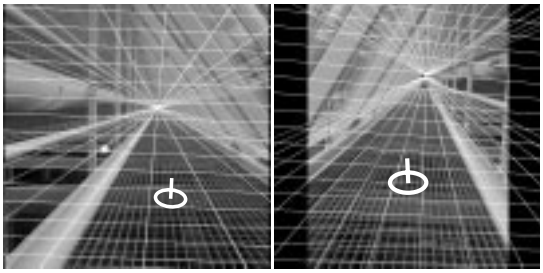
[16] Aloimonos J. Perspective approximations. *Image and Vision Computing*, 8(3):179–192, August 1990.

[17] Bracewell R.N., Chang K.-Y., Jha A.K., and Wang Y.-H. Affine theorem for two-dimensional fourier transform. *Electronics Letters*, 29(3):304, 1993.

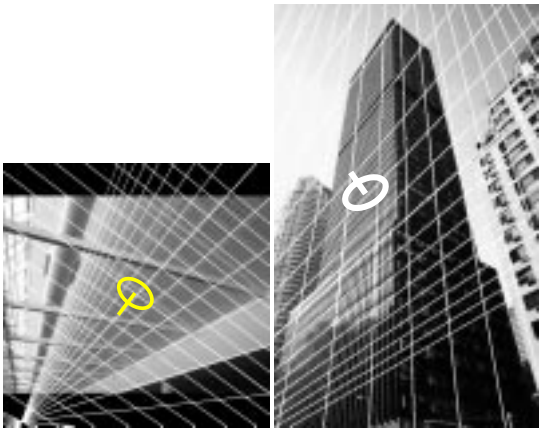
[18] Kay S.M. *Modern Spectral Estimation:Theory and Application*. Prentice Hall, 1988.

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

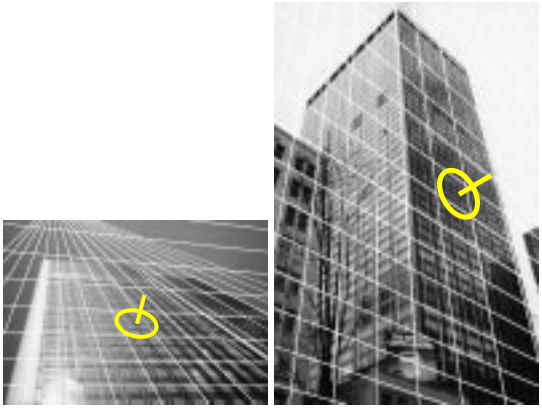
[20] Brodatz P. *Textures: A Photographic Album for Artists and Designers*. Dover, New York, 1966.



(a) (b)



(c) (d)



(e) (f)

Figure 7: *Lab scenes and skyscraper images.*