

Estimating the 3D orientation of texture planes using local spectral analysis

E. Ribeiro¹, E.R. Hancock*

Department of Computer Science, University of York, York YO1 5DD, UK

Received 9 December 1998; received in revised form 7 October 1999; accepted 24 October 1999

Abstract

This paper describes a new method for recovering the perspective geometry of textured planes from local spectral analysis. The novel contribution of the work is to use the angular distribution of texture moments to locate the vanishing point of textured planes viewed under perspective geometry. Our key observation is that lines of uniform spectral orientation radiate from the vanishing point. We exploit this property to develop a simple algorithm for estimating the slant and tilt of textured planes. This is a two-step process. It commences by computing the Fourier power spectrum at each of a series of local neighbourhoods and identifying contours which connect the local spectral moments of equal power. As noted by Sevens (1981), these contours are perpendicular to the tilt direction. The second step is to triangulate the vanishing point. We do this by using a correlation method to identify lines of uniform spectral orientation that originate from the vanishing point. This second property represents the novel contribution of the paper. We evaluate the new algorithm on both synthetic imagery with known ground-truth and on real-world data. © 2000 Elsevier Science B.V. All rights reserved.

Keywords: Texture planes; Local spectral analysis; Vanishing point

1. Introduction

Shape-from-texture was identified by Marr [2] as one of the key ingredients of the $2\frac{1}{2}$ D sketch. Moreover, the feasibility of the process is well-grounded in psychophysics [3]. From a computational perspective, the main obstacle to practical shape-from-texture is the fact that the identification of surface-markings is in itself insufficient for the recovery of shape from a single image. When posed in a monocular framework the shape-from-texture problem is ill-defined. The basic problem stems from the fact that the local distortions of the texture elements are not entirely attributable to variations in surface orientation. The process must also account for foreshortening effects due to global perspective geometry, together with natural variations in the texture pattern. Moreover, even if these additional factors can be controlled, there is ambiguity in the recovery of surface depth due to the fact that the overall scale of the texture elements may be unknown.

1.1. Related literature

It is for these reasons that the process of shape-from-texture must be constrained using certain model assumptions. The most powerful of these are the isotropy and homogeneity assumptions. Isotropy asserts that the density of texture elements has no detectable preferential direction when viewed in a fronto-parallel manner. Homogeneity requires that the density of texture elements is uniform. There are two distinct schools of thought as to how these constraints should be exploited. According to the structural school it is the shape or local geometry of texture primitives that should be used to infer the overall perspective geometry and hence determine changes in local surface orientation. By contrast, the spectral method attempts to infer shape from distortions in the frequency representation of texture.

Structural methods use the geometry of edges, lines or arcs to identify perspective distortion using shape cues [4–6]. Unfortunately, the segmentation of the texture primitives is a process of extreme fragility, and in most cases the techniques have only been demonstrated on synthetic or highly contrived data. Some of the difficulties associated with the need for accurate geometric information can be overcome by working with statistics or texture moments

*Corresponding author. Tel.: +44-1904-43-3374; fax: +44-1904-43-2767.

E-mail address: erh@cs.york.ac.uk (E.R. Hancock).

¹Supported by CAPES-BRAZIL, under grant: BEX1549/95-2.

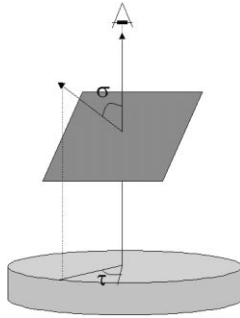


Fig. 1. Slant and tilt parametrization of the plane orientation.

defined either over single primitives or groups of primitives [7–10].

The use of frequency domain or spectral measurements represents a way of overcoming some of the restrictive requirements imposed by the need to work with accurately segmented shape primitives. Moreover, it increases the range of natural texture that can be accommodated. Frequency properties can be captured using a number of different representations including the Fourier transform, the wavelet function and Gabor wavelets [11–14]. Interrogation of the literature reveals that it was Bajcsy and Lieberman [15] who were among the first to use frequency domain methods for shape-from-texture in an application involving the recovery of 3D surface orientation in outdoor scenes. Stevens [1] provided a computational framework based on the perception of perspectivity from texture gradients. The more recent study of Krumm and Shafer [16] recovers slant angles for periodic surface textures in which there is a dominant peak frequency. The method has a limited geometrical model which only accommodates a unidimensional local frequency distortion. This restriction is lifted by the method of Super and Bovik [12,17] which employs 2D Gabor wavelets to detect the local projective distortion in the power spectrum. Krumm and Shafer [16] have shown how to relate the frequency distortion between local patches under an affine transform. Finally, Garding [18] has shown how the analysis of perspectivity can be simplified by searching in the direction perpendicular to the gradient direction, i.e. along the tilt axis.

1.2. Paper overview

The work reported in this paper provides a bridge between the geometrically intuitive structural methods and their more robust frequency domain counterparts. The aim is to develop a geometric algorithm that can be used to estimate perspective deformation from spectral distribution. This avoids the need for iterative optimisation or relaxation methods [16,19]. The method uses vanishing point estimation to determine planar orientation. In this respect it can be regarded as a frequency domain counterpart of several structural methods. For instance Kender's [20]

agglomerative method uses the Hough-transform to recover vanishing points from line-textures. Another structural approach to vanishing point recovery is to use the area gradient of the primitives. A simple version of this idea was reported by Ohta et al. [21] who places strong reliance on the ability to accurately estimate the primitive area. A more practical scheme has recently been reported by Kwon [22] who has shown how mathematical morphology can be used to determine the size of the texture primitives. However, this refinement is sensitive to the choice of the structuring element shape.

In this paper we present a geometrically intuitive algorithm which allows vanishing point, and hence perspective distortions, to be recovered from spectral information. Our method commences by searching for contours of equal spectral power on the image plane perpendicular to the tilt direction. For planar textures, these contours correspond to lines. They connect spectral moments of equal area but different orientation. The direction of the vanishing point is determined from the relative orientation of the spectral moments. Here we make use of the observation that lines that radiate from the vanishing point connect points which have identically oriented spectral distributions. Our proof of this new property is based on the affine projection of texture moments. Although the texture-moment representation has also been exploited by Kender [20] and by Kwon [22] among others, our novel contribution is to demonstrate angular spectral constancy along line radiating from a vanishing point. This new observation suggests a simple algorithm for triangulating the vanishing point. By correlating the angular power distribution we search for lines along which there is a uniform spectral orientation. Once a sufficiently large sample of such lines is to hand, the vanishing point may be reconstructed by intersection. Since spectral information is independent of texture primitive shape and size, this geometrically inspired method does not require any pre-segmented image information. It is therefore applicable to a wide diversity of real-world textures.

The outline of this paper is as follows. Section 2 outlines the perspective geometry involved in the viewing of 2D planes embedded in a 3D space. In particular, we concentrate on the relationship between the slant and tilt angles of the texture-plane and the position of the vanishing point on the image-plane. In Section 3, we describe how perspective projection transforms the local spectral content of a textured image. In Section 4, we focus on the issue of recovering vanishing point from local spectral analysis. In this section, we point out that lines originating from the vanishing point are associated with uniform spectral orientation. This new angular property is the key ingredient for developing a practical method for recovering the perspective geometry of textured planes. Section 5 details the steps of our algorithm. A number of experiments are presented in Section 6 aimed at demonstrating the practical potential of our method. Finally, Section 7 offers some conclusions and suggests directions for future work.

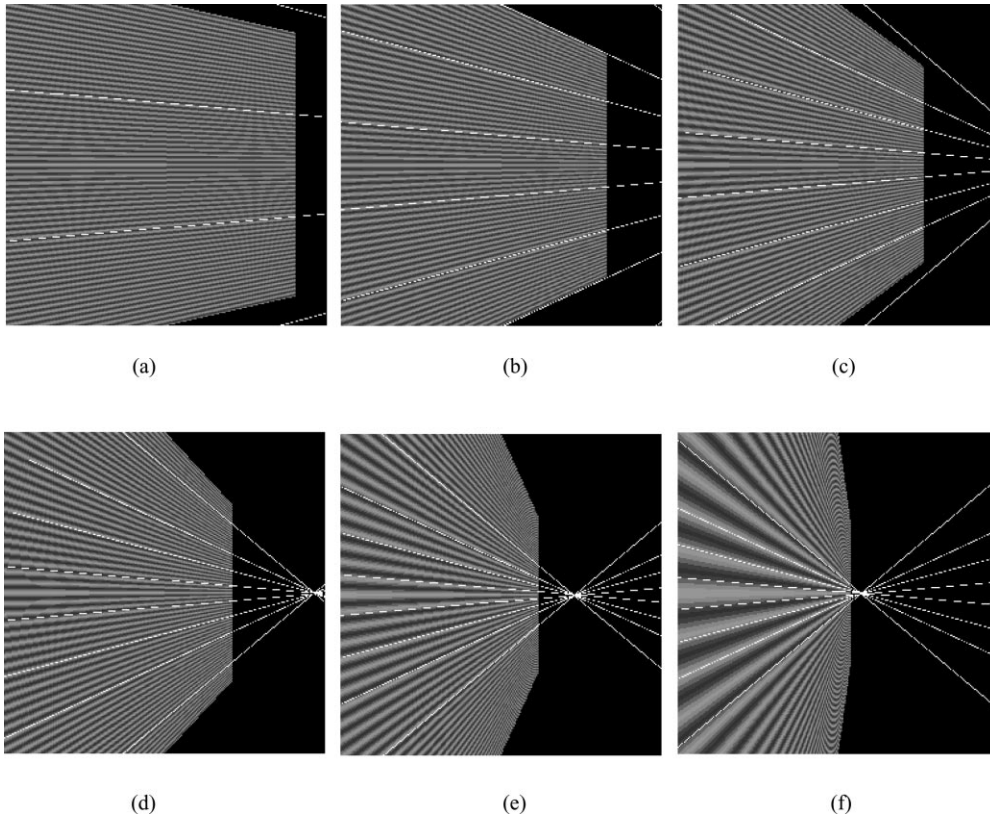


Fig. 2. Lines with uniform spectral orientation together with the sinusoid image for different slant angles. The model for the spectral lines matches accurately to the image response. The spectral component is $\mathbf{U}_i = (0, 0.125)$. (a) $\sigma = 10^\circ$, (b) $\sigma = 20^\circ$, (c) $\sigma = 30^\circ$, (d) $\sigma = 40^\circ$, (e) $\sigma = 60^\circ$, (f) $\sigma = 80^\circ$.

2. Geometric modelling

We commence by reviewing the projective geometry for the perspective transformation of points on a plane. 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. We assume that the two co-ordinate systems share a common origin. Both planes are assumed to be parallel to the x - and y -axis of their relevant system of co-ordinates. The image plane is taken as having zero z -intercept while the texture-plane resides at a height h above the common origin. We represent the orientation of the viewed surface plane using the slant angle σ and tilt angle τ . This parametrisation is a natural way to model local surface orientation and is illustrated in Fig. 1. In a viewer centred co-ordinate system the slant is the angle between line-of-sight and the normal vector to the plane. The tilt is the angle of rotation of the normal vector about the line-of-sight.

We are interested in the transformation between points on the texture plane and their counterparts on the image plane. Suppose that the texture plane is viewed with a camera of focal length f . We let $\mathbf{X}_t = (x_t, y_t)^T$ denote a point on the texture plane and $\mathbf{X}_i = (x_i, y_i)^T$ denote the corresponding point on the image plane. The transformation between the texture-plane and image-plane co-ordinate systems is given

by

$$\mathbf{X}_i = T_p \mathbf{X}_t \tag{1}$$

The inverse transformation is given by

$$\mathbf{X}_t = T_p^{-1} \mathbf{X}_i \tag{2}$$

The transformation matrix T_p is computed using the slant and tilt angles in the following manner:

$$T_p = \frac{f}{h - x_t \sin \sigma} \begin{bmatrix} \cos \sigma \cos \tau & -\sin \tau \\ \cos \sigma \sin \tau & \cos \tau \end{bmatrix} \tag{3}$$

The net effect of this transformation is to distort the viewer-centred parallel line patterns in the direction of the vanishing point $\mathbf{V} = (x_\infty, y_\infty)^T$ on the image plane. Suppose that the object-centred texture pattern consists of a family of parallel lines which are oriented in the direction of the vanishing point. When transformed into the image-centred co-ordinate system, this family of lines can be represented using the following set of parametric equations:

$$\begin{pmatrix} x_t \\ y_t \\ f \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 + \lambda \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} \tag{4}$$

where λ is the parametric variable of the individual lines. The three constants in $\mathbf{B} = (b_1, b_2, b_3)$ are the direction

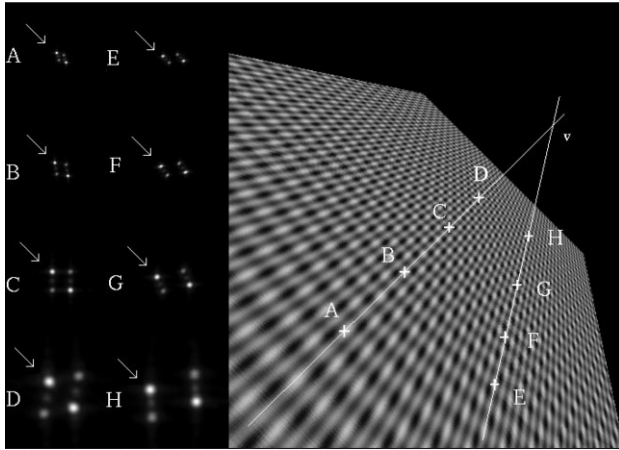


Fig. 3. Angular spectral components over lines pointing to a vanishing point.

cosines for the entire family. The individual lines in the family are each parametrised by the vector $\mathbf{A} = (a_1, a_2, a_3)^T$.

Under perspective projection, the direction cosines for the family of parallel lines are related to the position of the vanishing point (x_∞, y_∞) in the viewer-centred co-ordinate

system in the following manner:

$$\begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \frac{1}{\sqrt{x_\infty^2 + y_\infty^2 + f^2}} \begin{pmatrix} x_\infty \\ y_\infty \\ f \end{pmatrix} \tag{5}$$

Suppose that the vector $\mathbf{N} = (p, q, 1)^T$ represents the surface-normal to the texture-plane in the viewer-centred co-ordinate system of the image. Since every line lying on the texture-plane will be perpendicular to this normal vector,

$$\mathbf{N} \cdot \mathbf{B} = pb_1 + qb_2 + b_3 = 0 \quad \text{and} \quad \tan \tau = \frac{p}{q} \tag{6}$$

Using Eqs. (5) and (6), the 3D planar surface orientation can be completely recovered. In the next section we describe how the vanishing point coordinates \mathbf{V} can be estimated from local spectral information.

3. Projective distortion of the power spectrum

The spectral representation of a signal, normally termed the *spectral density function* or *power spectrum*, describes the energy distribution of the signal as a function of its

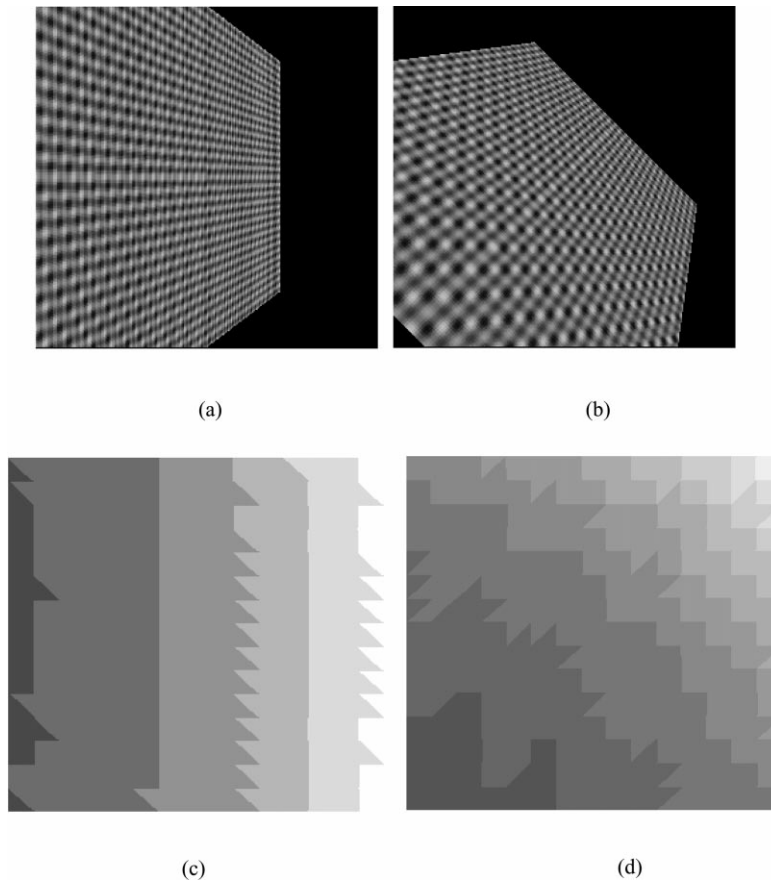


Fig. 4. Artificial textures and their gradient contour maps.

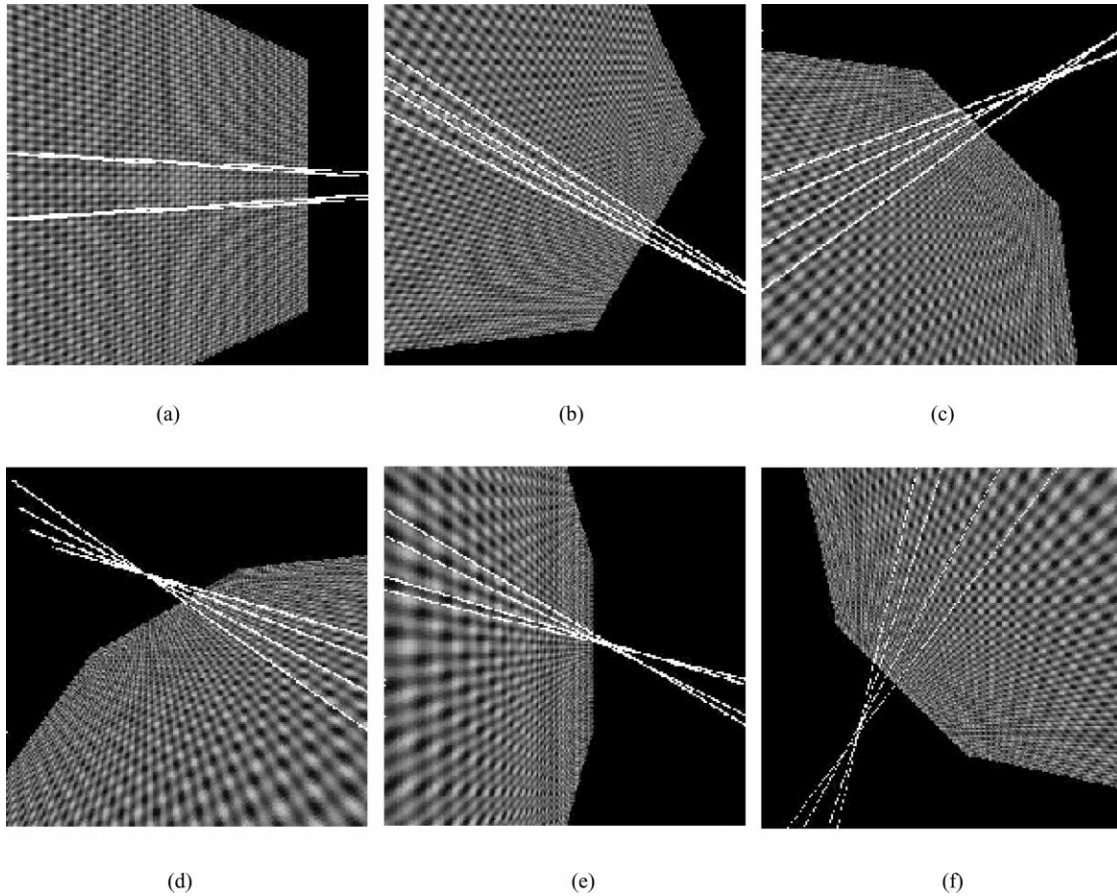


Fig. 5. Artificial texture images: Group 1—sinusoidal texture images. (a) $\sigma = 20^\circ$ and $\tau = 0^\circ$; (b) $\sigma = 30^\circ$ and $\tau = 330^\circ$; (c) $\sigma = 45^\circ$ and $\tau = 45^\circ$; (d) $\sigma = 60^\circ$ and $\tau = 120^\circ$; (e) $\sigma = 70^\circ$ and $\tau = 0^\circ$; (f) $\sigma = 50^\circ$ and $\tau = 225^\circ$.

frequency content. The power spectrum representation of a image $f(x,y)$ may be defined as the Fourier transform of image autocorrelation (which is always non-negative by definition). We commence by considering the power-spectrum in the object-centred co-ordinate system of the texture plane. If u_t and v_t are the frequencies in the x_t and y_t directions, then the power spectrum in the texture plane is given by

$$P_t(u_t, v_t) = T_1 T_2 \sum_{k=-p_1}^{p_1} \sum_{l=-p_2}^{p_2} r_x(k, l) \times \exp[-j2\pi(kuT_1 + lvT_2)] \quad (7)$$

here $r_x(k, l)$ is the autocorrelation function of the image.

Our overall goal is to consider the effect of perspective transformation on the power-spectrum. If $\mathbf{U}_i = (u_i, v_i)^T$ and $\mathbf{U}_t = (u_t, v_t)^T$ are, respectively, the image-plane and texture-plane frequencies, then they are related to one another by a frequency-domain Jacobian [12]. In practice, however, we will be concerned with periodic textures in which the power spectrum is strongly peaked. Although this is a restrictive assumption, the built environment contains numerous examples

of man-made and regularly textured planar surfaces. Examples include brick-walls, tiled roofs together with patterned floor and wall coverings. Subject to this restriction, we can confine our attention to the way in which the dominant frequency components transform. If the peaks are narrow, then we can restrict ourselves to considering the transformation of instantaneous frequency components rather than the complete power-spectrum. Under this assumption, the corresponding transformed viewer-centred frequency is given by

$$\mathbf{U}_i = J^t \mathbf{U}_t \quad (8)$$

where J^t is the transpose of the Jacobian of \mathbf{X}_i under the inverse perspective transformation T_p^{-1} given by Eq. (3). The Jacobian J is given by

$$J = J(\mathbf{X}_i) = \begin{bmatrix} \frac{\partial}{\partial x_i} x_t(x_i, y_i) & \frac{\partial}{\partial y_i} x_t(x_i, y_i) \\ \frac{\partial}{\partial x_i} y_t(x_i, y_i) & \frac{\partial}{\partial y_i} y_t(x_i, y_i) \end{bmatrix} \quad (9)$$

As a result, the instantaneous frequency peaks transform

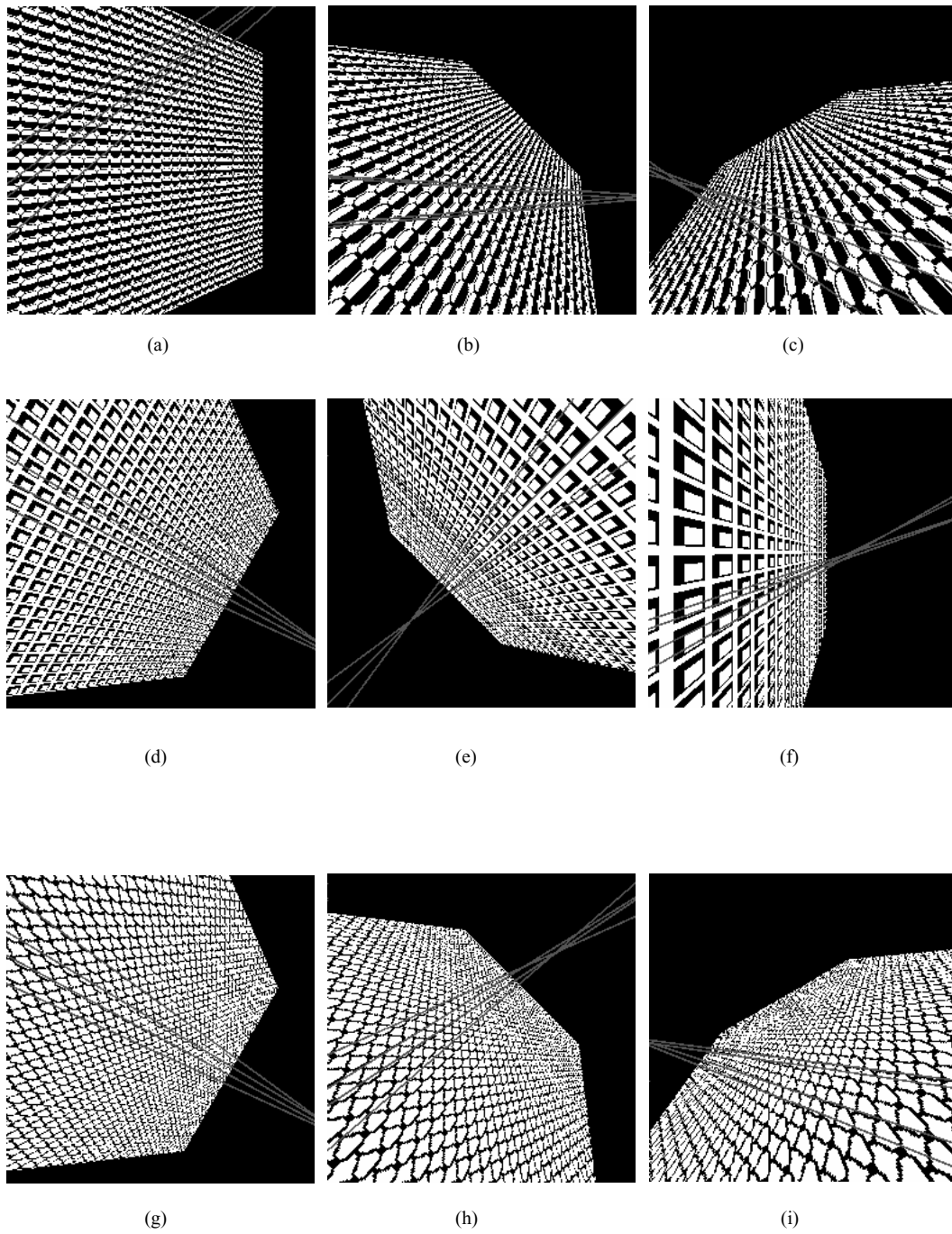


Fig. 6. Artificial texture images: Group 2—regular geometric textures. (a) $\sigma = 20^\circ$ and $\tau = 0^\circ$; (b) $\sigma = 45^\circ$ and $\tau = 45^\circ$; (c) $\sigma = 60^\circ$ and $\tau = 120^\circ$; (d) $\sigma = 30^\circ$ and $\tau = 330^\circ$; (e) $\sigma = 50^\circ$ and $\tau = 225^\circ$; (f) $\sigma = 70^\circ$ and $\tau = 0^\circ$; (g) $\sigma = 30^\circ$ and $\tau = 330^\circ$; (h) $\sigma = 45^\circ$ and $\tau = 45^\circ$; (i) $\sigma = 60^\circ$ and $\tau = 120^\circ$.

with the matrix [12]

$$J^t = \frac{-k^2 \sin \sigma}{h} \begin{bmatrix} x_i \cos^2 \tau + y_i \sin \tau \cos \tau & -x_i \sin \tau \cos \sigma \cos \tau + y_i \cos^2 \tau \cos \sigma \\ x_i \cos \tau \sin \tau + y_i \sin^2 \tau & -x_i \sin^2 \tau \cos \sigma + y_i \cos \tau \cos \sigma \sin \tau \end{bmatrix} + k \begin{bmatrix} \cos \tau & -\sin \tau \cos \sigma \\ \sin \tau & \cos \tau \cos \sigma \end{bmatrix} \quad (10)$$

where $k = h/(f \cos \sigma + \sin \sigma(x_i \cos \tau + y_i \sin \tau))$. In the next section of this paper we will use Eq. (10) to establish some properties of the projected spectral distribution in the viewer-centred co-ordinate system. In particular, we will show that contours of equal total power correspond to lines that are parallel to tilt direction. We will also demonstrate that lines radiating from the vanishing point connect points with identically orientated spectral distributions. In Section 5 we will exploit these two properties to develop an algorithm

for recovering the image-plane position of the vanishing point, and hence, estimating the orientation of the texture-plane.

4. Image-plane spectral moments

In this section, we are interested in the properties of

Table 1
Actual × estimated slant and tilt values (Artificial Group 1)

Image	Actual values		Estimated values		Absolute error	
	Slant (σ)	Tilt (τ)	Slant (σ')	Tilt (τ')	(σ')	(τ')
(a)	20	0	21.9	0.9	1.9	0.9
(b)	30	330	27.0	338.4	3.0	8.4
(c)	45	45	42.7	55.2	2.3	10.2
(d)	60	120	59.0	108.1	1.0	11.9
(e)	70	0	73.0	5.7	3.0	5.7
(f)	50	225	50.6	233.1	0.6	8.1

Table 2
Actual × estimated slant and tilt values (Artificial Group 2)

Image	Actual values		Estimated values		Absolute error	
	Slant (σ)	Tilt (τ)	Slant (σ')	Tilt (τ')	(σ')	(τ')
(a)	20	0	21.1	0.9	1.1	0.9
(b)	45	45	57.5	48.3	12.5	3.3
(c)	60	120	72.7	114.8	12.7	5.2
(d)	30	330	29.6	316.3	0.4	13.7
(e)	50	225	51.2	227.4	1.2	2.4
(e)	70	0	70.0	1.2	0.0	1.2
(e)	30	330	29.1	322.6	0.9	7.4
(e)	45	45	47.6	48.1	2.6	3.1
(e)	60	120	63.6	114.5	3.6	5.5

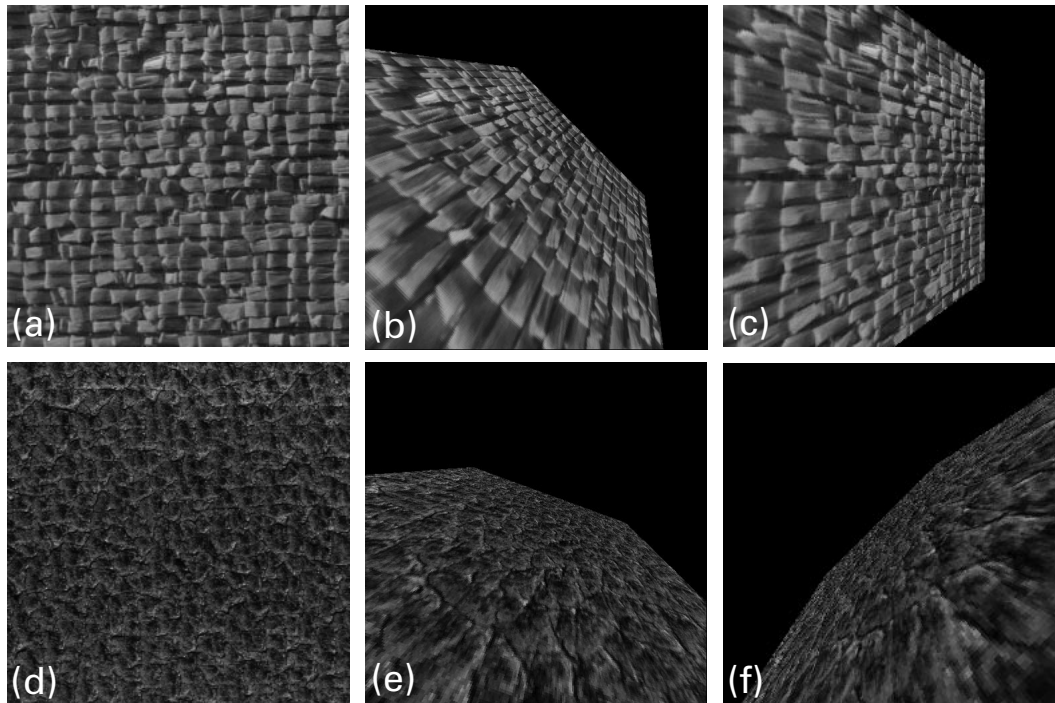


Fig. 7. Real texture images: Group 1—projected images. (a) $\sigma = 0^\circ$ and $\tau = 0^\circ$; (b) $\sigma = 45^\circ$ and $\tau = 45^\circ$; (c) $\sigma = 30^\circ$ and $\tau = 0^\circ$; (d) $\sigma = 0^\circ$ and $\tau = 0^\circ$; (e) $\sigma = 60^\circ$ and $\tau = 70^\circ$; (f) $\sigma = 75^\circ$ and $\tau = 135^\circ$.

Table 3
Actual \times estimated slant and tilt values (Real Group 1)

Image	Actual values		Estimated values		Absolute error	
	Slant (σ)	Tilt (τ)	Slant (σ')	Tilt (τ')	(σ')	(τ')
(b)	45	45	51.0	49.5	6.0	4.5
(c)	30	0	20.9	2.2	9.1	2.2
(e)	60	70	79.1	82.2	19.1	12.2
(f)	75	135	81.3	138.2	6.3	3.2

isotropic and homogeneous textures under perspective projection onto the image plane. This representation was used in the work of Kwon [22]. However, here we use it to establish a new geometric property of texture peaks, namely constancy of orientation in the direction of vanishing points on the image plane. To demonstrate this property, there are two quantities that interest us. The first of these is the distribution of power-density. The second quantity is the distribution of the spectral orientation. In this section we will demonstrate that the former allows us to estimate tilt direction, while the latter is related to the direction of the vanishing point.

4.1. Lines of equal power

We commence by considering the moment representation for the spectral content of an isotropic homogeneous texture [11,19]. Isotropic homogeneous textures have a circularly symmetric spectral representation when viewed in a fronto-parallel direction. When viewed under perspective projection the spectral representation becomes elliptical.

The second-order spectral moments m_{pq} ($p + q = 2$, $p = 0-2$, $q = 0-2$) are obtained by fitting ellipses to the 2D frequency distribution [19]. In the texture-plane co-ordinate system, the required moments are related to the power-spectrum in the following manner:

$$M_{pq} = \sum_{u_i} \sum_{v_i} u_i^p(x_i, y_i) v_i^q(x_i, y_i) P(u, v) \quad (11)$$

where the frequency components $u_i(x_i, y_i)$ and $v_i(x_i, y_i)$ are obtained using the inverse of the perspective transformation given by Eq. (8). The spectral data can now be represented using the moment-matrix

$$S_2 = \begin{bmatrix} M_{20} & M_{11} \\ M_{11} & M_{02} \end{bmatrix} \quad (12)$$

Table 4
Estimated slant and tilt values (Real Group 2)

Image	Slant (σ')	Tilt (τ')
(a)	7.8	-3.4
(b)	21.3	178.0
(c)	17.8	163.2
(d)	18.7	176.7

The eigenvalues and eigenvectors of the matrix S_2 are related to the geometry of the best-fit ellipse. The principal eigenvector is aligned along the major axis of the ellipse, while the second eigenvector is aligned along the minor axis. The two eigenvalues are equal to the lengths of the major and minor axes of the ellipse. In the case of an isotropic homogeneous texture, we have $S_2 = mI$ due to the circular symmetry for isotropic distributions where I is the 2×2 identity matrix. To commence we use the moments analysis to show that the density gradient is zero for points connecting lines which are perpendicular to the tilt direction. This property was suggested as an important key to recovering perspectivity by Stevens [1]. We evaluate the area of the ellipses along a line perpendicular to tilt direction. Furthermore, we will represent the perpendicular to the tilt directions as a straight line intercepting the y_i -axis at y_0 as follows:

$$y_i = \frac{y_0 \sin \tau - x_i \cos \tau}{\sin \tau} \quad (13)$$

To pursue our development, we note that the lengths of the major and minor axes, i.e. L and l , of the projected spectral ellipses are given by:

$$L = \frac{1}{2}(M_{20} + M_{02}) + \frac{1}{2}\sqrt{4M_{11}^2 + (M_{20} - M_{02})^2} \quad (14)$$

$$l = \frac{1}{2}(M_{20} + M_{02}) - \frac{1}{2}\sqrt{4M_{11}^2 + (M_{20} - M_{02})^2}$$

The image-plane power density is proportional to the area of the projected spectral ellipses

$$A_i = \frac{\pi}{4} L \times l = \frac{\pi}{4} (M_{20}M_{02} - M_{11}^2) \quad (15)$$

Substituting for the image plane spectral moments M_{20} , M_{11} and M_{02} , the density is equal to

$$A_i = \frac{(f \cos \sigma + y_0 \sin \sigma \sin \tau)^6 m^2 \pi}{4f^2 h^4 \cos^4 \sigma} \quad (16)$$

As a result, lines which are perpendicular to tilt direction are characterised by local spectral moments of equal area. In other words, we have proved that lines of uniform power-density on the image plane are orthogonal to the tilt direction. We can therefore recover the tilt axis by searching for the direction of maximum power uniformity.

Table 5
Estimated slant and tilt values (Real Group 3)

Image	Slant (σ')	Tilt (τ')
(a)	43.6	78.0
(b)	41.4	70.7
(c)	47.8	75.0

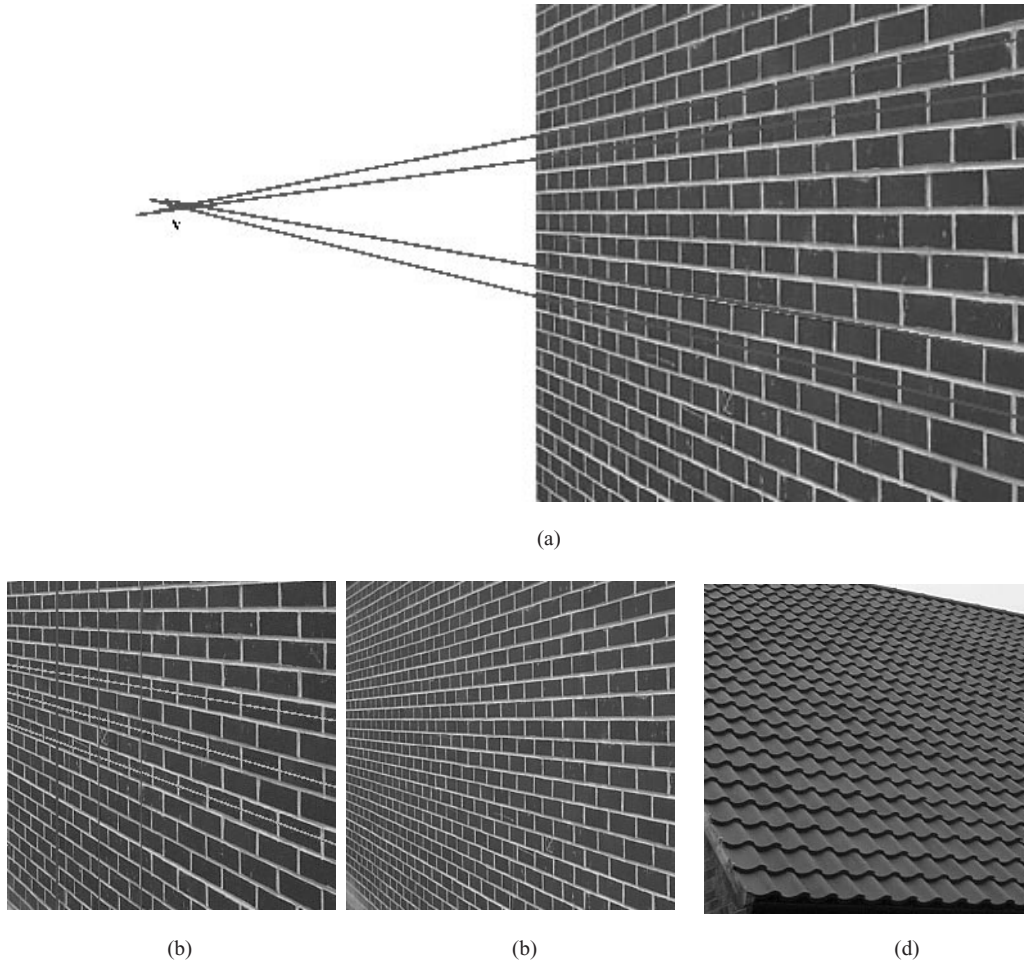


Fig. 8. Real texture images: Group 2—outdoor images.

4.2. Lines of constant spectral orientation

We next consider the angular structure of the local spectral distribution. Specifically we will show how the spectral angle can be used to estimate the vanishing point location and hence planar surface orientation. We commence by analysing the geometric nature of the 2D spectral information along lines radiating from the vanishing point. Consider a spectral peak with frequency vector $(u_t, v_t)^T$ at a point (x_t, y_t) on the texture plane. The direction of the spectral distribution is given by $\beta = \arctan(v_t/u_t)$. From Eq. (8), the orientation of the corresponding spectral components in the image plane is

$$\tan \alpha = \frac{v_i}{u_i} = \frac{u_t f \sin \tau + v_t(x_i \sin \sigma + f \cos \tau \cos \sigma)}{u_t f \cos \tau - v_t(y_i \sin \sigma + f \sin \tau \cos \sigma)} \quad (17)$$

For simplicity, we use a rotated system of coordinates for the image plane in which the x -axis is aligned in the tilt direction. In our rotated system of coordinates,

the above equation simplifies to

$$\tan \alpha = \frac{v_i}{u_i} = - \frac{(f \cos \sigma + x_i \sin \sigma)v_t}{y_i v_t \sin \sigma - f u_t} \quad (18)$$

Now consider a line in the image plane that radiates from the vanishing point. This line is the result of projecting one of a family of lines that are aligned in the direction of the line-of-sight on the texture plane. The texture-plane frequency peak associated with this family of lines can be represented by the vector $\mathbf{U}_t = (0, v_t)^T$. After perspective projection onto the image plane the family of lines can be represented by the equation

$$y_i = c(x_i \sin \sigma + f \cos \sigma) \quad (19)$$

where c is the parametric variable of the family. These lines meet at the vanishing point

$$\mathbf{V} = (-f \cos \sigma / \sin \sigma, 0)^T \quad (20)$$

Substituting for y_i from Eq. (19) into the simplified formula for the image-plane spectral angle given in

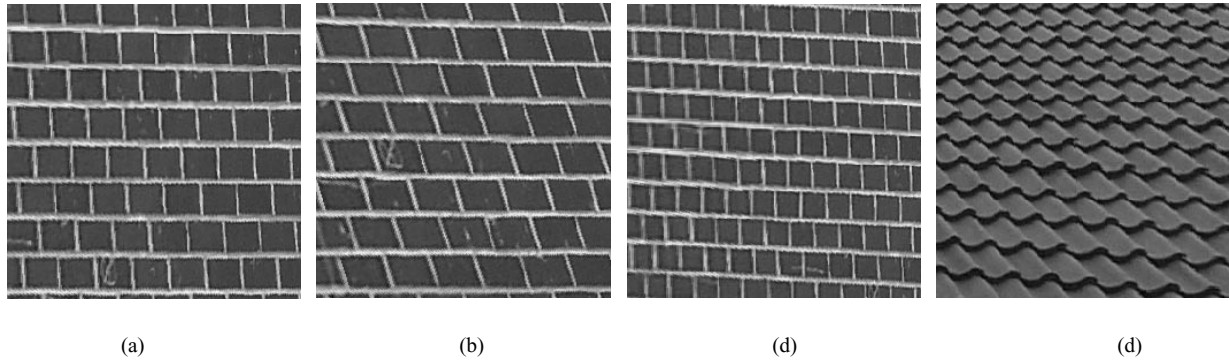


Fig. 9. Real texture images: Group 2—back-projected images.

Eq. (18), we obtain the result

$$\tan \alpha = \frac{v_i}{u_i} = -\frac{(f \cos \sigma + x_i \sin \sigma)v_t}{c(f \cos \sigma + x_i \sin \sigma)v_t \sin \sigma} = -\frac{1}{c \sin \sigma} = \text{constant} \quad (21)$$

Upon substituting for the resulting value of c into the equation of the line radiating from the vanishing point in Eq. (19), we find

$$y_i = -\frac{x_i}{\tan \alpha} + f \frac{u_t(\tan \alpha \cos \tau - \sin \tau) - v_t \cos \sigma(\tan \alpha \sin \tau + \cos \tau)}{v_t \tan \alpha \sin \sigma} \quad (22)$$

In other words the slope of the line depends only on the tangent of the image-plane spectral angle, i.e. $\tan \alpha$. Since the texture plane spectral frequency $U_t = (u_t, v_t)^T$ is constant due to the homogeneity assumption, each such line connects points on the image plane whose local spectral distribution have uniform spectral angle α . Moreover, these lines will intercept at a unique point which is a vanishing point in the image plane.

To provide an illustration, Fig. 2 shows some examples of the uniform angle lines computed using the perspective model developed in this section. In the figure we plot examples of the lines of uniform spectral angle for a sinusoidal texture and several slant angles. For simplicity, we illustrate for tilt equal to zero.

Fig. 3 shows some local power spectra sampled along two different lines that radiate from the vanishing point labelled V . The spectra are labelled by capital letters and are shown in the left-hand panel of the figure. In each case the dominant peak frequency is marked by a white arrow. We have two groups of spectra belonging to two different straight lines in the figure. The first group is labelled with the letters A–D while the second group is labelled with the letters E–H. The main feature to note from these examples is that while the dominant frequency peak appears at the same angle within each group, it appears at different angles in the different groups.

Based on this angular constancy property, we can locate the vanishing point coordinates in the image plane by connecting points which have a uniform spectral angle. We meet this goal by searching lines for which the angular correlation between the spectral moments is maximum. To proceed we adopt a polar representation for the power spectrum. Suppose $P_{r\phi}(r, \phi)$ is the power spectrum in polar coordinates where $r = \sqrt{u_t^2 + v_t^2}$ is the radial variable

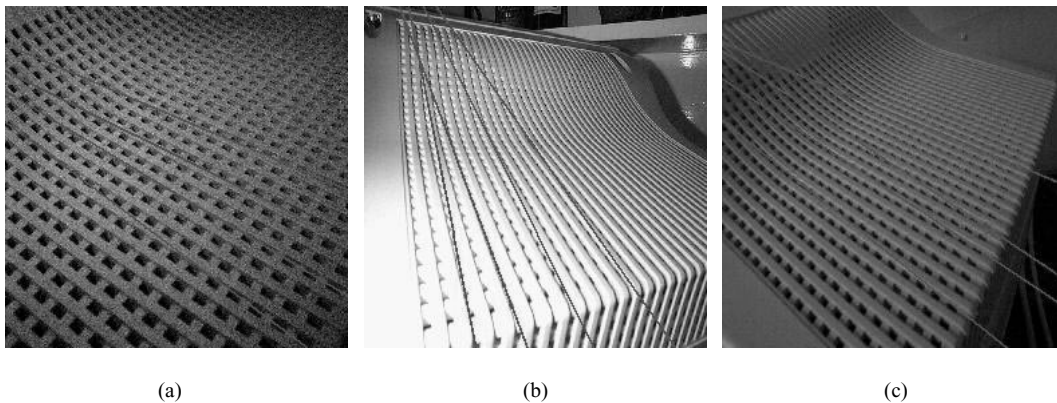


Fig. 10. Real texture images: Group 3—indoor images.

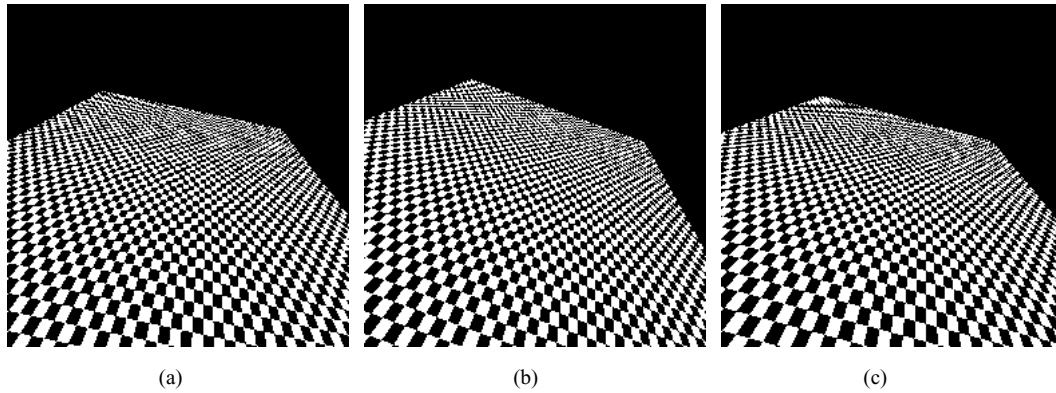


Fig. 11. Real texture images: Group 3—estimated planes.

and $\phi = \arctan(v_1/u_1)$ is the angular variable

$$P_\phi(\phi) = \int_r^0 P(r\phi) dr \quad (23)$$

The angular distribution of spectral power at any given image point can be matched against those of similar orientation by maximising angular correlation. For the purpose of matching we use the linear correlation:

$$\rho = \frac{\int_\phi P_\phi(\phi) P'_\phi(\phi) d\phi}{\int_\phi P_\phi(\phi) d\phi \int_\phi P'_\phi(\phi) d\phi} \quad (24)$$

Here $P_\phi(\phi)$ and $P'_\phi(\phi)$ are the two angular distributions being compared. The points with the highest value of ρ can be now connected to determine a line pointing to the direction of the vanishing point.

4.3. Estimating tilt direction

In order to estimate the surface normal direction from Eqs. (5) and (6), we need first to determine the tilt direction. We use the fact that for most homogeneous textures the density gradient is zero for points located at a same distance from the viewpoint. These points with zero density gradient interconnect on lines which are perpendicular to the tilt direction. Analogously, these lines connect points with constant total radial spectral energy. Once these lines are detected, the tilt direction can be easily estimated and used together with the vanishing point coordinates to estimate surface orientation. We use standard least squares method in order to fit the straight lines to the equal energy contours. Fig. 4 illustrates the contour maps of the density gradient of two artificial textures. Tilt direction can be estimated by fitting straight lines to these contours.

5. Vanishing point detection

The main contribution in this paper is to use the two properties outlined in the previous section to develop a geometric algorithm for estimating the vanishing point for

a textured plane viewed under perspective geometry. We commence by generating an isocontour map based on the local spectral energy. This isocontour map reflects the constancy of the zero density gradient over lines perpendicular to the tilt direction as described in Section 4.3. We fit straight lines over the spectral density isocontour map and determine the tilt direction τ . In the next step we choose any point on one of the best fitted isocontour lines. The angular descriptor of this point as given by Eq. (23) is correlated to the angular descriptor for a sample of points over the image and the points with maximum correlation are obtained. We again fit straight lines using these maximum correlated points to determine the equation of the line pointing to the vanishing point V . Now, we repeat the angular description-correlation procedure to determine another line pointing to the vanishing point. The intercept coordinates of these lines will be the coordinates of V . Finally, we determine the planar surface orientation based on the normal vector from V using Eqs. (5) and (6). The steps employed are summarised as follows.

Summary of algorithm

1. Generate an isocontour map for the local spectral radial energy.
2. Fit straight lines to the isocontour data.
3. Determine the tilt direction line using the best fitted lines.
4. Determine the vanishing point by triangulating the lines of maximum spectral angular correlation.

6. Experiments

Finally, we provide some results which illustrate the accuracy of planar pose estimation achievable with our new 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 both synthetic textures and projected real-world textures. The second part of our experimental study focuses on natural textures planes (brick-walls and roof-tiles) where the ground truth is unknown. In order to

give some idea of the accuracy of the slant and tilt estimation process, we back-project the texture onto the fronto-parallel plane. Since the textures are man-made and rectilinear in nature, the inaccuracies in the estimation process manifest themselves as residual skew.

6.1. Synthetic texture planes

We commence with some examples for a synthetic regular sinusoid texture which we have already used to illustrate the estimation of spectral moments. Fig. 5 shows the synthetic texture in a number of poses with the slant and tilt angles annotated. In each case the texture is oriented in the direction of the vanishing point. The estimated lines pointing to the vanishing point are presented superimposed on the image.

In Table 1 we list the actual and computed pose angles for the different orientations of the texture plane. A more natural way of representing the planar pose is to compute the orientation of the surface normal. However, we choose to display the slant and tilt dependence of the errors since our algorithm has separate steps for estimating these two angles. The average error in tilt angle is 7° while the average error in slant angle is 2° . For this group of artificial images, we observe that the vanishing point is accurately detected for most imaging geometries. In all six cases the slant angle is accurately estimated. However, the tilt angle is less well estimated. Since the tilt direction is computed by searching for lines of uniform spectral energy, it is the estimation of spectral energy that is suspected. The main difficulties occur when the tilt angle is large. Here the large error is attributable to the choice of scale used in the computation of the local texture spectra. In fact we use a fixed scale factor for the spectral estimator. When the slant angle is large this scale may be too large to capture the energy content of the frequency spectra at points on the plane close to the viewer. However, even without adapting the scale factor, the average error is relatively small.

A second group of artificial images is shown in Fig. 6. In this group we have taken three different artificial textures and have projected them onto planes of known slant and tilt. The textures are composed of regularly spaced geometric primitives of uniform size. Specifically, we have chosen elliptical, rectangular and lattice shaped primitives. However, it is important to stress that in this case the texture elements are not oriented in the direction of the vanishing point. Superimposed on the projected textures are the estimated lines of uniform spectral orientation. In Table 2 we list the ground-truth and estimated values of the slant and tilt together with the corresponding errors. The agreement between the estimated and ground-truth angles is good. Moreover, the computed errors are largely independent of the slant angle. The average slant-error is 4° and the average tilt-error is 5° . Because of angular aliasing, for this second group of textures there exist multiple vanishing points. Examples occur for the lattice texture shown in

Fig. 6(g) and (h). The reason for this aliasing effect is that the corresponding spectral peak varies according to the orientation of the texture plane and the spatial position of the spectral estimator over the image. However, the angular constancy property still holds.

We now turn our attention to more realistic imagery. The rows in Fig. 7 show three different poses of two natural textures. The first column shows the original texture in its fronto-parallel pose. The slant and tilt angles are again appended. Table 3 lists the actual and computed pose angles (see also Tables 4 and 5). The agreement is generally good, although there are problems with the example in Fig. 7(e), where we again encounter aliasing effects.

6.2. Real world examples

The final piece of experimental work focuses on real world textures with unknown ground-truth. The textures used in this study are three views of a brick-wall and a view of a York pantile roof. The images were collected using a Kodak DC210 digital camera and are shown in Fig. 8. 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 of uniform spectral orientation. In the case of the brick-wall images there closely follow the mortar lines. In the case of the pantile texture the lines follow the different layers of the roof. In Fig. 9 we show the back-projection of the textures onto the fronto-parallel plane. In the case of the brick-wall, any residual skew is due to error in the estimation of the slant and tilt parameters. It is clear that in the case of Fig. 9(a) and (c), the slant and tilt estimates are accurate. In the case of Fig. 9(b) there is a residual skew of 10° . Finally, in the case of the pantile texture, the back-projection is fairly good. The primitives are of uniform size, however, there is a residual skew of about 20° . Figs. 10 and 11 show the real texture images of indoor images and estimated planes, respectively.

7. Conclusions

We have described a simple algorithm for estimating the slant and tilt of textured planes viewed under perspective geometry. The method is based on a new angular property of local spectral distributions under perspective projection. This property can be regarded as a natural extension of the results reported by Stevens [1] and Garding [18].

The implementation of the method involves searching for two sets of lines. The first of these connect points of equal spectral power and are oriented in the tilt direction. The second set of lines connect points which have identically oriented spectral moments. These lines intercept at the vanishing point.

The method appears to produce reasonable accuracy. The

simplicity of the model and the absence of any type of ambiguity is another clear advantage. Methods based on vanishing point have been developed using direct measurements on the image plane [20–22]. This method is completely independent of such a direct procedure. There is no necessity to know the size or shape of the textural primitives.

There are a number of ways in which the ideas presented in this paper can be extended. Firstly, there are more elegant ways for exploiting the angular property to estimate vanishing point position and subsequently recovering the slant and tilt angles. As concrete examples we have recently developed practical methods which recover slant and tilt from multiple vanishing point locations [23] and for estimating the required vanishing point locations using Hough-based voting in the spectral domain [24]. Suffice to say, these alternative methods rely on a more comprehensive use of the property of angular constancy for corresponding spectral peaks. They can be employed to estimate different vanishing point locations. Moreover, they allow us to directly solve for the orientation parameters. This avoids the rather clumsy requirement to predetermine the tilt angle using energy gradient. Secondly, we note that the selection of the appropriate scale for local spectral analysis is a source of instability for the method. An adaptive scale selection could be applied to improve the quality of spectral estimation. The third line of investigation is to extend our ideas to curved surfaces, using the method to estimate local slant and tilt parameters. Studies aimed at developing these ideas are in hand and will be reported in due course.

References

- [1] K.A. Stevens, The information content of texture gradients, *Biological Cybernetics* 42 (1981) 95–105.
- [2] D. Marr, *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*, Freeman, New York, 1982.
- [3] J.J. Gibson, *The Perception of the Visual World*, Houghton Mifflin, Boston, 1950.
- [4] K. Ikeuchi, Shape from regular patterns, *Artificial Intelligence* 22 (1984) 49–75.
- [5] J. Aloimonos, M.J. Swain, Shape from texture, *Biological Cybernetics* 58 (5) (1988) 345–360.
- [6] K. Kanatani, T. Chou, Shape from texture: general principle, *Artificial Intelligence* 38 (1989) 1–48.
- [7] A.P. Witkin, Recovering surface shape and orientation from texture, *Artificial Intelligence* 17 (1981) 17–45.
- [8] J. Garding, Direct estimation of shape from texture, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 15 (1993) 1202–1208.
- [9] A. Blake, C. Marinos, Shape from texture: estimation, isotropy and moments, *Artificial Intelligence* 45 (3) (1990) 323–380.
- [10] R. Buse, H. Tsui, Estimating surface orientation from directional textures, *Pattern Recognition Letters* 14 (1993) 431–439.
- [11] L.G. Brown, H. Shvaytser, Surface orientation from projective foreshortening of isotropic texture autocorrelation, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 12 (6) (1990) 584–588.
- [12] B.J. Super, A.C. Bovik, Planar surface orientation from texture spatial frequencies, *Pattern Recognition* 28 (5) (1995) 729–743.
- [13] J. Malik, R. Rosenholtz, A differential method for computing local shape-from-texture for planar and curved surfaces, *IEEE Conference on Vision and Pattern Recognition* (1993) 267–273.
- [14] Ko Sakai, L.H. Finkel, A shape-from-texture algorithm based on human visual psychophysics, *IEEE Conference on Vision and Pattern Recognition* (1994) 527–532.
- [15] R. Bajcsy, L. Lieberman, Texture gradient as a depth cue, *Computer Graphics and Image Processing* 5 (1976) 52–67.
- [16] J. Krumm, S.A. Shafer, Texture segmentation and shape in the same image, *IEEE International Conference on Computer Vision* (1995) 121–127.
- [17] B.J. Super, A.C. Bovik, Filters for directly detecting surface orientation in an image, *SPIE Conference on Visual Communications and Image Processing* (1992) 144–155.
- [18] J. Garding, Surface orientation and curvature from differential texture distortion, *IEEE International Conference on Computer Vision* (1995) 733–739.
- [19] B.J. Super, A.C. Bovik, Shape from texture using local spectral moments, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 17 (4) (1995) 333–343.
- [20] J.R. Kender, Shape from texture: an aggregation transform that maps a class of texture into surface orientation, *Sixth IJCAI, Tokyo* (1979) 475–480.
- [21] Y. Ohta, K. Maenobu, T. Sakay, Obtaining surface orientation from texels under perspective projection, *Seventh IJCAI, Vancouver* (1981) 746–751.
- [22] J.S. Kwon, H.K. Hong, J.S. Choi, Obtaining a 3-d orientation of projective textures using a morphological method, *Pattern Recognition* 29 (1996) 725–732.
- [23] E. Ribeiro, E.R. Hancock, Improved orientation estimation for texture planes using multiple vanishing points, *Pattern Recognition* (2000) in press.
- [24] E. Ribeiro, E.R. Hancock, 3-D planar orientation from texture: estimating vanishing point from local spectral analysis, *IX British Machine Vision Conference* (1999) 255–264.