3-D Planar Orientation from Texture: Estimating Vanishing Point from Local Spectral Analysis

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

Abstract

This paper describes a novel method for recovering the perspective geometry of textured surfaces from local spectral moments. It commences by computing the Fourier power spectrum at each of a series of local neighbourhoods. We illustrate how the vanishing point of the global perspective geometry can be estimated from local spectral moments. This is a two-step process. We commence by identifying contours which connect the local spectral moments of equal power. These contours are perpendicular to the tilt direction at the image plane. The second step is to triangulate the vanishing point. We do this by using a correlation method to identify lines with identical oriented spectral moment. In this way we can use a sample of spectral moments estimate the vanishing point location.

1 Introduction

Shape-from-texture was identified by Marr as one of the key ingredients of the $2\frac{1}{2}$D sketch [4]. Moreover, the feasibility of the process is well-grounded in psychophysics [7]. 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, then there is ambiguity in the recovery of surface depth due to the fact that the overall scale of the texture elements may be unknown.

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 texture elements have no detectable preferential direction when viewed in a front-parallel manner. Homogeneity requires that the density of texture elements is uniform. There are two distinct schools of thought as to

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

BMVC 1998 doi:10.5244/C.12.33

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.

The structural method uses the geometry of edges, lines or arcs to identify perspective distortion using shape cues [5, 10]. 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 defined either over single primitives or groups of primitives [1, 6]. 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 textures that can be accommodated. The approach is exemplified by the work of Super and Bovik [2] which employs two-dimensional Gabor wavelets to detect the local projective distortion in the power spectrum.

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. Specifically, 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. For planar textures, these contours correspond to lines of equal distance from the vanishing point. 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 components. Here we make use of the observation that lines that radiate from the vanishing point connect points which have identically oriented spectral distribution. We search for the vanishing point by correlating the angular power distribution. 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.

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 axes 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 relative orientations of the two planes using the slant angle σ and tilt τ angles. If the focal length of the camera is f, then the perspective transformation between the corresponding points $\vec{X}_t = (x_t, y_t, h)^T$ on the texture plane and

$\vec{X}_i = (x_i, y_i, 0)^T$ on the image plane is:

$$\begin{bmatrix} x_i \\ y_i \end{bmatrix} = \frac{f}{h - x_t \sin\sigma} \begin{bmatrix} \cos\sigma\cos\tau & -\sin\tau \\ \cos\sigma\sin\tau & \cos\tau \end{bmatrix} \begin{bmatrix} x_t \\ y_t \end{bmatrix} \quad (1)$$

The inverse transformation is given by

$$\begin{bmatrix} x_t \\ y_t \end{bmatrix} = k \begin{bmatrix} \frac{\cos\tau}{\cos\sigma} & \frac{\sin\tau}{\cos\sigma} \\ -\sin\tau & \cos\tau \end{bmatrix} \begin{bmatrix} x_i \\ y_i \end{bmatrix} \quad (2)$$

where

$$k = \frac{h\cos\sigma}{f\cos\sigma + \sin\sigma(x_i\cos\tau + y_i\sin\tau)}$$

The net effect of this transformation is to distort the viewer-centred texture pattern in the direction of the vanishing point $\vec{V} = (x_\infty, y_\infty, 0)^T$ 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 set of parametric equations $\vec{X}_s = \vec{A} + \lambda\vec{B}$, where λ is the parametric variable of the individual lines. The three constants forming the vector $\vec{B} = (b_1, b_2, b_3)^T$ are the direction cosines for the entire family. The individual lines in the family are each parametrised by the vector $\vec{A} = (a_1, a_2, a_3)^T$.

The direction cosines for the family of parallel lines are related to the position of the vanishing point $(x_\infty, y_\infty, 0)$ in the viewer centred co-ordinate system in the following manner

$$\begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \frac{1}{\sqrt{x_\infty^2 + y_\infty^2 + f^2}} \begin{pmatrix} x_\infty \\ y_\infty \\ f \end{pmatrix} \quad (3)$$

Suppose the the vector $\vec{N} = (p, q, 1)^T$ 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, then

$$\vec{N}\cdot\vec{B} = pb_1 + qb_2 + b_3 = 0 \qquad and \qquad \tan\tau = \frac{p}{q} \quad (4)$$

Using Equations 3 and 4, the 3-D planar surface orientation can be completely recovered. In the next section we describe how the vanishing point coordinates $\vec{V}$ 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 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

$$P_t(u_t, v_t) = T_1 T_2 \sum_{k=-p1}^{p1} \sum_{l=-p2}^{p2} r_x(k,l) \times exp[-j2\pi(kuT_1 + lvT_2)] \tag{5}$$

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 $\vec{U_i} = (u_i, v_i)^T$ and $\vec{U_t} = (u_t, v_t)^T$ are, respectively, the image-plane and texture-plane frequencies, then they are related to one-another by a frequency-domain Jacobian [2]. In practice, however, we will be concerned with 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. 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

$$\vec{U_i} = J_S^t \vec{U_t} \tag{6}$$

where J_S^t is the transpose of the Jacobian of the inverse perspective transformation of co-ordinates. As a result, the instantaneous frequency peaks transform with the matrix [2]

$$J_S^t = \frac{-k^2 \sin\sigma}{h} \begin{bmatrix} x_i \cos\tau^2 + y_i \sin\tau\cos\tau & -x_i \sin\tau\cos\sigma\cos\tau + y_i \cos\tau^2\cos\sigma \\ x_i \cos\tau\sin\tau + y_i \sin\tau^2 & -x_i \sin\tau^2\cos\sigma + y_i \cos\tau\cos\sigma\sin\tau \end{bmatrix}$$

$$+k \begin{bmatrix} \cos\tau & \sin\tau\cos\sigma \\ \sin\tau & \cos\tau\cos\sigma \end{bmatrix} \tag{7}$$

In the next Section we will use (7) 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 connect points of equi-distance from the vanishing point. We will also demonstrate that lines radiated from the vanishing point connect points with identically orientated spectral distributions. In the next sections we will exploit these two properties to develop a geometric 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 isotropic and homogeneous textures under perspective projection onto the image plane. There are two properties, that interest us. The first of these is the distribution of power-density. The second property 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 position 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, 3]. Isotropic homogeneous textures have a circularly symmetric spectral representation when viewed in a front-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 two-dimensional frequency distribution [3]. In the texture-plane co-ordinate system, the required moments are related to the power-spectrum in the following manner

$$M_{pq} = \sum_{u_t} \sum_{v_t} u_t^p (x_i, y_i) \, v_t^q (x_i, y_i) \, P(u, v) \tag{8}$$

where the frequency components $u_t(x_i, y_i)$ and $v_t(x_i, y_i)$ are obtained using the inverse of the perspective transformation given by (6). The spectral-data can now be represented using the moment-matrix

$$S_2 = \begin{bmatrix} M_{20} & M_{11} \\ M_{11} & M_{02} \end{bmatrix} \tag{9}$$

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 L of the major and l of the minor axis of the ellipse. The area of the ellipse is equal to $\frac{\pi}{4} L \times l$. We will represent the tilt perpendicular as a straight line intercepting the y_i-axis at y_o as follows:

$$y_i = \frac{y_o \sin\tau - x_i \cos\tau}{\sin\tau} \tag{10}$$

For points lying on this line the area of the spectral ellipse and hence the power density is proportional to:

$$A_i = \frac{\pi}{4} L \times l = 1/4 \, \frac{(f \cos\sigma + y_o \sin\sigma \, \sin\tau)^6 \, m^2 \pi}{f^2 h^4 \cos^4\sigma} \tag{11}$$

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.

4.2 Lines of constant spectral orientation

We next consider the directional properties of the local power-spectra. 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 co-ordinates, J_S^t simplifies to:

$$J_S^t = k \begin{bmatrix} x_i \sin\sigma - 1 & y_i \sin\sigma \cos\sigma \\ 0 & -\cos\sigma \end{bmatrix} \tag{12}$$

where $k = -\frac{h}{(f \cos\sigma + x_i \sin\sigma)}$. From (12) and considering the relation given by Equation 6, the orientation of the dominant spectral component is given by:

$$\tan\alpha = \frac{v_i}{u_i} = \frac{(f\cos\sigma + x_i \sin\sigma)v_t}{fu_t - y_i v_t \sin\sigma} \tag{13}$$

From (13), the equation of the line of constant orientation, i.e., α is:

$$y_i = -\frac{x_i}{\tan\alpha} + f\frac{u_t \tan\alpha - v_t \cos\sigma}{v_t \sin\sigma \tan\alpha} \tag{14}$$

For a given value of α, this equation represents a line point to the vanishing point V. The texture plane spectral frequency $U_t = (u_t, v_t)$ is constant due to the homogeneity assumption. As a result, each line connects points on the image plane whose the local spectral distribution have uniform spectral angle α. These lines will intercept at a unique point which is the vanishing point in the image plane in terms of the local spectral representation. The vanishing point co-ordinates are

$$\begin{bmatrix} x_\infty \\ y_\infty \end{bmatrix} = \begin{bmatrix} -f\frac{\cos\sigma}{\sin\sigma} \\ \frac{u_t f}{v_t \sin\sigma} \end{bmatrix} \tag{15}$$

As a result, we can find 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 $r = \sqrt{u_i^2 + v_i^2}$ is the radial variable and $\phi = arctan(v_i/u_i)$ is the angular variable:

$$P_\phi(\phi) = \int_r^0 P_{r\phi} dr \tag{16}$$

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:

$$\rho = \frac{\int_\phi P_\phi(\phi) P'_\phi(\phi) d\phi}{\int_\phi P_\phi(\phi) d\phi \int_\phi P'_\phi(\phi) d\phi} \tag{17}$$

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

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 radial energy. This isocontour map reflects the constancy of the zero density gradient over lines

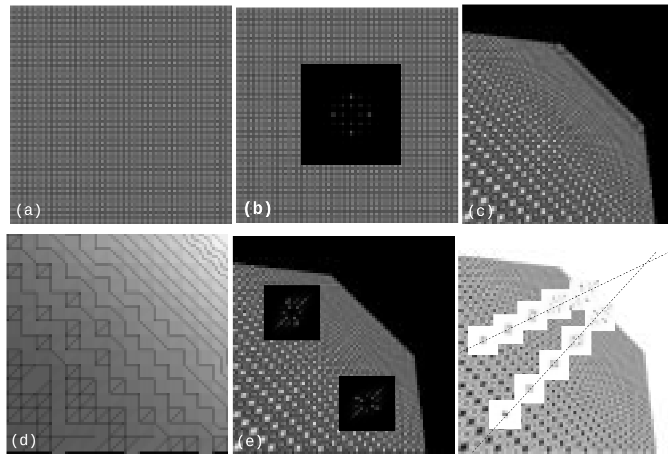

Figure 1: *(a) Unprojected homogeny texture; (b) Local power spectral content of (a); (c) Perspective projection of (a); (d) Projected Contour plot of lines of equal spectral power; (e) Two chosen reflected symmetric points lying at a line of same spectral power; (f) Lines of same spectral angle pointing to the direction of the vanishing point.*

perpendicular to the tilt direction as described in Section 4.1. 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 equation 16 is correlated to the angular descriptor for a sample of points over the image and the points with maximum correlation are obtained. We fit again 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 Equations 3 and 4. 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.

Figure 1 shows the main steps followed by our method to estimate planar surface orientation. In Figure 1a we show a homogeneous texture in the front-parallel view. Figure 1b shows the corresponding power-spectra at a representative image-plane location. In the case of the front-parallel texture, the dominant spectral components are associated with the repetitive structure of the texture in the x and y directions. There are also weaker peaks associated with both the harmonic structure of the texture and the diagonal patterns in the texture. Figure 1c presents the texture projected onto the image plane under perspective transformation. A contour plot is shown in Figure 1d describing the lines with same total spectral energy which are perpendicular to the tilt direction. In Figure 1e, the two spectra are located symmetrically about the axis of perspectivity. It is clear that the spectra themselves are reflected versions of one-another. In Figure 1f the constancy of the local angular spectral configuration can be clearly observed. The figure shows various local power spectrum responses over a line of same angular configuration. Moreover, the plot shows how the area of the spectrum decreases with increasing distance from the vanishing point.

6 Experiments

Finally, we provide some results which illustrate the accuracy of planar pose estimation achievable with our new shape-from-texture algorithm. We commence with some examples for the synthetic regular texture which we have already used to illustrate the estimation of spectral moments. Figure 2 shows the synthetic texture in a number of poses with the slant and tilt angles annotated. In Table 1 we list the actual and computed pose angles for the different orientations of the texture plane. The average error in tilt angle is 5 degrees while the average error in slant angle is 10 degrees The main systematic effect is that the errors are significant for both small and large slant angles. This is attributable to aliasing effects in sampling the spectral moments. There is no systematic structure to the tilt errors. Turning our attention to more realistic imagery, the rows in Figure 3 shows different poses of two natural textures. The slant and tilt angles are again appended. Table 2 lists the actual and computed pose angles. The agreement is generally good, although there are problems with row e, where we again encounter aliasing effects.

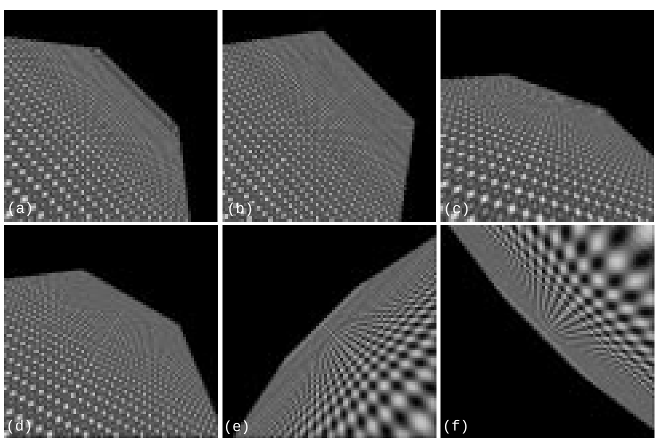

Figure 2: *Artificial texture images (a)* $\sigma = 45^\circ$ *and* $\tau = 45^\circ$*; (b)* $\sigma = 30^\circ$ *and* $\tau = 45^\circ$*; (c)* $\sigma = 60^\circ$ *and* $\tau = 70^\circ$*; (d)* $\sigma = 45^\circ$ *and* $\tau = 60^\circ$*; (e)* $\sigma = 75^\circ$ *and* $\tau = 135^\circ$*; (f)* $\sigma = 80^\circ$ *and* $\tau = 225^\circ$

TABLE 1 - Actual × Estimated slant and tilt values (Artificial Images)

	actual values		estimated values		absolute	error
$image$	slant (σ)	tilt(τ)	slant(σ')	tilt(τ')	σ'	τ'
(a)	45	45	38.79	46.6	6.21	1.6
(b)	30	45	21.2	49.9	8.8	4.9
(c)	60	70	57.9	71.9	2.1	1.9
(d)	45	60	45.2	64.9	0.2	4.9
(e)	75	135	68.9	134.5	6.1	0.5
(f)	80	225	76.1	232.2	3.9	7.2

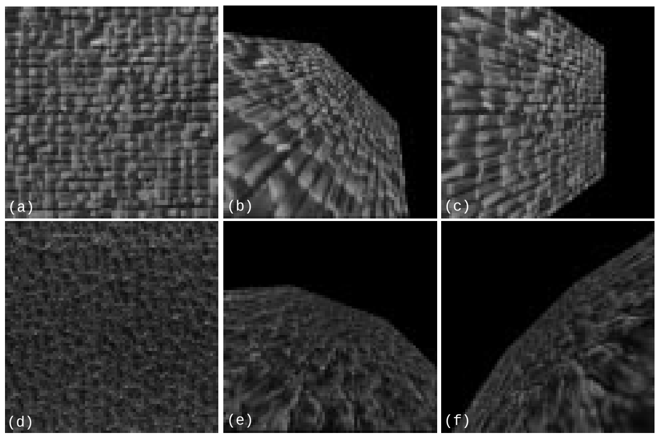

Figure 3: *(a) Natural texture image 1 (b) $\sigma = 45^o$ and $\tau = 45^o$; (c) $\sigma = 30^o$ and $\tau = 0^o$; (d) Natural texture 2 (e) $\sigma = 60^o$ and $\tau = 70^o$; (f) $\sigma = 75^o$ and $\tau = 135^o$*

TABLE 2 - Actual × Estimated slant and tilt values (Proj.Natural Images)

	actual values		estimated values		absolute	error
$image$	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

7 Conclusions

We have described a simple algorithm for estimating the slant and tilt of textured planes viewed under perspective geometry. The method searches 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 [8, 12, 9]. This method is completely independent on such 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. In the first instance, we are considering ways of improving the search for the two sets of lines. Specific candidates include Hough-based voting methods. The second 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] Witkin A. P. Recovering surface shape and orientation from texture. *Artificial Intelligence*, 17:17–45, 1981.

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

[3] Super B.J. and Bovik A.C. Shape from texture using local spectral moments. *IEEE Trans. on Patt. Anal. and Mach. Intelligence*, 17(4):333–343, 1995.

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

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

[6] Garding J. Direct estimation of shape from texture. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 15:1202–1208, 1993.

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

[8] 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.

[9] 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.

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

[11] Brown L.G. and Shvaytser H. Surface orientation from projective foreshortening of isotopic texture autocorrelation. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 12(6):584–588, 1990.

[12] Ohta Y., Maenobu K., and Sakay T. Obtaining surface orientation from texels under perspective projection. In *7th IJCAI, Vancouver*, pages 746–751, 1981.