MVA2000 IAPR Workshop on Machine Vision Applications, Nov. 28–30, 2000, The University of Tokyo, Japan

13—25 Estimation of Surface Location of a Textured Plane Using Spectral Moments

Toshio Tsuji Yasuo Yoshida *
Department of Electronics and Information Science
Kyoto Institute of Technology

Abstract

This paper proposes a method for estimating surface location of a textured plane using spectral moments. When a textured plane approaches a camera and slants, the 2-D power spectrum concentrates around lower frequencies and expands to the slant direction. The spectral moments also make change correspondingly. The distance from camera, slant angle and the slant direction can be derived from the comparison between the spectral moments of a standard image and those of an observed image. We show the validity of the method in simulations by using the synthesized images.

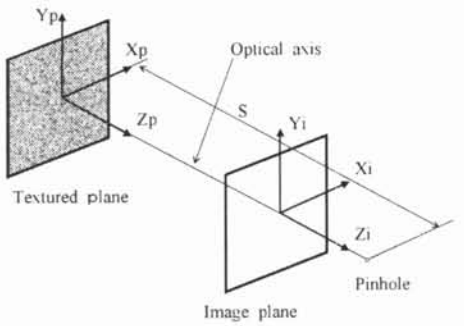
1 Introduction

Detection of the relative location between an objective plane and a camera is one of the most important problems in the 3-D motion estimation. We propose a method for locating the objective plane using the fineness of a texture pattern on its surface. Generally, all textures are classified into two large groups, i.e., structural and statistical ones. The surface location of a structural textured plane can be estimated by the size and distortion of its geometric pattern. In the case of a statistical one, however, it is nearly impossible to estimate it from the geometric pattern because of its indeterministic property. For that reason, we use the 2-D power spectrum of the textured plane image. If a textured plane approachs a camera, the image enlarges and its power spectrum concentrates around lower frequencies. Contrarily, if the plane goes away from the camera, the image shrinks and its power spectrum expands to higher frequencies. If the plane slants, the image shrinks along the slant direction and its power spectrum expands along the same direction.

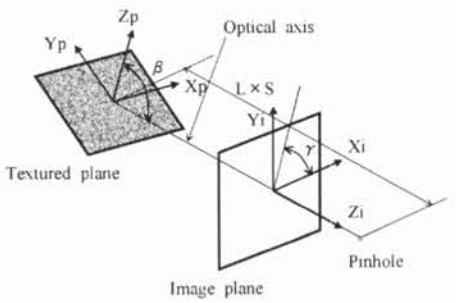
Our goal is to estimate the surface location, which is described by the distance from a camera, the slant angle that determines the degree of obliqueness of the surface, and the tilt angle that is the direction of slant.

To take advantage of the above properties in power spectrum, we use the secord-order spectral moments. The surface location can be derived by comparing the spectral moments of a standard plane image with those of an observed one.

This paper is organized as follows. Section 2 describes the theory of our method. Section 3 provides some experimental results. Section 4 terminates the paper with some discussion.



(a) Standard image location



(b) Test image location

Fig.1 Location of camera and textured plane

2 Theory of Estimation

2.1 Location and Power Spectrum

A textured plane in 3-D space is projected onto the image plane in a camera. We provide the simplified projection model with a pinhole camera. Relationship between a camera and a textured plane are shown in Fig.1.

In the case of a standard image, the textured plane locates at a certain distance S from the camera

*Address: Sakyou-ku, Kyoto-shi 606-8585 Japan. E-mail: yyoshida@dj.kit.ac.jp

and is perpendicular to the optical axis. The coordinate system of the textured plane (X_p, Y_p, Z_p) is parallel to that of the image plane (X_i, Y_i, Z_i) and the origins of those two coordinate systems are set on the optical axis. In the case of an observed image, the textured plane not only slants but also locates at L times longer or shorter distance from the camera. The slant angle β is defined as the angle between Z_p and Z_i, which takes nonnegative values between 0° and 90°. Furthermore, the angle between X_i and the projection of Z_p onto the image plane is defined as the tilt angle γ, which takes values between 0° and 180°.

It is assumed that the size of the visual field on a textured plane is much smaller than the distance from a camera L. The gray level relationship between a standard image $F(X_i, Y_i)$ and an observed image $F_T(X_i, Y_i)$ is given by

$$F_T(X, Y) = F(X', Y') \tag{1}$$

$$\begin{pmatrix} X \\ Y \end{pmatrix} = \frac{1}{L} \cdot \mathbf{R}_1(\gamma)\mathbf{R}_2(\beta)\mathbf{R}_1(-\gamma) \begin{pmatrix} X' \\ Y' \end{pmatrix} \tag{2}$$

where

$$\mathbf{R}_1(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \ \mathbf{R}_2(\theta) = \begin{pmatrix} \cos\theta & 0 \\ 0 & 1 \end{pmatrix} \tag{3}$$

On the other hand, the relationship between the power spectrum of a standard image $S(\omega, \nu)$ and that of an observed image $S_T(\omega, \nu)$ is given by

$$S_T(\omega, \nu) = \frac{\cos\beta}{L^2} \cdot S(\omega', \nu') \tag{4}$$

$$\begin{pmatrix} \omega \\ \nu \end{pmatrix} = L \cdot \mathbf{R}_1(\gamma)\mathbf{R}_2^{-1}(\beta)\mathbf{R}_1(-\gamma) \begin{pmatrix} \omega' \\ \nu' \end{pmatrix} \tag{5}$$

where ω and ν are the frequency components along X_i and Y_i axes, respectively.

As shown in (4) and (5), the distribution of a power spectrum makes change in accordance with the location parameters, β, γ and L.

2.2 Spectral Moment

The $(p+q)$th order spectral moment is defined by

$$\mu_{p,q} = \int\int_{(\omega,\nu)\in D} \omega^p \nu^q S(\omega, \nu) d\omega d\nu \tag{6}$$

$$p, q = 0, 1, 2, \cdots$$

where D denotes the integral domain. The integral domain of an observed image must include the same part as that of a standard one. If a textured plane shifts along the optical axis and dose not slant, those two integral domains may be within the concentric circles that include a certain rate of the whole image power[1]. But when the textured plane slants, the power spectrum expands to the slant direction and the integral domain must be within an ellipse. So we decide the ellipse with the variance and covariance matrix of a power spectrum shown as

$$\begin{pmatrix} \omega & \nu \end{pmatrix} \begin{pmatrix} \mu_{2,0} & \mu_{1,1} \\ \mu_{1,1} & \mu_{0,2} \end{pmatrix}^{-1} \begin{pmatrix} \omega \\ \nu \end{pmatrix} \leq K \tag{7}$$

where K is set as the total power within the ellipse is a certain rate of the whole power. In this experiment, we set 85% at the first time. If we can not get the solution, namely, location papameters, we reduce the rate by every 5 points in percent by the time we get the solution. The problem is that we can not decide the ellipse at one time, because the moments $\{\mu_{p,q}\}$ in (7) are obtained by (6) and the integral domain must be decided by (7), where the moments $\{\mu_{p,q}\}$ are also used. Therefore, we iteratively determine the moments; first we set a circle domain to (6) and then modify it to the elliptic domain given by (7) and repeat this procedure until the moments converge.

2.3 Rotational Moment

Rotational moment is convenient for estimating parameters and defined by [3] [4]

$$D_{n,l} = \int\int_{(r,\theta)\in D} r^n e^{il\theta} S(r\cos\theta, r\sin\theta) r dr d\theta \tag{8}$$

where $S(r\cos\theta, r\sin\theta)$ is a power spectrum on the polar coordinate system. We use $D_{2,2}$ and $D_{2,0}$ expressed by the spectral moments of a standard image $\{\mu_{p,q}\}$ and those of an observed image $\{\mu'_{p,q}\}$ shown as

$$\begin{aligned} D_{2,2} &= L^2 e^{i2\gamma}\Big(\frac{1}{\cos^2\beta}\hat{\mu}_{2,0} - \hat{\mu}_{0,2} + i\frac{2}{\cos\beta}\hat{\mu}_{1,1}\Big) \\ &= \mu'_{2,0} - \mu'_{0,2} + i2\mu'_{1,1} \end{aligned} \tag{9}$$

$$\begin{aligned} D_{2,0} &= L^2\Big(\frac{1}{\cos^2\beta}\hat{\mu}_{2,0} + \hat{\mu}_{0,2}\Big) \\ &= \mu'_{2,0} + \mu'_{0,2} \end{aligned} \tag{10}$$

where

$$\hat{\mu}_{2,0} = \mu_{2,0}\cos^2\gamma + \mu_{0,2}\sin^2\gamma + \mu_{1,1}\sin 2\gamma \tag{11}$$

$$\hat{\mu}_{0,2} = \mu_{2,0}\sin^2\gamma + \mu_{0,2}\cos^2\gamma - \mu_{1,1}\sin 2\gamma \tag{12}$$

$$\hat{\mu}_{1,1} = (\mu_{0,2} - \mu_{2,0})\frac{\sin 2\gamma}{2} + \mu_{1,1}\cos 2\gamma \tag{13}$$

D_{22} and D_{20} vary corresponding to the changes in distance L, slant angle β and tilt angle γ of the observed plane.

2.4 Solution of Location Parameters

The ratio D_{22}/D_{20} is a complex value and invariant to the distance L. The real part and imaginary part of D_{22}/D_{20} are shown below.

The real part of D_{22}/D_{20}

$$\frac{M^2\hat{\mu}_{2,0}\cos 2\gamma - 2M\hat{\mu}_{1,1}\sin 2\gamma - \hat{\mu}_{0,2}\cos 2\gamma}{M^2\hat{\mu}_{2,0} + \hat{\mu}_{0,2}} = \frac{\mu'_{2,0} - \mu'_{0,2}}{\mu'_{2,0} + \mu'_{0,2}} \quad (14)$$

The imaginary part of D_{22}/D_{20}

$$\frac{M^2\hat{\mu}_{2,0}\sin 2\gamma + 2M\hat{\mu}_{1,1}\cos 2\gamma - \hat{\mu}_{0,2}\sin 2\gamma}{M^2\hat{\mu}_{2,0} + \hat{\mu}_{0,2}} = \frac{2\mu'_{1,1}}{\mu'_{2,0} + \mu'_{0,2}} \quad (15)$$

where $M = 1/\cos\beta$.

The slant angle β and the tilt angle γ can be derived from (14) and (15). The distance L also can be obtained by substituting β and γ to (9) or (10).

2.5 Iteration

In principle, we can estimate the location parameters as metioned above. But some errors occur in estimation, because the visual field of an observed image is not coincident with that of a standard image. It is desired that the moments of both images should be calculated from the same image parts. Therefore, we propose the iterative method for improving parameter estimation by simulating the image with the estimated parameters from the standard image. Here, it is assumed that the center of a standard image is coincident with that of an observed one. First, we calculate the spectral moments of a standard image $\{\mu_{p,q}\}$ and those of an observed image $\{\mu'_{p,q}\}$, and estimate the location parameters, i.e., $\beta(1)$, $\gamma(1)$ and $L(1)$. Then we simulate the image with the estimated parameters from a standard image by interpolation. Second, we calculate the spectral moments of the simulated image $\{\bar{\mu}_{p,q}\}$ and transform them to those of the standard image $\{\mu_{p,q}\}$ by

$$\begin{pmatrix} \mu_{2,0} & \mu_{1,1} \\ \mu_{1,1} & \mu_{0,2} \end{pmatrix} = \mathcal{R}^{-1} \begin{pmatrix} \bar{\mu}_{2,0} & \bar{\mu}_{1,1} \\ \bar{\mu}_{1,1} & \bar{\mu}_{0,2} \end{pmatrix} (\mathcal{R}^T)^{-1} \quad (16)$$

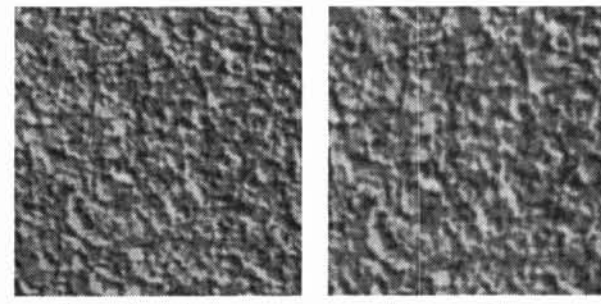
where T is a transposed matrix.

$$\mathcal{R} = L \cdot \mathbf{R}_1(\gamma)\mathbf{R}_2^{-1}(\beta)\mathbf{R}_1(-\gamma) \quad (17)$$

we also estimate the location parameters $\beta(2)$, $\gamma(2)$, $L(2)$ by using $\{\mu'_{p,q}\}$ and newly obtained $\{\mu_{p,q}\}$. When we iterate a series of estimation as metioned above, the visual field of the standard image approaches to that of the observed image, and the estimation error become smaller. As the criterion of estimation error, we use the square of power spectral difference between the observed image and the simulated image. In this experiment, we iterate this procedure at least four times and we decide the location parameters when we get the first minimum of the spectral difference. An exceptional case is that when the minimum occurs at the first or second iteration, we trace the second minimum and adopt the smaller one.

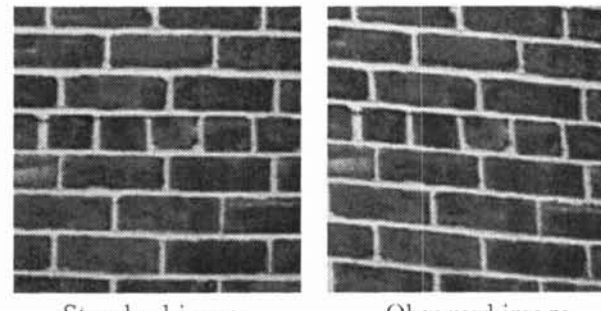
3 Experiments

We demonstrated the proposed method in simulations using the cork and brick images in Brodatz album [2]. All images are simulated by using Keys interpolation [5]. Power spectrum is obtained with Hanning window and FFT. The examples of standard and observed images are shown in Fig.2 and Fig.3.



Standard image (β=0°, L=1.0) Observed image (β=30°, γ=160°, L=0.8)

Fig.2 Cork images



Standard image (β=0°, L=1.0) Observed image (β=30°, γ=45°, L=0.95)

Fig.3 Brick images

In this case, the estimation results of cork were $\beta = 31.29°$, $\gamma = 166.60°$ and $L = 0.788$. Those of brick were $\beta = 29.90°$, $\gamma = 39.70°$ and $L = 0.963$.

We also demonstrated other cases and the estimation results are shown in Fig.4 (Cork) and Fig.5 (Brick). In these figers, a dot and an arrow indicate the average and standard deviation of the errors, respectively, where γ is set to 0, 5, 10, 15, —, 175. So an average and a standard deviation are calculated from 36 samples. The error of β and γ is absolute one and the error of L is relative one.

As shown in Fig.4(a) and Fig.5(a), the smaller the slant angle β is, the bigger the estimation error of β and γ are. Especially, we could not estimate the tilt angle γ in case of $\beta = 10$.

As mentioned above, when a textured plane slants, the image shrinks and the spectral moment expands along the slant direction by $\cos\beta$ times. If the slant angle β is small, the change of $\cos\beta$ is small in the order of β^2. For that reason, it is difficult to esitmate β and γ at small β.

As shown in Fig.4(b) and Fig.5(b), the estimation for $L = 1.0$ is the best and as L goes apart from 1.0,

the results get worse.

Generally, the estimaion errors of cork are smaller than brick, and we could get good result for distance L in all cases.

4 Conclusion

We have proposed the method for estimation of surface location of a textured plane. We showed our method is effective for textured planes, especially statistical texture such as cork.

We are also trying the case that an observed image not only shifts and slants but also rotated.

References

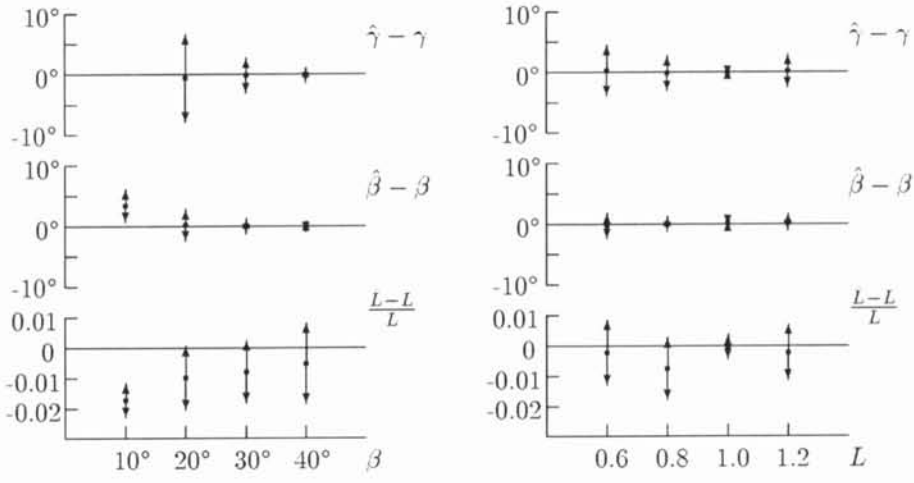
[1] Y.Wu and Y.Yoshida, " Detection of scaling factor and rotation angle of textured image based on spectral moments"IEICE Trans. D-II vol.J78-D-II, no.10, pp.1553-1559, Oct. 1995.

[2] P. Brodatz, " Textures, " Dover Pub. 1966.

[3] C.Teh, and R.T.Chin, " On image analysis by the methods of moments, " IEEE Trans. on Pattern Analysis and Machine Intelligence, vol.10, no.4, pp.496-531, July. 1988.

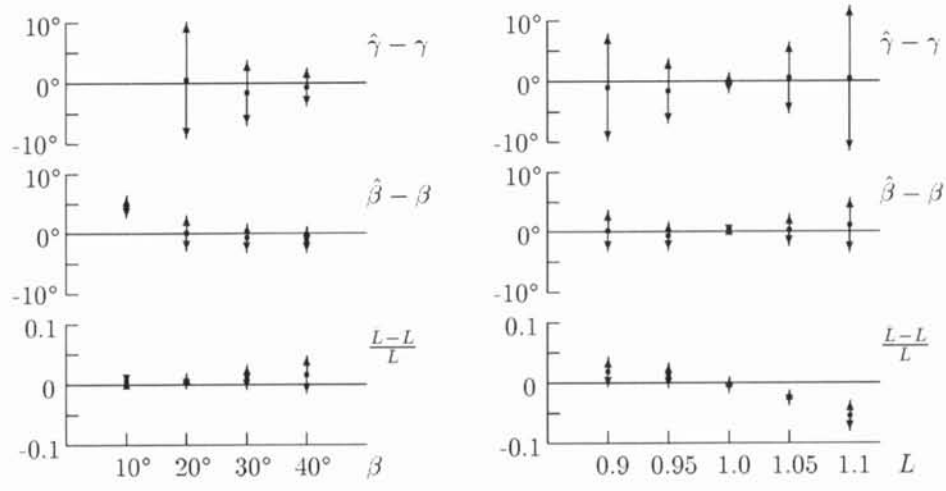
[4] Y.Yoshida, and Y.Wu, " Classification of rotation and scaled textured images using invariants based on spectral moments, " IEICE Trans. vol.E81-A, no.8, pp.1661-1666, Aug. 1998.

[5] R.G.Keys, "Cubic Convolution Interpolation for Digital Image Processing, " IEEE Trans. on Acoustics, Speech, and Signal Processing, vol.29, no.6, pp.1153-1160, Dec. 1981.



(a)distance fixed at $L = 0.8$ (b)slant angle fixed at $\beta = 30°$

Fig.4 Errors for estimation of distance L, slant angle β and tilt angle γ (cork image)



(a)distance fixed at $L = 0.95$ (b)slant angle fixed at $\beta = 30°$

Fig.5 Errors for estimation of distance L, slant angle β and tilt angle γ (brck image)