基於連續小波轉換之脊點一致性的複雜紋路表面方向計算*

RIDGE POINTS CONSENSUS OF CONTINUOUS WAVELET TRANSFORM FOR SHAPE FROM COMPLEX TEXTURE

呂俊賢†, 黃文良‡, 詹寶珠†

† 成功大學電機所
‡ 中央研究院資訊所

摘要

我們提出一個在透射投影模式下,利用連續小波轉換之脊平面來計算複雜紋路影像的表面方向.由於是處理複雜紋路,所得的資訊不完全可信.因此,我們使用robust regression和fitting技術,找出一組脊點之最佳子集來代表複雜紋路之支配頻率變化;從而求出表面方向.

關鍵字: 複雜紋路　連續小波轉換　脊平面

Abstract

A new method is proposed to estimate the surface orientations of planar complex textures under perspective projection through the ridge surface of 2D continuous wavelet transform (CWT). The ridge surface derived to be a parabola are used as the surface cue. A robust regression technique and a fitting technique are combined to choose the best subset of ridge points from a single ridge surface to represent the dominate frequency variations of complex textures. From the selected subset of ridge points, the tilt ans slant angles are computed. The performance of our method is demonstrated on several complex textures. The simplicity of this method in dealing with complex textures is emphasized.

Keywords: complex texture continuous wavelet transform ridge surface

1. INTRODUCTION

It had been pointed out that texture variations provide important cues in visual perception of three dimensional structures from a monocular image. The surface orientation of a planar texture are the slant angle, which determines the degree of obliqueness of the surface, and the tilt angle, which is the direction of slant. The main difficulty in solving the shape from texture problem is how to characterize the texture variations of perceived images. In the literature, many authors [1][4] used texture features such as the texels or edges as the representation of texture variations. However, it is well-known that the detection of texels and edges are not much reliable. The textural variations is better to be characterized by the space/frequency representation [9][15][16]. We chose the continuous wavelet transform tuned to various dilations and rotations to describe the projected images. The Morlet wavelet optimizing the spatial and the frequency resolutions simultaneously is adopted in this paper. In our previous papers [9][12], analytical solution had been proposed to estimate the planar surface orientation of a strongly ordered textur, i.e., regular texture. The dominate frequency variations of textured images are represented by the ridge surfaces of continuous wavelet transform. The ridge surfaces mark the places in the spatial frequency domain where the energy distribution (squared magnitude) of wavelet transform is mostly concentrated.

In this paper, a new method is proposed for complex textures (weakly ordered or disordered textures) described in Rao's book [14]. The strongly ordered textures may be approximately regarded as the image of single component, whereas the complex textures are considered to be composed of multiple components. Difficulties will arise when space/frequency representation is used to characterize complex textures. Linear transform such as short time Fourier transform (STFT) and wavelet transfrom (WT) or bilinear transfrom such as Wigner distribution (WD) are usually adopted for image analysis. For these linear transforms, the energy distributions (squared magnitude) of STFT and WT called spectrogram and scalogram are often used to represent images. However, these energy distributions shown to be nonlinearity [3] (similar to WD) will contain "cross terms" while analyzing multicomponent images. The cross terms of spectrogram and scalogram are found comparable with those of WD [10]. In our previous papers [9][12], it was shown that main information of strongly ordered textures are easy preserved along

*This work is supported by National Science Council under the Grant NSC 86-2745-E-001-006

dominate frequencies based on the scalogram of continuous wavelet transform. However, for complex textures, the ridge surfaces will be inadequate to describe the local frequency variations due to the energy distribution of wavelet transform is interfered with cross terms. Some techniques of cross terms reduction have been proposed [3][17] to improve the readability of space/frequency representations. Since our estimation of surface orientation is based on a parabola fitting, we focus on investigating the combination with a robust regression approach [8] to choose the dominate frequencies as the representatation of texture variations. Although the energies of the image pixels corresponding to these chosen frequencies are not exactly equal to the idealized auto-terms. Empirical experiments show that this choice is simple and powerful in estimating the surface orientation of a complex texture.

In the past, to our knowledge, only Kube [11] devoted to estimate the slant angle of random textures. He used orthographical projection model and assumed isotropic textures. Our method differs from his in that we use perspective projection model and no texture type is assumed. In addition, the used texture models are different: we used sinusoidal texture model whereas they used fractional Brownian plane. Furthermore, our method can estimate tilt angle while his method can not.

2. CHARACTERIZATION OF TEXTURES BY CONTINUOUS WAVELET TRANSFORM (CWT)

Let $\psi_{(\mathbf{b},s,\theta)}(\mathbf{x})$ be obtained by the translation, scaling, and rotation of $\psi(\mathbf{x})$: $\psi_{(\mathbf{b},s,\theta)}(\mathbf{x}) = \frac{1}{s^2}\psi(\mathbf{r}_{-\theta}\frac{(\mathbf{x}-\mathbf{b})}{s})$, where $\mathbf{b} \in \mathcal{R}^2$, $s > 0$, and $\theta \in [0, 2\pi)$ are the translation, scaling, and rotation parameters, respectively, and $\mathbf{r}_\theta$ is the rotation matrix of angle θ. In our implementation, we use the $2D$ Morlet wavelet, defined as $\psi_M(\mathbf{x}) = e^{j\mathbf{k_0}^T\mathbf{x}}e^{-|\mathbf{x}|^2/2}$ in the spatial domain, which is $\hat{\psi}_M(\mathbf{w}) = e^{-|\mathbf{w}-\mathbf{k_0}|^2/2}$ in the frequency domain, where $\mathbf{k_0}$ is the center frequency of the Morlet wavelet. Also, following the conventional usage [2], the scale parameter s takes discretized values with $s = 2^{o+\frac{v}{n}}$, where o is the octave, v is the voice, and n is the number of voices per octave. The Morlet wavelet optimizes both the spatial resolution and the frequency resolution simultaneously; therefore, it is well adapted to characterize the local spatial frequency.

We adopt the idealized or monochromatic texture model [5] with amplitude A and phase p constant:

$$g(\mathbf{x}) = A\cos(\Omega^T\mathbf{x} + p), \qquad (1)$$

where $\Omega^T\mathbf{x}$ denote frequencies. For an image $f(\mathbf{x})$ with N components, it is modeled as

$$f(\mathbf{x}) = \sum_{k=1}^{N} g_k(\mathbf{x}) = \sum_{k=1}^{N} A_k\cos(\Omega_k^T\mathbf{x} + p_k). \qquad (2)$$

The CWT of an image $f(\mathbf{x})$ in the frequency domain is

$$(\mathcal{W}f)(\mathbf{b},s,\theta) = \sum_{k=1}^{N}\frac{A_k}{2}\hat{\psi}_M(s\mathbf{r}_{-\theta}\Omega_k)e^{-j(\Omega_k^T\mathbf{b}+p_k)}.$$

Since $\hat{\psi}_M(\mathbf{w})$ is concentrated at the frequency $\mathbf{k_0}$, the k-th frequency component Ω_k of $f(\mathbf{x})$ will be concentrated around $\mathbf{r}_{\theta_k}\mathbf{k_0}/s_k$, where θ_k is the angle between $\mathbf{k_0}$ and $\Omega_\mathbf{k}$, and s_k^{-1} is the magnitude multiplier in order to scale $||\mathbf{k_0}||$ to $||\Omega_\mathbf{k}||$. In such case, the texture energy at $(\mathbf{b},s,\theta)$ is approximately the summation of the energy of each Auto-terms plus *Cross terms*:

$$|(\mathcal{W}f)(\mathbf{b},s,\theta)|^2 = \sum_{k=1}^{N}\frac{A_k^2}{4}\left|\hat{\psi}_M(s\mathbf{r}_{-\theta}\Omega_k)\right|^2 + \textit{Cross terms}, \qquad (3)$$

where the *Cross terms* defined as

$$\sum_{k,l=1,k\neq l}^{N}\frac{A_lA_k}{4}\hat{\psi}_M(s\mathbf{r}_{-\theta}\Omega_l)\hat{\psi}_M(s\mathbf{r}_{-\theta}\Omega_k)\cos((\Omega_l-\Omega_k)^T\mathbf{b}+p_l-p_k)$$

comes from interference between different frequency components. The *Cross terms* will not be completely absent provided that there are multiple frequency components in an image. However, they will be small as long as the wavelet $\hat{\psi}(\mathbf{w})$ is well localized in frequencies and the different frequency components are separated far enough such that

$$\forall l\ \forall k \neq l\ \forall s\ \forall\theta\ \ \hat{\psi}_M(s\mathbf{r}_{-\theta}\Omega_l)\hat{\psi}_M(s\mathbf{r}_{-\theta}\Omega_k) \approx 0.$$

In this case, the texture energy at $(\mathbf{b},s,\theta)$ is approximately the summation of the energy of each Auto-terms:

$$|(\mathcal{W}f)(\mathbf{b},s,\theta)|^2 \approx \sum_{k=1}^{N}\frac{A_k^2}{4}\left|\hat{\psi}_M(s\mathbf{r}_{-\theta}\Omega_k)\right|^2. \qquad (4)$$

Eqs. 3 and 4 indicate that the texture energy at $\mathbf{b}$ is concentrated around N different components centered at points $(\mathbf{b}, s = \frac{||\mathbf{k_0}||}{||\Omega_k||}, \theta_k)$ in CWT. They will be referred to as *ridge points* hereafter. The ridge points at the given $\mathbf{b}$ can be extracted by selecting the squared-modulus global maxima [12] or the squared-modulus local maxima [9] among the neighborhoods of θ and s at $\mathbf{b}$. One can then read off from these points about the dominate frequencies Ω_k. Another way to extract the local frequencies by phase stationary, we refer to [7]. In the following, we will explain why to use the single ridge surface.

Fig. 1(a) shows a strongly ordered texture with almost one component. Its single ridge and multiridge are shown in Figs. 1(b) and (c), respectively. We found that the multiridge without small fragmentations is prefect than the single ridge. This is because the single ridge contains some image patches where the small amounts of pixels usually come from local structures of textures, whereas the multiridge delete these patches by a simple "continuity" criterion. It should be emphaseized that the detected multiple ridges is *continuous*. This is also why the results of [9] based on multiple ridges are better than those of [12] based on a single ridge, both for strongly ordered texture. Also, almost all the existing methods could work well for strongly ordered textures. However, in practice, there are often complex textures (weakly ordered and disordered texture) [14] encountered in our world. Complex textures are always composed of multiple frequency components. Even the space/frequency representations [9][15][16] have been widely used to characterize textures, the overall space-frequency or space-scale represetnations are not clear to describe complex texture variations. An example of a complex texture and its detected single ridge and multiridge are shown in Figs. 1(d), (e), and (f), respectively. Note that only one of the multiple ridges having the maximal number of pixels are shown. It is observed that neither the single ridge nor the multiridge (base on only simple "continuity" criterion) is adequate to represent local frequency variations of the complex texture if no extra process is made for the detected ridges. It is the goal of this paper to investigate how to select the useful subset of data from the ridge surfaces that efficiently describe the local frequency variations of complex textures by incorporating a robust regression technique. Because the multiple ridges may be obtained by arbitrarily connecting ridge points, it is thus suitable to use single ridge as the surface cue.

Dut to the limit of space, the detail description of estimating surface orientation can be found in [9][12].

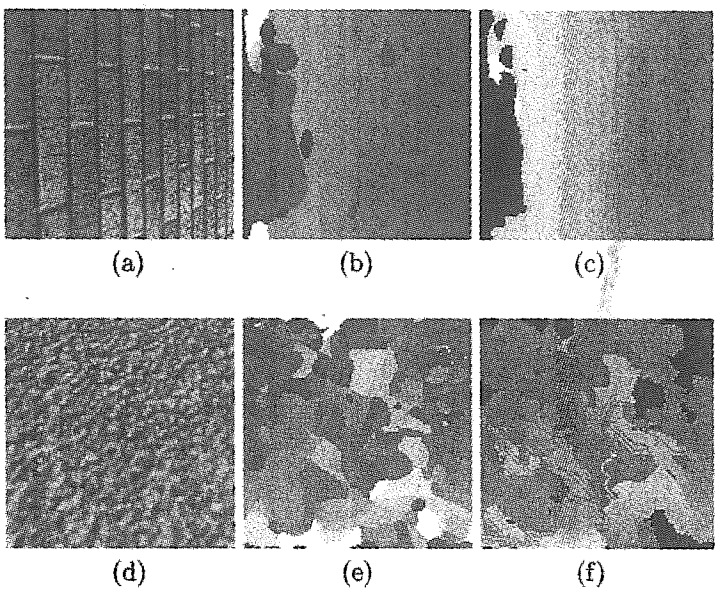

Figure 1: Single ridges (b) and (e), and multiridge ridges (c) and (f) of a strongly ordered texture (a) and a complex texture (d).

3. SHAPE FROM COMPLEX TEXTURE BY RIDGE POINTS CONSENSUS

3.1. Observations of Frequency Variations of Various Textures

Fig. 2(a) shows an image of sinusoid containing only one frequency. Fig. 3(a) shows the modulus of the wavelet transform at various rotations and scales at a given location. It is found that the energies appear in one orientation and concentrated on a narrow band of frequency. The peak frequency of this location is fairly apparent. It is enough to represent the dominate frequency at this location by the peak frequency.

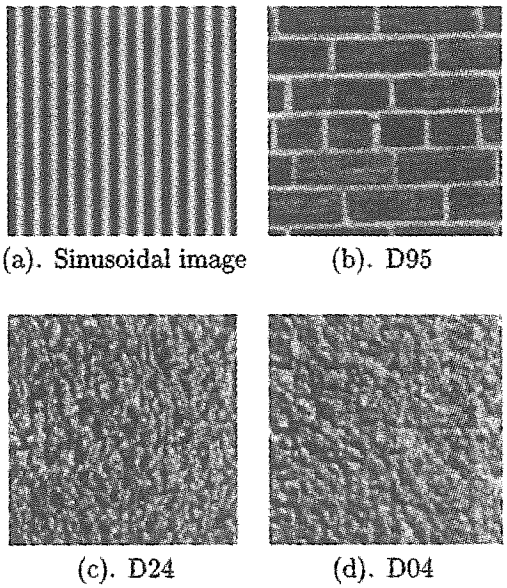

Figure 2: Various structured images: (a) Sinusoid image; (b) Strongly ordered texture; (c) Weakly ordered texture; (d) Disordered texture.

Unfortunately, there is few image like sinusoid image encountered in the real-world. According to Rao's book [14], the textured images can be roughly categorized into three classes: strongly ordered, weakly ordered, and disordered textures. All these textures can be regarded as the images composed of multiple components. Fig. 2(b) shows an strongly ordered texture D95. Similarly, the modulus responses with respect to scale at one location are shown in Fig. 3(b). We see that the energy appears more than one orientations and concentrated on some frequency bands. But, it is still easily to choose peak frequencies as the "signature" of the local texture variations. Moreover, the neighboring pixels in regular textures usually have similar frequency responses and, thus, the multiple ridge algorithm gives us good continuous ridge surfaces which tend to separate different frequency components. However, for the weakly ordered and disordered textures shown in Figs. 2(c) and (d), the spatial frequency distributions become much more complex and the neighboring pixels tends to have different frequency distributions in each orientation (see Figs. 3(c) and (d)). In such cases, the multiridge detection algorithm is unable to determine "good" continuous ridge surfaces. Recall that the ridge points of single ridge have the maximal response in the CWT domain. The most important information, even partial, are still preserved. Also, our estimation method is based on a fitting method [9][12]. Using the least-squares technique to fit a parabola from the ridge points will not satisfy for complex textures due to the spurious ridge points resulting from various frequency components. It is therefore important to choose the partial ridge points from a single ridge to characterize the dominate frequency variations.

3.2. Selection of Useful Ridge Points

Because a ridge point is represented with a rotation and a scale parameters in addition to the position parameter [9][12]. We found that two constraints about the consistency of rotations and the small fitting errors have to be imposed.

Constraint 1 (Consistency of orientation): When a perceived textured image is slanted, the representative orientation of the selected ridge points should be consistent. The ridge points violate the consistency are rejected as outliers from the ridge surface.

Constraint 2 (Small fitting errors): Owing to our estimation of surface orientation is based on a fitting method, the best subset of ridge points should fit a parabola within a certain error. In order to satisfy this constraint, a robust regression technique (RANSAC) [8], is appropriate because they have the advantage of finding a subset of points that best sat-

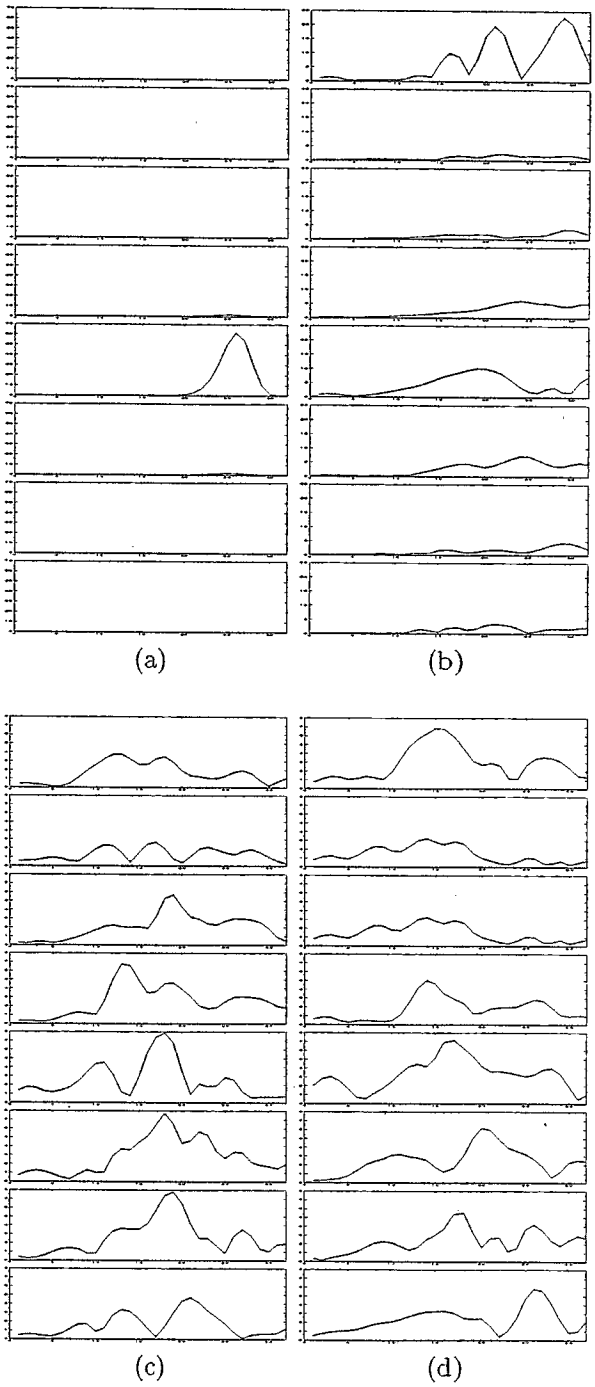

Figure 3: The modulus of wavelet transform with respect to scale at various orientations of a location: (a) Sinusoidal image, (b) D95, (c) D24, and (d) D04. The scale increases from the left to the right horizontally, modulus is indicated vertically; and the subplot boxs show the orientations from $-90°$, $-67.5°$, ..., $67.5°$(top to bottom)

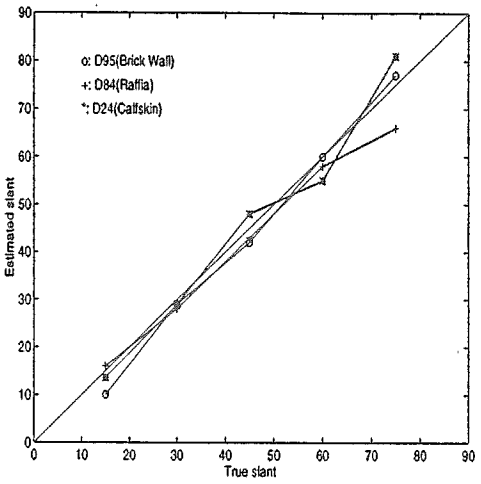

Figure 4: Estimations for three natural textures D95, D84, and D24 slanted 15°, 30°, 45°, 60°, and 75°.

isfy some constraints and reject the rest of points as outliers. The selection of useful ridge points by RANSAC can be found in [13].

4. EXPERIMENTAL RESULTS

The complex texturs used in our experiments are mainly synthesized from Brodatz's album [6] and obtained from MIT VisTex database. We first perform a similar experiment as Turner's [16] by our ridge points consensus method. The estimated slant angles are shown in Fig. 4. It is found that estimated slant angles are not deviated from the true slant angles too much even the slant angle is up to 75°. However, Turner's results go down when the slant angle is large and the textures become complex (See Fig. 8 in [16]).

In addition, another complex textures used for experiments are shown in Fig. 5(a1)-(l1). It is seen that the regular structures of these images are not apparent. The estimated results can be found in [13]. The average error of tilt is within 4.46° and those of slant angles using voting and curve fitting are within 2.56° and 3.89°, respectively. Subfigures 2 and 3 of Fig. 5 show some of the single ridges and the inliers extracted from subfigures 1. Comparing each pair of the single ridge and the inliers, it is observed that the preserved information reveal the tendency of frequency variations to some extent.

From the experimental results, we found that the average errors for complex textures are larger than those of strongly ordered textures [9][12]. But they are still lower than 5° averagely. The most importance is that the estimation of surface orientation for complex textures are become possible by the new yet simple method.

5. CONCLUSIONS

We have proposed a new methods for estimating the surface orientations of complex textures under the perspective projection model. The dominate frrequency variations of complex texture are effectively characterized by the ridge points consensus in the CWT domain. The performance of our method had been demonstrated on several complex images. It is emphasized that the proposed method is simple and powerful in dealing with complex textures without using any techniques of reducing interference terms. Compared with our previous results for strongly ordered textures [9], some results for complex textures are inferior. But, this paper has at least made the estimation of surface orientation for complex textures possible.

6. REFERENCES

[1] J. Aloimonos, "Shape from texture", *Biological Cybernetics*, Vol. 58, 1988, pp. 345-360.

[2] J. -P. Antoine, P. Carrette, R. Murenzi, and B. Piette, "Image analysis with two-dimensional continuous wavelet transform", *Signal Processing*, Vol. 31, 1993, pp. 241-272.

[3] F. Auger and P. Flandrin, "Improving the readability of time-frequency and time-scale representations by the reassignment method", *IEEE Trans. Signal Processing*, Vol. 43, 1995, pp. 1068-1089.

[4] D. Blostein and N. Ahuja, "Shape from texture: Integrating surface element extraction and surface estimation", *IEEE Trans . Pattern Anal. and Machine Intell.*, Vol. 11, 1989, pp. 1233-1251.

[5] A. C. Bovik, "Analysis of multichannel narrowband filters for image texture segmentation", *IEEE Trans. Signal Processing*, Vol. 39, 1991, pp. 2025-2043.

[6] P. Brodatz, "Textures: A photographic album for artists and designers", *Dover Publications*, 1966.

[7] N. Delprat, B. Escudie, P. Guillemain, R. Kronland-Martinet, P. Tchamitchian, and B. Torrésani, "Asymptotic wavelet and Gabor analysis: Extraction of instantaneous frequencies", *IEEE Trans. Inform. Theory*, Vol. 38, No. 2, 1992, pp. 644-664.

[8] M. A. Fischler and R. C. Bolles, "Random Sample Consensus: A Paradiam for Model Fitting with Applications to Images Analysis and

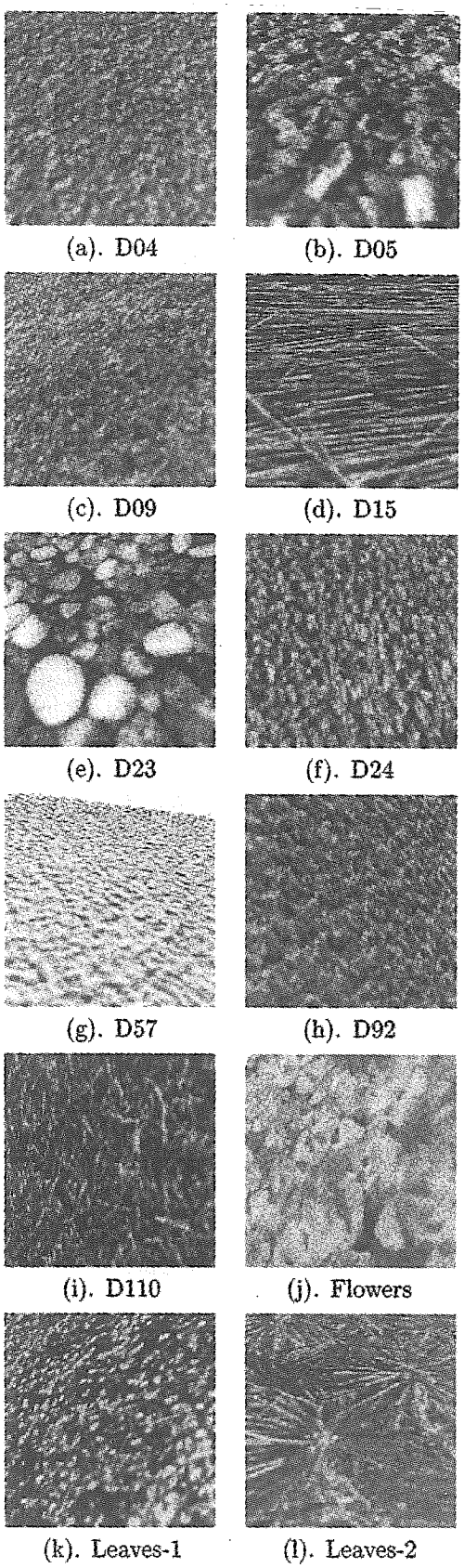

Figure 5: Complex textures: (a)-(i) Synthesized Brodatz's textures; (j)-(l) *MIT VisTex* textures.

Automated Cartography", *Communications of ACM*, Vol. 24, 1981, pp. 381-395.

[9] W. L. Hwang, C. S. Lu and P. C. Chung, "Shape From Texture: Direct Estimation of Planar Surface Orientation Using Continuous Wavelet Transform", *To appear in IEEE Trans. Image processing*, 1997.

[10] S. Kadambe and G. Faye Boudreaux-Bartels, "A comparison of the existence of cross terms in the Wigner distribution and the squared magnitude of the wavelet transform and the short time Fourier transform", *IEEE Trans. Signal Processing*, Vol. 40, 1992, pp. 2498-2517.

[11] P. Kube, "Using Frequency and Orientation Tuned Channels to Determine Surface Slant", *Eight Annual Conference of the Cognitive Society, Amherst, Massachusetts*, 1986, pp. 235-244.

[12] C. S. Lu, W. L. Hwang, H. Y. Mark Liao, and P. C. Chung, "Shape from texture based on the ridge of continuous wavelet transform", *IEEE Conf. Image processing*, Vol. I, 1996, pp. 295-298.

[13] C. S. Lu, W. L. Hwang, and P. C. Chung, "Toward A General Shape From Texture system:", *IPPR Conf. CVGIP*, 1997, pp. 122-129, Taiwan, ROC.

[14] A. Ravishankar Rao, "A Taxonomy for Texture Description and Indetification", *Springer Verlag*, 1990.

[15] B. Super and A. C. Bovik, "Planar surface orientation from texture spatial frequencies", *Pattern Recognition*, Vol. 28, No. 5, 1995, pp. 729-743.

[16] M. R. Turner, G. L. Gerstein, and R. Bajcsy, "Underestimation of visual texture slant by human observers: a model", *Biological Cybernetics*, Vol. 65, 1991, pp. 215-226.

[17] W. J. Williams, "Reduced Interference Distributions: Biological applications and Interpretations", *Proc. of the IEEE*, Vol. 84, 1996, pp. 1264-1280.